Quarterly Progress Reports To iDigBio Submitted By Active Thematic Collections Networks (TCNs)

May 2021

CONTENTS:

- ~~Google Analytics across ADBC~~
- Reports from the following **active** TCNs:

☒ CAP	☒ GLOBAL	☒ PILSBRY
☒ Cretaceous World	☒ LepNet & SCAN	☐ SoRo
☒ DigIn	☒ MiCC	☒ TORCH
☒ Endless Forms	☐ oVert	☒ TPT
☐ EPICC	☒ PCC	
☒ ESB		

- Reports from the following **retired** TCNs are no longer included:

~~FIC~~	~~MaCC~~	~~SERNEC~~
~~GLI~~	~~MAM~~	~~TTD~~
~~InvertNet~~	~~MHC~~	~~VACS~~
~~InvertEBase~~	~~NEVP~~	
~~LBCC~~	~~Paleoniches~~	

iDigBio is funded by a grant from the National Science Foundation's Advancing Digitization of Biodiversity Collections Program (Cooperative Agreement EF-1115210). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

CALIFORNIA PHENOLOGY TCN – QUARTERLY REPORT – APRIL 2021

Assembled by Katie Pearson, 29 April 2021

PROGRESS IN DIGITIZATION EFFORTS

Figure 1 shows our progress in imaging, transcribing, georeferencing, and phenologically scoring the target specimens for the original 22 CAP institutions, explained more in detail in the following sections.

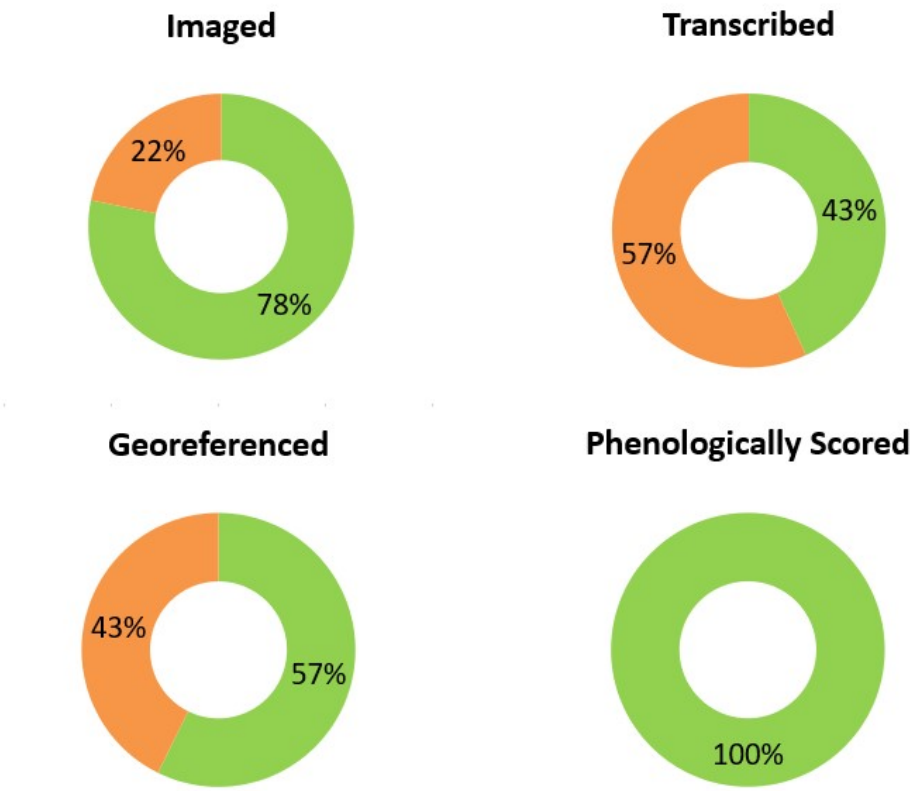


Figure 1. Progress in each of our four major digitization goals. Totals represent the original goals of the CAP grant: 902,400 specimens imaged and phenologically scored, and 300,000 transcribed and georeferenced. Additional specimens to be digitized by the PEN grant are tabulated in the PEN section below.

TRANSCRIPTION

An estimated 129,000 specimen records have been transcribed across the CAP Network since the beginning of the project. This is approximately 43% of our goal.

Transcription has largely been accomplished by institutional volunteers in CCH2 and online volunteers in Notes from Nature, including those who attended WeDigBio events (see Education & Outreach).

GEOREFERENCING

We have georeferenced 171,890 specimen records, which is 57% of our goal. Georeferencing is conducted by trained staff and students at HSC, OBI, and SD, by naturalist volunteers are part of the "100 Club," and by undergraduate students in the cross-institution herbarium digitization course led by Cal Poly. The CAP 100 Club currently has 28 active members who have collectively georeferenced over 8,000 specimens since September 2020. The digitization course consists of 18 undergraduate students who have georeferenced over 260 records since March 30, 2021.

We recognized that thousands of specimens had township, range, section (TRS) data in the verbatim coordinates field of CCH2, yet these same records did not have latitude and longitude values. We created an R script to convert these TRS data into coordinates with uncertainty radii of 969 m (1 mi^2), using a conversion table provided by Nelson Rios at Yale. With this code, we were able to add georeferences to 35,465 specimens in CCH2. The code is publicly available on Zenodo (Pearson 2021, http://doi.org/10.5281/zenodo.4507032).

IMAGING

Eleven of our 22 herbaria (50%) have accomplished their imaging goals (Figure 2). Of the remaining herbaria, eight have been able to resume imaging since the COVID-19 shutdowns. The other herbaria have used this time offsite to process images, georeference specimens, and transcribe specimens. Figure 2 shows the current state of CAP imaging as of April 29, 2021.

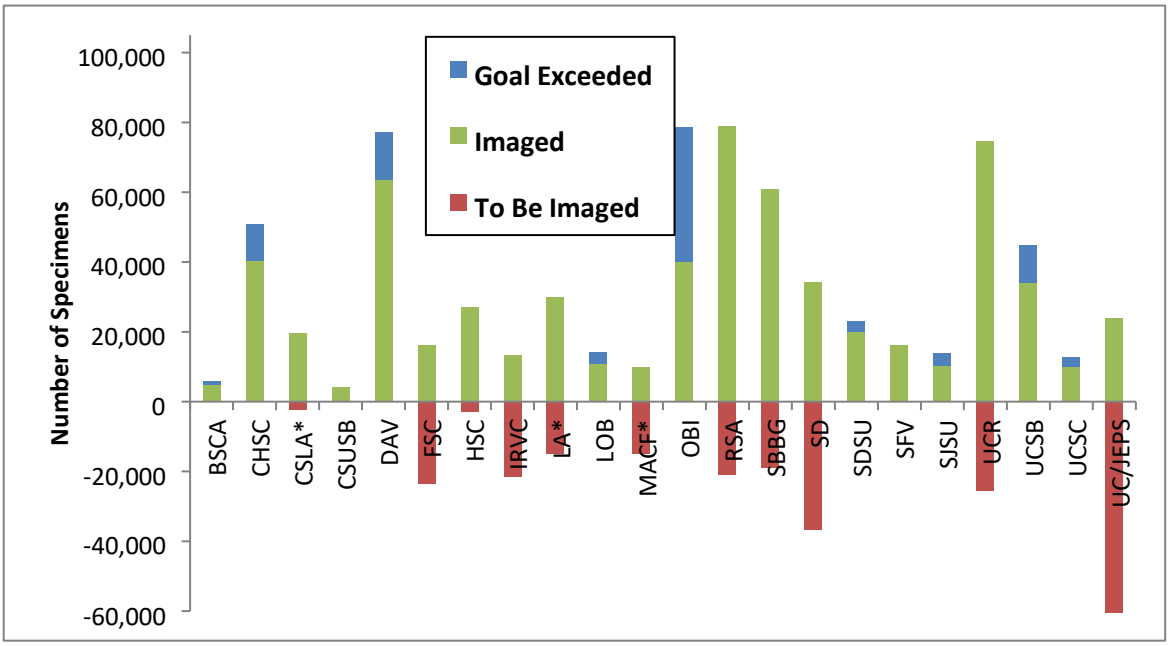


Figure 2. Herbarium specimen imaging progress. Green portions represent the number of specimens that have been imaged, while blue portions represent the number of specimens that have been imaged beyond the expected target specimens. Red bars below the zero line indicate the number of target specimens that have not yet been imaged. Asterisks indicate herbaria that are currently unable to image due to COVID-19-related prohibitions.

PEN PROGRESS

Specimen data from PUA were successfully extracted from an RBase database, coalesced by the PM, and are now managed live in CCH2. Of the 71,333 specimens, 70,163 (98%) are already georeferenced and none are imaged. We used the trait coding from text tool to phenologically score 55,633 (78%) of the specimens, which brings our total number of phenologically scored specimens to 1,430,615 as of April 29, 2021, which is 136% of our goal even with the additional PEN institutions.

Imaging has begun at SFSU, albeit with limited hours due to COVID-19 precautions. Imaging is ongoing at UNLV and OSC, which have achieved 50% and 22% of their imaging goals, respectively. CDA was approved as a subcontract as part of the NSF award as of February, but funds are still being handled by their respective offices and have not yet been liberated for purchasing equipment. CDA will begin imaging once they have acquired their equipment. The remaining two PEN institutions will begin imaging later in 2021 (SHTC) and 2022 (PUA), as previously scheduled. Figure 3 shows the current imaging progress at PEN institutions.

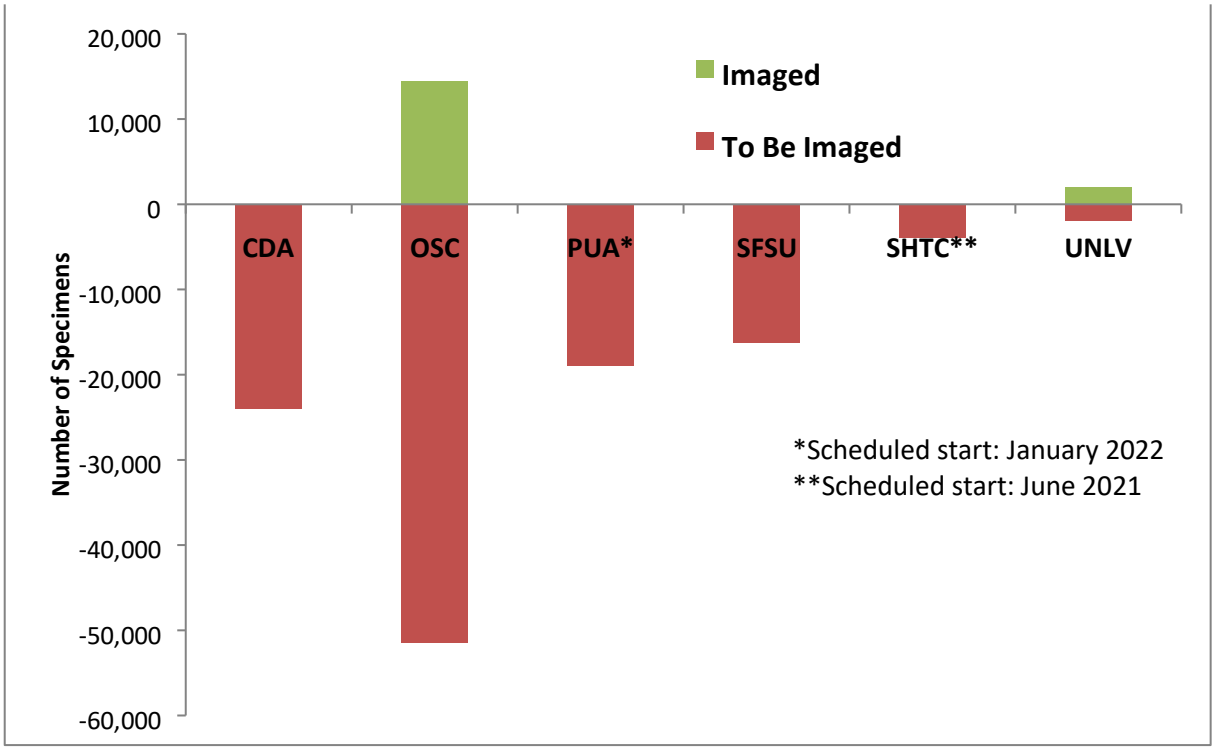
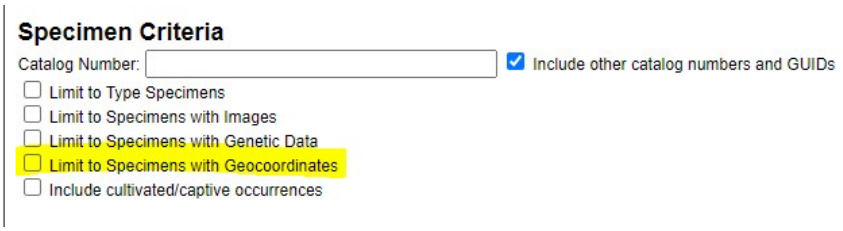


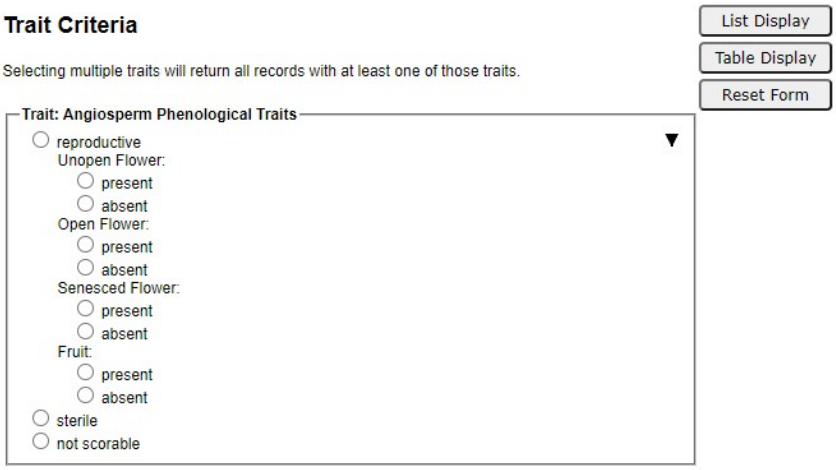
Figure 3. Herbarium specimen imaging progress for the seven PEN institutions. Green portions represent the number of specimens that have been imaged, while red bars below the zero line indicate the number of target specimens that have not yet been imaged. For SD, the total number of specimens to be imaged, including those added as part of the PEN grant, is indicated in Figure 2; therefore, SD is not included in this figure.

PORTAL DEVELOPMENT

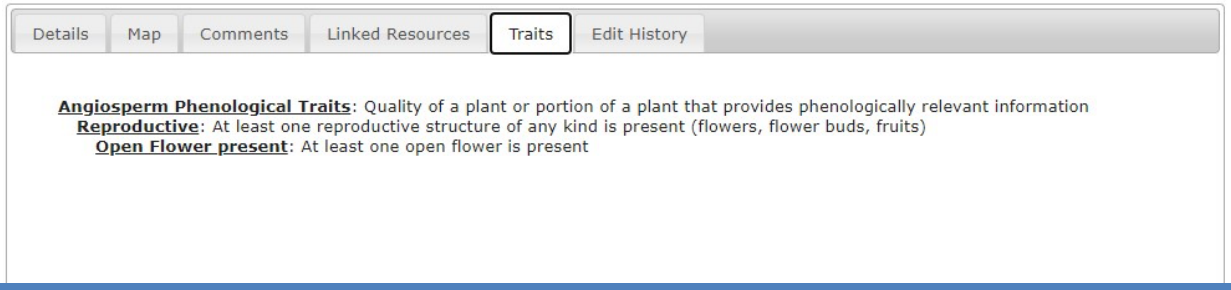
With the help of Chris Tyrrell (Milwaukee Public Museum), we are continuing to build new functionality to the CCH2 portal that can be integrated into other Symbiota portals. Chris developed a way to search the portal by specimens that are georeferenced, which is now functional in our portal (see screenshot below).



Chris also developed a way to search specimens by trait criteria, specifically by phenological traits in our portal. This functionality is versatile; it can be used by any portal to search for any trait that the portal has enabled. This tool will be added to the CCH2 portal once it has been fully vetted by Ed Gilbert, Symbiota developer (see screenshot below).



We pushed forward the production of a traits tab that can be viewed on the public page of any specimen. The format of this tab will undergo future development, but it currently displays all traits associated with an occurrence (see screenshot below).



SHARE AND IDENTIFY BEST PRACTICES AND STANDARDS (INCLUDING LESSONS LEARNED)

As previously described, we created an R script to convert these TRS data into coordinates to the precision of 969 m (1 mi^2) uncertainty, using a conversion table provided by Nelson Rios at Yale. With this code, we were able to add georeferences to 35,465 specimens in CCH2. The code is publicly available on Zenodo (Pearson 2021, http://doi.org/10.5281/zenodo.4507032).

We have continued to hone the use of the code to extract georeference data from botanical duplicates and have created a new version with documentation, again publicly available on Zenodo (Pearson 2021, https://zenodo.org/record/4637000).

OBI beta tested a process for quickly acquiring georeferences for specimens from prolific collectors. We downloaded all specimens collected by Dave Keil (>20,000 specimens) from the OBI collection and sorted the spreadsheet according to collector number. We were then able to apply coordinates from georeferenced specimens with similar locality strings to specimens that previously lacked coordinates. We also applied the metadata associated with these coordinates to the newly georeferenced specimens. OBI georeferenced 2,273 records using this technique. We plan to apply a similar technique to prolific collectors of other herbaria.

We are working with the California botanical community at large to develop a CCH policy on protecting specimen data for sensitive taxa (e.g., threatened species). As part of these efforts, we have solicited statements from the botanical community about their opinions on this matter, i.e., whether specimen data should be redacted for some or any taxa, and collated their responses here (https://www.capturingcaliforniasflowers.org/sensitive-taxa.html). In May, we will hold a community informational meeting in which we hear statements from key stakeholders and present currently available best practices documentation (e.g., the GBIF Best Practices for Generalizing Sensitive Species Occurrence Data).

IDENTIFY GAPS IN DIGITIZATION AREAS AND TECHNOLOGY

Feedback from members of our georeferencing "100 Club" has indicated that the batch georeferencing tools could use more development to be more efficient.

We are developing a way to search for specimens based on taxon-level traits, such as CNPS rarity ranking.

SHARE AND IDENTIFY OPPORTUNITIES TO ENHANCE TRAINING EFFORTS

In March 2021, the ASU Symbiota team developed a new label-making tool for Symbiota portals. We created a video tutorial and written guide for using this new tool, available here (https://youtu.be/ult9QAWax9I) and here (https://www.capturingcaliforniasflowers.org/uploads/1/6/3/7/16372936/newlabelprintingguide.docx), respectively. We also updated the Guide to Using a Symbiota-based Portal with these and other new portal adjustments (https://www.capturingcaliforniasflowers.org/symbiota.html).
We have continued to support the work of our "100 Club" of naturalist georeferencers through email and by sending them promotional "I Digitize California Plants" stickers. On April 8, we held a 100 Club virtual co-working session in which we georeferenced and socialized. We will hold similar events monthly to promote member engagement.

During the winter 2021 quarter, we led an online digitization course in which 30 students from 9 institutions learned how to transcribe specimens and work in CCH2. We are leading a similar course for the spring 2021 quarter in which many of the same students have been trained to georeference specimens in the GEOLocate Collaborative Georeferencing (CoGe) platform. This course meets synchronously via Zoom, once per week for 2

hours. As part of the course, the students also give a short presentation on an herbarium-related topic. The 18 students in this course have georeferenced 265 specimens as of 29 April 2021.

In preparation for the spring 2021 version of this course, we developed an alternative georeferencing training course that includes a CoGe module (https://www.capturingcaliforniasflowers.org/georeferencingcourse-coge.html) featuring a new "Georeferencing in GEOLocate CoGe" video (https://youtu.be/h1JfJuSC-eg).

Oscar Vargas, herbarium director at digitizing institution HSC, is leading a class of students in georeferencing in the CCH2 portal. The students were trained using the online georeferencing course followed by a hands-on session with the PM.

SHARE AND IDENTIFY COLLABORATIONS WITH OTHER TCNS, INSTITUTIONS, AND ORGANIZATIONS

PM Pearson has been able to apply the georeference duplicate searching code to institutions of the GLOBAL TCN as part of her duties in that role. This has led to some of the improvements in the georeferencing code described previously in this report.

After finding our digitization documentation on our website, digitization managers at North Dakota State University and Polly Hill Arboretum independently contacted us about each acquiring one of the custom camera mounts that we developed to fit the Ortery Photosimile 50 lightbox. We manufactured these mounts and provided them to NDSU and Polly Hill at cost.

We have continued to increase the quality of our data, especially georeferences, by examining records that had previously been flagged by Calflora. California CCH2 data will soon be ingested by Calflora and therefore be searchable on their site.

SHARE AND IDENTIFY OPPORTUNITIES AND STRATEGIES FOR SUSTAINABILITY:

On February 16th, CAP held a meeting of PIs to discuss how to engage undergraduates more consistently in herbarium-related activities. The goal is not only to enhance education and student involvement in herbaria, but also increase the visibility and value of the herbaria to administrators and colleagues.

As we near the last year of CAP funding, CAP leadership has begun to plan for the sustainability of tasks currently completed by the PM beyond 2022. We are drafting necessary documentation to enable the continuation of certain tasks (e.g., Notes from Nature expeditions) and planning training opportunities for PIs and collaborators. One of our primary strategies is to build up the Consortium of California Herbaria as an organized, democratic society with regular meetings and task groups.

SHARE AND IDENTIFY EDUCATION AND OUTREACH (E&O) ACTIVITIES:

The PM shares updates on the project and phenology-related news via the network Twitter account (@CalPhenologyTCN).

Three blog posts were written and published to the CAP website: https://www.capturingcaliforniasflowers.org/blog-recap. Blog posts are publicized via Twitter and the "Herbarium Junkies" Facebook page.

On February 16th, CAP held a meeting of PIs to discuss how to incorporate herbarium-related activities into undergraduate courses or create herbarium-focused independent study and research courses. PIs shared their current approaches, and potential improvements and other approaches were discussed as a group. Participants agreed to share their developed educational materials via a Google Drive folder.

As previously described, we led an online digitization course in which 30 students from 9 institutions learned how to transcribe specimens and work in CCH2. In the spring quarter, we are leading a similar course which many of the same students (and some new students) have learned to georeference specimens in the GEOLocate Collaborative Georeferencing (CoGe) platform. Currently, 16 students and 2 volunteers from seven institutions are participating in this course.

For winter quarter 2021, PI Yost led a phenological research course at Cal Poly similar to our spring 2020 course, assisted by the new CAP-funded postdoctoral student (Natalie Love), a graduate student at UCSB (Tadeo Ramirez Parada), and the PM. Ten students completed this course, including analyzing changes in flowering time with climate, writing a research paper, and presenting a research poster virtually. PI Mazer (UCSB) is now running this course with 15 students at UCSB for the spring 2021 quarter.

The UC Riverside Botanical Garden Newsletter published an article about an indispensable herbarium volunteer at the UC Riverside herbarium, thereby publicizing UCR's digitization work (https://drive.google.com/file/d/1BzkomSBifYfsmtkIf-tPsXrKUyhg9j5j/view).

The CAP Network participated in the spring WeDigBio event by leading two virtual events on April 8th ("Herbarium Happy Hour") and April 10th ("Botany Brunch"). These events each included a presentation on the use of herbaria for research, education, and conservation and a virtual tour of the Robert F. Hoover Herbarium at California Polytechnic State University. Participants then learned to transcribe specimen records in Notes from Nature and played transcription games. The Thursday event engaged 55 participants, and the Saturday event engaged 42 participants. Over the WeDigBio weekend (April 811th), volunteers transcribed 1,149 specimens.

Four Notes from Nature expeditions were completed during this quarter involving 8,642 specimens from UC Los Angeles, Humboldt State, Cal State LA, and Fresno State. The data were cleaned and imported into CCH2. We launched three new expeditions in February and March 2021 consisting of 5,943 specimens from Cal Poly, Cal State LA, and UC Los Angeles.

TCN Quarterly Progress Report

Prior to each IAC meeting, TCNs are asked to complete a quarterly progress report in the areas outlined below. The TCN Lead PI or Project Manager collects information from all collaborators and compiles them into one overall progress report for the TCN. The TCN Lead PI or Project Manager then submits the quarterly reports via an email to Cat Chapman. An archive of previously submitted reports is available on the Internal Advisory Committee wiki page.

TCN Name

Cretaceous World

Person Completing the Report

Name and role of the person completing the report: Bruce S. Lieberman on behalf of Chris Beard

Share Progress in Digitization Efforts

Regarding the **University of Kansas** portion of the project, formerly led by PI Bruce S. Lieberman (BSL) and now led by Chris Beard, and with major involvement from collections manager Natalia Lopez Carranza (NLC), we have databased 136,899 fossil specimens total, 423 new specimens since the last reporting period. 111,156 of these specimen records are also georeferenced. In addition, we have georeferenced a total of ~ 9,989 localities associated with this project. This work has been performed by NLC. All KU digital data have been shared with iDigBio and GBIF.

One of the tasks that NLC has also been focusing on in conjunction with graduate student Rene Martin (RM) is adding new archival numbers to specimens, and thus far a total of 3,927 specimens have been numbered, 2,650 new specimens since the last reporting period.

Regarding the **University of Colorado** portion of the project, led by PI Talia Karim:

They have databased 20,827 Cretaceous specimens total, with no new specimens databased since the last reporting period. 14,906 of these specimen records are also

georeferenced. In addition, they have georeferenced 316 localities since the last reporting period and have now georeferenced (and in entered in Specify) a total of 1,294 Cretaceous localities associated with this project. They georeferenced an additional 140 localities during the reporting period that are in the process of being entered into Specify. They generated no new images.

Two digitization assistants have been working remotely on verbatim locality transcription, georeferencing, and Specify data entry during the reporting period. This work has led to the creation of over 80 new localities during the reporting period and the transcription of verbatim locality data for several hundred more localities, all of which has been entered into Specify or is in the process of being entered into Specify. Written workflows for these tasks have been updated with the input of the digitization assistants.

Regarding the **University of Texas** portion of the project, led by Lisa Boucher and Rowan Martindale with major participation from Liath Appleton and Matthew Brown

Digital record updates for iDigBio:

They have 24,530 Cretaceous cataloged records, representing ~80,000 specimens total. Of the total number of specimen records, 24,194 have been georeferenced.

In addition, they have now georeferenced a total of 5,822 Cretaceous localities (out of 5,897 associated with this project).

The total number of Specify database images generated is 11,359.

Due to COVID-19, their imaging and georeferencing progress has slowed as the collection is only open to a limited extent. Students and volunteers have been contributing remotely, but their capacity to take new images is limited. Some of the remaining Cretaceous localities to georeference are challenging and those that can be resolved require access to physical records.

Share Best Practices, Standards, and Lessons Learned

N/A

Share Identified Gaps in Digitization Areas and Technology

N/A

Share Opportunities to Enhance Training Efforts

Regarding the **University of Kansas** portion of the project, we will be supporting two additional graduate students who will be work this summer to digitize our invertebrate paleontology collections: one is of Afro-Caribbean descent, the other is female. They will join female graduate student Rene Martin. Another graduate student will be supported to digitize our vertebrate paleontology collections.

Share Collaborations with other TCNs, Institutions, and/or Organizations

N/A

Share Opportunities and Strategies for Sustainability

N/A

Share Education, Outreach, Diversity, & Inclusion (EODI) Activities

Regarding the **University of Kansas** portion of the project, one of our important areas of outreach has involved the updated second edition of the *Digital Atlas of Ancient Life app* available for iOS. It has received a total of more than 9,970 downloads, with 1,528 downloaded in February 2021, a new record, and more than 550 downloaded in March 2021.

Another key outreach development was creating a new version of the *Digital Atlas of Ancient Life app* available for Android; it has now received a total of 4,882 downloads.

Lastly, the *Digital Atlas Identify app*, available for iOS, has received a total of more than 3,520 downloads, with more than 950 new downloads in February and March 2021. In addition, it has been receiving good and increasing usage, with an average of about 100 images of fossil specimens submitted for identification each week. BSL provides identifications for the images which are then transmitted to users via the app.

Please also see the next section on our *Digital Atlas of Ancient Life* website, which plays a key role in our education and outreach activities.

Share Information About Your Website and/or Portal Usage

Regarding the **University of Kansas** portion of the project, the *Digital Atlas of Ancient Life* website continues to receive major usage. Note that the development of this website and its continued maintenance, the addition of content, etc., has been led by Jonathan Hendricks at the Paleontological Research Institution. During the period February - March the site has a monthly average of more than 30,100 clicks and 3.54 million impressions, while the *Cretaceous Atlas* itself has a monthly average of more than 1,250 clicks and 66,400 impressions. In both cases the primary users were based in the U.S.A. The former also has high numbers of users based in India and the Philippines. The latter also has high numbers of users based in Canada and the UK.

Share Other Activities and/or Progress

Regarding the **University of Kansas** portion of the project, one paper describing results from research was published:

Byrum, S., and **B. S. Lieberman**. 2021. Phylogeny and biogeography of Cretaceous spatangoid echinoids with special emphasis on taxa from the Western Interior Seaway. *Journal of Paleontology* 95, 613-623 https://doi:10.1017/jpa.2020.102 .

NLC, graduate student RM, post-doc LaVine and BSL continue to work on research using digitized trilobite specimens to conduct morphometric analyses for the purposes of producing a phylo-morphospace. We are trying to develop this as a possible case study to show one type of research that can be done during a pandemic.

Finally, this will be our last quarterly progress report to iDigBio as we are concluding the project effective June 30, 2021. We thank everyone at iDigBio for all of the assistance they have provided us over the years, we've enjoyed getting the chance to work with you, and we wish you all the best.

TCN Quarterly Progress Report

Prior to each IAC meeting, TCNs are asked to complete a quarterly progress report in the areas outlined below. The TCN Lead PI or Project Manager collects information from all collaborators and compiles them into one overall progress report for the TCN. The TCN Lead PI or Project Manager then submits the quarterly reports via an email to Cat Chapman. An archive of previously submitted reports is available on the Internal Advisory Committee wiki page.

TCN Name

Digitization TCN: Collaborative Research: Documenting Marine Biodiversity through Digitization of Invertebrate Collections (DigIn)

Person Completing the Report

Regina Wetzer (Lead PI)

Share Progress in Digitization Efforts

ALMNH: Kevin Kocot: We have been working to associate collection data with 'mystery numbers' some of our specimens are labeled with. I've uploaded some sample mystery labels and aim to get more involved with the Expedition Working Group to seek out help here: https://alabama.box.com/s/nyxeetyytso6driawr3y4kpgxv10he2z
We have formatted our recent Antarctic field collection data into an Arctos bulkloader template and need to cross-reference some apparent errors with the physical specimens before uploading the data to Arctos. We have updated our website and the Invertebrate Zoology collection now has its own webpage: https://collections.museums.ua.edu/collections/

AMNH: Fani Rodriguez, Chris Johnson, Lily Berniker: AMNH hired & trained (in EMu) via Zoom two digitizers. We continue to upload our scanned card catalog images (4,619), create skeletal catalog records and complete transcription of the information (4,164). We are doing quality control on our locality records (569 new; 1,878 updated) in preparation to submit these records for georeferencing. Taxon (1,401 new; 2,926 updated), party (235 new; 357 updated), and expedition (16 new; 14 updated) records are being created, updated and verified. Senior Personnel, Lily Berniker, is investigating, updating and creating previously undigitized accession records (151 new; 240 updated). We have started imaging and databasing specimens onsite in the collection of undetermined Brachiopoda.

ANSP: Paul Callomon: 4,634 new database records created representing lots in wet collection. New interface was created to link inventoried lots to scans of original card indexes and ledgers. New system of prefixes developed to separate lots in different phyla that have same catalog number. Prefixes represent current phyla for all taxa, whereas original organization of collection was only into groups such as Polychaetes, Echinoderms and Crustaceans. 40 dry type lots linked to original and subsequent publications in database and rehoused.

BPBM: Holly Bolick: We have 3 people (volunteer, staff, and intern) trained for digitization project (specimens and images); we have precurated over 130 lots that are ready for digitization; we have 85 new records in our specimen database (70 test records from previous report were QC'd and completed); we have over 120 in situ invertebrate images organized (inventoried, categorized, relabeled) and moved from expedition folders into specimen image folders (by taxon) but not yet matched to specimen records; we have approximately 250 specimen images (that are linked by catalog number) transferred from accession folders into specimen folders (by taxon and catalog number); we are developing a new database structure enabling images and other media (labels, pdfs) to be connected directly specimens in database.

CAS: Christina Piotrowski: NfN-Invertebrate Time Machine (ITM): This quarter CAS staff scanned 3,628 card catalog labels and renamed and parsed undigitized records for 18,883 cards. Total progress this quarter: transcriptions of >9,800 catalog cards, each transcribed independently by 3 volunteers (during >29,400 independent volunteer classifications). CAS staff trained/engaged volunteers via >860 online NfN Talk comments. CAS staff continued QA/QC/reformatting of transcribed data to prep for database ingestion. CAS staff did not scan any content from Station List Files this quarter.

FWRI: Paul Larson: Continued curating field data to confirm digitize-able lots.

MCZ: Adam Baldinger: 970 uncataloged lots (mostly echinoderms) were databased this quarter from spreadsheet data; 1,219 records in database were cleaned/vetted for accuracy, and of these, 1,185 with verified georeferences.

NHMLA: Regina Wetzer: Since we are still not able to access our collections physically because of COVID restrictions, we have been focusing on processing specimen records with DNA sequence data, and associated images. Barcode quality sequences for 854 specimens were uploaded to BOLD and are publicly available. These are ready for uploading to iDigBio. We have been enhancing our digital assessment management system and have added ~250 specimen images previously not associated with specimen records.

RSMAS: Nikki Traylor-Knowles: Our digitization efforts have been focused on scanning all of the notebooks and cards that had specimen identifications. We have scanned most of these and are starting to input data into a spread sheet. We have also been trying to decide on what database platform to use, as Specify seems to be upgrading and there is debate about what folks should be do going forward, so we are waiting to hear what the consensus is.

SBNHM: Daniel Geiger: 3,555 lots digitized. During lockdown, we worked a lot on taxonomic dictionary. Unfortunately, Specifiy does not permit query for those changes, so they cannot be quantified.

SIO-BIC: Charlotte Seid: 197 digital records created as part of routine collection management activities.

SIO-PIC: Linsey Sala: Preparation of our specimen-based catalog fields in FilemakerPro and developing interrelated tables (ongoing) for loan and to use for tracking, label generation, web & iDigBio exports.

UCM: Leanne Elder: Imaging set up is together, and tanks are being made to image wet specimens over the summer. No student workers this semester, so no additional

specimens have been catalogued since last quarter. Some pre-digitizing curation has been done on shrimp, including topping off ethanol in evaporated jars to stabilize specimens.

UF: John Slapcinsky: Florida Museum cataloged and uploaded 2,130 new records of which 1,269 were georeferenced. We uploaded 8,885 images and are currently preparing additional images for upload.

VIMS: Jennifer Dreyer: 500 records have been entered into Excel this quarter. These specimens are from research projects in the Benthic Ecology Lab at VIMS and these data have been organized so that those specimens are in the queue to be entered into Excel next. In many cases these specimens are being georeferenced outside of GEOLocate portal but will eventually be uploaded there as well.

NOTE that a selection of quantitative progress measures has also been reported:

	Grant proposal commitments		Commitments completed		Records ready to upload	Georeferencing			Curation	Specimen photography		Label or catalog data capture				Direct capture from specimens		
Institution	DigIn lots to digitize for upload	DigIn images to create or mobilize for upload	Specimen records uploaded to iDigBio	Images uploaded to iDigBio	Specimen records fully prepared for upload to iDigBio	Records prepared for georeferencing	Records uploaded to CoGeo	Records QCed and repatriated after georeferencing	Specimens precurated for digitzation	Specimen images taken	Specimen images QCed and databased	Records imaged from ledgers, cards, or labels	Records OCRed or transcribed	Records QCed	Records databased	Specimen records directly captured from specimen labels	Specimen records QCed and databased	(If you have other quantities you're working on, add columns)
ALMNH	5,250	0																
AMNH	56,708	7,000	0	0	0	2447	0	0	0	0	0	4691	200	2000	4164	0	0	0
ANSP	22,060	1,600																
AUMNH	10,000	5,000																
BPBM	6,238	3,900			85				130		250							
CAS	59,616	3,500	0	0	0	0	0	0	0	0	0	3628	9800	3000	0	0	0	0
FMNH	1,140	50																
FWRI	33,582	150								0	0					0	0	0
HBOM	10,000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MCZ	31,564	4,631	970	0	1,219	1,185	0	0	0	0	0	0	0	0	0	0	0	0
NCSM-NMI	31,283	675	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
NHMLA	320,000	2,572	0	0	854					0	250							
RSMAS	55,000	0				888						3000	2006					
SBNHM	100,000	4,500	3555	0	3555	n/a	n/a	n/a	not captured	0	0	0	0	1475	1475	2070	2070	0
SIO-BIC	29,300	30,000	0	0	0	0	0	0	0	0	0	0	0	0	0	197	197	0
SIO-PIC	34,371	0																
UCM	3,285	1,000	0	0	0	0	0	0	225	46	0	225	0	0	0	0	0	0
UF	20,000	400,000	2130	8885	1269	0	0	0	0	0	0	0	0	0	0	2130	0	0
VIMS	6,000	125	0	0	0	0	0	0	0	0	0	0	0	0	0	500	0	0

The quantitative table can be accessed here:
https://docs.google.com/spreadsheets/d/1kf1HHIzN4RH69unAdrcjgU0Pc8Fi3B8As5nm8ikAGis

Share Best Practices, Standards, and Lessons Learned

AMNH: Fani Rodriguez, Chris Johnson, Lily Berniker: The team meets weekly to discuss progress, issues and develop standards for digitizing. We created a live spreadsheet to share our standard practices with our teams members. We agreed on a water body and geographical hierarchy. We have increased our efficiency by generating skeletal catalog records that correspond to the uploaded card catalog images in batch. The volunteers have moved away from our data entry form to a spreadsheet, which they now find easier to populate. The form was a good starting point, but requires data to be re-entered over and over.

ANSP: Paul Callomon: Whenever a jar is first handled for inventory it should also be overhauled - fluid checked, container changed or inspected – as it will subsequently sit somewhere else in the revised system.

BPBM: Holly Bolick: We are using the "Expedition" field in our database to track records being added through the grant which also tags the work to the individual user; we have a condition reporting field (RESPRO) in our database used to track the current condition, physical upgrades to storage containers, and presence of specimens with date.

CAS: Christina Piotrowski: CAS staff provided DigIn Workflows WG presentation on ITM-Notes from Nature project towards potential future label transcription project. We shared our process, basics for building and running a NfN project, and lessons learned. Staff also served on DigIn Steering Committee to help plan future activities/direction, collective needs, and help with administrative work.

MCZ: Adam Baldinger: Permanent staff involved in project participate in Steering, Expedition and Georeferencing committees/work groups. Information is then shared with others in MCZ's Invertebrate Zoology and Malacology departments, including those working on other TCN's (ESB and PILSBRy).

NHMLA: Regina Wetzer: Significant time was spent on Steering Committee, All Hands Meeting, and overall grant administrative tasks as well as development of the Project Manager Job Description and Advertisement which benefitted tremendously from Steering Committee member and iDigBio personnel discussion and input. These documents are currently circulating through the administrative hierarchy at NHMLA. NHMLA staff are actively participating in Expedition, Georeferencing, and Workflow Working Groups.

RSMAS: Nikki Traylor-Knowles: The scanning has taken a little time to get used to. We purchased a portable scanner and have finally figured out how to use it properly and import all data.

SBNHM: Daniel Geiger: With survey specimens getting integrated into the main collection, some storage areas are getting too crowded, and portions of the collection (Echinodermata) needed to be shifted. Those sorts of time sinks were not included in the original time estimate.

UCM: Leanne Elder: Participating in Georeferencing, Specify, and Workflows Working Groups, monthly meetings to establish best practices and standards.

VIMS: Jennifer Dreyer: Still figuring out what all the best practices will be for implementing efficient workflows that are consistent with other DigIn institutions.

Share Identified Gaps in Digitization Areas and Technology

ALMNH: Kevin Kocot: Arctos is challenging to use, and it is time-consuming getting trainees up to speed.

AMNH: Fani Rodriguez, Chris Johnson, Lily Berniker: We don't yet have an imaging workflow, because we are only just now starting to work with specimens onsite.

ANSP: Paul Callomon: Inadequate database management resources at institutional level hamper digitization projects. Too much is still done ad-hoc by individual PIs.

BPBM: Holly Bolick: We have a straightforward way of tracking newly added specimen records easily by person and date, but we do not have that capability for tracking record updates yet (it would have to be a manual and labor-intensive tracking method for now).

CAS: Christina Piotrowski: CAS had difficulty with Open Refine installation this quarter, but we may be able to use it after locating a Java substitute. We've identified a future roadblock related to Specify appending records (post-georef and during workflows to expand skeletal records). We have additional tech gaps regarding specimen/label photography protocols and best practices.

MCZ: Adam Baldinger: We are in the process of setting up a macrophotography work station to image specimens.

NHMLA: Regina Wetzer: Anxious to get onsite and to optimize data capture from specimens on shelves (our biggest need). We will do this with Collections Managers first, and plan to expand work force with Work Study in late August thereby seriously cranking up volume.

RSMAS: Nikki Traylor-Knowles: The biggest issue we are running into is what database to use. Since we are at the beginning we want to use a database that will continue to be used –and not have to be upgraded etc etc etc every few years. So this is the biggest hurdle right now, as we are ready to input this data.

SIO-PIC: Linsey Sala: Need to prepare two computer workstations for on-site student workers.

UCM: Leanne Elder: Still setting up Specify database. Working out digital asset management: where to store them, workflow to link to specimen records etc.

VIMS: Jennifer Dreyer: There are still big gaps for me in learning Specify and using the workbench. Based on what I have learned from other institutions, I may have to reconsider the Specify forms to capture data I had not considered when I initially created our database. Since we have not started uploading data from Excel yet, I know we need to have the form fields determined first before we move forward with importing our records into Specify.

YPM: Nelson Rios: Lack of standards within Darwin Core to describe complex collecting events temporally and geospatially.

Share Opportunities to Enhance Training Efforts

CAS: Christina Piotrowski: CAS is participating in the DigIn Specify users group to share workflows and expertise using this CMS; CAS is available to help advise potential users of NfN for DigIn if funding can be secured for DigIn to use their services. Also can advise on Zooniverse project building. All institutions will participate in share sessions to train and share best practices and troubleshoot any inefficiencies in our workflows.

NHMLA: Regina Wetzer: We offered Narrative Training for thirty collections and other staff members across participating institutions. This training is designed to prepare participants to be much more effective communicators in any medium.

RSMAS: Nikki Traylor-Knowles: I think that as a community we need to think about what database we are going to use for the long term. I think that it should be free, and I don't see why this would be hard to do. I wonder if we could partner with a tech company who may be interested in helping our cause. But we just need to have consistency so that groups like ourselves which are new are not recreating the wheel!

SIO-BIC: Charlotte Seid: Conducted 4 FileMaker Pro training/collaboration sessions (6 hrs) with SIO-PIC.

SIO-PIC: Linsey Sala: Identified & edited daily work log tracking sheet that will best suit the digitization workflows for SIO-PIC in preparation for tracking future on-site student worker progress.

VIMS: Jennifer Dreyer: I am on the Workflow and Specify working groups and these meetings have been very helpful in terms of learning from other institutions workflow protocols in handling specimens but especially working in Specify.

Share Collaborations with other TCNs, Institutions, and/or Organizations

ALMNH: Kevin Kocot: I am planning a course with Gustav Paulay and Jon Norenburg at Friday Harbor Labs for Summer 2022 on Integrative Biodiversity and Taxonomy of Invertebrates and a meiofauna diversity and taxonomy workshop with Ashleigh Smythe at the Smithsonian Marine Station in Fort Pierce, FL later in Summer 2022.

AMNH: Fani Rodriguez, Chris Johnson, Lily Berniker: Increased networking with other institutions working on similar taxa, which is leading to a more cohesive marine invertebrate community. Working with NHMLA on an initiative with a citizen scientists group that reports sightings of a possibly invasive species of anemone that is starting to make a more pronounced appearance on the Pacific Coast.

BPBM: Holly Bolick: There are multiple grants at Bishop Museum currently (some TCNs) focused on digitizing specimen collections so the back end database structure changes are resulting from the needs and grant requirements of multiple funding sources; uber checklist of Hawaii marine invertebrates being developed through other funding; will connect to specimen database and digital reference library.

CAS: Christina Piotrowski: ESB TCN, Zooniverse/Notes from Nature, WoRMS, GeoLocate, TDWG for marine georeference localities (see WG); potential collaborations involving new CAS initiatives as they pertain to grant-generated biodiversity records, California National Marine Sanctuary collaborations re: newly digitized specimen records.

MCZ: Adam Baldinger: Information is shared among permanent MCZ staff working on other TCN's: ESB and PILSBRy, and an NSF CSBR cryocollections grant.

NHMLA: Regina Wetzer: Definitely benefiting from increased networking among DigIn institutions as well as the cross fertilization with the ESB TCN. Notably, three working groups (Expedition Data, Georeferencing, and Nomenclature) and two user help groups (Slack Help and Specify User Group) are now completely shared between both DigIn and ESB TCNs.

RSMAS: Nikki Traylor-Knowles: Eastern Seaboard Mollusks TCN, and FLMNH.

SIO-PIC: Linsey Sala: Collaborating with Charlotte Seid in SIO-BIC to develop our FilemakerPro specimen-based catalog and future loan tracking capabilities.

UCM: Leanne Elder: UCM has concurrent PEN to InvertEBase for terrestrial invertebrates. L. Elder is communicating with InvertEBase PIs for advice on best practices and lessons learned.

UF: John Slapcinsky: Collaborating with Eastern Seaboard TCN by sharing several committees.

YPM: Nelson Rios: ESB & DigiLeap

Share Opportunities and Strategies for Sustainability

AMNH: Fani Rodriguez, Chris Johnson, Lily Berniker: We have just recently moved our EMu database to a new high-powered server with increased space.

CAS: Christina Piotrowski: We captured data records for extinct species and potential minor geographic range expansion of at least one extant species. Historical ITM records provide baseline for past marine species distributions.

RSMAS: Nikki Traylor-Knowles: My biggest concern about sustainability is money once this grant is over, I worry that our collection could be in trouble. I am hoping that we can start some fundraising soon and would love to learn from anyone in the group who has experience with this. We need donors!

Share Education, Outreach, Diversity, & Inclusion (EODI) Activities

ALMNH: Kevin Kocot: I have used the Invertebrate Zoology collection extensively in my Invertebrate Zoology course this semester.

AMNH: Fani Rodriguez, Chris Johnson, Lily Berniker: Several zoom presentations to school groups on our collection and databasing activities. We are preparing to create a webpage for the AMNH project to increase the visibility of our marine invertebrate collection.

ANSP: Paul Callomon: The principal worker in the wet collection will leave in September to become a graduate student in veterinary science, and the curation of the Joseph Leidy Collection of animal and human parasitic worms is proving a useful primer.

CAS: Christina Piotrowski: CAS staff answered more than 900 questions this quarter about data and inspired transcribers' curiosity about associated taxa, localities, and collectors. We provided content supporting the general value of historical marine collections data. We have at least several known students working among our volunteer base.

NHMLA: Regina Wetzer: We are organizing and will be hosting an Educator Workshop: Dive Into Marine Biodiversity 28 Jun-2 July for California State University Dominguez Hills STEM teachers which will involve hands-on, multi-day immersion with visits to the Museum, the shore, sampling from a

reasearch vessel, and exposure to traditional and modern approaches for studying marine biodiversity. This workshop funded by Boeing will lay ground work for our Summer 2022 outreach training for the CSUDH STEM teachers participating in this grant.

RSMAS: Nikki Traylor-Knowles: Currently we are training 1 undergraduate intern, and one professional masters student in how to digitize, catalogue and photograph samples.

SBNHM: Daniel Geiger: Due to COVID 19, no external persons have been permitted in the collection. We are exploring opportunity for externally funded graduate student intern in Winter 2021.

SIO-BIC: Charlotte Seid: Conducted 6 virtual E&O presentations (5 hrs) for elementary through graduate students, highlighting invertebrate biology and the value of digitized museum collections.

SIO-PIC: Linsey Sala: Conducting virtual presentations with incoming and first year graduate students to highlight SIO-PIC resources and importance of natural history collections for scientific research and education. Utilizing SIO-PIC reference specimen materials for virtual spring quarter graduate level Marine Zooplankton course lab.

UCM: Leanne Elder: Participated in Girls at the Museum Engaging in Science program over Zoom. UCM IZ group led two sessions with ~10 middle school girls in each. Session was on algae symbionts in Invertebrates, focused on corals, anemones and mollusks. Discussed with students how museums use specimens, sent packs to students with paper microscopes and macroalgae samples in lugols solution for them to put on a slide and view.

UF: John Slapcinsky: Florida Museum shares tweets and photos about invertebrates and lab happenings on Spineless Science @ UF which can be accessed @UFInvertZoo.

VIMS: Jennifer Dreyer: Answered questions from a local 8th grader who is working on a capstone project and needed to interview someone about horseshoe crab specimens in Chesapeake Bay and those in our Invertebrate Collection.

Share Information About Your Website and/or Portal Usage

The skeletal project website is up at https://digin-tcn.org. The Symbiota-based InvertEBase is being prepared to receive records from DigIn, particularly for use as a georeferencing intermediate processing platform, but is not yet ready for broad-based data input.

Share Other Activities and/or Progress

ANSP: Paul Callomon: Cleaning and rebottling is integrated with the digital inventory in the General Invertebrates collection. A large body of material that has collecting data and identifications, but that has never been cataloged is now separated and ready for digitization de novo.

BPBM: Holly Bolick: We recently purchased a microscope camera through other funds and will be able to use this to image small wet specimens in the lab on the fly if necessary (saving time and reduces the need to reserve shared imaging equipment).

CAS: Christina Piotrowski: Database infrastructure: final preparations for CAS database migration and field standardization work that will permit us to better standardize our data and to host grant images and other files.

FWRI: Paul Larson: Our institute is beginning to transition back to in-office work – will enable hiring of techs to begin data capture.

MCZ: Adam Baldinger: Part-time Curatorial Assistant position has been advertised. Candidate screening is still in progress.

NHMLA: Regina Wetzer: Significant time has been invested in preparing job descriptions for upcoming Project Manager hires and refining them based on input from DigIn Steering Group members. Job search is expected to begin in the next few months.

RSMAS: Nikki Traylor-Knowles: Scanning and backing up the data from the museum.

VIMS: Jennifer Dreyer: Researched and purchased a new microscope camera to replace our current model that is not compatible with Windows 10, which was a required upgrade at our institution. While this grant will not contribute to the camera cost, most of the images we take are from research specimens that will be added to our Invertebrate Collection.

YPM: Nelson Rios: Integration of Bathymetric Maps into GEOLocate. Almost finished incorporating complex geometries.

Additional costs due to COVID pandemic

AMNH: Fani Rodriguez, Chris Johnson, Lily Berniker: We had to purchase Barcode Filer to have a license for our interns to use offsite.

CAS: Christina Piotrowski: 1. Non-grant CAS staff time spent on Digitization and NfN (we cannot yet hire grant staff due to remote work). We're unable to train and supervise on site and cannot physically prepare for workflow ramping up/hire/train new grant staff. 2. Supply, equipment expenses to support remote scanning of cards and documents that would normally be scanned on site. 3. Home internet bandwidth upgrades (partially subsidized by CAS). IN GENERAL, we remain uncertain re: the full impact of COVID on our project budget over the long term. We're currently spending significant CAS staff time on "simple" grant student planned scanning work by necessity to support a remote transcription project, so we'll require proportionally more hours of highly trained staff later in the funded period to complete less simple tasks. The current higher load of grant supporting activities prevent staff from attending to other non-grant related projects.

SBNHM: Daniel Geiger: No additional direct cost. Work is less efficient, so most likely cataloging rate is slower than anticipated/ideal. Due to less efficient work, PI has approved all time-off request for CA and CM. That time will be made up in 2022.

UCM: Leanne Elder: Still unable to image ledgers as the scanner that is free to use at the university library is unavailable. Will likely have to buy a book scanner for ~ 600 dollars to keep the time consistent with original plan for the grant (this was reported in the first quarter report as well).

AUMNH	Auburn University, Auburn, AL – no submission
HBOM	Florida Atlantic University, Fort Pierce, FL – no submission
NCSM-NMI	North Carolina Museum of Natural Sciences, Raleigh, NC – no submission

==========

Key to Abbreviations

NHMLA	Natural History Museum of Los Angeles County (lead), Los Angeles, CA (R.Wetzer)
ANSP	Academy of Natural Sciences, Philadelphia, PA (P. Callomon)
AMNH	American Museum of Natural History, New York, NY (C.Johnson)
AUMNH	Auburn University, Auburn, AL – no submission
ASU	Arizona State University, Tempe, AZ [subaward]
BPBM	Bishop Museum, Honolulu, HI – no submission
CAS	California Academy of Sciences, San Francisco, CA (C.Piotrowski)
VIMS	College of William & Mary—Virginia Institute of Marine Science, Gloucester Point, VA (J.Dreyer)
FMNH	Field Museum of Natural History, Chicago, IL [subward]
HBOM	Florida Atlantic University, Fort Pierce, FL (D. Hanisak)
FWRI	Fish and Wildlife Research Institute, St. Petersburg, FL (P.Larson)
MCZ	Harvard University, Cambridge, MA (A.Baldinger)
NCSM-NMI	North Carolina Museum of Natural Sciences, Raleigh, NC (M.McCuller)
Q-Quatics	Q-Quatics, Laguna, Philippines [subaward]
SBNHM	Santa Barbara Museum of Natural History, Santa Barbara, CA (D.Geiger)

SIO-BIC	Scripps Institution of Oceanography, University of California San Diego, CA (C.Seid)
SIO-PIC	Scripps Institution of Oceanography, University of California San Diego, CA (L.Sala)
ALMNH	University of Alabama Tuscaloosa, AL (K.Kocot)
UCM	University of Colorado, Boulder, CO (L.Elder)
UF	University of Florida, Gainesville, FL (J.Slapcinsky)
RSMAS	University of Miami, Rosenstiel School of Marine & Atmospheric Science, Miami, FL (N.Traylor-Knowles)
YPM	Yale University Peabody Museum of Natural History, Boston, MA [subaward] (N.Rios)

TCN Quarterly Progress Report

Prior to each IAC meeting, TCNs are asked to complete a quarterly progress report in the areas outlined below. The TCN Lead PI or Project Manager collects information from all collaborators and compiles them into one overall progress report for the TCN. The TCN Lead PI or Project Manager then submits the quarterly reports via an email to Cat Chapman. An archive of previously submitted reports is available on the Internal Advisory Committee wiki page.

TCN Name

Digitizing "endless forms": Facilitating Research on Imperiled Plants with Extreme Morphologies (Endless Forms / EF)

Person Completing the Report

Matthew Pace (Lead PI)

Share Progress in Digitization Efforts

Herbarium	# specimens barcoded	# specimens imaged	# specimens fully transcribed	# specimens georeferenced
New York Botanical Garden (NY)	177367	147252	78244	21068
Illinois Natural History Survey (ILLS/ILL)	44542	44547	38041	6856
Marie Selby Botanical Gardens (SEL)	49,262	52,961	12,100	35,002
Philadelphia Academy of Science (PH)	44538	44538	42944	26348
University of Massachusetts (MASS)	10,587	10,000	10,000	1,167

California Academy of Sciences (CAS)	66,326	39,738	66,326	13,265
Field Museum (F)	135,222	119,372	135,222	39246
University of Wisconsin Madison (WIS)	57,077	54,976	50,672	28,203
Harvard University (GH/A/AMES/NEBC/ECON)	168184	137550	168184	9178
Missouri Botanical Garden (MO)	229,893	92,356	229,893	174,283
Botanical Research Institute of Texas (BRIT)	33,177	33177	36,490	0
University of Michigan (MICH)	51346	51346	47887	16142
University of Minnesota (MIN)	16325	16325	9776	2000
University of Texas (TEX)	5185	37071	56,524	13826
University of California (UC)	19760	19760	7685	2052
California Botanic Garden (RSA)	26,505	7,988	10,824	2,996
TOTAL	**1135296**	**908957**	**1000812**	**391632**

.

Share Best Practices, Standards, and Lessons Learned

Share information here. You can also embed graphics if desired.

Share Identified Gaps in Digitization Areas and Technology

Share information here. You can also embed graphics if desired.

Share Opportunities to Enhance Training Efforts

Share information here. You can also embed graphics if desired.

Share Collaborations with other TCNs, Institutions, and/or Organizations

At NYBG, the Ferns & Endless Forms TCNs collaborated for a joint WeDigBio transcription event and seminar series.

Share Opportunities and Strategies for Sustainability

Share information here. You can also embed graphics if desired.

Share Education, Outreach, Diversity, & Inclusion (EODI) Activities

The Endless Forms TCN was highlighted during NYBG's WeDigBio seminar seminars and training events. Talks focused on epiphytic plant families (Orchid, Ferns, and Bromeliads), and the speakers discussed how digital herbarium data helps augment and facilitate their research. The transcription events associated with WeDigBio also focused on epiphytic genera. Approximately 500 people participated across all the events.

*P. Fritsch presented (9 April 2021): "Evolutionary origins and species diversity in the wild blueberries tribe Vaccinieae" for the BRIT Lunchtime Lecture series, in which he highlighted the importance of access to digitized herbarium records (images and data) for his research and highlighted the current and past funded TCNs. In particular he included a discussion of the complementary nature of iNat records and the value of specimen images, particularly when specimens have been damaged post-imaging (https://brit.org/research/research-resources/research-lecture-series/; https://www.youtube.com/watch?v=LEIkOpYHufs&list=PLFlQggX9zTKQ3QuZ7lgnJ9zIMW997bIZq&index=2)

• *T. Rehman presented (2021-04-17): "Collections, Conservation, and Community Science: A Collaboration to Conserve our Planet's Natural Heritage" - for the Fort Worth Public Library Earth Day 2021 Events in which currently funded TCNs were highlighted and participants were encouraged to explore their local natural history collections and participate in the weekly Armchair Botany Zoom sessions (https://www.fortworthtexas.gov/departments/library/events/earth-day-2021?BestBetMatch=rehman|a720ba43-6954-4033-b1ac-a72478fc15dc|f3d8342e-e792-4add-b724-7420243aa350|en-US; https://www.youtube.com/watch?v=0945UpOxnFw)

• *T. Rehman, A. Bordelon, D. Barroso, J. Best, C. Barrett presented (8, 10, 11 April 2021): three WeDigBio Transcription Blitzes including discussions about currently funded TCNs, a highlight presentation about the TORCH TCN a behind-the-scenes tour of

the BRIT herbarium (8 April), a behind-the-scenes tour of the Sumner Molecular and Structural Lab at BRIT (10 April; led by BRIT botanist Alejandra Vasco), and a presentation about the collecting of herbarium specimens, including live tour of BRIT specimen receiving room and collecting supplies (12 April). Attendees from multiple countries were in attendance (brit.org/wedigbio)

Share Information About Your Website and/or Portal Usage

Share information here. You can also embed graphics if desired, such as from Google Analytics.

Share Other Activities and/or Progress

Share information here for things that do not fit into the above categories. You can also embed graphics if desired.

TCN Quarterly Progress Report

Prior to each IAC meeting, TCNs are asked to complete a quarterly progress report in the areas outlined below. The TCN Lead PI or Project Manager collects information from all collaborators and compiles them into one overall progress report for the TCN. The TCN Lead PI or Project Manager then submits the quarterly reports via an email to Cat Chapman. An archive of previously submitted reports is available on the Internal Advisory Committee wiki page.

TCN Name

Eastern Seaboard: Mobilizing millions of marine mollusks

Person Completing the Report

Rüdiger Bieler, FMNH, Lead PI

Share Progress in Digitization Efforts

Share information here. You can also embed graphics if desired.

COVID-19 restrictions continued to impact staff access to the collections and widespread hiring freezes for new staff caused unavoidable delays. However, where possible under the circumstances, work is underway. Members of the collaborative found creative ways to prepare for, and begin, digitization.

FMNH ESB**:** Collections staff tested the new, institution-wide database structure in anticipation of ESB data entry.

FMNH ESB subaward **HMNS:** Work has begun on digitizing lots from the Dr. John "Wes" Tunnell collection. To date, 62 lots representing 561 specimens have been digitized. This collection includes scientifically important specimens collected in the Gulf of Mexico by Dr. Thomas Pulley (HMNS), Dr. Henry Ladd (USNM), and Dr. William Clench (MCZ) in the 1950's. Work in the database to bring taxonomic records up to date with the currently accepted phylogenetic hierarchy has also begun. Updating taxonomic records has begun at the family level and above because we will want to create new specimen cards as we update genus and species level taxonomy. Since January approximately 8,144 catalog records representing 120,538 ESB specimens have had this taxonomic work done on them in our database.

FMNH subaward **FWRI:** Have been examining scanned field data records from projects that will be digitized (or partially are) and querying the database to identify which field data already is digitized and which is not for our undigitized specimens.

FMNH ESB subaward **HBOM:** We finished the mold remediation that was necessary in HBOM and have begun an inventory of specimens that will be digitized. We hired a new Collections Manager who started in April who is critical to these efforts. No digitization of samples has occurred yet.

ANSP ESB**:** 28,833 lots linked to the ESB project in our internal database. Of these, 128 are old lots newly digitized as they were missed by previous efforts and 287 are new lots added to the collection since the start of the project.

BMSM ESB**:** Pre-processed / cleaned about 1,115 lots from Don Swenson collection (Martin and Palm Beach counties, Florida) for digitization; started job search for digitizer cataloging sponsored by ESB.

CM ESB**:** 24 new ESB marine lots digitized

DMNH ESB**:** Standardization of coastal marine locality names is our primary focus. **DMNH** ESB subaward **RSMAS:** Our digitization efforts have been focused on scanning all of the notebooks and cards that had the specimen identifications. We have scanned most of these and are starting to input data into a spreadsheet. We have been also trying to decide on what database platform to use, as Specify seems to be upgrading and there is debate about what folks should do going forward, so we are waiting to hear what the consensus is.

MCZ ESB**:** 3,607 existing records in our database with New England and mid Atlantic localities were flagged for ESB inclusion. Coordinates for 3,598 of these records confirmed. Taxonomy, agents and data in other fields vetted as well. No new records were cataloged.

LACM ESB**:** Continued to georeference Museum stations as well as discussed the "Vitality," "Means of Inference," and "Project" fields for EMu with our EMu support staff.

NCSM ESB**:** All the card catalogues of the Duke and Charleston collection have been scanned and are in the process of being renamed in order to link to the specimen(s). All of this information had been transcribed by the public and has been imported into our database. We have fully databased, including georeferencing and updating taxonomic names according to MollucsBase, 211 lots. Totals lots to date are: 87 Bivalvia, 47 Cephalopods, 77 Gastropoda, 3 Polyplacophora, and 1 Scaphopoda.

UF ESB**:** Entered 409 new specimen records, georeferenced 24077 previously cataloged specimens from Florida, uploaded 154 photos mostly from the Greenlaw Collection.

UMMZ ESB**:** Specimen records from 131 lots (Polyplacophora & Gastropoda; 547 specimens) were newly entered since January 2021. In addition, 987 gastropod lots were prescreened and are ready for data entry. Newly reported specimens are from Massachusetts (3 lots), North Carolina (7), Georgia (2), Florida (115), Texas (2), Gulf of Mexico (2). Totals: 369 lots, 1990 specimens; Polyplacophora: 68 lots, 236 specimens; Bivalvia: 35 lots, 236 specimens; Gastropoda: 266 lots, 1518 specimens.

Share Best Practices, Standards, and Lessons Learned

Share information here. You can also embed graphics if desired.

Working groups have been established (e.g., workflow, georeferencing) and hold regular meetings. Several of these groups are joint efforts with the DigIn-TCN.

FMNH ESB**:** nothing to report.

FMNH ESB subaward **HMNS:** nothing to report.

FMNH ESB subaward **FWRI:**Have presented FWRI practices to Workflows working group and discussed such in Specify working group.

FMNH ESB subaward **HBOM:** Other than learning from others on the projects as we go, nothing to report.

ANSP ESB**:** Enhanced document of standard view for photography of mollusks and made it available to the ESB group. **BMSM** ESB**:** nothing to report

CM ESB**:** nothing to report

DMNH ESB**:** nothing to report

DMNH ESB subaward **RSMAS:**The scanning has taken a little time to get used to. We purchased a portable scanner and have finally figured out how to use it properly and import all data.

MCZ ESB**:** Participation in working groups/committees: Steering, Expeditions, Georeferencing, Workflows, Outreach.

LACM ESB**:** nothing to report.

NCSM ESB**:** nothing to report.

UF ESB**:** nothing to report

UMMZ ESB**:** nothing to report

Share Identified Gaps in Digitization Areas and Technology

Share information here. You can also embed graphics if desired.

FMNH ESB**:** nothing to report.

FMNH ESB subaward **HMNS:** Due to limitations in our previous collections database many of the fields in our current database did not exist. This lack of proper fields led to information in our taxonomy and localities

modules to be inputted incorrectly or omitted. These modules are currently being updated but it is a time consuming process. **FMNH** ESB subaward **FWRI:** Nothing to report yet

FMNH ESB subaward **HBOM:** We are developing an image analysis station and plans to barcode the HBOM specimens.

ANSP ESB**:** Developed recommendations for field and conventions for reporting live versus dead condition. Required fields are Vitality and Means of inference and recommended is presence of tissue. Required values for vitality are "Live", "dead" and "can't tell", with recommended values of "live and dead mixed", "fresh dead" and "not evaluated".

BMSM ESB**:** nothing to report

CM ESB**:** nothing to report

DMNH **:** Live/dead determinations have been paused until we institute changes to the Specify database to accommodate Vitality and Inference fields.

DMNH ESB subaward **RSMAS:**The biggest issue we are running into is what database to use. **MCZ** ESB**:** work station for macro-photography is being set up.

LACM ESB**:** nothing to report.

NCSM ESB**:** Nothing to report as of yet.

UF ESB**:** nothing to report

UMMZ ESB**:** nothing to report

Share Opportunities to Enhance Training Efforts

Share information here. You can also embed graphics if desired.
FMNH ESB**:** nothing to report.

FMNH ESB subaward **HMNS:**

FMNH ESB subaward **FWRI:** nothing to report yet

FMNH ESB subaward **HBOM:** nothing to report

ANSP ESB**:** nothing to report

BMSM ESB**:** nothing to report

CM ESB**:** High school volunteer becoming adept at using part of EMu database

DMNH ESB**:** nothing to report

DMNH ESB subaward **RSMAS**:I think that as a community we need to think about what database we are going to use for the long term.

MCZ ESB**:** nothing to report

LACM ESB**:** nothing to report.

NCSM ESB**:** nothing to report.

UF ESB**:** nothing to report

UMMZ ESB**:** nothing to report

Share Collaborations with other TCNs, Institutions, and/or Organizations

Share information here. You can also embed graphics if desired.

Much effort has been spent to facilitate exchanges and avoid duplication with other relevant TCNs. InvertEBase is the main partner for our Symbiota portal – with close coordination with DigIn-TCN, PILSBRY-TCN, and Smithsonian's Panama project (which is also in the process of joining the InvertEBase platform). Shared working groups (with members of multiple TCNs) have been established and are beginning to meet regularly. ESB and DigIn, in particular, share many upcoming needs in authority file development, georeferencing needs, and workflow development, and the two projects have a strong overlap in their steering committees to assure the best-possible flow of ideas and information.

FMNH ESB**:** Coordinating above efforts.

FMNH ESB subaward **HMNS:** Nothing to report.

FMNH ESB subaward **FWRI:** Participate in joint working groups with DigIn **FMNH** ESB subaward **HBOM:**

ANSP ESB**:** Participated in joint working groups with Digin.

BMSM ESB**:** nothing to report

CM ESB**:** nothing to report

DMNH ESB**:** Gave a presentation at the most recent ESB meeting about ongoing collaborations with BCEENET (RCN-UBE) to develop CUREs. DMNH is organizing a panel presentation at the upcoming BCEENET meeting (June) to introduce novice ecology faculty to TCN data including ESB and DigIn.

DMNH ESB subaward **RSMAS:** DigIn

MCZ ESB**:** ADBC TCN DigIn and PILBSRy - shared working groups; NSF CSBR cryopreservation of MCZ types **LACM** ESB**:** nothing to report.

NCSM ESB**:** nothing to report.

UF ESB**:** Participate in joint working groups with DigIn **UMMZ** ESB**:** nothing to report

Share Opportunities and Strategies for Sustainability

Share information here. You can also embed graphics if desired.
FMNH ESB**:** nothing to report.

FMNH ESB subaward **HMNS:** nothing to report.

FMNH ESB subaward **FWRI:** nothing to report yet

FMNH ESB subaward **HBOM:** Our HBOM remediation and hiring of a Collections Manager are key to sustainability of our efforts.

ANSP ESB**:** nothing to report

BMSM ESB**:** nothing to report

CM ESB**:** nothing to report

DMNH ESB**:** Ongoing collaboration with BCEENET is a means to develop new audiences for digitized data. We are working with undergraduate faculty who have no previous collection experience to show them the breadth and depth of the digital data available, and how it can be used for research in a variety of undergraduate lab conditions.

DMNH ESB subaward **RSMAS:** My biggest concern about sustainability is money. I am hoping that we can start some fundraising soon.

MCZ ESB**:** nothing to report

LACM ESB**:** nothing to report.

NCSM ESB**:** nothing to report.

UF ESB**:** nothing to report

UMMZ **:** nothing to report

Share Education, Outreach, Diversity, & Inclusion (EODI) Activities

Share information here. You can also embed graphics if desired.

Our Social Media Committee (spearheaded by BMSM, with FMNH, MCZ, and FWRI) defined parameters and spearheaded creation of media resources for the project. ESB now has media presence on: Facebook https://www.facebook.com/groups/easternseaboardmollusks; Instagram @eastern_seaboard_mollusks; Twitter @EMollusks; iNaturalist https://inaturalist.org/projects/eastern-seaboard-mollusks

FMNH ESB**:** as above.

FMNH ESB subaward **HMNS:** We have hired a graduate intern from the University of North

Texas and she will begin work with us in May. This graduate student will be working with our Inventory Manager to verify catalog records match actual specimens, update taxonomy, collection localities, and learn proper collections care and preservation methods.

FMNH ESB subaward **FWRI:** nothing to report yet

FMNH ESB subaward **HBOM:** nothing to report

ANSP ESB**:** Identified 7590 observations on iNaturalist Eastern Seaboard Project since November 2020 and scored more than 20,000 as live or dead collected. Hosted Susan Hewitt who gave Zoom talk to Philadelphia Shell Club on iNaturalist and ESB on January 21, attended by 37 people.

BMSM ESB**:** participated in outreach committee, posting on Facebook, which rolls out to Instagram and potentially other media. Working on the curation and set-up of a temporary exhibition ("In Focus: Precision Photography of Extraordinary and Uncommon Shells") of largeformat, high-resolution images of Florida and ESB specimens from BMSM catalog to be launched on Memorial Day weekend.

CM ESB**:** nothing to report

DMNH ESB**:** nothing to report

DMNH ESB subaward **RSMAS:**Currently we are training 1 undergraduate intern, and one professional masters student in how to digitization, catalogue and photograph samples.

MCZ ESB: continued enhancement of iNaturalist page for ESB project.

LACM ESB**:** nothing to report.

NCSM ESB**:** Currently I am training one undergraduate on how to digitize and identify mollusks and I plan on bringing in a volunteer soon.

UF ESB**:** Identified and posted observations on iNaturalist Eastern Seaboard Project. Filmed an episode of Stingray Tom's Florida on museum collections of mollusks https://youtu.be/BCkiYUeApCQ

UMMZ ESB**:** nothing to report

Share Information About Your Website and/or Portal Usage

Share information here. You can also embed graphics if desired, such as from Google Analytics.
FMNH ESB**:** nothing to report.

FMNH ESB subaward **HMNS:** nothing to report.

FMNH ESB subaward **FWRI:** Not tracked - inquired with FWRI Communications team

FMNH ESB subaward **HBOM:** nothing to report

ANSP ESB**:** nothing to report

BMSM ESB**:** Requested access to catalog portal from BMNSM IT personnel

CM ESB**:** nothing to report

DMNH ESB**:** nothing to report

DMNH ESB subaward **RSMAS:** We are just starting to renovate our website, but will need to help in finishing this. We have only gotten the template together and areas that we want to include.

MCZ ESB**:** Twitter account @Emollusks currently has 69 followers **LACM** ESB**:** nothing to report.

NCSM ESB**:** All databased material will be uploaded monthly to our website for others to enjoy.

UF ESB**:** nothing to report

UMMZ ESB**:** nothing to report

Share Other Activities and/or Progress

Share information here for things that do not fit into the above categories. You can also embed graphics if desired.

Communication infrastructure and file storage has been organized via a dedicated Slack workspace and file storage on Google Drive. We also are sharing relevant Slack channels with the DigIn TCN. Cost-Reimbursement Subaward Agreements have been executed.

FMNH ESB**:** Coordination as described above.

FMNH ESB subaward **HMNS:** nothing to report. **FMNH** ESB subaward **FWRI:** nothing to report yet **FMNH** ESB **s**ubaward **HBOM:** nothing to report.

ANSP ESB**:** nothing to report

BMSM ESB**:** Finally managed to have our most recent dataset (Specify 7 Cloud) ingested by iDigBio. Published an annotated list of the BMSM types including material from ESB at Type specimens in the molluscan collection of the Bailey-Matthews National Shell Museum, Florida, USA | LEAL | Zootaxa (mapress.com)

CM ESB**:** nothing to report

DMNH ESB**:** nothing to report

DMNH ESB subaward **RSMAS:** Scanning and backing up the data from the museum **MCZ** ESB**:** Part time Curatorial Assistant position has been advertised; candidate screening currently in progress.

LACM ESB**:** nothing to report.

NCSM ESB**:** nothing to report.

UF ESB**:** nothing to report

UMMZ ESB**:** nothing to report

TCN Quarterly Progress Report

TCN Name

Building a global consortium of bryophytes and lichens: keystones of cryptobiotic communities (GLOBAL)[1]

Person Completing the Report

Miranda Zwingelberg (GLOBAL Project Manager)

Share Progress in Digitization Efforts

This report covers progress completed during the period of January 1 - March 31, 2021.

Digitization progress at all GLOBAL institutions continued to be limited by COVID restrictions during 2021-Q1. A few collaborators saw slight increases in access during this period, allowing more collections to begin some digitization activities. Several others are still waiting to start GLOBAL work until conditions are more favorable.

Imaging Equipment

Additional progress was reported in purchasing and setting up imaging equipment during 2021-Q1. COLO purchased basic equipment (color square, blackout paper, etc.) for use with an existing light box. NY purchased a new camera and lens. MSC bought a new PC and began researching camera options. TENN tested out a new camera and lens to use in a second imaging station. F set up two remote label imaging stations and one on-site light box. YU finalized their imaging set up and CINC & MU set up a second imaging station.

[1] Throughout this report, herbaria are referred to by their Index Herbariorum acronyms, which correspond to institutional names as follows: ALA = University of Alaska, Fairbanks, ASU = Arizona State University, BRY = Brigham Young University, CINC & MU = University of Cincinnati & Miami University, COLO = University of Colorado, DUKE = Duke University, F = The Field Museum, FLAS = University of Florida, ILL & ILLS = University of Illinois at UrbanaChampaign & Illinois Natural History Survey, LSU = Louisiana State University, MICH = University of Michigan, MIN = University of Minnesota, MO = Missouri Botanical Garden, MSC = Michigan State University, NY = New York Botanical Garden, OSC = Oregon State University, PH = The Academy of Natural Sciences of Drexel University, TENN = University of Tennessee, Knoxville, UC = University of California, Berkeley, WIS = University of Wisconsin, YU = Yale University

Workflow Development

Many institutions spent time in 2021-Q1 continuing to draft, update, and refine their imaging workflows (ASU, F, LSU, MO, MSC, NY, TENN, WIS, and YU). CINC & MU were able to take advantage of a student's photography background to improve their imaging set up. Additional activities included cleaning data (DUKE, MO, and UC)

and pre-curation of specimens (F). ASU worked on developing a new lichen and bryophyte label format that can be used to print specimen labels directly onto packets from within the portals. MO reestablished their IPT to allow updates to the Bryophyte Portal. The GLOBAL leadership team organized a meeting of the Specimen Imaging Working Group in February to facilitate sharing between the collaborators.

Essential progress was also made in mapping out centralized georeferencing workflows, especially for those collections shared to the portals as "snapshots." The GLOBAL Project Manager, Portal Manager, and Georeferencing Manager met with the ASU team to review options in February. They then began to schedule individual consultations with collaborators to discuss the flow of coordinate data between the portals and internal databases (DUKE, F, LSU, and MIN). These meetings will continue into 2021-Q2.

Image Uploading

Important progress was made by the IT team at ASU to establish image hosting services and workflows for the submission of GLOBAL images to the Bryophyte and Lichen Portals. External storage was made available to the portals via a mounted directory. An Apache server and the portals were configured to utilize the server mount. A Dropbox group was established to manage requests for each institution and collection type. A manual workflow was developed for transferring images from Dropbox request folders to processing directories within the ASU server system. Scripts were setup and configured to create image derivatives, transfer files to web server space, and link the images to occurrence records within databases. Lichen and bryophyte processing profiles were set up for each institution. Trial uploading of images with this manual work-around were completed successfully (ASU, COLO, DUKE, TENN, WIS) and the option was made available to the entire GLOBAL group in March.

Personnel

The GLOBAL Portal Manager (ASU) and a Lead Digitizer at NY started working on the project in January 2021. F trained employees for their remote stations and an on-site technician. PH hired a full-time co-op student for barcoding and imaging work, who started at the end of March. MSC and YU made progress hiring undergraduates for summer work. FLAS brought on 4 work-study students.

Digitization

Twelve institutions (BRY, CINC & MU, COLO, DUKE, F, FLAS, ILL & ILLS, LSU, NY, PH, TENN, and WIS) reported progress on digitization deliverables, with a total of 20,062 specimens barcoded (6,680 bryophytes and 13,202 lichens), 10,209 labels imaged (4,009 bryophytes and 6,200 lichens), 5,956 specimens imaged (1,559 bryophytes and 4,397 lichens), 5,881 specimen records uploaded to the portal (2,356 bryophytes and 3,525 lichens), 15,361 skeletal records created (2,618 bryophytes and 12,743 lichens), 14,494 labels fully transcribed (12,831 bryophytes and 1,663 lichens), and 16,702 specimens georeferenced (1,649 bryophytes and 15,053 lichens).

Table 1: Digitization progress by GLOBAL collaborators in 2021-Q1, separated by Bryophyte (B) and Lichen (L) specimens.

	# Barcodes Added		# Labels Imaged		# Specimens Imaged		# Uploaded to Portal		# Skeletal Records Created		# Fully Transcribed		# Georeferenced	
	B	L	B	L	B	L	B	L	B	L	B	L	B	L
ALA														
ASU														
BRY		350												
CINC & MU	254						254				6,138	45	125	3,250
COLO		2,279		2,279				3,048		3,048		3		
DUKE	1,016		1,231		668	1,834	572	50	329		89		10	
F	1,500	2,624	339	2,624					1,500	1,663				
FLAS	3,765	577	1,551	461			1,258	367	367		62	20		
ILL & ILLS											5,000			
LSU	47	8			3		47	8	8			503		483
MICH														
MIN														
MO														
MSC														
NY	157	7,042	542	514	542	514			157	7,042	1,289	101	757	28
OSC														
PH	121		121		121									
TENN			225		225		225	52	257		253	1	62	
UC														
WIS		322		322		2,049				990		990	695	11,292
YU														
Totals	6,860	13,202	4,009	6,200	1,559	4,397	2,356	3,525	2,618	12,743	12,831	1,663	1,649	15,053
B+L Totals		20,062		10,209		5,956		5,881		15,361		14,494		16,702

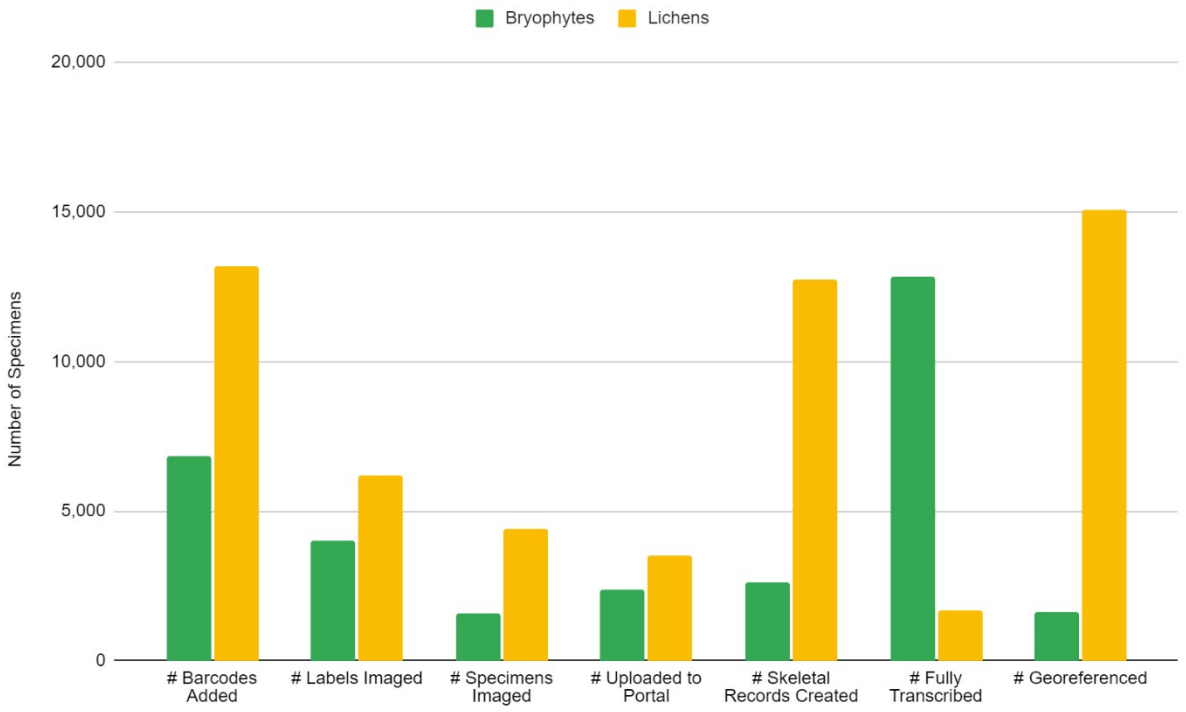


Figure 1: Digitization progress for the GLOBAL collaboration in 2021-Q1, separated by Bryophyte and Lichen specimens.

Share Best Practices, Standards, and Lessons Learned

Flexible Workflows

COVID restrictions continue to require flexible digitization workflows. Some institutions are still using remote imaging stations (F, DUKE), and others are making use of off-site activities like label transcription (COLO, CINC & MU, DUKE, F, and LSU). COLO and F continue to image only labels in order to provide remote transcribers with material. NY is prioritizing barcoding and skeletal data capture along with imaging to take advantage of limited time in their imaging lab. Those collaborators who are imaging specimens have also demonstrated the importance of flexibility. Two main strategies have emerged: photographing the label and specimen separately or removing the specimen and capturing both label and specimen in a single image. CINC & MU originally chose to take two images, but changed to removing the specimen onto archival paper for a single image, which takes less time overall for them. Alternately, PH began by taking one image, but found that leaving messy specimens with loose dirt inside their packets and photographing them separately from the labels was more efficient. TENN continues to image specimens and labels separately to decrease handling of delicate specimens. DUKE has updated their original workflow for multiple packets per sheet to take one full-sheet photograph and use post-processing to crop to individual specimen labels. NY uses black felt and paper to isolate individual packets on the same sheet for imaging and makes use of an auto-cropping function afterwards. This range of solutions

shows that the "optimal workflow" will vary depending on the institution's set up and the storage, composition, and fragility of the specimens. Rates and workflows are expected to continue to be updated as experience and COVID restrictions progress.

Collaboration

Basecamp, Zoom, and email were used to collaboratively problem-solve workflow challenges and share alternative methods and tools across institutions. The working group for Specimen Imaging held one Zoom meeting in 2021-Q1 where participants demonstrated additional workflows and discussed OCR, post-processing, and file formats. There was a preliminary meeting to begin discussion of Skeletal Data standards and tools. The Georeferencing Working Group met twice, once with a smaller focused group and a second time open to all GLOBAL collaborators to discuss tools, resources, and plans. These meetings allow the group to ask questions, offer feedback, and discuss their own challenges.

The first Management Committee Meeting was held in February open to all GLOBAL team members to review quarterly and cumulative grant progress. The GLOBAL Project Manager (TENN) completed initial meetings with several collaborators in early 2021-Q1 (ILL/ILLS, MIN, NY, OSC, YU) to meet the teams and discuss plans and challenges. She also met weekly with the Portal Manager (ASU) to coordinate project activities and needs.

Leveraging Institutional Expertise

Collaborators continued to post test images of specimens on Basecamp for feedback on image content and quality (CINC & MU, LSU, MSC, NY, PH, WIS, and YU). The GLOBAL Portal Manager assessed shared images for appropriate quality and size and GLOBAL participants offered feedback and advice.

Duplicate Matching

Transcription and georeferencing are time consuming steps of digitization. Duplicate specimens, duplicate collecting events, and exsiccate series are common within and across GLOBAL collections. Utilizing these existing records can reduce the number of specimens that need individual attention. In order to take advantage of this resource, all GLOBAL collaborators have been encouraged to use the duplicate matching tools available in the portals. The GLOBAL Portal Manager also shared a script she developed during the Georeferencing Meeting in March. The script matches possible duplicates in the portal and produces a list of duplicate records with coordinates that could be transferred to those that have not yet been georeferenced. The Portal and Georeferencing Managers began discussing use of the script with each collaborator to confirm those interested in this resource.

Another source of possible duplicates that was explored during 2021-Q1 was international collections data. The Project Manager, Lead PI, and Portal Manager worked with the GLOBAL

team to compile a list of publically shared international collections that could include transcribed and geolocated specimens duplicated in GLOBAL collections. Seventeen international collections were successfully added into the Bryophyte and Lichen Portals during 2021-Q1, providing over 780,000 bryophyte and 540,000 lichen specimens as possible duplicates.

Exsiccate sets are common in bryophyte and lichen collections, so an accurate list of titles is important to help record and match specimens. Lead PI Budke (TENN) worked to update the Bryophyte Portal list and ASU PI Bungartz cleaned up the library of exsiccate in the Lichen Portal, linking exsiccate titles with IndExs (http://www.botanischestaatssammlung.de/DatabaseClients/IndExs/index.jsp). Best practices were shared on how to link exsiccate specimen records during image and skeletal metadata upload. Exsiccate management tools were reconfigured and improved for advanced management and presentation of the exsiccate datasets (ASU).

Share Identified Gaps in Digitization Areas and Technology

Imaging Uploading

Batch image uploading was flagged as the highest priority during 2020-Q4. In 2021-Q1, ASU continued developing a tool to batch upload images into the portals and a method to automate the transfer of images from Dropbox to the ASU server. The method is expected to use the Dropbox API and portal controls are being developed to automate or semi-automate the transfer of images into the ASU system. ASU is also developing use of a third party, open source JavaScript library to establish an additional method for batch submission of images directly within the user interface. An interim image uploading workflow was put in place and tested with five institutions during 2021-Q1. It was made available to the whole GLOBAL team in March and will remain in place until the automated version is available.

Barcode Renaming

Another challenge identified was renaming image files with the specimen barcode. Automated renaming is problematic for images taken inside a specimen packet, in which the barcode may not be visible. Additional improvements were made to the “BarcodeRenamer” tool in 2021-Q1 to improve efficiency with different barcode formats and orientations. Development continued on the “PhotoWatcher,” a program that will rename images as part of the image capturing routine and can be used to harvest skeletal occurrence metadata. The program is being beta-tested and, once a stable and reliable version is available, it will be shared with all other institutions participating in the TCN (ASU).

Georeferencing

Connecting new specimen records with existing coordinates from duplicate specimens, collecting events, or localities offers a wealth of data to save time on georeferencing for the GLOBAL project. However, making those

connections, and assessing the quality of existing georeferences, poses challenges to efficient and effective use of this data resource. In response, ASU worked on developing a georeferencing tool that will harness GEOLocate web services to rapidly georeference batches of specimens with little human effort, as well as a tool to optimize the harvesting of data from specimen duplicates and specimens from duplicate collection events. WIS continued to strategize georeferencing actions across multiple institutions in various stages of progress. Difficulties include inconsistent global geography, completeness of label transcription, and data management among institutions. Questions remain on standards for some geographic fields.

Taxonomic Filters

Specimen data outside of the Bryophyte and Lichen Portals offers the opportunity for additional information as well as possible duplicates to update incomplete records on existing specimens. One challenge is separating the particular records of interest from a larger collection of data. The GLOBAL Portal Manager and ASU team worked to develop taxonomic filters to isolate bryophyte records from plant data and separate lichenized, lichenicolous and allied fungi from occurrence records of non-lichenized fungi.

Share Opportunities to Enhance Training Efforts

Symbiota

ASU created new versions of the Symbiota checklist tutorial video and a tutorial on using the new label generation tools in Symbiota portals. All videos now available on the Symbiota YouTube channel (https://www.youtube.com/channel/UC7gIMVLRnTA6ES3VTsci7iQ). A guide for use of the portals' Skeletal Data Entry tool was developed between the GLOBAL Portal Manager (ASU) and the Project Management team at TENN.

Transcription

The GLOBAL Project Manager began compiling transcription resources during 2021-Q1 to share on Basecamp and the project website. Resources were submitted by several collaborators including CINC & MU, COLO, LSU, and WIS. A student worker at CINC recorded a video on techniques for deciphering and translating non-English, hand-written herbarium labels (https://www.youtube.com/watch?v=M_rPFzPfZTM&feature=youtu.be). LSU shared a guide to Brazilian road markers and TENN uploaded a copy of a guide to Japanese geography developed by a former student. COLO project manager Allen recorded a series of videos on how to install and use the Imaging Workflow Application from the LBCC TCN. Several institutions under the project are reviewing this system for the project. He also recorded a video to help students transcribe from images in the portal.

Georeferencing

The GLOBAL Management team led a Georeferencing Meeting with the GLOBAL group to review available georeferencing resources and tools. WIS continued to post resources on shared pages for georeferencing help

sites and training manuals and began meeting with participating institutions to discuss georeferencing readiness and strategies.

Share Collaborations with other TCNs, Institutions, and/or Organizations

F connected with members of the TCN's External Advisory Committee to confirm their invitations, roles, and availability to participate. All responses were positive and the Executive Committee plans to meet to discuss how to best engage this resource.

ASU, in collaboration with the International Union for the Conservation of Nature (IUCN), established an agreement to share a Global Checklist of Red-Listed Lichen Species through the portal. That list is now online at: https://lichenportal.org/cnalh/checklists/checklist.php?clid=1430&pid=0. Species listed as part of the Global Fungal Red-List Initiative will be added as soon as their assessments are published. Detailed occurrence records of red-listed species in the portal are protected and accessible only to a select group of data managers and conservation specialists.

Lead PI Budke (TENN) sent emails out to contacts at twenty international collections chosen to utilize as data snapshots for duplicate matching in the portals. The correspondence introduced the GLOBAL project and made some initial connections between the project and the broader international community. All feedback received was very positive and collections will be acknowledged on the GLOBAL project website.

The GLOBAL Project Manager (TENN) reached out to the NSF Program Officer to discuss and confirm reporting requirements and COVID related extension protocols. She also participated in the BIO-wide Virtual Office Hours: COVID-19 Response on March 2.

The GLOBAL Lead PI and Project Manager (TENN) participated in the February iDigBio Quarterly IAC meeting to connect with other active TCN's.

Share Opportunities and Strategies for Sustainability

Portal Management

Katie Pearson was hired as the GLOBAL Portal Manager and was introduced to the community. She was provided with necessary logins and data management access and reached out to existing portal contributors. She will be responsible for providing support and monitoring the help emails of the Lichen and Bryophyte Portals (ASU).

ASU established SSL security certificates for the Bryophyte and Lichen Portals and three associated portals (Arctic Lichen Flora, Frullania, and Líquenes en América Latina), which involved remapping some of the internal images to use the https protocols. Various image URLs were also remapped to https protocol for institution that have recently installed SSL certificates. They will continue to encourage and assist other institutions hosting external images mapped into the portal to install SSL certificates.

ILL & ILLS transitioned the MyCo Portal to the same code base, Symbiota-light, that is used for the Bryophyte and Lichen Portals.

Taxonomy

A meeting to discuss plans for Bryophyte taxonomy was held in February involving participants from ASU, DUKE, F, MO, and TENN. An updated version of the nomenclature backbone for the Bryophyte Portal was completed by MO and is in the process of being integrated by ASU. Data maintenance was completed on the taxonomic thesaurus in the lichen consortium, and new names continued to be added and linked with synonyms and basionyms (ASU).

Share Education, Outreach, Diversity, & Inclusion (EODI) Activities

F and TENN worked on an abstract (on behalf of all TCN participants and institutions) to be submitted in April 2021 to the online conference: Bryophytes, lichens, and northern ecosystems in a changing world (July 6-9, 2021). The Québec B(ryophytes) and L(ichens) (BL2021) will bring together four major bryological, lichenological and botanical societies: the
International Association of Bryologists (IAB), the American Bryological and Lichenological Society (ABLS), the Canadian Botanical Association (CBA-ABC) and the Société québécoise de bryologie (SQB).

ASU participated in discussions on how to best present the new portal to the international research community (with the goal to have more international communities join). Data management agreements were developed with international collections joining the Lichen Portal, particularly collections from South America. ASU PI Bungartz continued outreach to lichenologists from Latin America by organizing a Symbiota workshop (in Spanish) for the Consorcio de Herbarios de Líquenes en América Latina during the 9th Symposium of the International Association for Lichenology, in Brazil August 1-6; the two day workshop is being planned as post-congress event: https://doity.com.br/ial9/blog/workshops.

The GLOBAL Project Manager met with TENN OIT staff for training on website development and management in February. She began building the GLOBAL WordPress website in March. Project information, progress, and resources have been added and the site is to be launched publicly in April.

Share Information About Your Website and/or Portal Usage

The GLOBAL Project Manager built the GLOBAL WordPress project website in March, to be launched publicly in April.

Share Other Activities and/or Progress

Imaging Tagging

ASU made the first steps towards developing an online glossary and image library as part of comprehensive revision of character matrix that drives identification keys, with the goal of tagging images with controlled vocabulary of key characters.

Lepidoptera of North America Network &
Symbiota Collections of Arthropods Network (SCAN)
Quarterly Report

May 4, 2021
Neil Cobb

Progress in Digitization Efforts:

This is a joint report for the two Thematic Collections Networks (TCNs) SCAN and LepNet. Many museums are involved in both SCAN and LepNet, including collections that have received funding from both TCNs, collections that are unfunded for one TCN and funded by the other, and some collections that are providing data to both and are unfunded by the ADBC program. Both TCNs share the same database https://scan-bugs.org/portal/, which depending on the context we refer to as the SCAN-LepNet database or the LepNet-SCAN database. We will also serve arthropod data for InverteBase and will serve Terrestrial Parasite Tracker TCN data when it becomes available (See TPT TCN report for details). Summary statistics presented here were compiled from data accessed on the SCAN portal, May 3, 2021. **Table 1** shows the key statistics of Lepidoptera (LepNet) and non-Lepidoptera (SCAN) records to date. These consist of all records and images, including records and images from data providers who have allowed us to post their data on the SCAN/LepNet portal. Providing data from these additional providers increases our ability to georeference, add to taxonomic tables, and more accurately assess the total digitization effort for any given taxon. We provide data specific to institutions that received direct funding from the NSF-ADBC program in the annual reports to NSF.

Table 1. Records in SCAN/LepNet database, "all data" reflects all arthropod taxa, "Non-Lep" includes all non-Lepidoptera arthropod data, and Lepidoptera includes only Lepidoptera taxa.

	All data	Non-Lep (SCAN)	Lepidoptera (All Leps)
Specimen Records	27,723,542	22,222,586	5,500,956
# Georeferenced	24,028,873	19,098,198	4,930,675
# Imaged	6,669,125	4,135,818	2,533,307
# Identified to species	17,199,052	11,892,259	5,306,793

The SCAN network started in 2012 and the TCN funding has ended, but SCAN continues to support PEN projects. The LepNet grant was initiated on July 1, 2016 and there are currently 27 ADBC funded museums and one non-funded museum (Oklahoma State University). The museums comprising the NSF-ADBC LepNet are all serving records and images on the LepNet Portal and are serving data directly to iDigBio via IPT or through DwC archives on the LepNet-SCAN portal. Twenty museums are serving DwC archives to iDigBio and six museums are serving data snapshots with the LepNet portal. We have set up the SCAN Portal to serve all arthropod data from North America as well as all data from North American arthropod collections where specimens were collected outside of North America.

LepNet - The LepNet ADBC-funded museums are still on target to meet goals for records and images. An additional 59 collaborators (non-ADBC funded museums that use our data portal to serve their data) have also provided additional records for Lepidoptera. There are 49 collections (referred to as added-value) that have allowed us to harvest their data via IPT to serve

lepidopteran records. Although most of the Lepidoptera imaged are from iNaturalist, 352,841 are specimen images **Table 2** shows the top 10 families of Lepidoptera in terms of total occurrences digitized.

What is most encouraging about the lepidopteran records is that 97% of the records are identified to species, which is higher than any of the other major orders. Thus, the primary factor limiting the production of "research-ready" data is due to georeferencing. For Lepidoptera 87% of the records are research-ready (i.e., identified to species and georeferenced) and by georeferencing existing records we should increase that percentage to 90% over the next three years. We realize that many records represent misidentified specimens and we also need to seek additional non-ADBC funding to review as many specimen identifications as possible. We are sponsoring three LepNet Partners to Existing Networks (PEN) grants (San Diego Natural History Museum, University of Wisconsin, and University of New Hampshire).

Table 2. The number of occurrence records for the top 10 families of Lepidoptera that have been digitized.

Family	Specimens	Georeferenced	Species ID	Georeferenced & Species ID
Nymphalidae	1,149,175	94%	99%	94%
Noctuidae	699,835	92%	97%	90%
Erebidae	595,209	87%	97%	84%
Geometridae	472,292	89%	96%	85%
Pieridae	410,921	86%	99%	85%
Hesperiidae	394,167	88%	98%	86%
Lycaenidae	323,336	94%	98%	93%
Papilionidae	227,740	88%	99%	88%
Crambidae	214,199	90%	97%	88%
Tortricidae	158,562	85%	95%	80%

Symbiota Collections of Arthropods Network (SCAN) - We have surpassed our overall TCN/PEN goals for the network and have been very successful in supporting data mobilization for unfunded museums and cooperation by larger collections that have allowed there data to be used to help mobilize data from other museums. We are sponsoring one SCAN PEN proposal, one through the American Museum of Natural History, focusing on several ground-dwelling families. **Table 3** shows data for the five major taxa we targeted in SCAN. All five groups have enough data to produce several papers, despite only 60% of the records with species-level identifications, accounting for 51% of the records being research-ready when you factor in percent records that are georeferenced.

Share and Identify Opportunities to Enhance Training Efforts: We are developing resources on the WordPress site http://www.scan-all-bugs.org/ .

Share and Identify Best Practices and Standards (including Lessons Learned):

Table 3 Number of records for the five focal SCAN taxa groups.

Taxa	# Specimen Records	# Georeferenced	# Specimen Identified to species	# Georeferenced & Ided to species
Formicidae	1,307,785	1,194,586	658,273	601,327
Carabidae	732,485	623,238	444,749	382,368
Acrididae	533,401	437,668	314,041	266,459
Araneae	485,564	461,544	443,291	325,573
Tenebrionidae	250,976	219,603	129,508	115,732

We share best practices on the SCAN/LepNet project website https://scan-all-bugs.org/ .

Images for Research - We developed a new and efficient process for uploading images to the database https://scan-bugs.org/portal/profile/index.php?refurl=/portal/imagelib/imagebatch.php? .We are participating in a TDWG-sponsored working group to develop standards for specimen images, including definition of morphological traits.

Identify Gaps in Digitization Areas and Technology: We are supporting the "LightingBug" project https://lightningbug.tech/, which will exponentially increase transcription rate of labels and produce specimen images comprising 360-degree image suites. The production of images will be transformational in terms of extending our capabilities to provide automated identifications and examine morphological traits.
We continue to seek out occurrence data to better understand the biogeography of the focal SCAN taxa and Lepidoptera. For most groups there is not enough data to talk about gaps. We are meeting this need by incorporating additional collections into the SCAN-LepNet database, and harvesting observational records from iNaturalist, Pollardbase, Buguide, LepSoc inventories, and smaller observation sets provided by individual lepidopterists.

Share and Identify Collaborations with other TCNs, Institutions, and Organizations:
We are currently only working with the Terrestrial Parasite Tracker TCN. We are also generally collaborating with a variety of individuals, projects and organizations to extend the ability to mobilize biodiversity data and promote the use of data in research. We are serving data from 246 collections, we continue to add one collection per month.

Share and Identify Opportunities and Strategies for Sustainability: Nothing to report

Other Progress (that doesn't fit into the above categories):

Focus on North American Arthropods We continue to provide North American data obtained from any credible sources to increase the quantity of data available to SCAN and LepNet users. We have added 29 new collections since the update.

GBIF Registration - There are 49 Live collections on SCAN that are now registered with GBIF and 93 other entomology collection datasets from the North America being served on GBIF for a total of 181 datasets. This leaves approximately 28 collections in North America that still need to register on GBIF.

Publications - We have published an overview of the LepNet project (Seltmann et al 2017), and several LepNet participants collaborated on a publication below (Belitz et al., 2018). Our review of North American entomology collections has been published in PeerJ. We are now developing a follow up review on completeness in bee data for the United States.

Belitz, M.W., Hendrick, L.K., Monfils, M.J., Cuthrell, D.L., Marshall, C.J., Kawahara, A.Y., Cobb, N.S., Zaspel, J.M., Horton, A.M., Huber, S.L. and Warren, A.D., 2018. Aggregated occurrence records of the federally endangered Poweshiek skipperling (*Oarisma poweshiek*). *Biodiversity data journal*, (6).

Cobb, N.S., L. Gall, J.M. Zaspel, L.M. McCabe, N.J. Dowdy. and A.Y. Kawahara. 2019 Assessment of North American Entomology Collections: Prospects and Challenges for Addressing Biodiversity Research. PeerJ, 7, p.e8086.

Google Analytics: Our Google Analytics data are dynamically shown https://datastudio.google.com/u/0/reporting/1VvEU4pM2LGqQXY0hVCTf98VvGmM7T_bu/page/cLZN for the SCAN portal, http://scan-bugs.org/portal/index.php .

TCN Quarterly Progress Report

Prior to each IAC meeting, TCNs are asked to complete a quarterly progress report in the areas outlined below. The TCN Lead PI or Project Manager collects information from all collaborators and compiles them into one overall progress report for the TCN. The TCN Lead PI or Project Manager then submits the quarterly reports via an email to Cat Chapman. An archive of previously submitted reports is available on the Internal Advisory Committee wiki page.

TCN Name

The Microfungi Collections Consortium: A Networked Approach to Digitizing Small Fungi with Large Impacts on the Function and Health of Ecosystems (MiCC)

Person Completing the Report

Andrew Miller (Lead PI)

Share Progress in Digitization Efforts

- BPI has imaged 40,000 of their 120,000 exsiccati. These images will be uploaded to the MyCoPortal over the summer. Another 39,387 images were uploaded.
- The University of Gotherburg (GB) has been added to the MyCoPortal along with their 117,684 records.
- The ILLS team has completed georeferencing 4491 records from CLEM, 4356 records from NCSLG, 2013 records from VPI, 1078 records from UCSC, and 3040 records from IND.
- PH georeferenced 12,733 records.
- ISC georeferenced 11,629 records.
- The MyCoPortal now contains 7,515,986 records, 59% of which are georeferenced, and 2,301,828 images.

Share Best Practices, Standards, and Lessons Learned

- Nothing new to report.

Share Identified Gaps in Digitization Areas and Technology

- Nothing new to report.

Share Opportunities to Enhance Training Efforts

- Nothing new to report.

Share Collaborations with other TCNs, Institutions, and/or Organizations

- The MyCoPortal has successfully transitioned from Symbiota-standard code base to Symbiota-light and continue working with Ed Gilbert and the GLOBAL TCN to update our taxonomic thesaurus.

Share Opportunities and Strategies for Sustainability

- Nothing new to report.

Share Education, Outreach, Diversity, & Inclusion (EODI) Activities

- Nothing new to report.

Share Information About Your Website and/or Portal Usage

Please see attached report from Google Analytics.

Share Other Activities and/or Progress

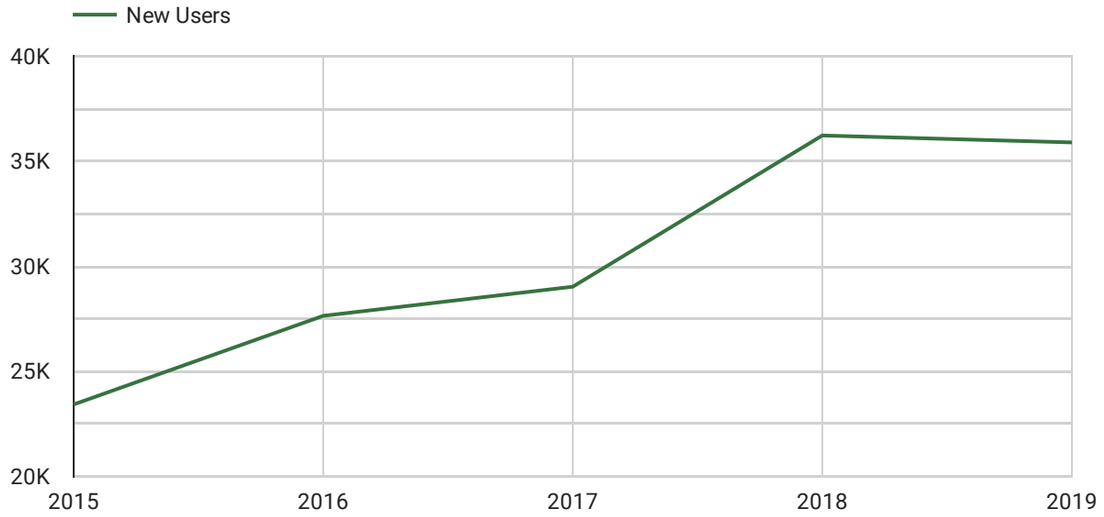
- MyCoPortal has been cited 102 times - 16 times in 2021, in peer-reviewed journal publications.

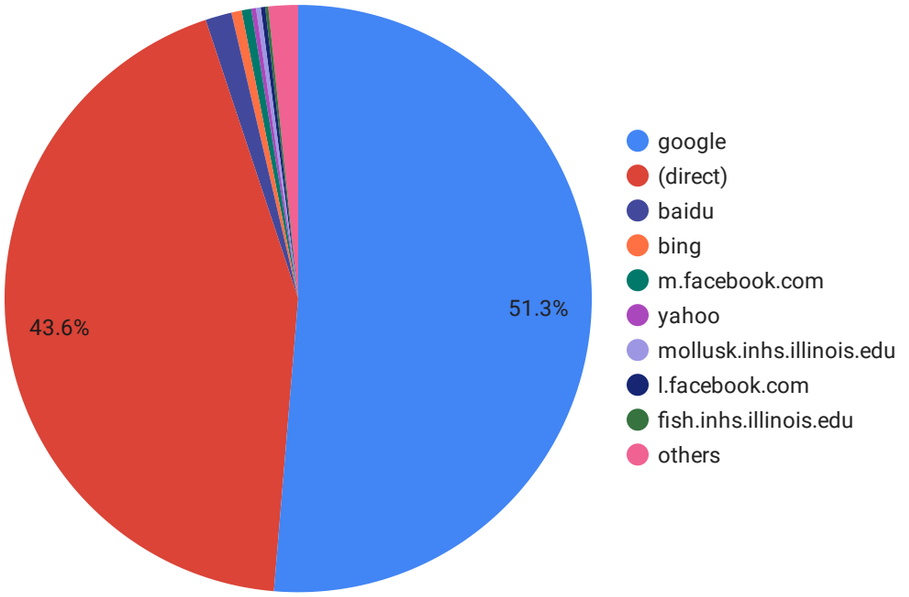
MyCoPortal Data Portal Statistics

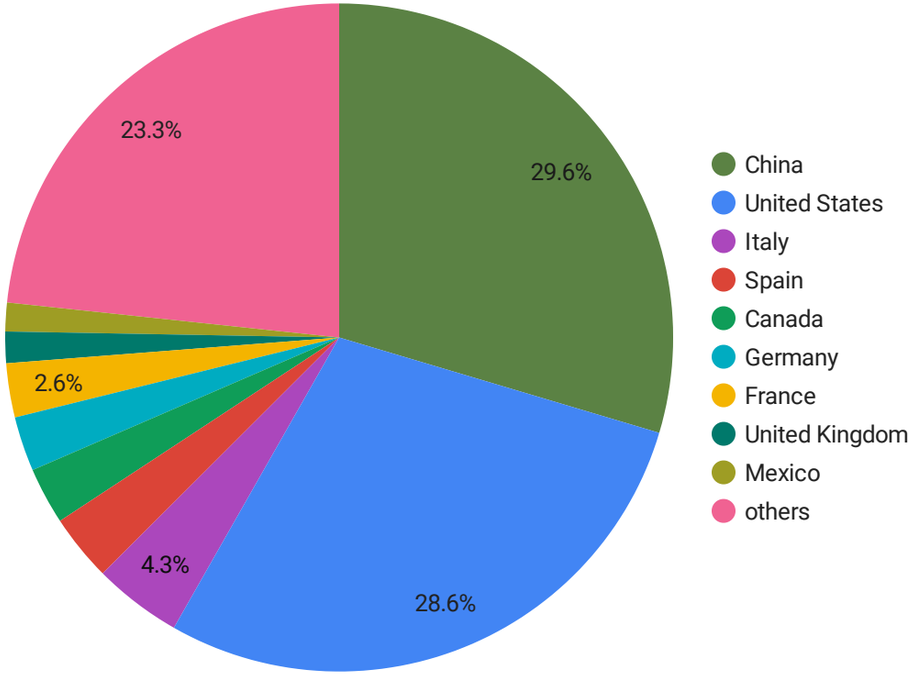
www.mycoportal.org Data from Google Analytics

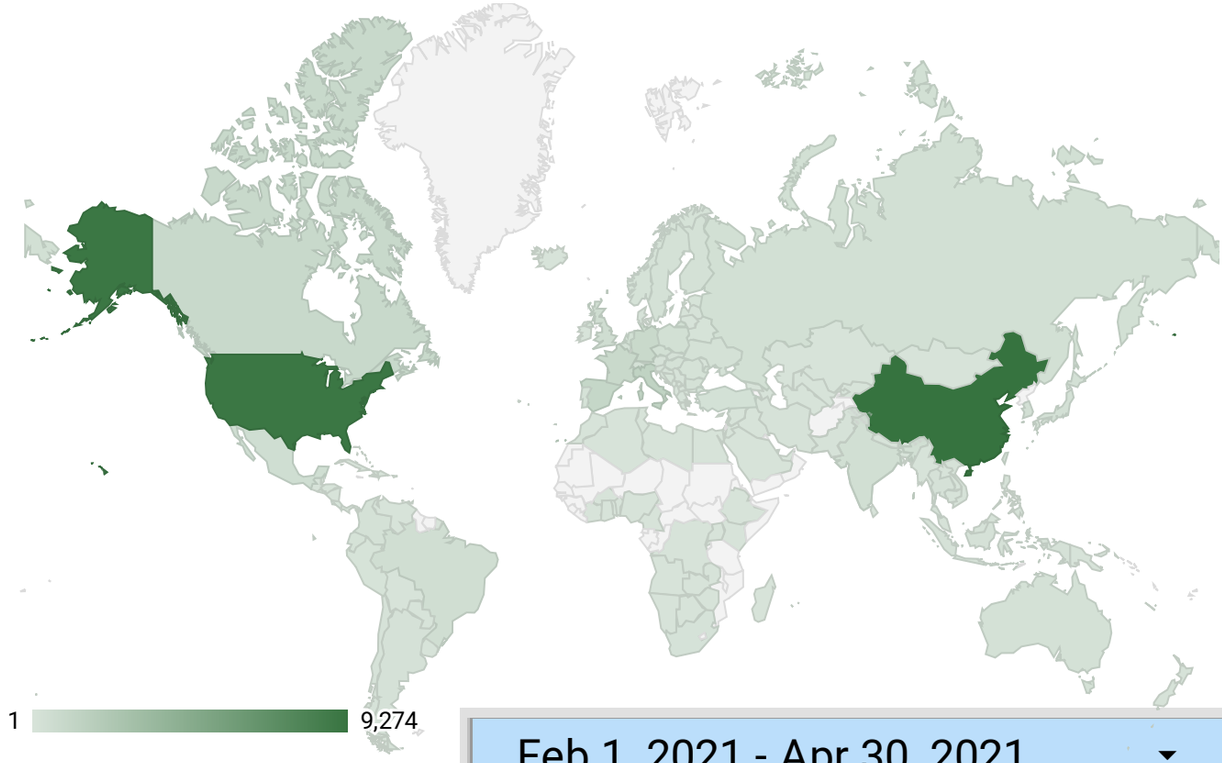
MYCOLOGY COLLECTIONS PORTAL

Users	New Users	Sessions	Number of Sessions per User	Pageviews	Pages / Session	Avg. Session Duration	Bounce Rate
20,541	19,388	31,322	1.52	82,682	2.64	00:02:42	69.99%









Feb 1, 2021 - Apr 30, 2021

iDigBio
Integrated Digitized Biocollections

TCN Quarterly Progress Report

Prior to each IAC meeting, TCNs are asked to complete a quarterly progress report in the areas outlined below. The TCN Lead PI or Project Manager collects information from all collaborators and compiles them into one overall progress report for the TCN. The TCN Lead PI or Project Manager then submits the quarterly reports via an email to Cat Chapman. An archive of previously submitted reports is available on the Internal Advisory Committee wiki page.

TCN Name

The Pteridological Collections Consortium: An integrative Approach to Pteridophyte Diversity Over the Last 420 Million Years (PCC).

Person Completing the Report

Amy Kasameyer (Project Manager).

Share Progress in Digitization Efforts

For extant specimen progress during this reporting period, Pteridophyte Collections Consortium members created skeletal records for **3,293** specimens, fully transcribed **28,527** specimens, imaged **9,385** specimens, and geo-referenced **10,277** specimen records. The total pteridophyte extant specimen progress including work done prior to the start of the grant is **598,738 (36% of goal)** skeletal records created, **1,095,661 (66% of goal)** extant specimens imaged, **1,066,756 (65% of goal)** extant specimens fully transcribed, and **277,224 (17% of goal)** extant specimens geo-referenced.

In the Pteridoportal (http://pteridoportal.org), we currently have **1,640,507** extant specimen records, **1,290,078** (79%) of which are imaged and **415,247** (25%) of which are georeferenced.

For fossil specimen progress during this reporting period, Pteridophyte Collections Consortium members databased **2,962** specimens, imaged **694** specimens, and geo-referenced **909** specimen records. The total pteridophyte fossil specimen progress including work done prior to the start of the grant is **34,280** (39% of goal) specimens databased, **30,412** (35% of goal) specimens imaged, and **15,663** (18% of goal) specimen records geo-referenced.

UF UNIVERSITY of FLORIDA
FLORIDA STATE UNIVERSITY 1851
FM FLORIDA MUSEUM
NSF

In the Pteridoportal (http://pteridoportal.org), we currently have **11,991** fossil specimen records, **9,085** (76%) of which are imaged and **5,275** (44%) of which are georeferenced.

Portal Manager Joyce Gross loaded the Bernice P. Bishop Museum fern data (45,752 specimen records) and the Lundell and University of Texas at Austin herbaria fern data (25,011 specimen records) into the Pteridoportal and updated the iNaturalist data in the Pteridoportal.

Collections continue to be impacted by COVID. Onsite activities at most institutions are limited, safety procedures are requiring work to be done at a slower rate, and at most institutions in-person staff are still restricted and volunteers are interns are not able to work in-person.

Institutions continue to make significant progress through remote work, for example, at **CHRB** transcription and georeferencing of PCC TCN specimens is being done virtually by their large group of just under 40 undergraduate students.

Share Best Practices, Standards, and Lessons Learned

N/A

Share Identified Gaps in Digitization Areas and Technology

N/A

Share Opportunities to Enhance Training Efforts

ALA: We have trained a new curatorial assistant, Stephanie Zaborac-Reed. She has done an amazing job in georeferencing 1900 records since January.

CHRB: Training was provided in March to a select few credit interns for georeferencing of specimens. PCC TCN specimens were georeferenced as a result of this!

UCMP: Diane Erwin provided training to two UCB students: Jaemin Lee (UCB PhD candidate, paleobotany focus) and James Pinto (participant on the project through the Undergraduate Research Apprentice Program). Both students were given the opportunity to work on collections with pteridophytes that allowed them to become immersed in specimen identification, databasing, georeferencing, literature searches, and imaging.

Vermont: Two new students are being trained who will replace students that are graduating.

Share Collaborations with other TCNs, Institutions, and/or Organizations

Partners are collaborating with the following TCNs: Endless Forms, TORCH, CAP, Cretaceous World. Specific collaborations include:

BRIT: As a result of another funded TCN (TORCH TCN funding) BRIT has digitized additional Pteridophyte records from three other institutions that hold pteridophyte specimens. These are (with herbarium code and counts representing images and digital records): Abilene Christian University (ACU; 118), University of North Texas (NTSC; 198), and Tarleton State University (TAC; 72). We expect to add two other collections (also funded by the TORCH TCN) that may also hold Pteridophyte specimens: Texas Women's University (TCSW) and Hardin-Simmons University (HSU).

FIELD: In contact with subawards as well as with University of California, Berkeley with their Carboniferous Citizen Science event that we hosted on their behalf with our Collection Club volunteers.

NYBG: Staff funded by the TORCH and Endless Forms TCNs are working to georeference collection localities for these projects. Some of the localities overlap with collections for the PCC TCN, resulting in georeferenced specimens for this project

Share Opportunities and Strategies for Sustainability

N/A

Share Education, Outreach, Diversity, & Inclusion (EODI) Activities

ALA: Used specimen images to create a key of Lycopodiaceae of Alaska as part of a student project in Systematic Botany BIOL331 taught by Steffi Ickert-Bond.

BRIT: T. Rehman presented (2021-04-17): "Collections, Conservation, and Community Science: A Collaboration to Conserve our Planet's Natural Heritage" - for the Fort Worth Public Library Earth Day 2021 Events in which currently funded TCNs were highlighted and participants were encouraged to explore their local natural history collections and participate in the weekly Armchair Botany Zoom sessions (https://www.fortworthtexas.gov/departments/library/events/earth-day-2021?BestBetMatch=rehman|a720ba43-6954-4033-b1ac-a72478fc15dc|f3d8342e-e792-4add-b724-7420243aa350|en-US; https://www.youtube.com/watch?v=0945UpOxnFw)

T. Rehman, A. Bordelon, D. Barroso, J. Best, C. Barrett presented (8, 10, 11 April 2021): three WeDigBio Transcription Blitzes including a short presentation about the TORCH TCN, a behind-the-scenes tour of the BRIT herbarium (8 April), a behind-the-scenes tour of the Sumner Molecular and Structural Lab at BRIT (10 April; led by BRIT botanist Alejandra Vasco), and a presentation about the collecting of herbarium specimens, including live tour of BRIT specimen receiving room and collecting supplies (12 April). Attendees from multiple countries were in attendance (brit.org/wedigbio)

FIELD: Hosted the second of four quarterly Collections Club events on Thursday, Friday and Saturday April 8, 9, and 10, 2021. The event had a lot to celebrate recognizing Global Citizen Science Month, Earth Day and WeDigBio. The event continued its virtual format which brought new opportunities as volunteers were able to participate from Indiana, Seattle and even as far as Indonesia. 129 community scientists collectively contributed to an astonishing figure of over 44,299 curated records. Volunteers worked on several projects including transcribing scientific label data from fern specimens (1039 fern records using the crowd sourcing module), cleaning and curating data for bibliographic records, and host/parasite data for the Insect Collection. For this Collections Club, we also had a special project from the University of California Berkeley. Cindy Looy, Benjamin Muddiman, and Ivo Duijnstee showed participants how to transcribe data of plant remains encased in limestone concretions or "coal balls", which preserve plant tissue 100's of millions of years old.

NYBG: Hosted a WeDigBio 2021 event on April 8-11. Approximately 325 guests attended live virtual events including an herbarium collections tour, a specimen digitization workshop, botanical research seminars, and a specimen transcription event. Participants assisted in transcribing ±4,700 specimen records for multiple TCNs including the PCC TCN, and recordings of the events are now freely available online through the NYBG Lecture Library.

TEX: George Yatskievych gave one online lecture to an Economic Botany class on herbaria that mentioned the PCC efforts.

UCB: Cindy Looy and Carl Rothfels taught two Jepson Workshops open to the public: F*iddleheads! Become fern fluent* and *Life in the Carboniferous swamps*. Diane Erwin and the UCMP's PCC fern team made regular postings to the project's Twitter, Facebook, and Instagram.

Share Information About Your Website and/or Portal Usage

During this reporting period, we had 4500 visitors to the Pteridoportal with an average session duration of 17 minutes. We had 178 visitors to our project website with an average session duration of 49 seconds.

Share Other Activities and/or Progress

N/A

TCN Quarterly Progress Report

Prior to each IAC meeting, TCNs are asked to complete a quarterly progress report in the areas outlined below. The TCN Lead PI or Project Manager collects information from all collaborators and compiles them into one overall progress report for the TCN. The TCN Lead PI or Project Manager then submits the quarterly reports via an email to Cat Chapman. An archive of previously submitted reports is available on the Internal Advisory Committee wiki page.

TCN Name

TCN Name and short code, such as: Enhancing Access to Taxonomic and Biogeographical Data to Stem the Tide of Extinction of the Highly Imperiled Pacific Island Land Snails (PILSBRy)

Person Completing the Report

Name and role of the person completing the report, such as: Norine Yeung (Lead PI)

Share Progress in Digitization Efforts

Share information here. You can also embed graphics if desired.

- All 6 collections have uploaded their Pacific Island land snail specimen records onto the PILSBRy symbiota portal for taxonomy clean up and addition of georeferenced data.
    - 376123 lots, 3681391 specimens, 109935 georeferenced records.
- The taxonomic authority file now contains 4080 names including authorship information.

Share Best Practices, Standards, and Lessons Learned

Share information here. You can also embed graphics if desired.

We are continuing to tweak workflow protocols for taxonomic and georeferencing cleanup and imaging protocols.

Share Identified Gaps in Digitization Areas and Technology

Share information here. You can also embed graphics if desired.

COVID19 continues to impact access to collections and recruitment of staff and volunteers. For example, per MCZ: work from home; staff occupancy in MCZ limited to 25 -50%. (= 1 staff member 1 day per week). This is similar at the Field Museum and UMMZ.

Share Opportunities to Enhance Training Efforts

Share information here. You can also embed graphics if desired.

Similar to the last report, COVID19 continues to impact physical training within the collection. However, we have created ppt presentations, protocol documents, videos, and organized virtual sessions to facilitate training within the project. Discussion are through Slack and meetings are recorded via Zoom and shared via Google Doc through Slack.

Share Collaborations with other TCNs, Institutions, and/or Organizations

Share information here. You can also embed graphics if desired.

We are continuing to work with local conservation agencies to incorporate species data and GPS information. Additional geographic areas besides Hawaii are now parsed out to the various collections to start tackling GPS data.

Share Opportunities and Strategies for Sustainability

Share information here. You can also embed graphics if desired.

Share Education, Outreach, Diversity, & Inclusion (EODI) Activities

Share information here. You can also embed graphics if desired.

Most of the collections are still constrained by the COVID19 pandemic. But any recruitment of staff or volunteers/interns are cognizant in providing opportunities to minorities and reaching out to local residents. However, we have been able to recruit women in science and Pacific Islanders as interns and volunteers for this TCN.

We have been able to post youtube #collectionbattles to educate the public the about the TCN project and general knowledge of molluscs. Will ask the collective at the next meeting on how to spread the word more broadly. Requesting potential battles across TCNs also.

Share Information About Your Website and/or Portal Usage

Share information here. You can also embed graphics if desired, such as from Google Analytics.

Share Other Activities and/or Progress

Share information here for things that do not fit into the above categories. You can also embed graphics if desired.

TCN Quarterly Progress Report

Prior to each IAC meeting, TCNs are asked to complete a quarterly progress report in the areas outlined below. The TCN Lead PI or Project Manager collects information from all collaborators and compiles them into one overall progress report for the TCN. The TCN Lead PI or Project Manager then submits the quarterly reports via an email to Cat Chapman. An archive of previously submitted reports is available on the Internal Advisory Committee wiki page.

TCN Name

Using Herbarium Data to Document Plant Niches in the High Peaks and High Plains of the Southern Rockies - Past, Present, and Future (SoRo)

Person Completing the Report

J Ryan Allen Project Manager

Share Progress in Digitization Efforts

Collectively for the current quarter roughly February 2021-April 2021 we have entered 10,758 new records into databases, barcoded 28,678 new specimens, imaged 33,099 new specimens and georeferenced 65,315 new records.

Our overall project totals are: 431,231 new database records, 944,441 newly barcoded specimens, 948,025 new images and 357,366 new georeferences.

The project after ~44 months (out of 48) has completed.

Data Entry 78.1%

Barcodes 108.6%

Images 110.7%

Georeferencing 58.0%

Most collections still digitizing are on reduced or hybrid access with a limited number of in person hours. Our data entry and georeferencing can be completed remotely, but in person tasks like imaging and

barcode have limited availability. We were fortunate that most collections were moving into the data entry and georeferencing phases when access became limited.

CO-PI Duran left ALAM for other employment. We attempted to find someone to take over the project but have been unsuccessful. At the request of the ALAM grants office COLO is now in the process of transferring the remaining ALAM subaward budget back to the COLO budget. All vascular plant specimens were imaged at COLO after being barcoded at ALAM. ALAM estimated 8,300 specimens for the project, but the herbarium only contained 5,161 vascular plants (the balance of the accessions were lichens and bryophytes). 294 transcriptions were completed at ALAM. Once the funds are back on the COLO budget we will complete as much of the transcription as we can with the remaining budget.

COLO hit a bit of a bottleneck for the project this quarter. Our database was moved to a new server and was not available for transcription work for most of April. Rather than stopping work transcribers worked on the WSC database completing 1844 specimens. COLO also was running low on transcribed specimens to georeferenced in our centralized workflow and worked on georeferencing CSCN records. 2745 CSCN specimens were completed this quarter. COLO plans to assist subawards where needed to complete as much digitization as possible. This progress is being recorded with the original scope of work not with progress at COLO.

HUH is working through a technical issue that is making it difficult to calculate the number of new images captured. These specimens will be added to the next report.

Share Best Practices, Standards, and Lessons Learned

SJNM: When the images were first placed in our server, they were placed in folders in 5000 image increments. It was suggested that we change the structure so that it followed the same format at other institutions. This was done with the folders renamed as SJNM-V-00X (where X=1 to 7) with up to 10,000 images per folder.

UNM: Productivity has been low. Student employees are working from home on laptops after spending hours on screen for their course work. Georeferencing on a laptop is not as efficient as working with the double-screened computers we purchased for this grant. We will not reach our goals within the original project period.

Share Identified Gaps in Digitization Areas and Technology

GREE: We have yet to upload any images to symbiota. We are in the process of moving images online via Cyverse as recommend by Ryan. This is in process by the intent is to have all images available online ASAP.

HUH: Georeferenced data has not yet been ingested into out local database or shared with the Southern Rockies portal. In the next quarter, we plan to complete the development necessary to ingest data from GeoLocate into the HUH Specify database.

We are still working on local infrastructure (or local control) for images captured under the project for several collections. As we are finishing up digitization budgets, making all of the data collected under the project available is a primary focus of the last part of the project.

Share Opportunities to Enhance Training Efforts

HUH: Because many of our staff are new georeferencers, we would be interested in whether the lead has time to QC a random sample of our georeferences, particularly those that were unable to be corrected and marked "skipped."

Share Collaborations with other TCNs, Institutions, and/or Organizations

COLO is also on the GLOBAL TCN, we have been sharing resources and tips from the SoRo TCN to help the project.

Share Opportunities and Strategies for Sustainability

We have been working with local collections to get project data backed up on local servers. SJNM has been working to get images available online on local servers. CSCN is working to build backups on local servers. A central goal of SoRo was to enable infrastructure at local institutions putting funding for backups on collection budgets so future projects also have infrastructure.

Share Education, Outreach, Diversity, & Inclusion (EODI) Activities

BHSC: We are currently planning a workshop for high school teachers through Oglala Lakota College to occur on May 15th 2021 on the Wingsprings preserve in Indian Trust Land.

COLO: PM Allen created a module using project metrics from the SoRo TCN to teach project management skills in a Business Administration classes (~76 students). The module started with building a project budget and timelines using GANTT charts and culminated in building an Excel dashboard complete with graphs to track project progress.

CSCN: I will be provided another georeferencing workshop for our summer hires, none of which have georeferencing experience and will try to streamline the process based on lessons learned from the first attempt in August 2020 (such as not waiting for students to fill out 100 georeferences on a spreadsheet before checking them).

HUH: Member of the HUH published on our photostation design in use by the digitization team: https://doi.org/10.1002/tax.12459

Share Information About Your Website and/or Portal Usage

Google Analytics for our SoRo site are still showing similar usage with a small decrease in users, sessions and pageviews. (Report attached).

This Quarter (February 1st2020-April 30th 2021) had 2114 users over 1,833 sessions and 24,370 pageviews.

We had a strange spike in users at the start of the quarter that I suspect represent some kind of automated (bot) activity. Roughly half of our userbase for the quarter happened the first week of February.

Last Quarter (November 1st2020-January 31st 2021) had 900 users over 1,233 sessions and 16,813 pageviews.

We suspect that most of the data use is through the primary SEINet portal.

Share Other Activities and/or Progress

NYBG hosted a WeDigBio 2021 event on April 8-11. Approximately 325 guests attended our live virtual events including an herbarium collections tour, specimen digitization workshop, botanical research seminars, and specimen transcription event. Participants assisted in transcribing ±4,700 specimen records, and recordings of the events are now freely available on the NYBG Lecture Library.

COLO had planned to participate in WeDigBio, but was in the middle of a server migration and our image server went offline prior to the event and we were forced to abandon. We hope to participate in the fall.

UNM: We've provided curatorial and botanical training for five graduate and eighteen undergraduate students, employing six graduate and ten undergraduate students, and training others from UNM (University of New Mexico) and CNM (Central New Mexico Community College).

Audience Overview

All Users
100.00% Users

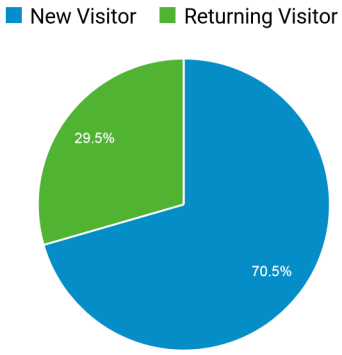
Feb 1, 2021 - Apr 30, 2021

Overview

Users	New Users	Sessions
2,114	1,833	3,509

Number of Sessions per User	Pageviews	Pages / Session
1.66	24,370	6.94

Avg. Session Duration	Bounce Rate
00:07:19	71.50%



	Language	Users	% Users
1.	en-us	1,124	53.17%
2.	en-gb	325	15.37%
3.	zh-cn	67	3.17%
4.	de-de	39	1.84%
5.	fr-fr	39	1.84%
6.	es-es	38	1.80%
7.	de	36	1.70%
8.	en	30	1.42%
9.	en-ca	28	1.32%
10.	it-it	27	1.28%

TCN Quarterly Progress Report

TCN Name

American Crossroads: Digitizing the Vascular Flora of the South-Central United States (TORCH)

Reporting Period: February 1st, 2021 - April 30th, 2021
Assembled by BRIT on May 4th, 2021, for May 5th IAC meeting

Person Completing the Report

Diego Barroso, TORCH TCN Project Manager <dbarroso@brit.org>

Reporting Institutions

BRIT – Botanical Research Institute of Texas
BAYLU – Baylor University
HUH – Harvard University
KANU – University of Kansas
MO – Missouri Botanical Garden
NOSU – Northeastern State University
NY – New York Botanical Garden
OKL – University of Oklahoma
OKLA – Oklahoma State University
SHST – Sam Houston State University
TAES – Texas A&M University-College Station
TAMUCC – Texas A&M University-Corpus Christi
TEX/LL – University of Texas at Austin [and subaward institutions]
TTC – Texas Tech University
UTEP – University of Texas at El Paso

Share Progress in Digitization Efforts

- Number of skeletal records created:

  BRIT = 0

  BAYLU = 0

  HUH = 0 new (1,829 total)

  KANU = 0

MO = 0

NOSU = 0

NY = 611 (project total: 29,091)

OKL = 1,136

OKLA = 2,419 (state only) (2,419 total)

SHST = 0

TAES = 0

TAMUCC = 30

TEX/LL & Subawards:

TEX-LL = 6,888
JWC = 3,272
SRSC = 0
SAT = 110
HPC = 0
OLLU = 0
SWT = 0
TLU = 0
PAUH = 0
RUNYON = 0
FWNC = 0

TTC = 0

UTEP = 0

Total skeletal records created this quarter: 14,466

- Number of fully-transcribed records created:

BRIT = 26,310

BAYLU = 0

HUH = 0 new (39,732 total)

KANU = 338 (total # fully transcribed records from OK & TX = 27,301)

MO = 179

NOSU = 0

NY = 1,098 (project total: 53,583)

OKL = 1,136

OKLA = 230 (55,049 total, including import from Oklahoma Vascular Plants Database)

SHST = 300

TAES = 0

TAMUCC = 0

TEX/LL & Subawards:

TEX-LL = 2,317
JWC = 461
SRSC = *ca*. 800
SAT = 110
HPC = 0
OLLU = 0
SWT = 0
TLU = 44
PAUH = 0
RUNYON = 0
FWNC = 0

TTC = 0

UTEP = 0

[***** Notes from Nature transcriptions: 6,804 subjects were uploaded into 6 Expeditions on BioSpex. Of these, 6,608 have been completely transcribed. From prior completed Notes from Nature Expeditions, 9,986 partial transcriptions were repatriated into the TORCH Symbiota portal.]

Total fully-transcribed records created this quarter: 33,323

- Number of specimens imaged:

BRIT = 10,733

BAYLU = 43

HUH = 39,034 total (* Imaging is ongoing, but we are able to report only on total images that have also been databased)

KANU = 17,389 (total # imaged specimens from OK and TX = 23,511)

MO = 83

NOSU = 0

NY = 5,158 (project total: 29,460)

OKL = 100

OKLA = 15,942 (69,241 total)

SHST = 0

TAES = 64,502
TAMUCC = 250

TEX/LL & Subawards:

TEX-LL = 6,888
JWC = 3,272
SRSC = *ca*. 800
SAT = 110
HPC = 0
OLLU = 0
SWT = 0
TLU = 0
PAUH = 0
RUNYON = 0
FWNC = 0

TTC = 731
UTEP = 7,000

Total number of specimens imaged this quarter: 172,035

- Number of specimens georeferenced:

BRIT = 94

BAYLU = 0

HUH = 4,988 georeferenced with coordinates;
9,199 reviewed but "skipped" with a reason given.
ca. 22,000 remaining to be reviewed (of records that have been databased).

KANU = 356 (total # georeferenced from OK & TX = 27,085)

MO = 133
NOSU = 0

NY = 5,150 (project total: 57,503)

OKL = 1,136

OKLA = 0

SHST = 0
TAES = 53

TAMUCC = All specimens included in the digitization project have georefererenced metadata attached.

TEX/LL & Subawards:

TEX-LL = 194
JWC = 276
SRSC = 0
SAT = 0
HPC = 0
OLLU = 0
SWT = 0
TLU = 44
PAUH = 0
RUNYON = 0
FWNC = 0

TTC = 1,725
UTEP = 1,124

Total number of specimens georeferenced this quarter: 46,472

- Other digitization or pre-digitization efforts:

BRIT:
Data cleaning of records generated from Notes from Nature, as well as data cleaning of locality-related field in records in preparation for georeferencing. Image processing (renaming to barcodes, image editing, and upload to TACC servers) has been a priority for our lead digitizers.

We continue to physically locate project specimens in BRIT collection (Oklahoma in BRIT/SMU North America folders) for digitization, and in the NLU collection for transcription (most NLU specimens had already been imaged but not transcribed when they were accessioned at BRIT in 2017).

We continue skeletal transcriptions of images from image sets containing a mix of project and non-project specimens (Texas and Oklahoma in VDB North America folders) in the VDB collection at BRIT in order to prioritize TORCH complete transcriptions. Utilizing the crowd sourcing module in Symbiota, providing training (outside of business hours) and ongoing support (via email, Zoom, and Google Docs) has resulted in 28,142 skeletal transcriptions (scientific name, country, state, county) in since 23 Jan 2021. These skeletal transcriptions have provided us with details to prioritize complete transcription of TORCH specimens.

The ACU herbarium has been imaged and transcribed and will be physically returned to the University on 11 May 2021. The collection holds 5,201 specimens of which 3709 specimens were collected in TX and OK.

Plastic panels for the custom TORCH lightboxes were completed by the vendor and will allow the immediate completion of the third and fourth light boxes to be delivered to TAES and OKL. Enough panels were purchased to build an additional five light boxes to be completed in the coming months.

KANU: We processed mounted, barcoded, databased, and imaged) about 40 of the 179 unmounted specimens from TX referenced in the last report. Work is continuing on the remainder of the unmounted specimens.

NY: OCR has been performed on all specimen labels to aid in transcription.

OKLA: Completed making genus folders for all Texas specimens (not using TORCH funds). Specimens repaired as needed (not using TORCH funds). Curation conducted as needed (not using TORCH funds). 13,588 images from were processed (13,583 unique). Of these 4,965 have been uploaded to TACC. Not all of these have been ingested into the TORCH Symbiota portal, since we are waiting for the Accession Number to be extracted so they can be linked to existing records.

SHST: Students performing data entry for backlog of uncatalogued specimens. 2,000 specimens in database corrected.

TAMUCC: At the Ruth O'Brien Herbarium we have worked extensively in our pre-digitization efforts in areas of specimen organization, intake processes, and quarantine methods.

TEX/LL: We retreived the herbarium of Saint Edward's University (a new data provider), and are in the process of freezing 48 boxes. We will begin imaging the specimens once they are stable and bug-free. We ordered additional barcodes for St. Edward's and Texas State.

- Comments about the digitization process:

BRIT: A small number of records that were modified/edited through our data-cleaning process were mistakenly recorded as complete transcriptions when they were in fact record corrections –this number was likely to be insignificant and thus we qualify but do not quantify this (reported in Section I., "Number of fully-transcribed records created")

HUH: Digitization staff are sharing time across multiple projects. At present, staff are focusing on georeferencing activities for the TORCH and SoRo TCNs. Georeferenced records currently reside in the GEOLocate website and are planned to be ingested back into our system next quarter (data to be subsequently shared via existing IPT). HUH also recently completed necessary changes to our image processing pipeline and staff are expected to resume allocating partial time to new record transcription.

KANU: We databased 338 seed packets (all barcoded; most with associated pressed specimens in the collection). Expect for specimens yet to be mounted and occasional specimens that were missed when we flagged specimens, we believe we have transcribed the label data for all of our OK and TX holdings. We have a small number of specimens (<150) that we might yet be able to georeference. Otherwise, our efforts now are focused on imaging of specimens. We do not plan to image the seed packets but will image all pressed and boxed specimens from OK and TX. During the reporting period, we imaged 87.2% of those specimens. We expect to finish imaging specimens in approximately two weeks. We then need to process the RAW images in Lightroom and upload the processed images to our attachment server. They will be available on GBIF and iDigBio when we upload a new instance of our database.

MO: As of this reporting, access to our collections remains restricted to our science staff, who are authorized to be on-site only one or two days per week. Presently, there is no provision for access to our facilities by students, volunteers and visitors. This has severely impacted our ability to make any significant progress on this project, and has precluded recruiting, training and adding new people. We are hopeful that in the coming months this will change, but we have no sense of when that might be.

NOSU: Still waiting on light boxes.

NY: Due to state and city restrictions, we are limited to only a small number of onsite staff each day, and each staff member can only work two days per week.

OKL: Things were slow because the graduate student on the project was on a TA-ship this semester. Things will move more quickly in the summer when she will be working on this project again and will also be training an undergraduate and another Master's student.

OKLA: All going well.

SHST: Problems setting up software for digitizing. Waiting on IT. Students working on backlog of uncatalogued specimens in the meantime.

TAES: We continue to make the imaging of our specimens our primary task.

TAMUCC: The digitization process has gone by rather smoothly with our only downfall being that we do not have an entire team on the project. We are aiming our focus on finding other dedicated individuals who would be interested in training to help out with our efforts so that we can ensure a constant flow of imaged specimens. Occasionally, we have run into problems with our camera (lens will not auto focus) and are required to readjust its position. This can lead to a slight difference in photo appearance from session to session. We re-digitized our first 100 specimens to allow for a standardized look within our photos but know that this is not a practical solution moving forward. Besides these minor obstacles, we have been headstrong about overcoming them and have since picked up some momentum in our flow. Now, we have reached our organizational goals and are able to digitize smoothly moving forward.

TEX/LL: Progress on label transcription has been slower this quarter than previously. Progress at data provider institutions has been especially slow.

TTC: We have now manually georeferenced all specimens collected in Texas, and we are now proceeding to our Oklahoma specimens.

UTEP: We have limited ability to have students work in the herbarium due to COVID-19. However we have been able to employ a couple work study students starting in January and to have been able to make more progress.

- Number of records available in iDigBio portal (cumulative):

BRIT (as of 2021-04-29, collected in TX and OK, without taxonomic constraints):

BRIT-SMU-VDB-NLU: 170,801

TAC:	7,065
NTSC:	0
ACU:	0
TOTAL for BRIT Lead:	177,866

BAYLU = 0

HUH = 36,073

KANU = A new instance of our database is uploaded to GBIF and iDigBio at the beginning of each month. This continues to be done for all transcribed records. New images will be uploaded beginning in about 20-30 days. [For this count, assumed 27,085 as above].

MO = Nothing to report.

NOSU = 0

NY = 88,241 (also includes Bryophytes and Fungi)

OKL = 0

OKLA = 0

SHST = 0

TAES = 0

TAMUCC = 0

TEX/LL & Subawards = 28,021 (includes TEX-LL + data providers). This has not changed during the quarter, but mostly because reports from Subawards were not received). Note that the TEX-LL specimens are an old upload that has been static (under the institutional acronym PRC) and we are working to set up the dynamic

relationship with iDigBio under the correct institutional acronym, at which point the PRC data should disappear.

TTC = 0

UTEP = 82,383

Total number of records available in iDigBio portal (cumulative): 439,669

- Number of records available in TORCH Symbiota portal (cumulative):

BRIT (as of 2021-04-29, collected in TX and OK, without taxonomic constraints):

BRIT-SMU-VDB-NLU: 184,825
TAC: 7,063
NTSC: 8,960
ACU: 3,709
TOTAL for BRIT Lead: 204,557

BAYLU = approximately 10,000 (non-transcribed)

HUH = 41,556

KANU = All records uploaded to GBIF and iDigBio should be accessible on TORCH portal. [For this count, assumed 27,085 as above].

MO = Nothing to report.

NOSU = 0

NY = 79,791 (same as previous report – IPT has apparently not been updated)

OKL = 135,617

OKLA = 57,466

SHST = 0

TAES = 234,583

TAMUCC = 0 (First batch will be available at the 500 digitized specimens mark.)

TEX/LL & Subawards:

Symbiota 4/30/2021

	CumulativeData	CumulativeImages	CumulativeGeoref
TEX-LL	229,019	128,471	50,974
JWC	461	3,272	276
SRSC	23,494	23,480	2,294
SAT	38,809	38,395	4,395
HPC	22,897	4,412	2,265
OLLU	0	0	0
SWT	0	0	0
TLU	6,795	0	6,626
PAUH	6,418	0	110
RUNYON	0	0	0
FWNC	1,918	1,918	273

TTC = 21,615

UTEP = 85,504

Total number of records in TORCH Symbiota portal (cumulative): 1,227,585

Share Best Practices, Standards, and Lessons Learned

Best Practices and Standards (Lessons Learned):

OKLA: Revised image correction workflow to better handle dark specimens. Began developing transcription workflow protocol in Symbiota.

SHST: Students record specimen data independently and turn in their progress to supervisor who compiles new data into main spreadsheet.

TAMUCC: The best practice we have learned this quarter is organization. We entered this project with very little insight or direction regarding our herbarium space and had about 7 years of dust that needed to be cleared before we could really begin with our digitization efforts. Unfortunately, there was no guide left for us outlining the ins and outs of the previous management practices, so we've started from the ground up, writing everything down as we go. Since the last quarter, we have phased out all of our old specimen cabinets and installed brand new, rubber-sealed cabinets that will prevent

any further infestation of our collection. Additionally, we conducted an extensive review of each of our available specimens and decided to remove over 300 specimens from the collection due to damaged plant tissue. This was another great practice we learned, as we know that the specimens located in the cabinets that are “ready-for-digitization” have already been examined for pests and tissue damage. This has helped immensely with our digitization efforts in recent weeks. Another standard we have implemented has been to clean table areas and floor areas after the handling of specimens and especially after the handling of intake specimens. We have since seen little to no bugs or pests in our herbarium space.

Share Identified Gaps in Digitization Areas and Technology

Identify Gaps in Digitization Areas and Technology (issues preventing progress):

BRIT: BRIT continues to struggle with establishing its internal protocols for georeferencing. Although we will wait until we have the majority of BRIT-responsible specimens transcribed, we have had small opportunities afforded us to commence this process for TORCH and larger responsibilities to do this for our two other funded TCNs, Endless Forms and Ferns. We find ourselves struggling to (1) take advantage of already-georeferenced localities in Symbiota (in a manner that allows us to vet georeferenced localities before we apply them to our specimens in an efficient and transparent manner), and (2) to effectively utilize the Geolocate CoGe and Symbiota connection to batch and collaboratively georeferenced our records (our current bottle neck is encountering that certain Symbiota fields may not be visible in CoGE, i.e. “Municipality”).

BAYLU: Since December of 2020, both technician and student worker both left their positions leaving no active digitization. After recruiting through the spring (within the limits of the COVID pandemic). Have hired new technician and worker who will both begin on May 3, 2021.

NY: McKenna Coyle suspended her on site work during January and February due to increasing COVID rates, and her need to commute via public transportation. Starting in March she resumed her two days per week onsite. Alex Patrascu completed her internship in February. We still have one internship on the TORCH grant, but due to Garden restrictions, we are not yet able to hire a second intern. We are hoping to begin advertising for a second intern in June, with a September start date.

OKLA: Need segmentation/OCR of accession stamp to link existing database records to images as they are obtained—we are close to trialing this workflow, and are drafting documentation of Best Practices for our process.

SHST: Herbarium is not suitable for wifi connection. IT is working on setup so that digitization can occur.

TAES: We are still waiting for two imaging stations. When these are received, we can begin working with our provider herbaria.

TAMUCC: We feel that our digitization efforts would greatly benefit from the ability to collaborate with more herbaria that are participating in this project. We understand that all participating herbaria are of different size and stature, and we would love to know if other small herbaria have faced some of the same obstacles we have. We'd also love to learn from the larger herbaria as they move forward in their digitization efforts.

TEX/LL: Our oldest imaging box is on its last legs and needs to have the lamps replaced.

Share Opportunities to Enhance Training Efforts

Opportunities to Enhance Training Efforts; Training and Professional Development Opportunities you offered and/or participated in (e.g., webinars, student digitizer training, etc.):

BRIT: Continued to host weekly Zoom conversations with the Armchair Botanist program to engage Notes from Nature volunteers transcribing project specimens.

Offered an additional and advanced training (via Zoom) for BRIT volunteers interested in transcribing directly into Symbiota (including those previously not contributing to online Notes from Nature transcriptions).

NY: Digitization staff have taken advantage of a wide range of webinars and online conferences sponsored by New York Botaical Garden, iDigBio, Royal Botanical Gardens, Kew, etc.

OKLA: Trained one new undergraduate assistant in imaging and transcription.

SHST: Students will be participating in group training from BRIT to master the digitization process.

TAMUCC: We have built strong step-by-step procedures and protocols as we have solidified them throughout our experience. We plan to leave these written protocols in our herbarium as well as in our databases to serve as a roadmap for future herbarium technicians we take on. We believe our written instructions will allow for their success and also provide as a foundation for them to continue to build off of.

If there are any public webinars that are available to us, we'd love to be made aware of them. We are always looking for new ways to learn and increase our efficiency at the Ruth O'Brien Herbarium.

UTEP: Georeferencing training has been offered to several student volunteers.

Two students have used georeferenced specimens in projects to develop floral lists for Hueco Tanks State Park and the Castner Range of the Franklin Mts.

Share Collaborations with other TCNs, Institutions, and/or Organizations

Collaboration with other TCNs, Institutions, and Organizations:
The TORCH TCN Project held two virtual Executive Committee meetings (with all 5 lead PI's in attendance), on March 17th, 2021, and April 26th, 2021. One of the agreements reached at the April 26th meeting is that we will plan for a TORCH Virtual Digitization Workshop to be held on Saturday, July 24th.

BRIT:

Joined Project Manager Diego Barroso (on April 27th) in discussions with representatives from the Witte Museum in San Antonio (Leslie Ochoa and Helen Holdsworth) and from Sam Houston State University Herbarium (Will Godwin), regarding how we might assist the Witte Museum in digitizing their herbarium collections (ca. 4700 sheets). These include important historical collections, as well as the Bexar Co. Herbarium, gifted to the Witte by the Native Plant Society of Texas.

Engaged with members of multiple chapters of the Texas Master Naturalists through email, presentations, or weekly Zoom event, including members of the following chapters: Alamo Area, Big Country, Blackland Prairie, Bois d'Arc, Capital Area, Central Texas, Coastal Prairie, Cradle of Texas, Cypress Basin, El Camino Real, Elm Fork, Galveston Bay Area, Goodwater, Guadalupe, Gulf Coast, Hays County, Heart of Texas, Indian Trail, Lindheimer, Lost Pines, Mid Coast Chapter, North Texas, Panhandle Chapter, Prairie Oaks, Red River, Rio Brazos, South Texas Border.

Project Manager Diego Barroso held two meetings with TAES representatives (with D. Spalink and D. Kruse, on March 18th; and with D. Spalink and K. Simpson on April 29th), to discuss processing and cleanup of accession numbers in the TAES/TAMU dataset.

D. Barroso, P. Fritsch, J. Best, and G. Yatskievych met with Jonathan Amith from the DEMCA project (Documenting Ethnobiology in Mexico and Central America) to discuss

the possibility of a collaboration between TORCH and DEMCA (two separate meetings in April 2021), using an Extended Specimen approach.

SHST: Imaging equipment is on loan from the Botanical Research Institute of Texas (BRIT), having been picked up by Will Godwin on April 7th. Also, negotiating an agreement with The Witte Museum (San Antonio, TX), to fold approximately 5,000 herbarium specimens from there into our digitization project. Spoke to Helen Holdsworth and Leslie Ochoa.

TAMUCC: We've kept in close contact with TAMU-College Station colleagues and have learned so much from them. We have been able to acquire all of our needed supplies for the continuation of our digitization thank to their guidance. They have proved to always be a helping hand to us and we have greatly appreciated their input whenever we have asked for it. Specifically, Kyle Simpson with TAMU has collaborated with us on how to capture the best quality image and we've sent plenty of photos back and forth discussing the small details that create a stellar capture. We've learned together that not every camera is the same, and sometimes it takes plenty of small adjustments to obtain the desired quality of image. We've also extensively discussed the angle of the camera and multiple methods of correcting an offset camera.

Share Opportunities and Strategies for Sustainability

Opportunities and Strategies for Sustainability:

KANU: We continue to carry out strategic curatorial and collection management work on specimens, especially those from OK and TX. This work is handled by the Collection Manager and/or a student employee not working on the TORCH grant. It has resulted in significant collateral collection improvements, such as replacing several thousand worn genus folders, refoldering specimens to reduce compaction, repairing damaged specimens, annotating specimens, and updating the Specify taxonomy tree.

SHST: Recycling paper waste associated with record-keeping.

TAMUCC: We have phased out our use of single-use plastic in the herbarium and are finding sustainable ways of disposing of damaged specimens (but have yet to find the best way).We have also used freeze quarantining methods to prevent the use of any chemical pesticides in our herbarium. Lastly, we obtain all of our collection newspaper from the library on campus, to prevent any unnecessary waste of material.

Share Education, Outreach, Diversity, & Inclusion (EODI) Activities

Methods of disseminating results to communities of interest (presentations, lectures, etc.):

BRIT:

T. Rehman presented (28-Jan-2021): "Herbaria: Centers for scientific collections and expertise engaged in the study of our natural heritage" for the Amon Carter Museum of American Art (ACMAA) monthly virtual lecture series, Experts Talk on the Carter (link), highlighting the value of herbaria, the TORCH TCN digitization project and looking at ACMAA piece "William Richards Trost's "Woodland Glade". (https://www.cartermuseum.org/events/virtual-experts-talk-carter-art-nature; recording may be posted at a future date)

T. Rehman (with Montana Williams, BRIT Director of Volunteers; 4 Feb 2021): “Armchair Botanist: Digitizing the Plants of Texas” for the February 2021 Texas Master Naturalist Virtual Volunteer Fair, showcasing this community science project (through Notes from Nature) to 340 attendees.

D. Barroso and T. Rehman presented (9 Feb 2021): “Armchair Botanist” program for the BRIT Green Revolution Teen Tuesdays programming.

T. Rehman presented (18 March 2021): "Aggregating our collections: A team effort to share information about the Texas flora" for the Houston Chapter of the Native Plant Society of Texas, highlighting the TORCH TCN project progress, instructing audience members on the use of the portal to access records, and showcasing some hidden treasures from TORCH TCN collections. (https://npsot.org/wp/houston/events-info/monthly-meetings/; https://www.youtube.com/watch?v=5J-0wiJyKuc)

P. Fritsch presented (9 April 2021): “Evolutionary origins and species diversity in the wild blueberries tribe Vaccinieae” for the BRIT Lunchtime Lecture series, in which he highlighted the importance of access to digitized herbarium records (images and data) for his research and highlighted the current and past funded TCNs. In particular he included a discussion of the complementary nature of iNat records and the value of specimen images, particularly when specimens have been damaged post-imaging (https://brit.org/research/research-resources/research-lecture-series/; https://www.youtube.com/watch?v=LEIkOpYHufs&list=PLFlQggX9zTKQ3QuZ7lgnJ9zIMW997bIZq&index=2)

T. Rehman, A. Bordelon, D. Barroso, J. Best, C. Barrett presented (8, 10, 11 April 2021): three WeDigBio Transcription Blitzes including a short presentation about the TORCH TCN, a behind-the-scenes tour of the BRIT herbarium (8 April), a behind-the-scenes tour of the Sumner Molecular and Structural Lab at BRIT (10 April; led by BRIT botanist Alejandra Vasco), and a presentation about the collecting of herbarium specimens, including live tour of BRIT specimen receiving room and collecting supplies (12 April). Attendees from multiple countries were in attendance (brit.org/wedigbio)

T. Rehman presented (2021-04-17): "Collections, Conservation, and Community Science: A Collaboration to Conserve our Planet's Natural Heritage" - for the Fort Worth Public Library Earth Day 2021 Events in which currently funded TCNs were highlighted and participants were encouraged to explore their local natural history collections and participate in the weekly Armchair Botany Zoom sessions (https://www.fortworthtexas.gov/departments/library/events/earth-day-2021?BestBetMatch=rehman|a720ba43-6954-4033-b1ac-a72478fc15dc|f3d8342e-e792-4add-b724-7420243aa350|en-US; https://www.youtube.com/watch?v=0945UpOxnFw)

SHST: Hosted tours of Warner Herbarium to local students and organizations.

TAMUCC: We have worked extensively with our Life Sciences Department to discuss the importance and value of not just herbarium specimens but the digitization of them as well. This quarter, we worked with the Plant Ecology classes to offer them a tour of the herbarium space as well as a chance to become a collector. We began by lecturing the class about the general overview of herbaria and specimen digitization, and then we offered an in person tour. Then we discussed all of the materials that go into specimen collecting and provided a roadmap for students to follow when conducting plant surveys. We believe that our method of allowing students to learn in a hands-on, experiential way, is what will lead them to be strong ecologists and collectors in the future. We are also confident that this method of disseminating our results to the university community has been highly effective. We are also working with our student newspaper in every way that we can to make sure that our progress is known outside of just the herbarium community.

TEX/LL: We prepared an abstract for a poster on TEX-LL digitization efforts for the Botany 2021 conference (BSA).

UTEP: Castner Range Flora Project has been presented at Local meetings, SACS and an abstract for this Summer Botanical Society of America National Meeting.

Hueco Tanks project will be finalized and distributed to the State Park utilizing georeferenced material.

Other Education and Outreach activities:

BAYLU: Have begun investigation in the use of herbarium specimens for assessing stomatal density/index over the timeframe of collections within individual phyla. Independent Research (BIO 3V90) studied aspects of this in terms of documenting breadth of collections for various ferns and angiosperm species in addition to testing use of dental putty impression for non-invasive sampling.

HUH: Davis is working with the Harvard Museum of Natural History to create a phenology module using digitized herbarium collections to understand plant response to climate change. The module is targeted to meet middle school curriculum standards in Massachusetts.

MO: Spotlight on the Garden: Herbarium, 17 February 2021. This was a one-hour interactive webinar with about 50 people, that included an in-house video (15 minutes) and Q&A about the importance of herbaria. The program highlighted herbarium-based projects, including this TCN, and selected plants from our collection.

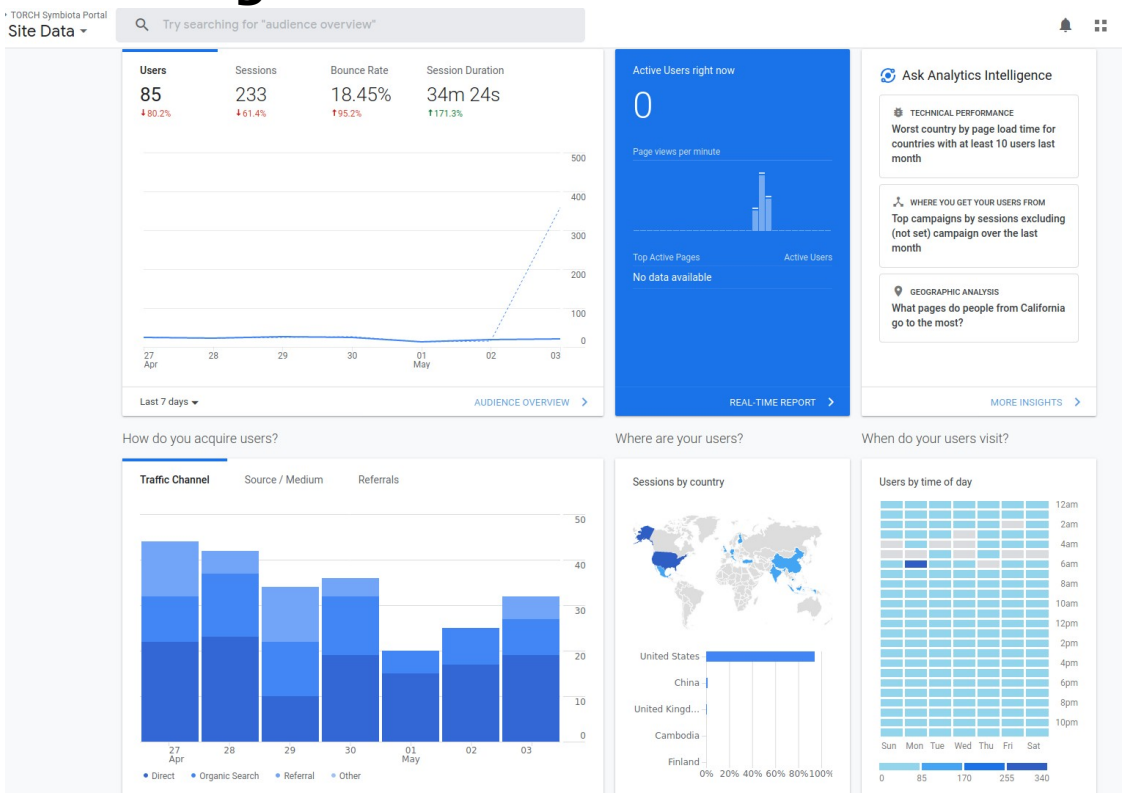
NY: Participated in WeDigBio in April 2021 by contributing records for transcription. We had several days of programming for iDigBio this year, including webinars, volunteer training sessions, and volunteer meet-and-greet sessions.

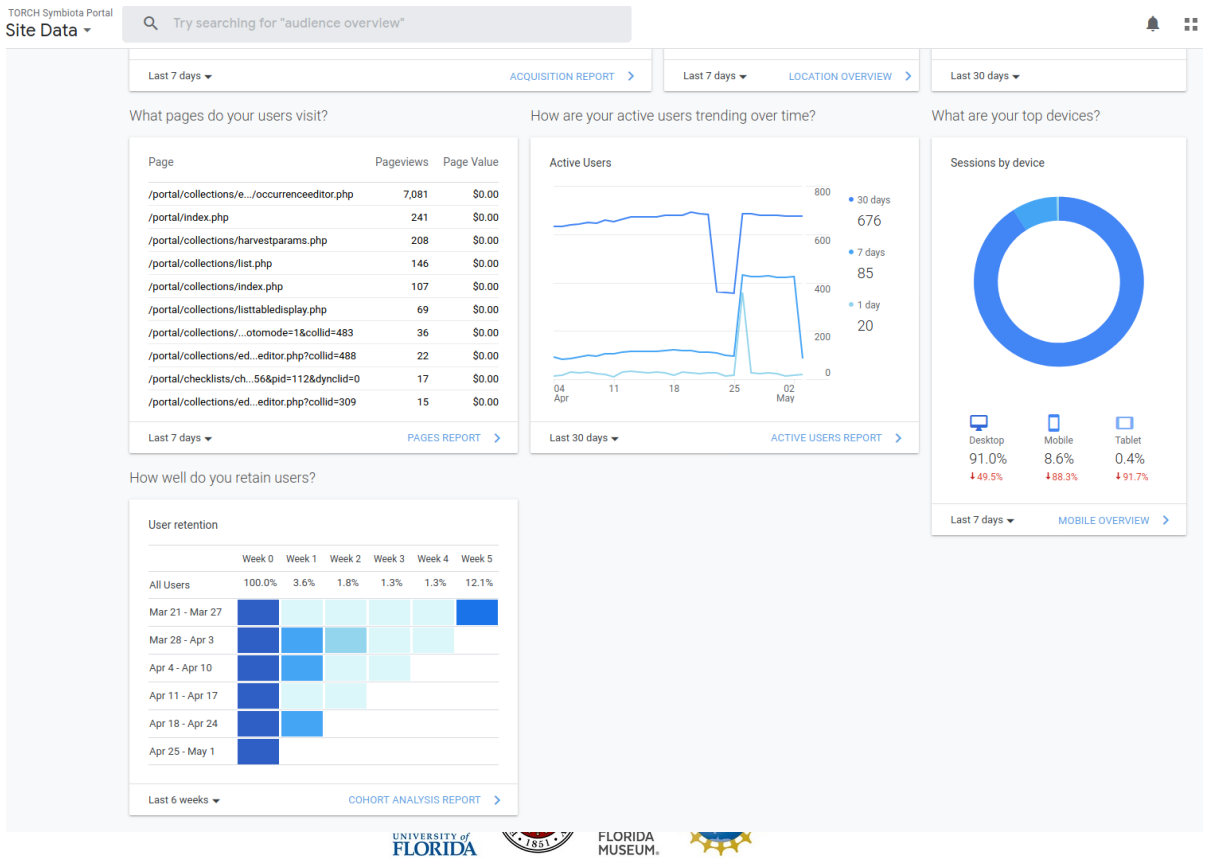
McKenna Coyle published the following Handlens article entitled, “Sister Mary Clare Metz: Faith and Flora” about a little-known but important collector in Texas: http://sweetgum.nybg.org/science/the-hand-lens/explore/narratives-details/?irn=7522

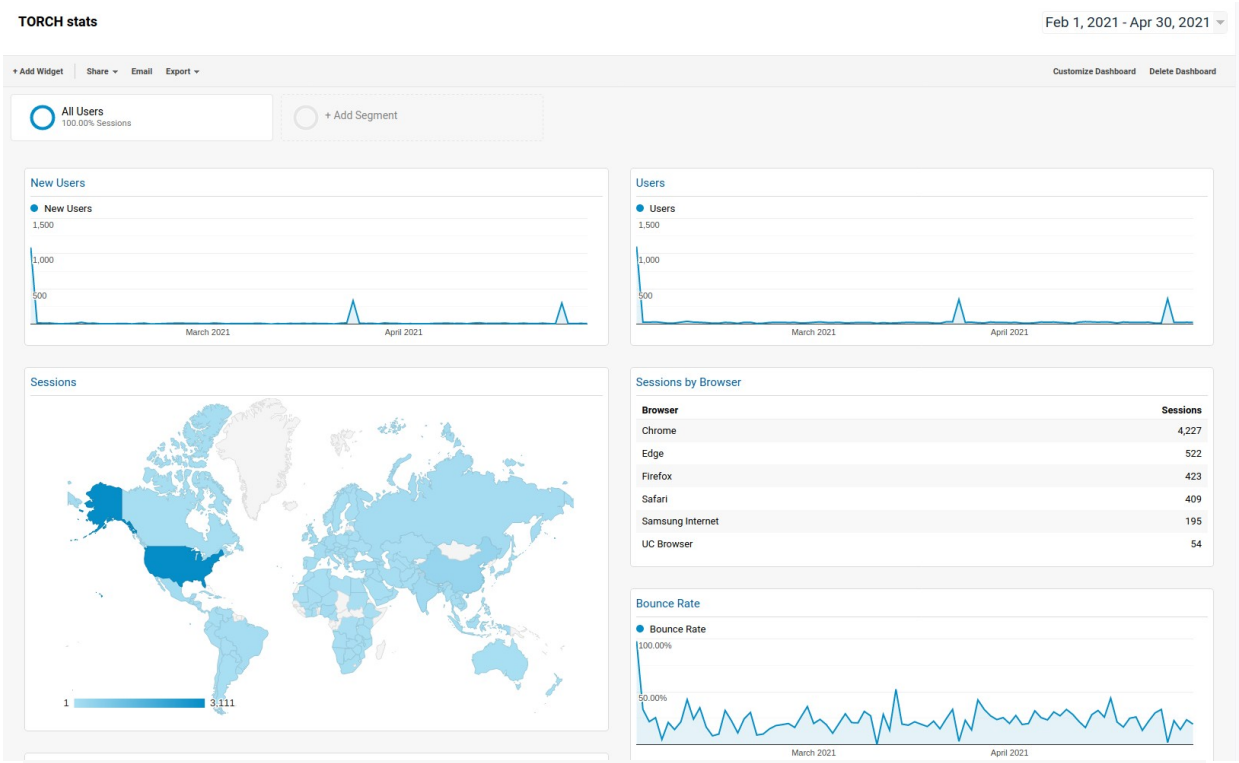
TAMUCC: We have worked with Oso Bay Wetland Preserve and Learning center as well as Pollywog Pond Preserve and wildlife sanctuary as partners in plant survey sites and we have since seen an influx in specimens from these locations. We’ve used these specimen collections as an educational opportunity for students to learn some of the practical uses of herbarium specimens, such as making physiological plant observations after an unusual freeze event that we observed in March.

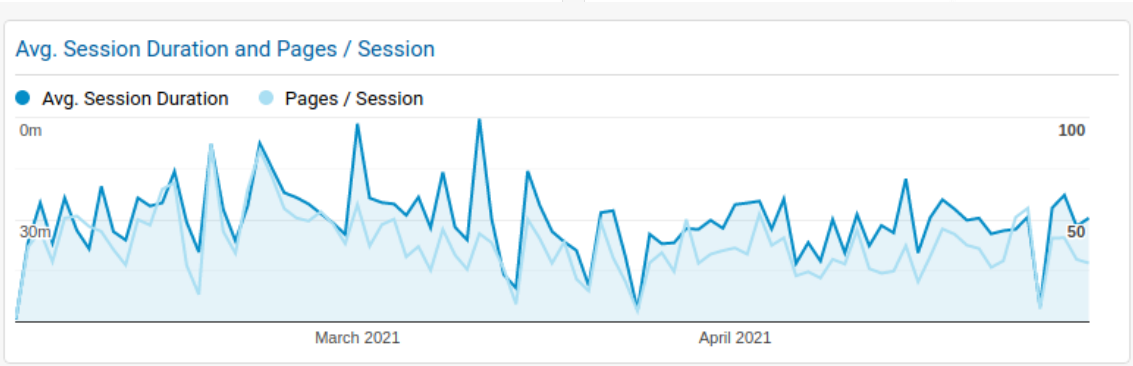
TEX/LL: George Yatskievych presented an online lecture to an Economic Botany class on the herbarium that included mention of TORCH.

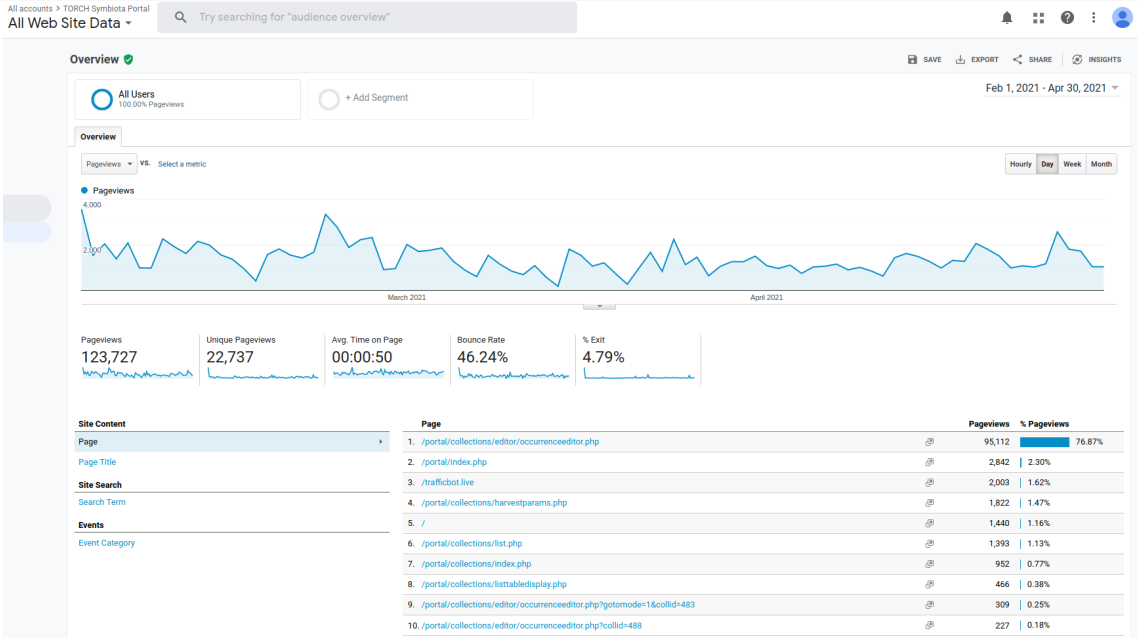
Share Information About Your Website and/or Portal Usage






TORCH stats
Feb 1, 2021 - Apr 30, 2021
+ Add Widget
Share
Email
Export
Customize Dashboard
Delete Dashboard
All Users
100.00% Sessions
+ Add Segment
New Users
New Users
1,500
1,000
500
March 2021
April 2021
Users
Users
1,500
1,000
500
March 2021
April 2021
Sessions
1
3,111
Sessions by Browser
Browser
Sessions
Chrome 4,227
Edge 522
Firefox 423
Safari 409
Samsung Internet 195
UC Browser 54
Bounce Rate
Bounce Rate
100.00%
50.00%
March 2021
April 2021


Avg. Session Duration and Pages / Session
Avg. Session Duration
Pages / Session
0m
30m
100
50
March 2021
April 2021


All accounts > TORCH Symbiota Portal
All Web Site Data
Try searching for "audience overview"
Overview
SAVE
EXPORT
SHARE
INSIGHTS
All Users
100.00% Pageviews
+ Add Segment
Feb 1, 2021 - Apr 30, 2021
Overview
Pageviews
vs.
Select a metric
Hourly
Day
Week
Month
Pageviews
4,000
2,000
March 2021
April 2021
Pageviews 123,727
Unique Pageviews 22,737
Avg. Time on Page 00:00:50
Bounce Rate 46.24%
% Exit 4.79%
Site Content
Page
Page Title
Site Search
Search Term
Events
Event Category
Page
Pageviews
% Pageviews
1. /portal/collections/editor/occurrenceeditor.php 95,112 76.87%
2. /portal/index.php 2,842 2.30%
3. /trafficbot.live 2,003 1.62%
4. /portal/collections/harvestparams.php 1,822 1.47%
5. / 1,440 1.16%
6. /portal/collections/list.php 1,393 1.13%
7. /portal/collections/index.php 952 0.77%
8. /portal/collections/listtabledisplay.php 466 0.38%
9. /portal/collections/editor/occurrenceeditor.php?gotomode=1&collid=483 309 0.25%
10. /portal/collections/editor/occurrenceeditor.php?collid=488 227 0.18%

Share Other Activities and/or Progress

Products generated (publications, conference presentations, technologies/techniques, websites, etc.):

HUH: Davis spoke on his research related to digitization at seminar at Washington University in February 2021. Also the following paper was published, documenting the HUH Digitization pipeline:

> **Davis CC, Kennedy JA, Grassa CJ. 2021.** Back to the future: a refined single-user photostation for massively scaling herbarium digitization. *Taxon*.

TAMUCC: We have generated an extensive binder full of resources for our herbarium technicians including how-tos and guidance documents for every aspect of pur herbarium, including the digitization aspect.

We plan to create YouTube videos verbalizing our efforts and techniques and will be excited to report on it during the next quarter.

Participants (especially those who have newly joined the project):
BRIT:

Ashley Bordelon, Digitization Coordinator (virtual engagement); abordelon@brit.org
Joe Lippert, Digitization Coordinator (imaging and image processing); jlippert@brit.org
Diego Barroso, TORCH TCN Project Manager; dbarroso@brit.org
Tiana Rehman, Collections Manager/Institutional Rep; trehman@brit.org
Jason Best, Dir. Biodiv. Informatics/Technovator; jbest@brit.org
Peter Fritsch, VP of Research/PI; pfritsch@brit.org
Jessica Lane, BRIT Herbarium Assistant; jlane@brit.org
Tessa Boucher, Digitization Technician, tboucher@brit.org
Rachel Carmickle, Herbarium Technician, rcarmickle@brit.org

BAYLU: Nothing to report.

HUH: Nothing to report.

KANU: No new participants.

MO: Nothing to report.

NOSU: Nothing to report.

NY:

McKenna Coyle, Lead Digitizer
Alexandra Patrascu, Intern (internship ended February 2021)

OKL: Nothing to report.

OKLA: Three undergraduate workers (one new) continued imaging, one began transcribing.

SHST:
Christina Harston
George Dewey III
Shannon Meade
Kate Chandler
Cole Turnage

TAES: Jennifer Deden (undergraduate student) was hired as a new technician.

TAMUCC:
Barnabas H. Daru – project supervisor and PI
Jordan Rodriguez – Lead Herbarium Technician
Jessica Palitza – Herbarium Technician in training

TEX/LL: Saint Edward's University is joining TORCH as a Data Provider.

TTC:
Undergraduate digitizers: Chase Bergeron, Jared Salzman
Graduate student mentor: Yanni Chen

UTEP:
UTEP Undergraduates:
- Ianellie Munguia (continuing)
- Justin Gaenzle (continuing-work study)
- Carlos Vargas (continuing)
- Muriel Norman (continuing – work study)
- Alexis Vallejo (continuing)
- Aparna Mangadu (continuing)

All georeferenced, trained and supervised by our collections manager Mingna "Vicky" Zhuang.

Other Progress not listed above (anything else to share):

KANU: We expect to complete most, if not all, of our work by the end of May 2021.

NY:
Remaining work to do on the project:

Skeletal records: 27,909 (out of a promised 57,000). We likely do not actually have this many specimens left from the TORCH region – this estimate was high.

Specimens imaged: 27,540 (out of a promised 57,000)

Full transcriptions: 23,417 (out of a promised 77,000)

Georeferences: 10,764 (out of a promised 68,267)

SHST: 240 new specimens were mounted.

TAMUCC: We have begun to disentangle the Kodak film photos left behind by the previous herbarium manager (Ruth O'Brien) and we hope to find ways to view them and utilize them in the future. We believe that these photos may give valuable insight to the undigitized specimens that exist within our collection.

TEX/LL: We have been focusing on Oklahoma-collected specimens, which have all been pulled from the general collection and are being imaged.

Questions/comments:

BAYLU: The Chair of the Department of Biology at Baylor University, Dr. Dwayne Simmons, has made an independent decision to move the herbarium from its current location to another room within the same building at Baylor in the middle of summer, 2021. I have protested this decision vigorously as being disrupted to no avail. I have also made it clear that the NSF funds provided for this project, in terms technician salaries are not budgeted for helping move, rather solely for digitization. I think it is prudent to report to the NSF perhaps the program manager that the Baylor administration has made this decision that will be disruptive for several months, just having recovered personnel in the middle of the pandemic. I am unsure whether this negates aspects of our agreed contract relating to facility and administrative support.

TAMUCC: Are there any group TORCH meetings we could be a part of? We'd love to hear more about the progress and obstacles that other herbaria have faced. Are there any upcoming webinars we should be made aware of?

SHST:

We submitted a proposal to the Witte Museum in San Antonio, TX, to collaborate on digitizing 4,000~5,000 of their Herbarium specimens. These are mostly from the Ellen Schulz Collection. I am interested in partnering with the Witte because our two institutions share a connection with Ellen Schulz and Roy Quillen. We proposed that in August 2021 we would travel to the Witte Museum with one or two students and bring the digitizing equipment; the work would be done locally. Our understanding is that the data has already been gathered into an Excel spreadsheet; therefore, Project Manager Diego Barroso and Data Manager Clay Barrett will be able to help ingest the data into the TORCH Symbiota portal. If the Witte could put us up somewhere then we could probably digitize 4,000 or 5,000 sheets in a few weeks.

TEX/LL:

TEX did not receive any reports from the data provider institutions this quarter, so the quarterly progress numbers are comparisons of the cumulative values for this quarter compared with last quarter; the SRSC figures are estimates. We will be more precise next quarter. Although we are still reporting on FWNC, they have been done for a while now. Their last set of specimens was returned to them during this quarter, and we also did some nomenclatural/taxonomic clean-up of their specimens (not using TORCH funds).

TTC:

We had some progress this quarter but we’re hoping to really ramp things up over the summer! We have also uploaded 731 new images to the TACC server from this quarter, so they are ready to be processed so they can be served to the TORCH portal.

TCN Quarterly Progress Report

Prior to each IAC meeting, TCNs are asked to complete a quarterly progress report in the areas outlined below. The TCN Lead PI or Project Manager collects information from all collaborators and compiles them into one overall progress report for the TCN. The TCN Lead PI or Project Manager then submits the quarterly reports via an email to Cat Chapman. An archive of previously submitted reports is available on the Internal Advisory Committee wiki page.

TCN Name

Digitizing collections to trace parasite-host associations and predict the spread of vector-borne disease (TPT)

Person Completing the Report

Jennifer Zaspel (Lead PI)

Share Progress in Digitization Efforts

	Transcribed records	High resolution images	Scanned vials	Scanned slides	Label images	Notes from Nature
ANSP		359		1624		
BPBM	7,000		73	2505	73	9,000
BYU	2,956					
CAS	7,587	591		207		
OSU	5,939					
FMNH	5,008			11,890		
INHS	6,162	348		670		
MPM	1,595	94		1,228		
MSU	2,527	17				
U Michigan	90,628	3,135		30,188		
U Minnesota				72,509		
U New H	48,611					
U New Mexico		465				
UNL	2,000	14,280				
UW Madison	9,345					
UW SP	1,063			563		

U Utah	8,186					2,000
Yale Peabody	4,560			2,797		
PERC	4,478					
Penn State	7,295		662	8081		
TAMU	6,149					
U Hawaii	2,500	1,500				

TPT launched 4 new NfN expeditions from 4 institutions. This far we have gotten >86,000 transcriptions accomplished for TPT.

Share Best Practices, Standards, and Lessons Learned

TPT is investing heavily in the development of taxonomy resources for our network and the greater arthropod collections community. PI Zaspel hired Drs. Nick Dowdy and Teresa Mayfield to help gather and format names lists for arthropod parasite groups. Dowdy started a repository for tools to clean Acari taxonomy, (https://github.com/njdowdy/tpt-acari-taxonomy.git). Vijay Barve compiled and harmonized master lists for Mammals and Birds from authoritative lists like ASM for mammals and IOC for birds. Rich Pyle from the Bishop Museum sent a list of Diptera names (~276K records) to Vijay Barve (the breakdown of names include 4,597 family records, 23,370 genus, 201,549 species, and 47,352 synonyms linked to their accepted name). The Bishop is also working with Catalog of Life to get the Diptera names updated in COL. TPT is planning to send a comprehensive list to Global Names at the end of March in order to facilitate and accelerate data entry for the TPT data providers (regardless of platform).

Share Identified Gaps in Digitization Areas and Technology

TPT network members continue to struggle with staffing issues and access to collections due to COVID19. PI Zaspel has reached out to all PIs and collaborators and we are confident that digitization work will ramp up this summer (many Co-PIs have summer staffing plans in place).

Share Opportunities to Enhance Training Efforts

TPT has organized a monthly webinar series to help the network revise workflows, adopt standards for data entry, review data, and discuss opportunities for future research projects that will use TPT data. Mark Smith, from Macroscopic Solutions, kicked off the series with a presentation on using the Macropod

imaging system for taking pictures of slide-mounted specimens in January (61 participants). Neil Cobb gave a webinar on the functionality of SCAN and shared updates on progress towards a new TPT synthetic data portal in February (30 participants), The TPT Research Advisory Board (RAB) gave a webinar on research uses and projects in April (40 participants), and Co-PI Seltmann and Jorrit Poelen gave a hands-on workshop focused on data use and review in GloBI in April (30 participants).

Share Collaborations with other TCNs, Institutions, and/or Organizations

TPT is collaborating with Vectorbase, NMNH, and Walter Reed to aggregate occurrence and observation data, deliver association data to GloBI, and provide taxonomy resources to the arthropod collections community.

TPT started an informal collaboration with the Denver Museum of Nature and Science PI Zaspel coordinated with curator Demboski to publish current digitized records to SCAN and tie them to the TPT project.

Share Opportunities and Strategies for Sustainability

TPT PIs hold regular meetings with external stakeholders (e.g., Sam Rund Vectorbase, David Pecor Walter Reed Systematics Laboratory, Jessica Bird NMNH Entomology Collections data manager). TPT was also focused on delivery of BI products (teaching modules) this quarter.

The TPT research advisory board met in December and January to discuss a vision and mission for the group. We have developed a spreadsheet with ongoing research projects using TPT data and are tracking progress monthly.

The TPT FieldGuide working group has prepared >2,500 poultry louse images for the field book and will be moving into the training and model building stage next quarter.

Share Education, Outreach, Diversity, & Inclusion (EODI) Activities

TPT PI Seltman is leading a Computer Science capstone using Global Biotic Interaction as part of the South Central Coast Data Science Partnership (https://centralcoastdatascience.org). Here, CS and Statistics students are creating reusable scripts for acquiring, filtering, and visualizing GloBI data. This initiative is funded by the NSF Harnessing the Data Revolution, DSC:1924205.

PI Seltmann (UC Santa Barbara) is also teaching a Collection Curation course about TPT this Winter to 11 UCSB undergraduate students. Students in the course are transcribing TPT records from Notes from Nature and extracting tick biotic interaction data from the literature. The literature extraction methods are incorporating PLAZI Taxonomic Treatments. Donat Agusti and Marcus Guidoti from Plazi are collaborating on this component, along with Jorrit Poelen. The resulting workflow will take a spreadsheet of extracted biotic interaction information that includes a DOI for the citation and share with both GloBI and PLAZI.

During the class, guest speakers included Jennifer Zaspel (TPT PI), Felicia Pugh (epidemiology), Andy MacDonald (tick disease ecologist). Students in the course engage in natural history collection data broadly, including issues of diversity, inclusion, repatriation, and colonialism. They have a Carpentries Python Bootcamp weekend during the course taught by the UCSB Library Collaboratory, and students are encouraged to develop a research project after the course to continue as a directed study.

TPT completed a new education module that incorporates NfN into an evolution module that includes museum specimens, data science, community science, phylogenetics, coevolution, and speciation concepts. The goal of this module is to demonstrate how natural history specimens are invaluable to research and theory and allow students to gather real biodiversity data as part of their coursework.

Share Information About Your Website and/or Portal Usage

TPT was granted permission from NSF to move funds out of travel and participant support in order to build a synthetic portal for the project on the SCAN site. This work is currently under development but we anticipate completion by May/June 2021.

Share Other Activities and/or Progress

PIs are continuing data reviews for TPT participant interaction data. This resulted in a new method for finding taxon names that are not present in any of the known taxonomic APIs (EOL, GBIF, etc.). TPT participants can see this information by going to the https://www.globalbioticinteractions.org/parasitetracker/ webpage and clicking on the names button next to their institution.

Taxonomy progress

- Name cleaning script completed - https://github.com/Jegelewicz/tpt-siphonaptera
  See also: https://github.com/njdowdy/ixodes-tpt.git
- Development of author string parsing and name parsing of DiscoverLife taxonomic lists
  https://github.com/njdowdy/discoverLife_Apoidea
- Development Contributions to TaxoTools R package, e.g.:
  https://github.com/vijaybarve/taxotools/issues/86
  https://github.com/vijaybarve/taxotools/issues/78
- Generated up-to-date local databases of GBIF and CoL taxonomic backbones and development of scripts for querying TPT taxonomic resources against these data aggregators
  https://github.com/njdowdy/tpt-donut-plot.git
- Development of a web resource to help track taxonomy changes over time and centralize access to curated taxonomy lists
  https://github.com/njdowdy/taxotracker
  https://www.taxotracker.com
- Advisory meeting with louse team on strategies for implementing computer vision-based specimen identification w/ Andre Poremski
- Gave the TPT taxonomy team an overview of my methodology for storing taxonomic data in a database structure
  https://drive.google.com/file/d/1gl-sZFEjji4MxyJEzdmAt2id-fEtX15z/view?usp=sharing
- Recorded short demo of setting up R projects with github version control
  https://www.youtube.com/watch?v=o8EnynFzW-Q

Taxon-specific progress

Siphonaptera

- Review of existing files and determination of which should be included

- Name cleaning script completed - https://github.com/Jegelewicz/tpt-siphonaptera
- Reviewed 600+ names not in agreement with GBIF
- 300 + names needing review sent to taxonomy expert

Acari

- Review of existing files and determination of which should be included
- Name cleaning script completed - https://github.com/Jegelewicz/tpt-acari

Phthiraptera

- Cleaned and merged the data from four sources
- Louse taxonomy V1
- NHM Louse database
- GBIF names
- Total names ~28K organized into accepted names and synonyms
- Generated formatted documents for each family for manual corrections

Diptera

- Continued with sorting data.
- Downloaded GBIF data

General

- Ongoing discussions with GBIF regarding taxonomy backbone
    - https://github.com/gbif/portal-feedback/issues/3346
    - https://github.com/gbif/portal-feedback/issues/3371
- Discussion with GloBI re Ixodes duplication - https://github.com/globalbioticinteractions/globalbioticinteractions/issues/602
- Sustaining the momentum of digitization in biodiversity collections Panel Participant - https://www.idigbio.org/content/sustaining-momentum-digitization-biodiversity-collections