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

February 2020

CONTENTS:

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.

Published on *iDigBio* (https://www.idigbio.org)

Home > Collaborators > TCN Quarterly Progress Report to iDigBio > Webform results > TCN Quarterly Progress Report to iDigBio

Submission #1598

Submission information

Form: TCN Quarterly Progress Report to iDigBio
Submitted by kds15e
Tuesday, February 4, 2020 - 10:08
75.128.64.143

TCN Name:
Capturing California's Flowers: Using Digital Images to Investigate Phenological Change in a Biodiversity Hotspot

Person completing the report:
katelin.d.pearson24@gmail.com

Progress in Digitization Efforts:
see attached

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

Identify Gaps in Digitization Areas and Technology:
see attached

Share and Identify Opportunities to Enhance Training Efforts:
see attached

Share and Identify Collaborations with other TCNs, Institutions, and Organizations:
see attached

Share and Identify Opportunities and Strategies for Sustainability:
see attached

Share and Identify Education and Outreach (E&O) Activities:
see attached

Google Analytics

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

Attachment 1
February2020QuarterlyReport.pdf

Attachment 2

Source URL: https://www.idigbio.org/node/564/submission/1598

CALIFORNIA PHENOLOGY TCN – QUARTERLY REPORT – FEBRUARY 2020

Assembled by Katie Pearson, February 4, 2020

PROGRESS IN DIGITIZATION EFFORTS:

IMAGING

All institutions are continuing to image specimens or have achieved their imaging goals and have moved on to image processing, transcription, and/or georeferencing. Figure 1 shows the distribution of unprocessed, barcoded/processed, and imaged target specimens per institution as of February 3, 2020.

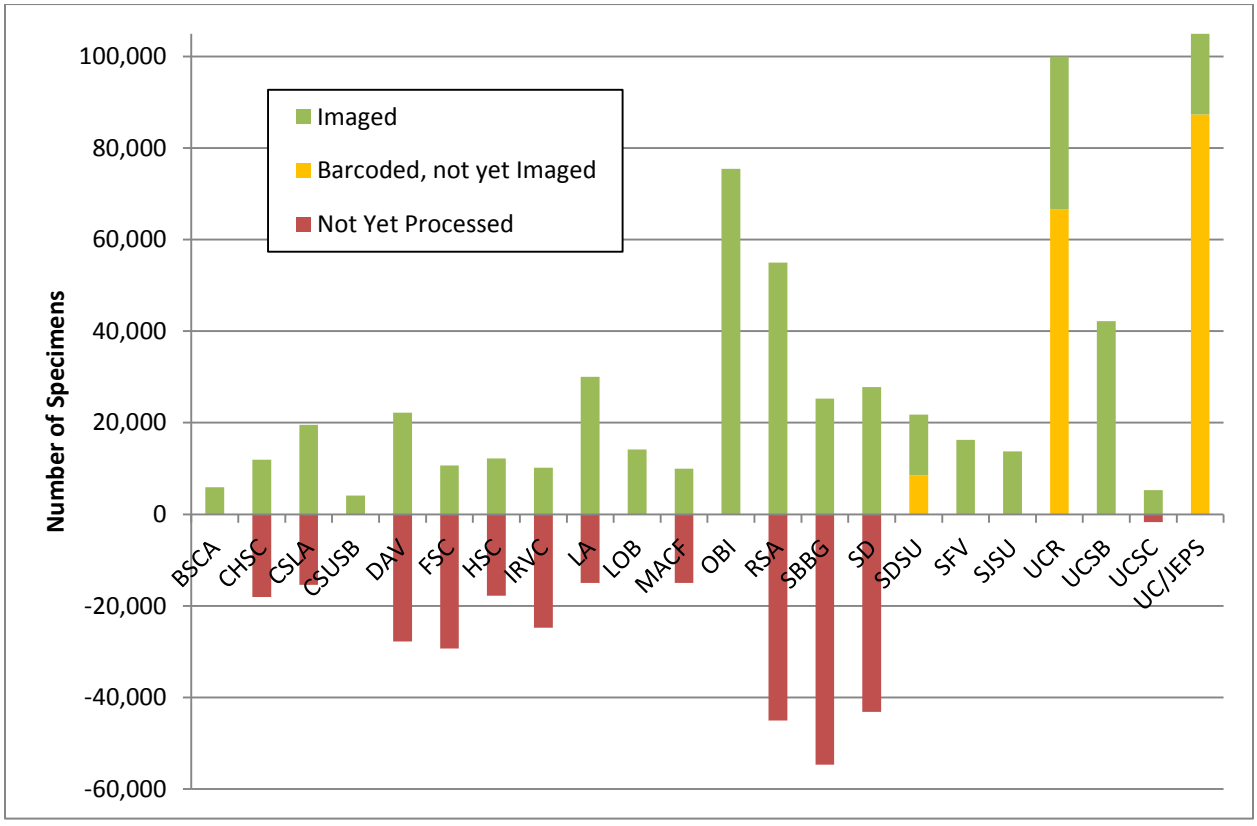


Figure 1. Digitization progress, in terms of number of specimens imaged, barcoded, or not yet processed. Bars above the zero line indicate specimens that have been processed in preparation for imaging or have been imaged. The green portions of these bars represent the number of specimens that have been imaged. Red bars below the zero line indicate the number of target specimens (i.e., specimens to be imaged as part of the CAP TCN) that have not yet been pre-processed or imaged.

TRANSCRIPTION

An estimated 30,000 specimen records have been transcribed across the CAP Network since July 2019. This is approximately one tenth of the goal number of transcriptions to be produced by this project.

Three Notes from Nature expeditions are currently live, representing specimens from four institutions (CSLA, FSC, IRVC, and LOB). These expeditions are 50% complete, representing 8,367 classifications (transcriptions) and resulting in 1,108 fully transcribed specimen records to date (each specimen must be transcribed three times to be considered "complete").

GEOREFERENCING

Since July 2019, 33,721 specimen records from CAP institutions have been georeferenced in CCH2. This is approximately one tenth of the goal number of georeferences to be produced by this project.

PHENOLOGICAL SCORING

Phenological scoring tools from images and text fields are now live in the CCH2 data portal but have not yet been implemented by CAP collaborators. Trial scoring at Cal Poly has resulted in phenological scores for 731 specimens.

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

Phenological scoring tools from images and text fields are now live in the CCH2 data portal. Protocols for using these tools are available on the project website: https://www.capturingcaliforniasflowers.org/documents.html. A new "Phenological Scoring" page is being developed to describe the phenological scoring schema, provide phenology-related protocols, and act as a repository for taxon-specific training materials as they are developed (https://www.capturingcaliforniasflowers.org/phenology.html).

Now that the phenological scoring tools are developed, we are creating a plan for scoring the intended >900,000 herbarium specimens. We plan to first prioritize California taxa that are easy to score for phenology and are well-represented in collections (e.g., *Eschscholzia, Lupinus*). Then we will move to less numerous taxa and taxa that will require more complicated training resources. The lead-PI, the PM, PI Susan Mazer, and graduate students Natalie Love and Tadeo Ramirez Parada are developing training materials for phenological scoring of specific taxa.

The PM led a pilot georeferencing training session for Cal Poly hired undergraduates. The questions that emerged from the training will help shape future georeferencing training efforts. Weekly georeferencing sessions are held at Cal Poly.

IDENTIFY GAPS IN DIGITIZATION AREAS AND TECHNOLOGY

Many CCH2 users, including CAP institutions, have indicated desire for new functionalities in CCH2, such as enabling searches by California Native Plant Society inventory ranking and native/non-native status. These requests, as well as bugs and problems, are now being tabulated and tracked in two GitHub repositories, the Wish List (https://github.com/CCH2-portal/CCH2-wish-list) and the Problems (https://github.com/CCH2-portal/CCH2-problems) repositories.

Currently, we have no solution for publishing data from "snapshot" institutions in CCH2 to GBIF. This is because not all snapshot institutions provide GUIDs for their records.

SHARE AND IDENTIFY OPPORTUNITIES TO ENHANCE TRAINING EFFORTS

Check-in calls with all institutions were held in November 2019. Year-end progress reports were issued to all CAP institutions in December 2019, which alerted some institutions to their need to accelerate digitization. The PM followed up with institutions that had questions about their progress or wanted to schedule trainings. An all-CAP conference call to discuss progress, updates, and the phenological scoring and georeferencing tools is scheduled for February 7th, 2020.

The PM visited CAP institutions during the month of January to provide training and troubleshooting help. Chico State (CHSC) had experienced significant delays in imaging and was experiencing a bottleneck in image processing. After the Northern California Botanists Symposium, where the PM gave a lightning talk, the PM spent a day with the CHSC PI and staff to help overcome these issues.

On a separate trip, the PM visited the herbaria at UCLA, Cal State LA, and UC Irvine, providing training in data entry, general use of the CCH2 data portal, image processing, and imaging, as needed.

The PM is working with PI Mare Nazaire (RSA) on a georeferencing training webinar on February 26th.

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

A workshop submitted to the Ecological Society of America meeting was not accepted; however, the PM will present a lightning talk at an ESA INSPIRE session at the ESA 2020 meetings.

Presentations about the CAP Network and community engagement were given to two separate chapters of the California Native Plant Society.

In mid-December, Faerthen Felix and Erica Krimmel brought the entire Sagehen Creek Field Station (SCFS) herbarium collection to Cal Poly. After establishing a workflow and receiving training, Felix and Krimmel used the Cal Poly digitization equipment to image all the specimens in three days (December 16-18), resulting in 1,069 specimen images that were immediately processed and made accessible online via CCH2. SCFS now manages their data live in CCH2.

SHARE AND IDENTIFY OPPORTUNITIES AND STRATEGIES FOR SUSTAINABILITY:

Following the sudden power outage at UF, the CCH2 data portal was transferred to servers at Arizona State University, which has considerable cyberinfrastructure that indicates future sustainability.

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

An announcement article about the CAP TCN was published in the peer-reviewed journal *Madroño* https://doi.org/10.3120/0024-9637-66.4.130.

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

Five blog posts were written and published to the CAP website: https://www.capturingcaliforniasflowers.org/blog-recap.

The PM and Cal Poly herbarium students led a workshop prior to the monthly meeting of the San Luis Obispo chapter of the California Native Plant Society on November 7th, 2020. The PM gave a short presentation on the project and Notes from Nature, and participants practiced using Notes from Nature to transcribe specimen labels.

The PM attended the Northern California Botanists Symposium January 13-14, 2020 and presented a lightning talk about using Notes from Nature to help California herbaria. The lightning talk was recorded and posted on the CNPS YouTube channel: https://www.youtube.com/watch?v=AQv3EXy03r8 (starting at 25:32). The PM also presented this talk to the executive board of the Mt. Lassen Chapter of the California Native Plant Society on January 15th. A higher quality version of this lightning talk will be recorded and posted on the CAP TCN YouTube channel in the near future.

As part of the training trip to UC Irvine, the PM also co-led a class of approximately 30 introductory biology students. Rebecca Crowe (collections manager at UC Irvine) and the PM introduced the students to herbarium curation, specimen-based research, and specimen digitization. The PM then trained half of the students in using Notes from Nature to transcribe label data, and Crowe trained the other half in pre-curation steps including barcoding.

The PM and lead-PI Jenn Yost, along with PI Katja Seltmann and graduate students Natalie Love and Tadeo Ramirez Parada at UC Santa Barbara, are developing a Course-based Undergraduate Research Experience curriculum that teaches students to use data from CCH2 to examine plant phenological responses to climate. Currently, the course is being iteratively developed with the help of four advanced undergraduate students. We aim to produce robust documentation such that the course could be easily adopted and run at other universities and colleges. Once final drafts are available, the course materials will be posted on the CAP website. A pilot session of the course will run in its entirety at Cal Poly during the spring quarter.

Published on *iDigBio* (https://www.idigbio.org)

Submission #1595

Submission information

Form: TCN Quarterly Progress Report to iDigBio
Submitted by neilscobb
Friday, January 31, 2020 - 18:46
47.215.133.118

TCN Name:
Southwest Collections of Arthropods Network (SCAN): A Model for Collections Digitization to Promote Taxonomic and Ecological Research

Person completing the report:
neilscobb@gmail.com

Progress in Digitization Efforts:
See attached

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

Identify Gaps in Digitization Areas and Technology:
See attached

Share and Identify Opportunities to Enhance Training Efforts:
See attached

Share and Identify Collaborations with other TCNs, Institutions, and Organizations:
See attached

Share and Identify Opportunities and Strategies for Sustainability:
See attached

Share and Identify Education and Outreach (E&O) Activities:
See attached

Google Analytics

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

Attachment 1

LepNet_SCAN_February_2020.docx

Attachment 2

Source URL: https://www.idigbio.org/node/564/submission/1595

Published on *iDigBio* (https://www.idigbio.org)

Submission #1596

Submission information

Form: TCN Quarterly Progress Report to iDigBio
Submitted by neilscobb
Friday, January 31, 2020 - 18:46
47.215.133.118

TCN Name:
Lepidoptera of North America Network: Documenting Diversity in the Largest Clade of Herbivores

Person completing the report:
neilscobb@gmail.com

Progress in Digitization Efforts:
See attached

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

Identify Gaps in Digitization Areas and Technology:
See attached

Share and Identify Opportunities to Enhance Training Efforts:
See attached

Share and Identify Collaborations with other TCNs, Institutions, and Organizations:
See attached

Share and Identify Opportunities and Strategies for Sustainability:
See attached

Share and Identify Education and Outreach (E&O) Activities:
See attached

Google Analytics

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

Attachment 1

LepNet_SCAN_February_2020.docx

Attachment 2

Source URL: https://www.idigbio.org/node/564/submission/1596

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

January 31, 2019
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, January 28, 2020. **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 (LepNet)
Specimen Records	22,676,043	18,419,006	4,257,037
# Georeferenced	19,086,256	15,362,486	3,723,770
# Imaged	3,903,414	2,461,799	1,441,615
# Identified to species	13,640,873	9,554,232	4,086,641

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 47 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, 170,854 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 87% 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 74% 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 the first LepNet Partners to Existing Networks (PEN) grant with the San Diego Museum of Natural History. This PEN project will focus on the Lepidoptera of Baja California, including a large number of historical records. They have already contributed over 13,000 records.

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

Taxa	# Specimen Records	# Georeferenced	# Specimen Identified to species	# Georeferenced & Ided to species
Nymphalidae	866,504	794,766	853,030	785,577
Noctuidae	551,907	497,764	532,520	484,671
Erebidae	409,709	364,190	391,397	350,357
Geometridae	354,158	311,771	338,290	298,221
Hesperiidae	343,517	290,141	335,880	283,861
Pieridae	341,709	285,533	337,471	282,400
Lycaenidae	271,896	242,957	267,095	239,354
Papilionidae	170,456	142,609	168,960	141,616
Crambidae	157,158	133,565	151,543	129,497
Tortricidae	134,876	107,454	125,529	100,934

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 scores of papers.

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	1,191,047	1,074,097	690,977	615,852
Carabidae	622,726	507,590	391,047	323,411
Araneae	252,097	198,201	208,854	169,032
Acrididae	431,679	218,036	368,783	203,830
Tenebrionidae	192,506	167,147	113,304	99,095

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 primarily working with other Symbiota TCNs and other Symbiota portals. 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 217 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 three new collections since the last update.

GBIF Registration - There are 47 Live collections on SCAN that are now registered with GBIF and 87 other entomology collection datasets from the North America being served on GBIF for a total of 181 datasets. This leaves approximately 30 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 North American arthropod data.

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 .

Published on *iDigBio* (https://www.idigbio.org)

Home > Collaborators > TCN Quarterly Progress Report to iDigBio > Webform results > TCN Quarterly Progress Report to iDigBio

Submission #1605

Submission information

Form: TCN Quarterly Progress Report to iDigBio
Submitted by mwdenslow
Wednesday, February 5, 2020 - 15:10
67.190.87.86

TCN Name:
SERNEC: The Key to the Cabinets: Building and Sustaining a Research Database for a Global Biodiversity Hotspot

Person completing the report:
michael.denslow@gmail.com

Progress in Digitization Efforts:
There are 117 collections serving data through the SERNEC portal. There are currently 4,903,407 specimen records and 450,080 (9%) of those records are georeferenced. There are currently 4,405,757 imaged specimen images available. There are currently 68 collections publishing to iDigBio.

Share and Identify Best Practices and Standards (including Lessons Learned):
The SERNEC – TCN protocols continue to be updated as needed and are posted on the SERNEC resources site (http://sernec.appstate.edu/resources).

Identify Gaps in Digitization Areas and Technology:
Nothing to report

Share and Identify Opportunities to Enhance Training Efforts:
Nothing to report

Share and Identify Collaborations with other TCNs, Institutions, and Organizations:
Nothing to report

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

Share and Identify Education and Outreach (E&O) Activities:
SERNEC continues to have a large on active presence on Notes from Nature (https://www.notesfromnature.org/active-expeditions/Herbarium) with many expeditions running
concurrently.

We have a new area of the SERNEC website dedicated to volunteer opportunities. https://herbarium.appstate.edu/sernec/volunteer-sernec

Google Analytics

Other Progress (that doesn't fit into the above categories):
All SERNEC – TCN PIs have now submitted their final reports. Appalachian State University has will remain active for an additional extension year to provide project support for data portal activities, citizen science and georeferencing.

Attachment 1

Attachment 2

Source URL: https://www.idigbio.org/node/564/submission/1605

Published on *iDigBio* (https://www.idigbio.org)

Submission #1606

Submission information

Form: TCN Quarterly Progress Report to iDigBio
Submitted by cskema
Wednesday, February 5, 2020 - 16:42
165.123.74.113

TCN Name:
The Mid-Atlantic Megalopolis: Achieving a greater scientific understanding of our urban world

Person completing the report:
cskema@upenn.edu

Progress in Digitization Efforts:
Please see attached pdf.

Share and Identify Best Practices and Standards (including Lessons Learned):
Please see attached pdf.

Identify Gaps in Digitization Areas and Technology:
Please see attached pdf.

Share and Identify Opportunities to Enhance Training Efforts:
Please see attached pdf.

Share and Identify Collaborations with other TCNs, Institutions, and Organizations:
Please see attached pdf.

Share and Identify Opportunities and Strategies for Sustainability:
Please see attached pdf.

Share and Identify Education and Outreach (E&O) Activities:
Please see attached pdf.

Google Analytics

Other Progress (that doesn't fit into the above categories):
Please see attached pdf.

Attachment 1

2020_02_MAM_Quarterly_Progress_Summary.pdf

Attachment 2

Source URL: https://www.idigbio.org/node/564/submission/1606

Mid-Atlantic Megalopolis TCN

Quarterly Progress Report[1]

November 2019 – January 2020

Progress in Digitization Efforts: Figure 1 shows progress over time for the MAM Project by changes in the number of both specimens entered into workflow and completely digitized specimens (= imaged + transcribed + georeferenced) against the number of specimens promised to NSF for the project. The current numbers for progress of digitization efforts by specimen category for each herbarium are shown in Table 1 and Figure 2. Although updated progress for HUDC and NY are no longer included in these reports as they closed their NSF grants on the MAM Project as of 31 August 2019, their total numbers of specimens completed in the MAM Project to that date are still reported in Figure 1 and Table 1 in the interest of showing project totals.

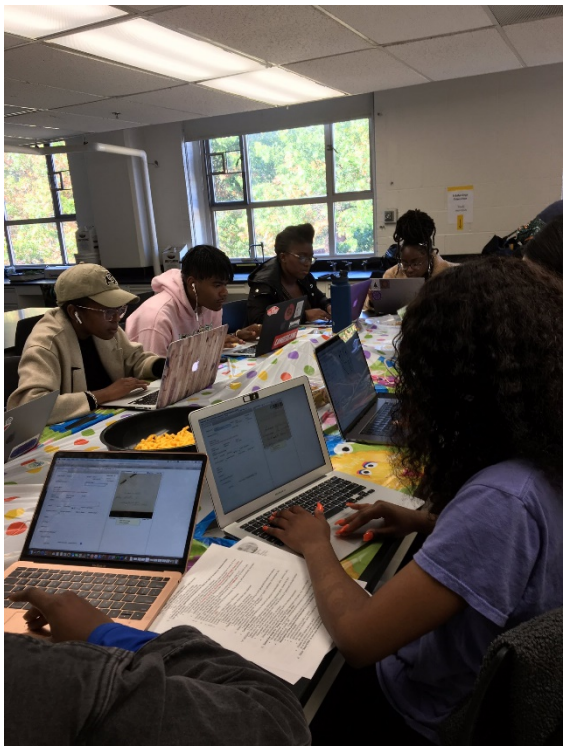


Figure 1. The Plant Diversity class at Howard University (HUDC) works on transcribing records for Delaware State University (DOV) with guidance from Dr. Janelle Burke.

Share and Identify Best Practices and Standards: Nothing to report.

Identify Gaps in Digitization Areas and Technology: Nothing to report.

Share and Identify Opportunities to Enhance Training Efforts: Nothing to report.

Share and Identify Collaborations with other TCNs, Institutions, and Organizations: Nothing to report.

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

Share and Identify Education and Outreach Activities: Students in the General Botany course at Delaware State University were introduced to the virtual herbarium of DOV and helped transcribe roughly 2,435 records as part of their coursework in the fall semester. CHRB has been doing herbarium outreach and tours with various groups, including the Rutgers University Plant Ecology course (50 students), Willowwood Arboretum, Morris County Park Commissions, and New Jersey Department of Environmental Protection. As part of WeDigBio, HUDC held a transcription blitz to help transcribe records for their fellow HBCU herbarium, DOV, with about 10 students transcribing about 250 records in total (see photo at left). PAC held three transcription events, one as a part of a regular workshop series, one for the Pennsylvania Native Plant Society, and a third in conjunction with WeDigBio. [The PAC and HUDC activities were missed in the last quarterly report.]

[1] Throughout this report, herbaria are referred to by their Index Herbariorum acronyms, which correspond to institutional names as follows: BALT = Towson University, CHRB = Rutgers University, CM = Carnegie Museum, DOV = Delaware State University, HUDC = Howard University, MARY = University of Maryland, MCA = Muhlenberg College, MOAR = Morris Arboretum of the University of Pennsylvania, NY = New York Botanical Garden, PAC = Pennsylvania State University, PH = The Academy of Natural Sciences of Drexel University, SIM = Staten Island Museum, TAWES = Maryland Department of Natural Resources

Other Progress: Nothing to report.

Figure 1. Progress over time for MAM Project.

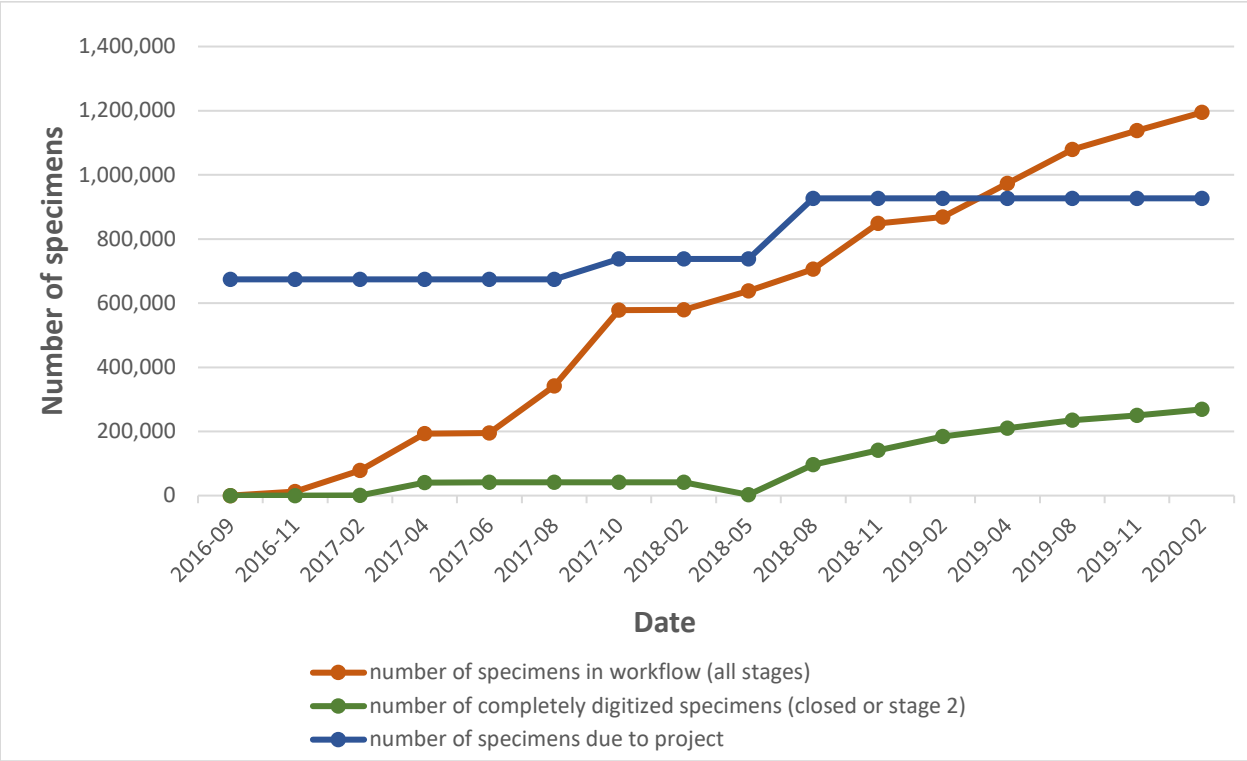
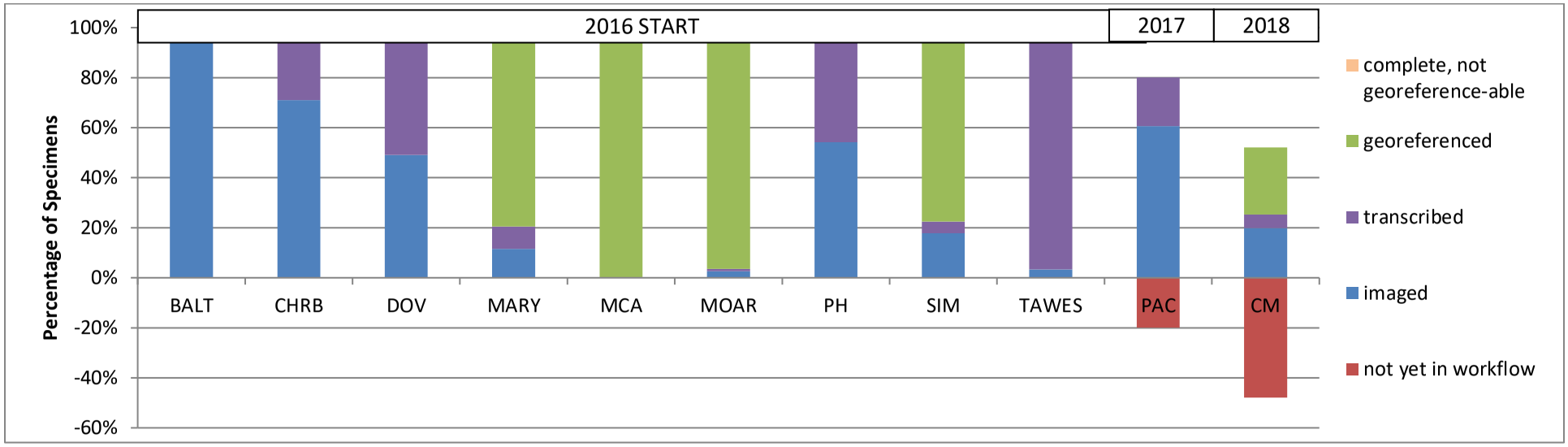


Table 1. Digitization of specimens by stage of completion and herbarium for MAM TCN.

	Herbarium											Totals
Specimen Stage	BALT	CHRB	DOV	MARY	MCA	MOAR	PH	SIM	TAWES	PAC	CM	
# specimens imaged[1]	30,000	1,003	0	0	0	0	0	0	0	615	37,456	**69,074**
# specimens as above and uploaded to Symbiota along with skeletal data; transcription/review may be in progress[2]	6,570	46,155	26,177	5,114	0	27	243,909	0	34	38,153	0	**515,900**
# specimens as above + completely transcribed and transcription reviewed[3]	0	19,277	25,044	4,036	0	185	201,055	1,037	3,362	12,485	10,257	**335,208**
# specimens as above + georeferenced[4]	0	0	2,079	35,763	51,009	20,335	5,268	15,997	149	0	50,436	**267,597**
# specimens that need special attention, e.g. go back to sheet[5]	127	0	0	124	2	534	351	3,857	86	0	32	**5,280**
# specimens imaged, uploaded, transcribed BUT not able to be georeferenced[6]	2	0	0	102	52	41	6	848	13	0	0	**1,202**
Totals	**36,699**	**66,435**	**53,300**	**45,139**	**51,063**	**21,122**	**450,589**	**21,739**	**3,644**	**51,253**	**98,181**	**1,194,261**

Processing Status in the MAM Portal: [1] No stage, not in Symbiota yet; [2] Unprocessed + Expert Required + Pending Review; [3] Stage 1; [4] Stage 2; [5] Stage 3; [6] Closed

Figure 2. Percentage of specimens by stage of completion and herbarium for MAM TCN. With this presentation of digitization progress, the final goal for each institution is to have a mostly green column above the X axis (could potentially have orange up to roughly 10%). Specimens not yet in workflow are set as negative numbers.



Published on *iDigBio* (https://www.idigbio.org)

Submission #1604

Submission information

Form: TCN Quarterly Progress Report to iDigBio
Submitted by jrallen99
Wednesday, February 5, 2020 - 13:40
128.138.130.189

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

Person completing the report:
james.allen@colorado.edu

Progress in Digitization Efforts:
Collectively for the current quarter roughly November 2019- January of 2020 we have entered 49,133 new records into databases, barcoded 63,818 new specimens, imaged 76,081 new specimens and georeferenced 49,364 new records.
In aggregate the project has now produced 315,186 new database records, 852,748 newly barcoded specimens, 864,556 new images and 120,599 new georeferences.
Georeferencing numbers for this quarter almost match total progress from the first two years of the project 49,364 georeferences for the quarter vs. 51,894 for (Sept 2017-July 2019). Data entry numbers are also trending higher compared to the first two years of the project. This represents a shift in the focus of the project away from imaging specimens and towards the goal of creating database entries and georeferencing specimens. Imaging and barcoding for collections that started in year one of the project collections designated as large scope and small scope are mostly complete.
The project after 29 months (out of 48) has completed. (Total for TCN and PEN)
Original TCN Progress
Data Entry 62.5%
Barcodes 101.9%
Images 106.0%
Georeferences 21.5%

Rancho Santa Ana Botanic Garden was added as a PEN on September 1st 2019. Starting with the next report we include data for the entire project including the PEN in a single metric. The PEN adds 42,600 new images, 49,000 new database entries, and 56,600 new georeferences. Coupled with existing data this project will add ~60,700 new SoRo specimens.
TCN plus PEN Progress

Data Entry 57.1%
Barcodes 97.1%
Images 101%
Georeferences 19.6%

The SoRo TCN has also brought several databases online over the past year of the project that were not previously available outside of local networks. At least 50,000 pre-existing database entries have been made available for the first time from CSCN, GREE and BHSC. This is in addition to the records reported above. We will add addition data from unfunded partners at federal institutions including several national park collections and the Smithsonian.

Share and Identify Best Practices and Standards (including Lessons Learned):
SJNM - We have been working at improving our method of scheduling student workers to take images. Initially we developed a weekly sign-up sheet. Students quickly signed up for all of the available times, but due to various issues, students were not completing their time allotments. This left the camera unused and other students not aware that it was available. We are also dealing with the end of semester issues with campus closed and students not available over the break. With the start of January, we hired a former student who starting to put in a full 20 hours per week.
ASC - Data cleaning has begun to track down all specimens without images and all images without specimen data. About 75 specimen records have been cleaned/fixed this quarter – most of these fixes involved duplicate or incorrect barcode fixes. Stats page states that we still have 38 images that are not mapping to specimen data.
GREE - Our 9000 specimens with a georeference update are all specimens that had existing lat/long on their labels but no other details. They had not been previously georeferenced as part of SoRo, but rather had existing coordinates. A large chunk of those were collected as part of a flora by a single collector, so I was able to reach out to him to learn about how you took points (uncertainly), what sort of GPS unit and datum he used, and other relevant details. With our next report we will have many specimens that were 'new' SoRo georeferences from last fall, but many of the DarwinCore fields were not present in our specify database so that info did not get recorded. COLO – This is a great example of using resources to improve the quality of legacy data. Reaching out to prolific collectors in our collections can go a long way to strengthening the quality of our data.
HUH - We switched to our new transcription app tool. This means that staff is barcoding and imaging specimens first. After the images are electronically processed and accessible from a queue, specimens are databased from the image. This has the advantage that staff does not need to look back and forth from an actual specimen sheet on the desk to the computer screen, but instead can stay in an ergonomically straighter position while databasing the specimens. Also, the specimens are only handled once.
CSCN - After the records were entered, I noticed an unnerving number of errors, especially involving missing or incorrect coordinates. Many of these errors appear to have been due to careless data entry by a former student worker, but there are also a number of other records with bizarre insertions of locality offsets that appears to date back to an earlier data migration within Specify. In any case, the large number of errors necessitates a different approach. As our collection is not especially widely distributed and has a lot of student collections I have decided the most efficient approach to begin is to target collections made by students who collected primarily or entirely from a few localities within the Nebraska part of the region. Once such a collector is identified, a query is run and arranged quantitatively so that any anomalous locality data can be recognized (in these cases I pull the specimen and often re-enter data if the specimen has a barcode). Since localities were entered in a very different format in our Specify database, batch

georeferencing is of limited use, as a single locality may appear in multiple different iterations. We have only just started this approach and will have to evaluate its effectiveness on a later date.
RSA - Our collection is organized regionally, with regions segregated into color-coded folders. Because our collection's emphasis is on California, California specimens are segregated from the remainder of North American specimens. We have found that it is much more efficient for staff to pull all folders of North American taxa, image, and then refile back into the collection than the time it takes to search and pull specimens specific to the SoRo region (although with caveats noted below). During the imaging process, when interns save an image of the specimen, they also enter the barcode and state from a drop down menu in a Google spreadsheet. Cells for the specimens in the areas within the SoRo region automatically color-code red, and those specimens are targeted for data entry in the SoRo project. During a data entry session, the intern pulls up the image from the RSA Herbarium's server and enters the data from the generated image. Specimens imaged from outside of the SoRo region are databased by Herbarium core staff (i.e., funded through institutional funding). As noted in the above workflow, in which we are pulling all of North American collections to image, including target and non-target specimens, might work better with smaller collections than a collection of our size; we are trying to determine if this approach works well for us or if it would be better to just pull the target specimens. By my calculations, we may image four times as much as the target specimens – we just want to be sure that we stay on track with our imaging progress. Our strategy for imaging and data entry is to prioritize selected taxa (e.g., Cactaceae, Astragalus, Lupinus, and Penstemon) that have strong representation in the SoRo region.

Identify Gaps in Digitization Areas and Technology:
SJNM - We are still having issues obtaining a server to host our images. Our IT department is looking into several options, including hosting the images on our own servers. SJNM will coordinate COLO and SEINet staff to work towards a setup to get the rest of their images online.
GREE - We have yet to upload any images to Symbiota. We are in discussion with the UNC Libraries about using the Digital Commons platform through our library portal https://digscholarship.unco.edu. The UNC Libraries already hosts photoplot data for various ecological projects, so this would be a good avenue for serving our specimen images. Other herbaria (University of Mary Washington, Utah State University – Type specimens, Wellesley College) are using digital commons and the UNC Library staff are enthusiastic about this collaboration.
HUH - We are organizing a georeferencing training session for our staff members and have been in contact about additional support with Ryan Allen from COLO. Our Director for Biodiversity Informatics will implement the Geolocate tool into our house-intern transcription tool. This way we will be able to add data georeferenced information to our database and push them out to SEINet through IPT.

HUH - The total of 97,564 specimens refers to SoRo data on the state level, not the county level. Our limitations to filter records on county level had been communicated with COLO at the beginning of the project and COLO had agreed to accept our specimen records based on the state.
COLO – Issues with separating collections to fit into a project is a common theme for large collections that we do not have in small and medium collections. It is important to understand the differences in scope for small (<25,000) medium (~25,000-100,000) and large collections. Small and medium collections tend to be very focused on their region and it is often possible to image all material and have the majority of the material be "to scope." Large collections often will need to image more collections to get to the ones in

scope and we would advise future projects to build in extra imaging for the sake of data discovery and filtering because it is more efficient to image entire folders of plants than it is to select individual specimens. This also minimizes fragmentation of the collection and the need for complicated tracking of what has and has not been completed. We have had some success in the project by focusing on specific families that are likely to have more "to scope" specimens than others.

Share and Identify Opportunities to Enhance Training Efforts:
UNM - Many of the specimens that we have mapped thus far have location information that is difficult to interpret. While this doesn't maximize the georeference numbers we report it has been a useful training exercise to start with some of the most difficult specimens. This process has allowed the student-employees to thoroughly explore the decision making process before we move on to specimens that can be mapped more easily. While it may make more sense to some to go for the low hanging fruit, we're happy with the student development that is taking place.
UNM - Each student chooses one New Mexico County and becomes familiar with that area. Rather than bouncing around from county to county in a very large state they are becoming quite familiar with the counties they have chosen to focus on.
COLO – Building geographic expertise is paramount to success in georeferencing. Assigning georeferencing by counties not only helps to build this expertise, but it also minimizes repeated work when georeferencers start to recognize similar localities and can build on previous records and experience.

Share and Identify Collaborations with other TCNs, Institutions, and Organizations:
HUH - We just opened a new, additional digitization posting for TORCH. This will benefit also the SoRo project since on a state level, TORCH and SoRo overlap.
RSA - We are part of the California Phenology TCN (CAP) to digitize target taxa for scoring phenology and using this as a proxy to investigate how climate change may be influencing flowering and fruiting times in California native plant species. California State Polytechnic University, San Luis Obispo is the lead institution. Because the CAP project has set up a CyVerse account where we load images for that project, it is much more efficient for us to load all images to CyVerse (CAP, SoRo, and all incoming, new acquisitions) where all of the images are served in Symbiota. To date we have loaded nearly 47,000 images to Symbiota via CyVerse.

FLD – I have begun a collaboration with the Ute Mountain Ute Tribe to be involved in rare plant research and a data sharing agreement has been agreed upon between the tribe and Fort Lewis College. As the herbarium is concerned we are also planning a series of targeted floristic surveys on tribal land to help identify areas of higher diversity with the vouchers for this project being filed at FLD.

Share and Identify Opportunities and Strategies for Sustainability:
RSA - All data (images, databased records, georeferenced coordinates) are entered directly into RSA's institutional database. This is maintained with institutional support and does not rely on external funding. All data generated as part of this project will become part of RSABG's digital assets, will be managed in accordance with RSABG's digital asset management plan and will persist indefinitely. RSABG has permanent curatorial staff tasked with management duties and is supported by RSABG IT staff.
COLO had its biggest turnover of digitization staff in the 8 years they have been digitizing. Developing training strategies to get new staff up and running are imperative. Creating documentation of workflows should be undertaken at each institution. These documents

are proving to be great resources to train new digitizers.

Share and Identify Education and Outreach (E&O) Activities:
COLO and FLD are exploring options to hold another Field to Digital Object workshop.
FLD - This semester my Plant Systematics course is completing a course-based research project to produce a floral checklist for the Ute Mountain Ute Tribal land based solely on databased and digitized collections. The tribe lacks any kind of plant species checklist for their lands and would like one as they expand on their focus in natural resource management. (This past summer they just hired for the first time a biologist). Many collections have been made from tribal lands but that information has never been shared with tribal authorities. My class is using databased information to generate a vouchered list of species, identify habitat types for species of interest, identify specific areas of interest based on past collections, and compile information on collection activities and collectors who have been active on tribal lands. This is giving students a lot of practice using digitized data and producing an important product based on the herbarium community's work mobilizing data.
CSCN - An exhibit covering the history of the herbarium is now on display as part of a showcase exhibit on campus. I have been invited to speak on the herbarium February and will discuss our current project progress.
RSA - Last semester we partnered with Dr. Ed Bobich at California Polytechnic Institute, Pomona to recruit undergraduate students from his Form & Function in Plants class to barcode as partial fulfillment of their course objectives. We had four barcoding sessions (informally called Barcoding Bonanzas) with the students. In total, 23 students barcoded ~4,000 specimens. Students learned about the specimens that they barcoded as well as reinforcing concepts that they learned in their class. At the end of the barcoding session students were given a tour of the herbarium to learn about the RSA collection. Our poll with the students illustrate that ~95% of the students have never been to the Garden before and ~98% have never set foot in a natural history collection before. We have found this to be a great strategy for recruiting interns on projects such as these, and have recruited two interns that began in January. Nazaire also wrote an article for Oak Notes, newsletter of the Rancho Santa Ana Botanic Garden Volunteers, to share some of the digitization efforts that are taking place in the RSA Herbarium. The SoRo project is mentioned in this article and is attached along with this report.

Google Analytics

Other Progress (that doesn't fit into the above categories):
COLO - All of the funded collections now have data in the SoRo Portal (SEINet infrastructure). The project is starting to work with collections that need to get data to iDigBio and GBIF, we are aggregating data from the quarterly reports and doing follow-ups to determine which collections need to be updated to share data. This is a goal for the next two quarters.
CSCN - Renovation of our building is scheduled to begin this summer, and we are to have everything moved out of the herbarium space by the end of May 2020. I have secured an adequate space for the herbarium in a nearby building so that progress can continue after the move. We will focus on georeferences during the period of the move.

Google Analytics for our SoRo site are showing an increase of usage.
This quarter saw 935 users over 1,466 sessions with a total of 19,190 pageviews.
Last quarter ending October of 2019, we had 647 users over 968 sessions with a total of 12,228 pageviews.
The quarter ending July of 2019 had 378 users over 546 sessions and 1,855 pageviews

These metrics represent an increase of usage of ~50% over the previous quarter. We are seeing an ~2.5 fold increase for users and sessions and 10 fold increase in pageviews compared to the quarter ending July of 2019 compared to this quarter. We do suspect that most of the use of SoRo data passes through the main SEINet page, but it is encouraging to see that the project site is also seeing good usage.

Attachment 1
SoRo_Analytics All Web Site Data Audience Overview 20191101-20200131.pdf

Attachment 2

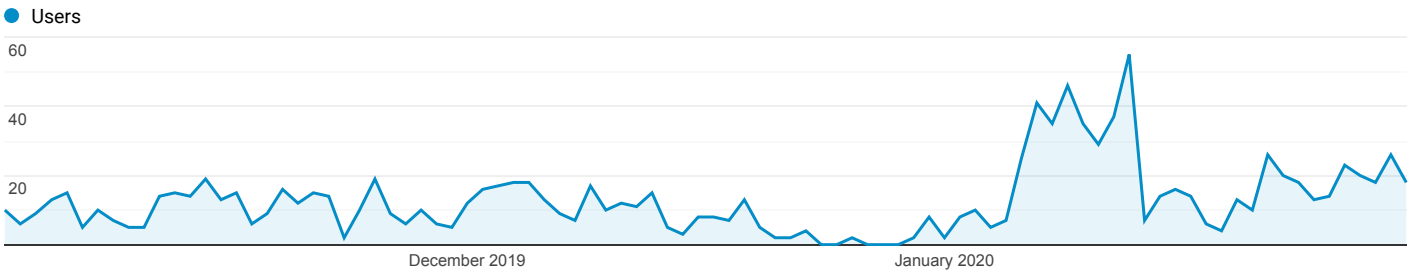
Source URL: https://www.idigbio.org/node/564/submission/1604

Audience Overview

All Users
100.00% Users

Nov 1, 2019 - Jan 31, 2020

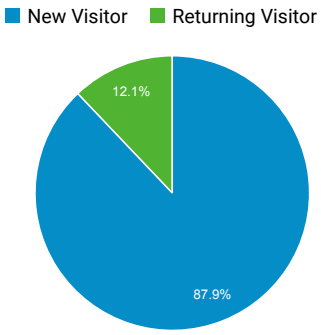
Overview



Users	New Users	Sessions
935	908	1,466

Number of Sessions per User	Pageviews	Pages / Session
1.57	19,190	13.09

Avg. Session Duration	Bounce Rate
00:18:57	44.13%



	Language	Users	% Users
1.	en-us	486	51.98%
2.	en-gb	74	7.91%
3.	zh-cn	39	4.17%
4.	es-es	30	3.21%
5.	de-de	28	2.99%
6.	de	17	1.82%
7.	fr	16	1.71%
8.	it-it	15	1.60%
9.	ru-ru	15	1.60%
10.	fr-fr	14	1.50%

Published on *iDigBio* (https://www.idigbio.org)

Submission #1599

Submission information

Form: TCN Quarterly Progress Report to iDigBio
Submitted by amiller
Tuesday, February 4, 2020 - 10:46
192.17.34.136

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

Person completing the report:
amiller@inhs.illinois.edu

Progress in Digitization Efforts:
• 6000 images from NYS have been uploaded to the MyCoPortal

Share and Identify Best Practices and Standards (including Lessons Learned):
• Nothing new to report.

Identify Gaps in Digitization Areas and Technology:
• Nothing new to report.

Share and Identify Opportunities to Enhance Training Efforts:
• Nothing new to report.

Share and Identify Collaborations with other TCNs, Institutions, and Organizations:
• The data snapshot for NYBG was updated, adding 21,057 occurrence records, ~14,200 georeferencings, and 11,803 images to MyCoPortal. (November 25th, 2019, and December 19th, 2019). Furthermore, 468,275 hot-links were added to the snapshot on MyCoPortal, creating a direct link back to the original records hosted at NYBG.
• Fdex, a simplified names database, primarily for projecting taxonomic hierarchy for fungal taxa has been added to the MyCoPortal thanks to a collaboration with MiCC and Scott Bates at Purdue University Northwest.

Share and Identify Opportunities and Strategies for Sustainability:
• MyCoPortal has been migrated from U of Florida servers to U of Illinois servers by Phil Anders, Illinois Natural History Survey biological informatician. (December 2019). Although a few snags were hit initially, all wrinkles are now ironed out and the MyCoPortal is back to normal functioning (January 2020).

• Taxonomy tree / Thesaurus on MyCoPortal is being cleaned/updated to better match MycoBank and Index Fungorum data, facilitating the interchange of information with the databases hosted on those portals. (December 2019 – January 2020)

Share and Identify Education and Outreach (E&O) Activities:
• Nothing new to report.

Google Analytics
iDigBio_Google_analytics_report_Q1_2020.pdf

Other Progress (that doesn't fit into the above categories):
• Please see MyCoPortal Data Portal Statistics generated from Google Analytics (attached).

• Diego Barroso has left his position as Project Manager of the MyCoPortal, to take up a new position as Project Manager for the TORCH TCN led by BRIT in Fort Worth, TX.
Publications

• MyCoPortal has been cited 54 times, 13 times in 2019, in peer-reviewed journal publications.

Presentations
• Nothing new to report.

Attachment 1

Attachment 2

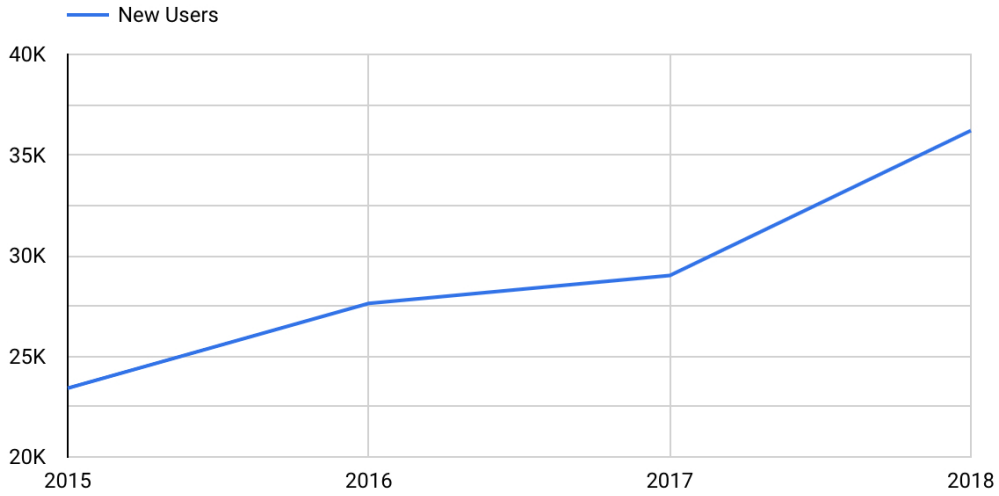
Source URL: https://www.idigbio.org/node/564/submission/1599

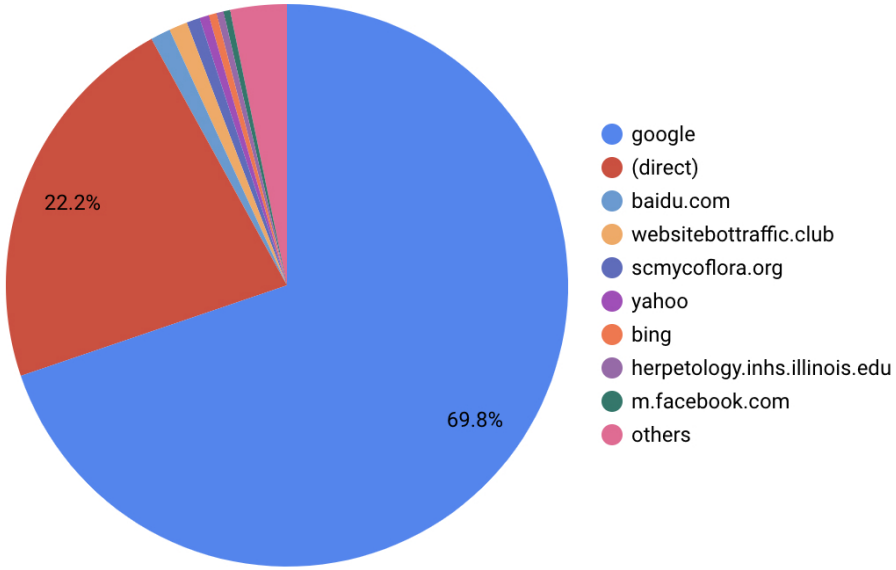
MyCoPortal Data Portal Statistics

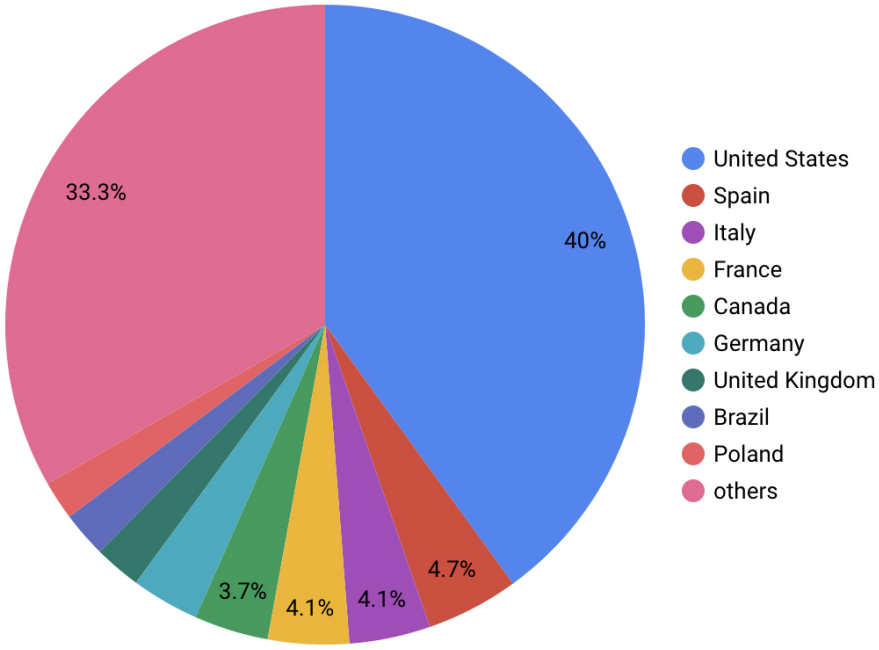
www.mycoportal.org **Data from Google Analytics**

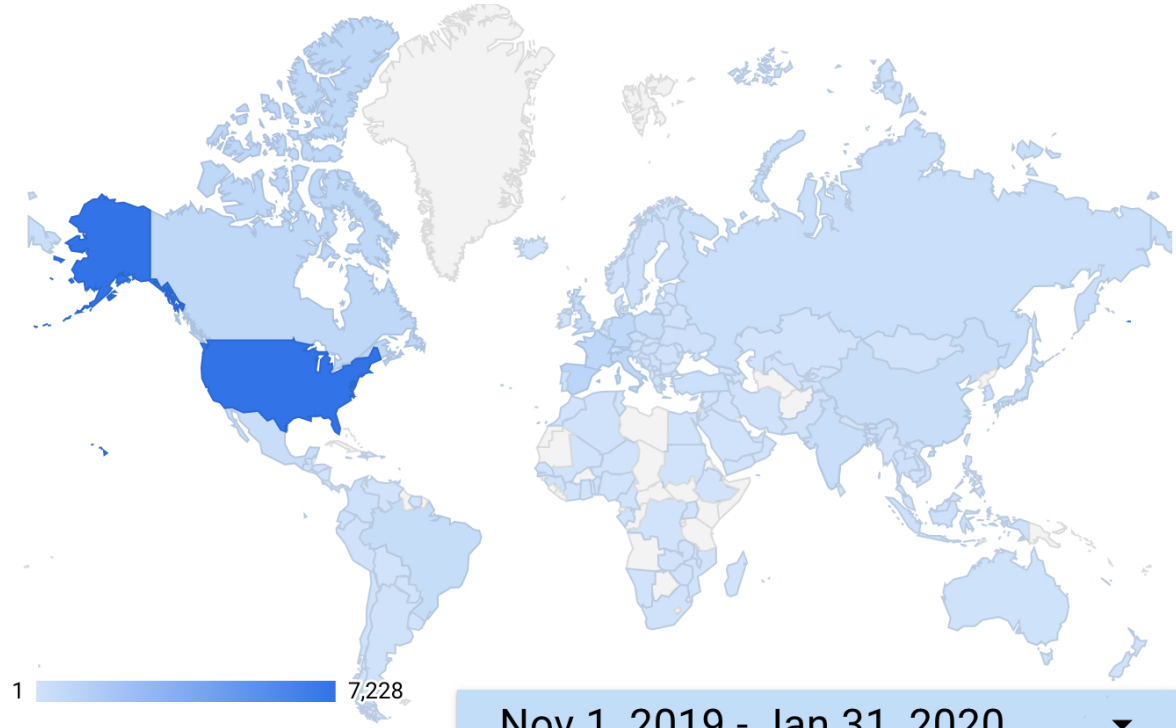
MYCOLOGY COLLECTIONS PORTAL

Users	New Users	Sessions	Number of Sessions per User	Pageviews	Pages / Session	Avg. Session Duration	Bounce Rate
11,004	9,957	18,065	1.64	46,955	2.6	00:02:53	60.18%









Nov 1, 2019 - Jan 31, 2020

Published on *iDigBio* (https://www.idigbio.org)

Submission #1601

Submission information

Form: TCN Quarterly Progress Report to iDigBio
Submitted by djbarroso
Tuesday, February 4, 2020 - 21:55
12.44.89.34

TCN Name:
TORCH: American Crossroads: Digitizing the Vascular Flora of the South-Central United States

Person completing the report:
diego.barroso@yahoo.com

Progress in Digitization Efforts:
• Number of skeletal records created:
BRIT = 0
BAYLU = 0
HUH = 29
KANU = 0
MO = 0
NOSU = 0
NY = 22,201
OKL = 0
OKLA = 0
SHST = 0
TAMUCC = 0
TAMU = 0
TEX/LL (Plant Resources Center) = 0
TEX/LL (Fort Worth Nature Center) = 0
TTC = 0
UTEP = 0

Total skeletal records created this quarter: 22,230

• Number of fully-transcribed records created:
BRIT = 436
BAYLU = 0
HUH = 3,556
KANU = 828

MO = 0
NOSU = 0
NY = 26,276
OKL = 0
OKLA = 100
SHST = 25,000
TAMUCC = 0
TAMU = 0
TEX/LL (Plant Resources Center) = 1,565
TEX/LL (Fort Worth Nature Center) = 1,824
TTC = 0
UTEP = 0

Total fully-transcribed records created this quarter: 59,585

• Number of specimens imaged:
BRIT = 5,181
BAYLU = 0
HUH = 3,528
KANU = 0
MO = 0
NOSU = 0
NY = Prior to the start of the project, we had imaged 51,995 specimens from the TORCH area. Since the start of the project, we have not imaged any additional specimens. The intern we are hiring next month will focus on imaging for this project.
OKL = 9,341
OKLA = 0
SHST = 0
TAMUCC = 0
TAMU = 0
TEX/LL (Plant Resources Center) = 0
TEX/LL (Fort Worth Nature Center) = 0
TTC = 0
UTEP = 0

Total number of specimens imaged this quarter: 18,050

• Number of specimens georeferenced:
BRIT= 0
BAYLU = 0
HUH = 1
KANU = 437
MO = 0
NOSU = 0
NY = 9,416
OKL = 0
OKLA = 0
SHST = minimal; will begin this year.
TAMUCC = 0
TAMU = 0
TEX/LL (Plant Resources Center) = 0

TEX/LL (Fort Worth Nature Center) = 0
TTC = 477
UTEP = 0

Total number of specimens georeferenced this quarter: 10,331

• Other digitization or pre-digitization efforts:
BRIT: Staging: Identification of TORCH specimens within folders of BRIT North America is being carried out by BRIT staff and volunteers by opening and evaluating cabinets and tagging folders containing TORCH specimens. After this tagging process, the herbarium folders containing project specimens can be pulled for imaging.
BAYLU: With the help of volunteers, primarily, we have added barcode labels to approximately 28,000 specimens for this reporting period. We have received and set-up Ortery Photobox with Nikon D810 camera, and Nikkon 50mm macro lens. Now awaiting camera mount from BRIT. In addition, have set up the HerbASAP python routine that we are now experimenting with for efficient post-processing of images presumably to work with Adobe Lightbox or Nikon Pro camera control. Have hired a student worker as of 01/27/20, who is currently barcoding specimens. The Central Texas Master Naturalist have been volunteering time to assist in barcoding. In addition, Baylor Biology students (primarily from the Plant Physiology course) have also contributed to barcoding. From this, 129 volunteer-hours have been recorded during the reporting period with average rate of 216 barcode labels affixed per hour by these groups.
HUH: Many TX and OK specimens have already been digitized through other digitization projects. Current total is 38,628 TX and OK records.
KANU: All OK and TX specimens in Asteraceae and Fabaceae and 60% of Poaceae specimens identified with drop tags in the cases in preparation for imaging. Many other genera and families also have been flagged. We now have: 19,697 fully-transcribed records from OK and TX; an estimated 3,000 total imaged specimens; and a total of 19,277 georeferenced records from OK and TX.
MO: We are in the process of developing a combined Texas/Oklahoma Checklist as a guide for searching for project material in the collection, which should be completed soon.
NY: When she started the job, McKenna Coyle made a sweep of the herbarium to look for species that had not been digitized for North America, and flagged these to search for specimens from the TORCH area. We estimate that there are only about 20,000 of our specimens from the U.S. & Canada that have not at least been skeletally digitized.
OKLA: 25% of Texas specimens have been located and segregated from other non-Oklahoma collections.
SHST: We have not yet begun digitization efforts. Most of our time was spent preparing for formal opening of this new facility which was on November 16, 2019. We need to complete check of accession numbers and eliminate approximately 700 duplicate numbers that were stamped on specimens decades ago. We have approximately 25,000 sheets that have had all label data entered. (“fully transcribed records”). We have about 20,000 more specimens in the “to-be-mounted” area. This was a large legacy inherited from the Department of Biological Sciences. We mount 200 per week and have about 2000 mounted. We have begun capturing their label data too.
We expect to finish correcting duplicate accession numbers by the end of February. We will require one year to mount the backlog and capture their data.
We have put a lot of effort into obtaining other legacy collections and datasets. These of Geyata Ajilsvji, Joe Liggio, Eric Keith.
If we do not use the inertia of this project to 1). push for processing of backlog and 2). capture images/specimens/data of local botanical legacies, then these important sources

of data are in danger of drifting away.
TAMU: Photography equipment, lightbox, barcodes, and computer have all been purchased or otherwise acquired. We have begun interviewing and hiring undergraduate student workers, and have also established a volunteer workforce for databasing and georeferencing blitzes. We are now waiting on the camera stand before we can commence digitization.
TEX/LL (Plant Resources Center): Acquired mobile imaging station to be assembled.
TTC: Mounted 45 specimens from 1973 collection from Guadalupe Mountains N.P.
UTEP: Ordered, received, and set up the lightbox for imaging.
• Comments about digitization progress:
BAYLU: Digitization station construction is still in progress.
MO: Progress during this quarter has been minimal. We are still waiting to obtain our lightbox and associated equipment. We are still in the process of recruiting and hiring staff for the project.
NOSU: We plans to do it all this summer with Abby Moore of OU.
NY: We are using our standard digitization practices, developed over the past 10 years.
OKL: Once the TORCH project hires its data manager, we will be able to upload all of our existing databased records into the TORCH Symbiota portal and will have more progress in the other areas. Currently most of the specimens that we are imaging have already been databased.
OKLA: Our imaging equipment has not yet been obtained and installed.
SHST: see above.
TAMUCC: We have not yet done anything regarding the project, as we only just received the last sets of equipment that we ordered for the project. The actual digitization will only start this month (February 2020).
TEX/LL: The Plant Resources Center assigned eight students to pull out Oklahoma specimens from the main collection. 60% of the herbarium has been reviewed, and Oklahoma plants have been aggregated for digitization. Our Lady of the Lake University is currently waiting on a mobile imaging station to begin processing specimens.
TTC: Main focus is georeferencing, to fix approximately 10,000 erroneous points.

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

BRIT = 69,638 (the discrepancy between this and the number of records in the Symbiota portal below is likely because the iDigBio total may include non-vasculars. This will be verified in case the discrepancy grows.)
BAYLU = 0
HUH = 1,285,191 total HUH records in iDigBio; 28,381 TORCH records. Note, iDigBio is months behind on ingesting HUH data.
KANU = we upload a new instance of our database to GBIF and iDigBio at the beginning of each month. All fully-transcribed records (19,697) from OK and TX will be available the first week of February.
MO = 0
NOSU = 0
NY = see answer below.
OKL = 0
OKLA = 0
SHST = 0
TAMUCC = 0
TAMU = 0
TEX/LL = 0
TTC = 0 (waiting to fix georeferencing)

UTEP = 0

• Number of records available in TORCH Symbiota portal (cumulative):
BRIT = 66,687
BAYLU = 0
HUH = 34,036 TORCH records
KANU = 0
MO = 0
NOSU = 0
NY = Answer to both previous questions: We have 51,995 (50,343 with images) TORCH vascular records available online in our Virtual Herbarium right now now. The TORCH portal is using data last updated in June 2018, but iDigBio is more recent - last updated Sept 2019. Our IPT is available for both portals to make updates anytime. Please contact us if there is any question about how to harvest data using our IPT for the TORCH Symbiota project.
OKL = 0
OKLA = 0
SHST = 0
TAMUCC = 0
TAMU = 0
TEX/LL = 0
TTC = 21,452
UTEP = 0

Share and Identify Best Practices and Standards (including Lessons Learned):
BRIT: Records of effort kept by staff are very important, and the importance of keeping these up to date must be impressed upon everyone. More efficient methods of tracking time could be evaluated.

NY: We have been experimenting with mass georeferencing using Geolocate for specimens from the TORCH region that were previously digitized through other projects. We are analyzing the results to see if word order in the locality string can improve georeferencing accuracy.

TTC: Created a guide for students who are georeferencing using Symbiota; it can be found here: http://mossmatters.com/protocols/protocol/Georeferencing_Guide.html

Identify Gaps in Digitization Areas and Technology:
BRIT: For the time being, it appears to be more efficient/cost-effective to image entire folders rather than to stage/pre-segregate only TX and OK specimens.
There are three points at which we are experiencing bottlenecks, (1) locating TORCH specimens in our cabinets (finding Oklahoma sheets in the BRIT-North America folders, where they are filed with all other states), (2) identifying Oklahoma and Texas specimens in folders of already imaged specimens in order to prioritize their transcription (in the VDB North America folders, where these two states are filed), and (3) in the image processing on-site, the upload to TACC servers, and the access to images of specimens through the Symbiota portal.
With the workflow changing for BRIT (particularly the image upload and ingest stages), we lack complete protocols and documentation. These will be established in Q1 2020 and shared with participants.
BAYLU: Big issue is finalization of the photobox and the file post-processing. Have been road-testing the camera and file transfer as well as assessing the HerbASAP routine. Will

need further guidance on image off-site storage (TACC) late Feb.
NY: We have greatly improved our ability to use OCR to group records in order to expedite data transcription. However, we could still use more advanced technology to improve OCR results.
OKLA: We have had difficulties in obtaining the imaging equipment. A solution appears to be almost in place.
SHST: No gaps; we are planning this first phase to get the backlog of legacy collections processed. These come to us from botanists like Guy Nesom, Michael Warnock, Geyata Ajilsvji.
TAMUCC: The main issue was that we did not get the recommended camera on time; i.e., the Nikon D81, along with the lightbox and camera stand, came very late.
TAMU: We are still procuring equipment. Some was previously back-ordered, which delayed our start.
TEX/LL: U. TX-Rio Grande Valley-Brownsville is currently on hold due to Brownsville planning on moving its collection to U. TX-Austin.
TTC: We are awaiting the newly-designed lightbox for imaging. ETA is Mid-February.
UTEP: Awaited equipment delivery.

Share and Identify Opportunities to Enhance Training Efforts:
Thanks to iDigBio funding, a TORCH Digitization Workshop will be held at BRIT on March 28th and 29th; it is estimated that the workshop will have approximately 40 attendees.
BRIT: Trained new digitizer, hired October 2019 (Rachel Carmickle), trained 6 new volunteers in digitization tasks. Lauren Snyder, a new student digitizer from TAMUCC, visited BRIT and was given a demonstration of the digitization workflow.
BAYLU: Led two training events early Nov 2019 for barcoding, one for the Texas Master Naturalist, and the other for students. Each approximately one hour long, with a social and bar-coding “blitz.” Now provide t-shirts for volunteers (iDig Texas Plants) and have service award pins for Texas Master Naturalist based on hours contributed (50 hour=bronze, 150=silver, 250=gold).
NY: We have hired a former intern, McKenna Coyle, to manage our TORCH project. McKenna is a recent college graduate who formerly worked on our SORO TCN project. She is receiving professional development in project management, supervision, and public outreach. Soon (by mid February) we will be adding an intern to the project. Interviews for this position are going on now. Our interns are recent college graduates who are paid for 1000 hours to work on the project.
OKL: We are training a new student digitizer.
SHST: Not begun yet.
TAMUCC: None yet, but there are plans to have undergraduate students participation and training.
TEX/LL (Plant Resources Center): Two student digitizers were trained in transcription.
TTC: Trained 6 undergraduates and 3 graduate students in georeferencing and mounting specimens. Students are a combination of volunteers, research credit hours, interns, and graduate assistants.

Share and Identify Collaborations with other TCNs, Institutions, and Organizations:
Jason Best, at BRIT, and George Yatskievych, at UT-Austin, were able to secure storage space with the Texas Advanced Computing Center (TACC), in order to house TORCH images. This repository will contain both archive images and derivatives for display on the web.
NY: We find a lot of synergy between our work on this project and our work on the SORO (Southern Rockies) TCN – the two projects are collaborating in digitization of groups that cross both areas.

OKL: We continue to collaborate with Bruce Hoagland, Amy Buthod, and Sam Basave.
SHST: We initiated contact with Helen Holdsworth at Witte Museum San Antonio. They have a small herbarium dating back to Ellen Schultz Quillen activity in 1920's. We have initiated discussions about cooperating to capture and share their digital data as part of our project.
TAMUCC: Student digitizer Lauren Snyder visited BRIT and learned about the digitization workflow being used there.
TEX/LL: We continue to collaborate with the Fort Worth Nature Center (a 2,000-specimen collection), and now have a new collaboration with St. Edwards University of Austin, TX (a 3,000-specimen collection).

Share and Identify Opportunities and Strategies for Sustainability:
OKLA: Permanent space for digitization studio allocated to OSU Herbarium.
TAMUCC: Faculty and students to grow the collection through field trips or class assignments in Plant Taxonomy course by collecting specimens for the herbarium and digitizing the collections.
Other Progress not listed above (anything else to share):
BRIT: A Google Group was established for all participants, which is actively being used to share progress, tips, and other information.
BAYLU: Currently still in phase one of the project including bar-coding, equipment preparation, and personnel hiring.
KANU: We have advertised for 4 additional hourly positions (the job advertisement closed on January 31st) and expect to hire those students the first week in February. Students will assist with all aspects of finding, databasing, and imaging specimens through the semester.
OKLA: Applications have been received for the Data Manager position, with interviews forthcoming.
UTEP: Given our budget, the collection manager here will not officially start until June, which will be when we will be most efficient.

Share and Identify Education and Outreach (E&O) Activities:
MO: 6 group tours of herbarium totaling 66 people. All tours include demonstration of imaging technology and explanation of TCN support. In addition, a special display on gathering electronic data on Orchidaceae in the Garden's herbarium, particularly focusing on the imaging process and TCN support, as well as other aspects of the orchid collection was presented at two opening events for the Garden's annual Orchid Show (30 and 31 January). 820 invited guests attended the two evening events. 886 total people involved in outreach activities.
NY: The TORCH project will be featured in a Valentine's Day Open House
OKL: Digitization was discussed in a class tour of our herbarium.

Methods of disseminating results to communities of interest (presentations, lectures, etc.):
Thanks to iDigBio funding, a TORCH Digitization Workshop will be held at BRIT on March 28th and 29th; it is estimated that the workshop will have approximately 40 attendees.
BAYLU: Provided an information seminar (30 minutes) to the Texas Master Naturalist on the importance of digital specimen collections and its relationship to studying biodiversity.
NY: All of our data are immediately available through our Virtual Herbarium website, and are made available for harvest by the TORCH Symbiota portal, iDigBio and GBIF through our IPT installation.
OKL: A slide on herbarium digitization was prepared for the University of Oklahoma library's digital display of campus research.
OKLA: Presentation of TORCH TCN activities to the Oklahoma Academy of Science in

November, 2019.
SHST: Native Plant Society local chapter is meeting in the museum now, and will be a vehicle for publicity about the herbarium.
TAMUCC: Presentations, news release, conference presentation, etc.

Google Analytics

Other Progress (that doesn't fit into the above categories):
PLEASE SEE ATTACHED PDF, WHICH CONTAINS THE TORCH TCN REPORT.

Products generated (publications, conference presentations, technologies/techniques, websites, etc.):
BRIT: Due to the lack of affordable light box equipment, BRIT developed a prototype light box fabricated from commonly-available parts which can be easily assembled by participants. The prototype was completed and tested and four units are expected to be delivered to participants in February 2020 with more units being delivered soon after that. Additionally, BRIT designed and built a low cost custom camera stand that mounts to the Ortery Photosimile 50. Two of these stands will be used by BRIT and an additional five units will be delivered to participants by February 2020 and more will be fabricated and delivered as needed.
NY: Nothing to report yet.
OKLA: Presentation of TORCH TCN activities to the Oklahoma Academy of Science in November, 2019.
TAMUCC: We have two conference presentations that are linked directly or indirectly to the digitization project as follows:
Rodriguez, J. & Daru, B.H. (2020) Mismatches and congruencies in plant sampling biases between observations and vouchered specimens. 46th conference of the South African Association of Botanists, Qwaqwa, Free State, South Africa (January 2020).
Snyder, L. & Daru, B.H. (2019) How digitization efforts at small local herbaria add to the scientific community and predicting the impact of the Ruth O'Brien herbarium. Global Change Symposium, Texas A&M University-Corpus Christi, December 2019.
TTC: Created a guide for students who are georeferencing using Symbiota; it can be found here: http://mossmatters.com/protocols/protocol/Georeferencing_Guide.html

Participants (especially those who have newly joined the project):
BRIT: Diego Barroso has been hired at BRIT as the new Project Manager for the TORCH TCN. Rachel Carmickle was hired as part-time digitizer, transcribing and imaging TORCH specimens.
BAYLU: Andrew Lewis, Baylor Biology student, student worker; Miriam (Morenike) Tajudeen, previous Baylor Biology student (graduate Dec 2019), now hired as part-time digitization technician; Melissa Mullins, education coordinator for the Center for Reservoir and Aquatic Systems Research at Baylor as the liaison with the Central Texas Master Naturalist
KANU: Maeve Hilgers – hourly student; assisting with finding specimens in cases, databasing, and, eventually, imaging.
NOSU: Potential students who will be helping this summer: Katie Worden, Emma Mills.
NY: McKenna Coyle – Project Coordinator.
OKL: Leann Monaghan will start as a Master's student in the Fall, but is now working 10 hours a week to learn all of the digitization protocols this Spring and Summer.
OKLA: Nimani Rathnasooriya Dummala Deniyalage, Victor Andreev, John Hodge (graduate assistants, supported by PBEE Dept.). Joseph Hogan, Alyssa Regier, Zachary

Shiever (undergraduate assistants, supported by TORCH TCN)
SHST: We currently have two student assistants working in the herbarium, and two volunteers.
TAMUCC: Lauren Snyder is joining the project as student digitizer.
TTC: Director: Matt Johnson; Lab Manager: Haley Hale; Graduate Assistant for Fall 2019: Yanni Chen; Graduate Assistants for Spring 2020: Zhiyuan Liu and Haoran Li; Undergraduate Assistants (paid on TCN grant): Chase Bergeron, Jennifer Mendez, Hayden Mathews; Undergraduate Assistants (not paid on TCN grant): Madeline Slimp, Cassidy Coker, Jared Salzman, Justin Dawsey, Howard Park.
UTEP: Two new student volunteers have been identified.

Attachment 1
TORCH-compiled-report-2020-02-04-Q1.pdf

Attachment 2

Source URL: https://www.idigbio.org/node/564/submission/1601

TORCH TCN — Quarterly Report

November 1st 2019 - January 31st, 2020

Assembled by BRIT on February 4th, 2020, for Feb. 5th IAC meeting

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

Institutions reporting:

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

TAMUCC – Texas A&M University-Corpus Christi

TAMU – Texas A&M University-College Station

TEX/LL – University of Texas at Austin

TTC – Texas Tech University

UTEP – University of Texas at El Paso

Progress in Digitization Efforts:

- Number of skeletal records created:
  BRIT = 0
  BAYLU = 0
  HUH = 29
  KANU = 0
  MO = 0
  NOSU = 0
  NY = 22,201
  OKL = 0
  OKLA = 0

SHST = 0
TAMUCC = 0
TAMU = 0
TEX/LL (Plant Resources Center) = 0
TEX/LL (Fort Worth Nature Center) = 0
TTC = 0
UTEP = 0

Total skeletal records created this quarter: 22,230

- Number of fully-transcribed records created:
  BRIT = 436
  BAYLU = 0
  HUH = 3,556
  KANU = 828
  MO = 0
  NOSU = 0
  NY = 26,276
  OKL = 0
  OKLA = 100
  SHST = 25,000
  TAMUCC = 0
  TAMU = 0
  TEX/LL (Plant Resources Center) = 1,565
  TEX/LL (Fort Worth Nature Center) = 1,824
  TTC = 0
  UTEP = 0

  Total fully-transcribed records created this quarter: 59,585

- Number of specimens imaged:
  BRIT = 5,181
  BAYLU = 0
  HUH = 3,528
  KANU = 0
  MO = 0
  NOSU = 0
  NY = Prior to the start of the project, we had imaged 51,995 specimens from the TORCH area. Since the start of the project, we have not imaged any additional specimens. The intern we are hiring next month will focus on imaging for this project.
  OKL = 9,341
  OKLA = 0
  SHST = 0
  TAMUCC = 0
  TAMU = 0
  TEX/LL (Plant Resources Center) = 0
  TEX/LL (Fort Worth Nature Center) = 0

TTC = 0
UTEP = 0

Total number of specimens imaged this quarter: 18,050

- Number of specimens georeferenced:
  BRIT= 0
  BAYLU = 0
  HUH = 1
  KANU = 437
  MO = 0
  NOSU = 0
  NY = 9,416
  OKL = 0
  OKLA = 0
  SHST = minimal; will begin this year.
  TAMUCC = 0
  TAMU = 0
  TEX/LL (Plant Resources Center) = 0
  TEX/LL (Fort Worth Nature Center) = 0
  TTC = 477
  UTEP = 0

  Total number of specimens georeferenced this quarter: 10,331

- Other digitization or pre-digitization efforts:

BRIT: Staging: Identification of TORCH specimens within folders of BRIT North America is being carried out by BRIT staff and volunteers by opening and evaluating cabinets and tagging folders containing TORCH specimens. After this tagging process, the herbarium folders containing project specimens can be pulled for imaging.

BAYLU: With the help of volunteers, primarily, we have added barcode labels to approximately 28,000 specimens for this reporting period. We have received and set-up Ortery Photobox with Nikon D810 camera, and Nikkon 50mm macro lens. Now awaiting camera mount from BRIT. In addition, have set up the HerbASAP python routine that we are now experimenting with for efficient post-processing of images presumably to work with Adobe Lightbox or Nikon Pro camera control. Have hired a student worker as of 01/27/20, who is currently barcoding specimens. The Central Texas Master Naturalist have been volunteering time to assist in barcoding. In addition, Baylor Biology students (primarily from the Plant Physiology course) have also contributed to barcoding. From this, 129 volunteer-hours have been recorded during the reporting period with average rate of 216 barcode labels affixed per hour by these groups.

HUH: Many TX and OK specimens have already been digitized through other digitization projects. Current total is 38,628 TX and OK records.

KANU: All OK and TX specimens in Asteraceae and Fabaceae and 60% of Poaceae specimens identified with drop tags in the cases in preparation for imaging. Many other genera and families also have been flagged. We now have: 19,697 fully-transcribed records from OK and TX; an estimated 3,000 total imaged specimens; and a total of 19,277 georeferenced records from OK and TX.

MO: We are in the process of developing a combined Texas/Oklahoma Checklist as a guide for searching for project material in the collection, which should be completed soon.

NY: When she started the job, McKenna Coyle made a sweep of the herbarium to look for species that had not been digitized for North America, and flagged these to search for specimens from the TORCH area. We estimate that there are only about 20,000 of our specimens from the U.S. & Canada that have not at least been skeletally digitized.

OKLA: 25% of Texas specimens have been located and segregated from other non-Oklahoma collections.

SHST: We have not yet begun digitization efforts. Most of our time was spent preparing for formal opening of this new facility which was on November 16, 2019. We need to complete check of accession numbers and eliminate approximately 700 duplicate numbers that were stamped on specimens decades ago. We have approximately 25,000 sheets that have had all label data entered. ("fully transcribed records"). We have about 20,000 more specimens in the "to-be-mounted" area. This was a large legacy inherited from the Department of Biological Sciences. We mount 200 per week and have about 2000 mounted. We have begun capturing their label data too.

We expect to finish correcting duplicate accession numbers by the end of February. We will require one year to mount the backlog and capture their data.

We have put a lot of effort into obtaining other legacy collections and datasets. These of Geyata Ajilsvji, Joe Liggio, Eric Keith.

If we do not use the inertia of this project to 1). push for processing of backlog and 2). capture images/specimens/data of local botanical legacies, then these important sources of data are in danger of drifting away.

TAMU: Photography equipment, lightbox, barcodes, and computer have all been purchased or otherwise acquired. We have begun interviewing and hiring undergraduate student workers, and have also established a volunteer workforce for databasing and georeferencing blitzes. We are now waiting on the camera stand before we can commence digitization.

TEX/LL (Plant Resources Center)**:** Acquired mobile imaging station to be assembled.

TTC: Mounted 45 specimens from 1973 collection from Guadalupe Mountains N.P.

UTEP: Ordered, received, and set up the lightbox for imaging.

- Comments about digitization progress:

BAYLU: Digitization station construction is still in progress.

MO: Progress during this quarter has been minimal. We are still waiting to obtain our lightbox and associated equipment. We are still in the process of recruiting and hiring staff for the project.

NOSU: We plans to do it all this summer with Abby Moore of OU.

NY: We are using our standard digitization practices, developed over the past 10 years.

OKL: Once the TORCH project hires its data manager, we will be able to upload all of our existing databased records into the TORCH Symbiota portal and will have more progress in the other areas. Currently most of the specimens that we are imaging have already been databased.

OKLA: Our imaging equipment has not yet been obtained and installed.

SHST: see above.

TAMUCC: We have not yet done anything regarding the project, as we only just received the last sets of equipment that we ordered for the project. The actual digitization will only start this month (February 2020).

TEX/LL: The Plant Resources Center assigned eight students to pull out Oklahoma specimens from the main collection. 60% of the herbarium has been reviewed, and Oklahoma plants have been aggregated for digitization. Our Lady of the Lake University is currently waiting on a mobile imaging station to begin processing specimens.

TTC: Main focus is georeferencing, to fix approximately 10,000 erroneous points.

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

BRIT = 69,638 (the discrepancy between this and the number of records in the Symbiota portal below is likely because the iDigBio total may include non-vasculars. This will be verified in case the discrepancy grows.)

BAYLU = 0

HUH = 1,285,191 total HUH records in iDigBio; 28,381 TORCH records. Note, iDigBio is months behind on ingesting HUH data.

KANU = we upload a new instance of our database to GBIF and iDigBio at the beginning of each month. All fully-transcribed records (19,697) from OK and TX will be available the first week of February.

MO = 0

NOSU = 0

NY = see answer below.

OKL = 0

OKLA = 0

SHST = 0

TAMUCC = 0

TAMU = 0

TEX/LL = 0

TTC = 0 (waiting to fix georeferencing)

UTEP = 0

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

  BRIT = 66,687

  BAYLU = 0

  HUH = 34,036 TORCH records

  KANU = 0

  MO = 0

  NOSU = 0

  NY = Answer to both previous questions: We have 51,995 (50,343 with images) TORCH vascular records available online in our Virtual Herbarium right now now. The TORCH portal is using data last updated in June 2018, but iDigBio is more recent - last updated Sept 2019. Our IPT is available for both portals to make updates anytime. Please contact us if there is any question about how to harvest data using our IPT for the TORCH Symbiota project.

  OKL = 0

  OKLA = 0

  SHST = 0

  TAMUCC = 0

  TAMU = 0

  TEX/LL = 0

  TTC = 21,452

  UTEP = 0

Best Practices and Standards (Lessons Learned):

BRIT: Records of effort kept by staff are very important, and the importance of keeping these up to date must be impressed upon everyone. More efficient methods of tracking time could be evaluated.

NY: We have been experimenting with mass georeferencing using Geolocate for specimens from the TORCH region that were previously digitized through other projects. We are analyzing the results to see if word order in the locality string can improve georeferencing accuracy.

TTC: Created a guide for students who are georeferencing using Symbiota; it can be found here: http://mossmatters.com/protocols/protocol/Georeferencing_Guide.html

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

BRIT: For the time being, it appears to be more efficient/cost-effective to image entire folders rather than to stage/pre-segregate only TX and OK specimens.

There are three points at which we are experiencing bottlenecks, (1) locating TORCH specimens in our cabinets (finding Oklahoma sheets in the BRIT-North America folders, where they are filed with all other states), (2) identifying Oklahoma and Texas specimens in folders of already imaged specimens in order to prioritize their transcription (in the VDB North America folders, where these two states are filed), and (3) in the image processing on-site, the upload to TACC servers, and the access to images of specimens through the Symbiota portal.

With the workflow changing for BRIT (particularly the image upload and ingest stages), we lack complete protocols and documentation. These will be established in Q1 2020 and shared with participants.

BAYLU: Big issue is finalization of the photobox and the file post-processing. Have been road-testing the camera and file transfer as well as assessing the HerbASAP routine. Will need further guidance on image off-site storage (TACC) late Feb.

NY: We have greatly improved our ability to use OCR to group records in order to expedite data transcription. However, we could still use more advanced technology to improve OCR results.

OKLA: We have had difficulties in obtaining the imaging equipment. A solution appears to be almost in place.

SHST: No gaps; we are planning this first phase to get the backlog of legacy collections processed. These come to us from botanists like Guy Nesom, Michael Warnock, Geyata Ajilsvji.

TAMUCC: The main issue was that we did not get the recommended camera on time; i.e., the Nikon D81, along with the lightbox and camera stand, came very late.

TAMU: We are still procuring equipment. Some was previously back-ordered, which delayed our start.

TEX/LL: U. TX-Rio Grande Valley-Brownsville is currently on hold due to Brownsville planning on moving its collection to U. TX-Austin.

TTC: We are awaiting the newly-designed lightbox for imaging. ETA is Mid-February.

UTEP: Awaited equipment delivery.

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

Thanks to iDigBio funding, a TORCH Digitization Workshop will be held at BRIT on March 28th and 29th; it is estimated that the workshop will have approximately 40 attendees.

BRIT: Trained new digitizer, hired October 2019 (Rachel Carmickle), trained 6 new volunteers in digitization tasks. Lauren Snyder, a new student digitizer from TAMUCC, visited BRIT and was given a demonstration of the digitization workflow.

BAYLU: Led two training events early Nov 2019 for barcoding, one for the Texas Master Naturalist, and the other for students. Each approximately one hour long, with a social and bar-coding "blitz." Now provide t-shirts for volunteers (iDig Texas Plants) and have service award pins for Texas Master Naturalist based on hours contributed (50 hour=bronze, 150=silver, 250=gold).

NY: We have hired a former intern, McKenna Coyle, to manage our TORCH project. McKenna is a recent college graduate who formerly worked on our SORO TCN project. She is receiving professional development in project management, supervision, and public outreach. Soon (by mid February) we will be adding an intern to the project. Interviews for this position are going on now. Our interns are recent college graduates who are paid for 1000 hours to work on the project.

OKL: We are training a new student digitizer.

SHST: Not begun yet.

TAMUCC: None yet, but there are plans to have undergraduate students participation and training.

TEX/LL (Plant Resources Center)**:** Two student digitizers were trained in transcription.

TTC: Trained 6 undergraduates and 3 graduate students in georeferencing and mounting specimens. Students are a combination of volunteers, research credit hours, interns, and graduate assistants.

Collaboration with other TCNs, Institutions, and Organizations:

Jason Best, at BRIT, and George Yatskievych, at UT-Austin, were able to secure storage space with the Texas Advanced Computing Center (TACC), in order to house TORCH images. This repository will contain both archive images and derivatives for display on the web.

NY: We find a lot of synergy between our work on this project and our work on the SORO (Southern Rockies) TCN – the two projects are collaborating in digitization of groups that cross both areas.

OKL: We continue to collaborate with Bruce Hoagland, Amy Buthod, and Sam Basave.

SHST: We initiated contact with Helen Holdsworth at Witte Museum San Antonio. They have a small herbarium dating back to Ellen Schultz Quillen activity in 1920's. We have initiated discussions about cooperating to capture and share their digital data as part of our project.

TAMUCC: Student digitizer Lauren Snyder visited BRIT and learned about the digitization workflow being used there.

TEX/LL: We continue to collaborate with the Fort Worth Nature Center (a 2,000-specimen collection), and now have a new collaboration with St. Edwards University of Austin, TX (a 3,000-specimen collection).

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

Thanks to iDigBio funding, a TORCH Digitization Workshop will be held at BRIT on March 28th and 29th; it is estimated that the workshop will have approximately 40 attendees.

BAYLU: Provided an information seminar (30 minutes) to the Texas Master Naturalist on the importance of digital specimen collections and its relationship to studying biodiversity.

NY: All of our data are immediately available through our Virtual Herbarium website, and are made available for harvest by the TORCH Symbiota portal, iDigBio and GBIF through our IPT installation.

OKL: A slide on herbarium digitization was prepared for the University of Oklahoma library's digital display of campus research.

OKLA: Presentation of TORCH TCN activities to the Oklahoma Academy of Science in November, 2019.

SHST: Native Plant Society local chapter is meeting in the museum now, and will be a vehicle for publicity about the herbarium.

TAMUCC: Presentations, news release, conference presentation, etc.

Other Education and Outreach activities:

MO: 6 group tours of herbarium totaling 66 people. All tours include demonstration of imaging technology and explanation of TCN support. In addition, a special display on gathering electronic data on Orchidaceae in the Garden's herbarium, particularly focusing on the imaging process and TCN support, as well as other aspects of the orchid collection was presented at two opening events for the Garden's annual Orchid Show (30 and 31 January). 820 invited guests attended the two evening events. 886 total people involved in outreach activities.

NY: The TORCH project will be featured in a Valentine's Day Open House

OKL: Digitization was discussed in a class tour of our herbarium.

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

BRIT: Due to the lack of affordable light box equipment, BRIT developed a prototype light box fabricated from commonly-available parts which can be easily assembled by participants. The prototype was completed and tested and four units are expected to be delivered to participants in February 2020 with more units being delivered soon after that. Additionally, BRIT designed and built a low cost custom camera stand that mounts to the Ortery Photosimile 50. Two of these stands will be used by BRIT and an additional five units will be delivered to participants by February 2020 and more will be fabricated and delivered as needed.

NY: Nothing to report yet.

OKLA: Presentation of TORCH TCN activities to the Oklahoma Academy of Science in November, 2019.

TAMUCC: We have two conference presentations that are linked directly or indirectly to the digitization project as follows:

Rodriguez, J. & Daru, B.H. (2020) Mismatches and congruencies in plant sampling biases between observations and vouchered specimens. 46th conference of the South African Association of Botanists, Qwaqwa, Free State, South Africa (January 2020).

Snyder, L. & Daru, B.H. (2019) How digitization efforts at small local herbaria add to the scientific community and predicting the impact of the Ruth O'Brien herbarium. Global Change Symposium, Texas A&M University-Corpus Christi, December 2019.

TTC: Created a guide for students who are georeferencing using Symbiota; it can be found here: http://mossmatters.com/protocols/protocol/Georeferencing_Guide.html

<u>Participants (especially those who have newly joined the project)</u>:

BRIT: Diego Barroso has been hired at BRIT as the new Project Manager for the TORCH TCN. Rachel Carmickle was hired as part-time digitizer, transcribing and imaging TORCH specimens.

BAYLU: Andrew Lewis, Baylor Biology student, student worker; Miriam (Morenike) Tajudeen, previous Baylor Biology student (graduate Dec 2019), now hired as part-time digitization technician; Melissa Mullins, education coordinator for the Center for Reservoir and Aquatic Systems Research at Baylor as the liaison with the Central Texas Master Naturalist

KANU: Maeve Hilgers – hourly student; assisting with finding specimens in cases, databasing, and, eventually, imaging.

NOSU: Potential students who will be helping this summer: Katie Worden, Emma Mills.

NY: McKenna Coyle – Project Coordinator.

OKL: Leann Monaghan will start as a Master's student in the Fall, but is now working 10 hours a week to learn all of the digitization protocols this Spring and Summer.

OKLA: Nimani Rathnasooriya Dummala Deniyalage, Victor Andreev, John Hodge (graduate assistants, supported by PBEE Dept.). Joseph Hogan, Alyssa Regier, Zachary Shiever (undergraduate assistants, supported by TORCH TCN)

SHST: We currently have two student assistants working in the herbarium, and two volunteers.

TAMUCC: Lauren Snyder is joining the project as student digitizer.

TTC: Director: Matt Johnson; Lab Manager: Haley Hale; Graduate Assistant for Fall 2019: Yanni Chen; Graduate Assistants for Spring 2020: Zhiyuan Liu and Haoran Li; Undergraduate Assistants (paid on TCN grant): Chase Bergeron, Jennifer Mendez, Hayden Mathews; Undergraduate Assistants (not paid on TCN grant): Madeline Slimp, Cassidy Coker, Jared Salzman, Justin Dawsey, Howard Park.

UTEP: Two new student volunteers have been identified.

Opportunities and Strategies for Sustainability:

OKLA: Permanent space for digitization studio allocated to OSU Herbarium.

TAMUCC: Faculty and students to grow the collection through field trips or class assignments in Plant Taxonomy course by collecting specimens for the herbarium and digitizing the collections.

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

BRIT: A Google Group was established for all participants, which is actively being used to share progress, tips, and other information.

BAYLU: Currently still in phase one of the project including bar-coding, equipment preparation, and personnel hiring.

KANU: We have advertised for 4 additional hourly positions (the job advertisement closed on January 31st) and expect to hire those students the first week in February. Students will assist with all aspects of finding, databasing, and imaging specimens through the semester.

OKLA: Applications have been received for the Data Manager position, with interviews forthcoming.

UTEP: Given our budget, the collection manager here will not officially start until June, which will be when we will be most efficient.

Published on *iDigBio* (https://www.idigbio.org)

Submission #1597

Submission information

Form: TCN Quarterly Progress Report to iDigBio
Submitted by EPICC
Monday, February 3, 2020 - 14:23
128.32.154.17

TCN Name:
Documenting Fossil Marine Invertebrate Communities of the Eastern Pacific - Faunal Responses to Environmental Change over the last 66 million years

Person completing the report:
aadineen@berkeley.edu

Progress in Digitization Efforts:
As of 1/28/2020, the TCN has fully curated and computer cataloged 1,871,507 specimens (114% of goal) and made 689,597 of these specimens (41% of goal) available in the iDigBio portal. The TCN has photographed 134,088 specimens (160% of goal) and georeferenced 32,911 localities (100% of goal). LACM reports that scanning is underway of their historic collection of taxonomic description cards. Similarly, ANSP recently got all of their original catalog ledgers scanned and saved digitally.

Share and Identify Best Practices and Standards (including Lessons Learned):
PRI has a new hire who has some knowledge of photogrammetry, and has started photographing specimens. This has made it much easier to train her in the standards and practices needed for this project. Her output is very good even after just a few weeks of taking photos. Being able to utilize people who have previous experience, helps cut down on training time and increases the speed of output. Likewise, UCMP had a student familiar with digital photography begin photographing EPICC specimens last fall. Her familiarity with the photography software has sped up the amount of photos UCMP is generating. LACM has started initial experimental use of a barcode scanner in collection, following user experience at USNM. LACMIP labels now include a printed barcode, derived from the LACMIP catalogue number. This can be read into Axiel-EMu and Excel spreadsheets, or for naming images from the label and from images containing the label. This should greatly speed up and minimize key-stroke and formatting errors in research data collection, image capture, and loan form completion, for example. Use at USNM indicates a 3-4x improvement in data capture rates, and improved ease of sharing quality data with USNM following visit.

Identify Gaps in Digitization Areas and Technology:

UCR reports that lack of physical space and software licenses limits the number of student interns that can work on digitization at any one time. ANSP is currently discussing moving to a modern museum-wide collections database (most likely Specify or Arctos), as they do not have a real, working, relational database at this point. In the meantime they are using Excel, as the Filemaker database we inherited is out of date, non-relational, and needs major work in terms of data-cleaning. They are waiting on updates from the museum president about a timeline for the move to a new database before they can make a decision on how to proceed.

Share and Identify Opportunities to Enhance Training Efforts:
Four high-school interns from CAS's Careers in Science program worked on a subset of the collection relevant to EPICC and presented their work at the American Geophysical Union 2019 Fall Meeting. Their work included locality verification via original CAS ledgers, specimen digitization, and a virtual field trip via the EPICC VFE to the Purisima outcrops in Santa Cruz Co., CA. PRI provided photography training to a new hire, while students at ANSP have been training with the staff of the ANSP Malacology department on specimen photography best practices. Likewise, UCR staff provided training to five new undergraduate interns. Lastly, LACM ran a workshop with several Los Angeles Unified School District teachers to train in the use of the “Marine Fossils of Southern California” sorting kit.

Share and Identify Collaborations with other TCNs, Institutions, and Organizations:
Two students at ANSP working on EPICC have partial funding from the NSF program Louis Stokes Alliance for Minority Participation at Drexel University (NSF HRD 1408052). In addition, LACM has been collaborating with USNM on remote identification from images; given assistance to ANSP on cleaning and quality control of locality records; and provided additional hand georeferencing of CAS, UCMP, and PRI southern California localities.

Share and Identify Opportunities and Strategies for Sustainability:
The TCN held their quarterly EPICC virtual meeting on December 13, 2019. Sixteen representatives from 10 institutions attended to follow-up on the topics discussed at the annual in-person meeting last September. Topics discussed include both inward and outward facing manuscripts focused on lessons learned during our TCN, upcoming meetings and workshops, logistics and changes that need to be made to ensure our data can be integrated smoothly, and potential additional no-cost extensions. In addition, several members of the TCN from different institutions are expected to attend the iDigBio workshop, “Georeferencing for Paleo”, in Salt Lake City in late April 2020.

Share and Identify Education and Outreach (E&O) Activities:
PRI staff, Rob Ross and Don Haas, and Lisa White from UCMP, are developing the third virtual field experience (VFE). The VFE is focused on collections from Pleistocene terraces in Palos Verdes Peninsula (southern Los Angeles County). A full draft is expected in Spring 2020. The team is also working toward the fourth and final VFE, focused on the Astoria Formation along the coast of Oregon. Colleagues from UCMP (White, Dineen, and Holroyd) and PRI (Ross and Haas) are providing initial field site information and images, and colleagues from the University of Oregon (Davis) and Portland State University (Granshaw) will meet in the field in February 23-25, 2020.
ANSP presented the EPICC project in December as part of an ANSP donor event featuring the Invertebrate Paleontology department, called "A Toast to the Collection." They showed EPICC specimens, presented the work of the EPICC group on "Quantifying the dark data in museum fossil collections as palaeontology undergoes a second digital

revolution," and spoke with donors about what digitization of collections means. LACM also presented (in part) on EPICC progress at Southern California Unified Malacologists meeting held January 25, 2020.

Google Analytics

Other Progress (that doesn't fit into the above categories):
LACM reports that they are nearly ready for iDigBio ingestion. A number of minor changes to DWC data have been reported to improve downstream usability. They have also been performing quality control assessment on taxonomic data. CAS also reports that many of the workflows created and implemented due to EPICC are now being employed in the CAS Mesozoic invertebrate collections.

Attachment 1

Attachment 2

Source URL: https://www.idigbio.org/node/564/submission/1597

Published on *iDigBio* (https://www.idigbio.org)

Submission #1602

Submission information

Form: TCN Quarterly Progress Report to iDigBio
Submitted by CatChapman
Wednesday, February 5, 2020 - 13:14
66.115.169.71

TCN Name:
TPT

Person completing the report:
kathrynsully95@gmail.com

Progress in Digitization Efforts:
All participating institutions have collections registered on TPT's data portal SCAN (https://scan-bugs.org/portal/). We are working with all data providers to be sure their database platforms can accommodate specimen host association data. We have made modifications to Symbiota, Specify, Arctos, and EMu. We are in the final stages of developing documentation for each platform and will share these workflows with iDigBio upon completion. It is our goal to develop a translation table with controlled vocabularies for host data entry at the TPT data integration workshop this month, and then begin delivering our transcribed records to SCAN/IPTs. We have plans to publish our process and resulting table upon completion of the workshop. Currently, TPT has 14,763 records transcribed, 508 specimens with research grade images, 9,620 scanned images of slide-mounted specimens, 385 scanned images of ethanol stored specimens, and over 70,000 association records on GloBI.

All 8 macropod systems have been delivered to imaging hub institutions (MPM, TAMU, UNL, CAS, UMMZ, UNH, ANS, BPBM) and optimized for imaging of slides and pinned specimens.

Fieldguide Developer Andre Poremski met with lead PI Jen Zaspel in December to establish a schedule of deliverables for the API. We have set the goal of testing our first batch of training images for April.

GloBI
GloBI integration profiles were developed for major platforms, EMu, Arctos, Specify, and Symbiota. GloBI is currently indexing association records from 9 data providing institutions and 1 non funded partner VectorBase. Jessica Light (TAMU) worked on uploading some

host-parasite associations to GloBI and will continue to work with GloBI over the course of the project. A github for TPT and GloBI has been established to submit data and ask questions.

Thus far, TPT has submitted 70,000+ associations representing 9,000+ taxa to GloBI (550,000 associations is our goal).

Share and Identify Best Practices and Standards (including Lessons Learned):
Imaging Workflows
Based on previous models from digitization projects (InvertNet/LepNet), MPM Exhibits and PI Zaspel and PM Kat Sullivan constructed trays to fit on standard flatbed scanners for slide-mounted specimens. These are used in a high-throughput imaging workflow for specimen imaging and data capture. During the quarter, 20 trays were assembled and distributed to 7 participating institutions. The second batch of 20 is under production and will be hand delivered to network members at the workshop this month.

MPM digitization project manager, Alyssa Caywood is working with Macropod developer Mark Smith to produce Macropod imaging workflows to be distributed at the TPT workshop in February. MPM PI and biodiversity informatics director, Chris Tyrrell has developed a detailed workflow and user guide for GeoLocate/georeferencing (to be distributed at the workshop this February).

In November, a standard reporting template was created for PIs to document grant activity according to the timeline of iDigBio quarters and NSF annual reporting. Each data providing institution also received standard task tracking and analysis sheets to document digitization progress.

PM Kat Sullivan has been working with TaxonWorks developer Matt Yoder to import network shared taxonomy files into TaxonWorks. Taxonomy files to be added in the next quarter include mites from Barry OConnor, fleas from Mike Hastriter at BYU, louse taxonomy from Species File, mosquitoes from VectorBase, and various diptera families from ITIS/SCAN. It is our goal to produce and distribute a standardized skeletal taxonomy file for all platforms in time for the February workshop.

Identify Gaps in Digitization Areas and Technology:
Many institutions are still working on hiring personnel and purchasing supplies as a consequence of late award notification and disbursement of funds.

Establishment of ‘resource relationship’ fields for host data entry and IPTs have prevented some data providers from publishing data. Data providers should have everything they need, including database modifications, between now and the end of the February TPT workshop.

Share and Identify Opportunities to Enhance Training Efforts:
New hires
BPBM: 2 techs, 1 volunteer, 1 summer intern 2020 hired
CAS: 1 volunteer, 1 full-time person
FMNH: trained 7 volunteers to inventory and clean slides
INHS: 3 students hired for databasing
MPM: 1 student technician hired, Project Manager Kat Sullivan started 9/3/19
PSU: 2 students hired for spring semester
PERC: Hired grad student to coordinate digitization

TAMU: Hired primary tech and 13 students for semester
UCSB: 1 intern hired
UHIM: 1 student hired for summer 2020
UM: 5 techs hired
MSB: hired 1 undergrad tech
UMN: 1 tech hired
Utah: 1 undergrad hired
UWM: 2 techs and 3 undergrads hired and trained
UWSP: 1 intern

Representatives from 23 out of 26 institutions attended the TCN kickoff meeting at the iDigBio Summit in Gainesville on October 1st. During the kickoff, presenters from the lead institutions (Zaspel, Tyrrell, Cobb, Seltmann, Allen, Sullivan) spoke on project management, reporting, taxonomy, data integration, and broader impacts. Additionally, a representative from each attending institution gave a five minute talk introducing their collection and digitization plan.

An additional workshop for imaging and data integration will take place February 24th-25th at the Field Museum with coordination assistance from iDigBio. In preparation for this workshop, numerous workflows and other documentation have been developed with the intent to distribute and demonstrate at the workshop. Workflows covering georeferencing, slide and vial scanning, Inselect, Macropod imaging, specimen data entry, taxonomy management, and GloBI integration are included in these workflows.

Katja Seltmann (UCSB) worked with Jorrit Poelen, a consultant on the TPT project and GloBI developer. Seltmann did user testing for the GloBI project, created documentation, and evaluated the "usability" of data from a biologist perspective. She discussed with TPT data providers about GloBI and how to share biotic interactions, developed examples of data ingestion to GloBI and data discovery through GloBI R and data download.

PI Mariel Campbell (MSB) attended CDC flea taxonomy ID training in Fort Collins Sept. 2019

Zaspel and Sullivan participated in 3 EMu working group meetings with collections staff and the IT team at the FMNH. The FMNH is sharing EMu mapping and modifications with MPM that will allow both institutions to store species association data more effectively for downstream linkages with GloBI and other applications.

PI Jim Boone (BPBM) is working with Specify User Support to develop import workflow into collection object table in Specify 7.

PSU:Speciesfile Group, University of Illinois are working on customizing image upload function to take composite slide scans. A Standard Operating Procedure for digitization of ethanol-preserved specimens was published to the Frost Entomological Museum website; DOI:10.26207/x4yb-bk89

Share and Identify Collaborations with other TCNs, Institutions, and Organizations:
Many PIs have participated or still participate in other TCNs, including InvertNet, InvertEBase, LepNet, SCAN, and Tri-Trophic. Best practices and workflows from those TCNs are shared with all participants in TPT.

Many outside institutions are developing collaborations with Parasite Tracker to share data

and integrate with Parasite Tracker portal. Walter Reed Biosystematics Unit and Vectorbase have provided taxonomies for the project. Additionally Vectorbase has shared blood meal/host association data with GloBI, which is integrated with other TPT biotic association data. Other institutions including NEON and Joel Hutcheson from the CFIA have started discussions around ways to collaborate.

Yale Peabody submitted a PEN proposal to partner with TPT in October (status is pending).

Golden Gate Raptor Observatory, (Golden Gate National Parks Conservancy) is collecting parasites from bird banding in collaboration with CAS.

Sarah Orlofske attended the Wisconsin Waterfowl Association Meeting. Hunters donated specimens for ectoparasite collecting. (UWSP)

Share and Identify Opportunities and Strategies for Sustainability:
We have started to create bidirectional linkages between data providing institutions and GloBI for harvesting biotic interaction data and linking back to the specimen record.

MPM has advanced its digital infrastructure by deploying a museum-specific instance of the Symbiota biodiversity data platform (symbiota.mpm.edu). In addition to establishing the external server and software, MPM was also able to construct and test the export process between the museum's internal database (Axiell EMu) and this new Symbiota instance. This web portal will now benefit all of MPM's natural history collections (not just those involved in TPT) and is served through the museum's Information Services department ensuring long term sustainability in data delivery.

Katja Seltmann is working with Neil Cobb and Evin Dunn at SCAN to add biotic interaction functionality to Symbiota. This builds off of prior work done by Ed Gilbert and Seltmann in 2017-18.

Share and Identify Education and Outreach (E&O) Activities:
● New TCN intro by Jen Zaspel (PU/MPM) at the iDigBio summit 10/2
● iDigBio IAC quarterly report given by Jen Zaspel (PU/MPM) on 11/6
● Nov. 2019 Sullivan, K., Caywood, A., Colby, J., Tyrrell, C., Zaspel, J., et al. (40 co-authors). Digitization TCN: Terrestrial Parasite Tracker: Digitizing collections to trace parasite-host associations and predict the spread of vector-borne disease. Entomological Collections Network Annual Meeting, St. Louis, MO.
● Jen Zaspel (PU/MPM) gave lecture during Milwaukee Museums week on Parasite Tracker 1/24
● Preliminary planning meeting with Brain Scoop team to film at MPM and FMNH
● Zaspel participated in interview for WUWM Lake Effect (Bonnie Noth) *received "Best of 2019" distinction
● Zaspel participated in interview with FOX6 (Ben Handleman)
● Zaspel provided content for numerous press releases for the Terrestrial Parasite Tracker project
● Abstract for TPT talks submitted for SPNHC and Digital Data meetings in 2020.
● Planning for Biology week programming in Milwaukee area schools and at MPM (JZ)
● Planning for MPM "Inside Out" members night (May 1, 2020 JZ)
● Planning for MPM behind the scenes tours for March adult sleepover "Unhealthy relationships: What is a parasite and how do you become one?" (JZ)
● PM Sullivan and MPM PI Tyrrell have developed background content for iDigBio wiki and

parasite tracker website (parasitetracker.org)
● (Thursday, 24 October 2019) Seltmann co-organized with Deborah Paul and Vincent Smith a symposium at Biodiversity Next conference, Leiden. The symposium titled "SP39 Increasing Opportunities to Align Data Initiatives for Bio/Geo Collections" was focused on connecting global biodiversity informatics initiatives so that we understand our collective visions and can work together. This symposium included discussion on the Extended Specimen, which prioritizes enhances specimen data, such as biotic interactions being captured in the TPT project.
● An outreach event for the TPT project was a musical performance on October 3, 2019 in Gainesville, Florida. The performance, titled "Our Pestest Pals," included songs about insect behavior and talked about the arthropods that live with us in our homes. Scientifically Speaking with Irene Moon is a pseudonym of Seltmann to bring entomology into rock clubs and art spaces. (https://en.wikipedia.org/wiki/Irene_Moon)
● PR article in UW-Madison College of Agriculture and Life Sciences (CALS) newsletter Article for the Newsletter of the Wisconsin Entomological Society used as the cover piece for the October 2019 Newsletter (UWM).
● Sarah Orlofske provided a public presentation related to ectoparasites for the UWSP Natural History Museum (UWSP)
● Orlofske developed Scientific Method lab for undergraduate students in the Introductory Animal Biology Course. Lab involved concepts of ecology and evolution of ectoparasites and included microscope examination of louse specimens.

Google Analytics

Other Progress (that doesn't fit into the above categories):
PI Zaspel and Branden Apitz designed a logo for TPT which has been shared across the network and branded on parasitetracker.org and the Parasite Tracker page on GloBI (https://www.globalbioticinteractions.org/parasitetracker)

Attachment 1
2020-02-05 TPT TCN iDigBio Qreport.docx

Attachment 2

Source URL: https://www.idigbio.org/node/564/submission/1602

TPT TCN-- Quarterly Report --September 2019-January 2020

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

Assembled by Jen Zaspel and Kat Sullivan, February 4, 2020

Progress in Digitization Efforts:

All participating institutions have collections registered on TPT's data portal SCAN (https://scan-bugs.org/portal/). We are working with all data providers to be sure their database platforms can accommodate specimen host association data. We have made modifications to Symbiota, Specify, Arctos, and EMu. We are in the final stages of developing documentation for each platform and will share these workflows with iDigBio upon completion. It is our goal to develop a translation table with controlled vocabularies for host data entry at the TPT data integration workshop this month, and then begin delivering our transcribed records to SCAN/IPTs. We have plans to publish our process and resulting table upon completion of the workshop. Currently, TPT has 14,763 records transcribed, 508 specimens with research grade images, 9,620 scanned images of slide-mounted specimens, 385 scanned images of ethanol stored specimens, and over 70,000 association records on GloBI.

	Transcribed records	Specimens imaged with Macropod	Scanned images	GloBI association records
ANS	411		385	
MPM	1026			
UM	12789		1851	67800
UMSP			7769	
MSB	537			537
BPBM		508		
Vectorbase (not on TCN)				12474

All 8 macropod systems have been delivered to imaging hub institutions (MPM, TAMU, UNL, CAS, UMMZ, UNH, ANS, BPBM) and optimized for imaging of slides and pinned specimens.

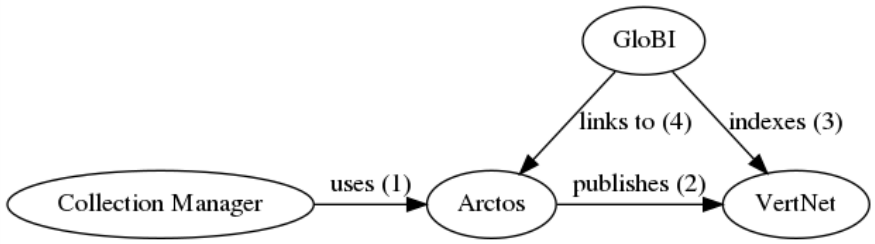
Fieldguide Developer Andre Poremski met with lead PI Jen Zaspel in December to establish a schedule of deliverables for the API. We have set the goal of testing our first batch of training images for April.

GloBI

GloBI integration profiles were developed for major platforms, EMu, Arctos, Specify, and Symbiota. GloBI is currently indexing association records from 9 data providing institutions and 1 non funded partner VectorBase. Jessica Light (TAMU) worked on uploading some host-parasite associations to GloBI and will continue to work with GloBI over the course of the project. A github for TPT and GloBI has been established to submit data and ask questions.

Arctos

Arctos managed collections are indexed by GloBI. This section explains how.

Arctos	GloBI Integration Profile
actors	Arctos, VertNet, GloBI
integration method	A collection manager uses Arctos to establish associations or relationships between records. Arctos periodically shares data with VertNet. VertNet uses GBIF IPT software to publish data archives. VertNet publishes a list of available datasets in the form of a RSS feed, including those shared by Arctos. Periodically, GloBI finds, and downloads, Arctos related data archives in VertNet. Then, GloBI indexes the associatedOccurrences fields of records in these Arctos data archives. The associatedOccurrences contain the association type (e.g., "eats") and a pointer to the occurrence id of the linked record.
diagram	
example collection	MSB-PARA

Thus far, TPT has submitted 70,000+ associations representing 9,000+ taxa to GloBI (550,000 associations is our goal).

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

Imaging Workflows

Based on previous models from digitization projects (InvertNet/LepNet), MPM Exhibits and PI Zaspel and PM Kat Sullivan constructed trays to fit on standard flatbed scanners for slide-mounted specimens. These are used in a high-throughput imaging workflow for specimen imaging and data capture. During the quarter, 20 trays were assembled and distributed to 7 participating institutions. The second batch of 20 is under production and will be hand delivered to network members at the workshop this month.

MPM digitization project manager, Alyssa Caywood is working with Macropod developer Mark Smith to produce Macropod imaging workflows to be distributed at the TPT workshop in February. MPM PI and biodiversity informatics director, Chris Tyrrell has developed a detailed workflow and user guide for GeoLocate/georeferencing (to be distributed at the workshop this February).

In November, a standard reporting template was created for PIs to document grant activity according to the timeline of iDigBio quarters and NSF annual reporting. Each data providing institution also received standard task tracking and analysis sheets to document digitization progress.

PM Kat Sullivan has been working with TaxonWorks developer Matt Yoder to import network shared taxonomy files into TaxonWorks. Taxonomy files to be added in the next quarter include mites from Barry OConnor, fleas from Mike Hastriter at BYU, louse taxonomy from Species File, mosquitoes from VectorBase, and various diptera families from ITIS/SCAN. It is our goal to produce and distribute a standardized skeletal taxonomy file for all platforms in time for the February workshop.

Identify Gaps in Digitization Areas and Technology:

Many institutions are still working on hiring personnel and purchasing supplies as a consequence of late award notification and disbursement of funds.

Establishment of 'resource relationship' fields for host data entry and IPTs have prevented some data providers from publishing data. Data providers should have everything they need, including database modifications, between now and the end of the February TPT workshop.

Share and Identify Opportunities to Enhance Training Efforts:

<u>New hires</u>

BPBM: 2 techs, 1 volunteer, 1 summer intern 2020 hired
CAS: 1 volunteer, 1 full-time person
FMNH: trained 7 volunteers to inventory and clean slides
INHS: 3 students hired for databasing
MPM: 1 student technician hired, Project Manager Kat Sullivan started 9/3/19
PSU: 2 students hired for spring semester
PERC: Hired grad student to coordinate digitization
TAMU: Hired primary tech and 13 students for semester
UCSB: 1 intern hired
UHIM: 1 student hired for summer 2020
UM: 5 techs hired
MSB: hired 1 undergrad tech
UMN: 1 tech hired
Utah: 1 undergrad hired
UWM: 2 techs and 3 undergrads hired and trained

UWSP: 1 intern

Representatives from 23 out of 26 institutions attended the TCN kickoff meeting at the iDigBio Summit in Gainesville on October 1st. During the kickoff, presenters from the lead institutions (Zaspel, Tyrrell, Cobb, Seltmann, Allen, Sullivan) spoke on project management, reporting, taxonomy, data integration, and broader impacts. Additionally, a representative from each attending institution gave a five minute talk introducing their collection and digitization plan.

An additional workshop for imaging and data integration will take place February 24th-25th at the Field Museum with coordination assistance from iDigBio. In preparation for this workshop, numerous workflows and other documentation have been developed with the intent to distribute and demonstrate at the workshop. Workflows covering georeferencing, slide and vial scanning, Inselect, Macropod imaging, specimen data entry, taxonomy management, and GloBI integration are included in these workflows.

Katja Seltmann (UCSB) worked with Jorrit Poelen, a consultant on the TPT project and GloBI developer. Seltmann did user testing for the GloBI project, created documentation, and evaluated the "usability" of data from a biologist perspective. She discussed with TPT data providers about GloBI and how to share biotic interactions, developed examples of data ingestion to GloBI and data discovery through GloBI R and data download.

PI Mariel Campbell (MSB) attended CDC flea taxonomy ID training in Fort Collins Sept. 2019

Zaspel and Sullivan participated in 3 EMu working group meetings with collections staff and the IT team at the FMNH. The FMNH is sharing EMu mapping and modifications with MPM that will allow both institutions to store species association data more effectively for downstream linkages with GloBI and other applications.

PI Jim Boone (BPBM) is working with Specify User Support to develop import workflow into collection object table in Specify 7.

PSU:Speciesfile Group, University of Illinois are working on customizing image upload function to take composite slide scans. A Standard Operating Procedure for digitization of ethanol-preserved specimens was published to the Frost Entomological Museum website; DOI:10.26207/x4yb-bk89

Share and Identify Collaborations with other TCNs, Institutions, and Organizations:

Many PIs have participated or still participate in other TCNs, including InvertNet, InvertEBase, LepNet, SCAN, and Tri-Trophic. Best practices and workflows from those TCNs are shared with all participants in TPT.

Many outside institutions are developing collaborations with Parasite Tracker to share data and integrate with Parasite Tracker portal. Walter Reed Biosystematics Unit and Vectorbase have provided taxonomies for the project. Additionally Vectorbase has shared blood meal/host association data with GloBI, which is integrated with other TPT biotic association data. Other institutions including NEON and Joel Hutcheson from the CFIA have started discussions around ways to collaborate.

Yale Peabody submitted a PEN proposal to partner with TPT in October (status is pending).

Golden Gate Raptor Observatory, (Golden Gate National Parks Conservancy) is collecting parasites from bird banding in collaboration with CAS.

Sarah Orlofske attended the Wisconsin Waterfowl Association Meeting. Hunters donated specimens for ectoparasite collecting. (UWSP)

Share and Identify Opportunities and Strategies for Sustainability:
We have started to create bidirectional linkages between data providing institutions and GloBI for harvesting biotic interaction data and linking back to the specimen record.

MPM has advanced its digital infrastructure by deploying a museum-specific instance of the Symbiota biodiversity data platform (symbiota.mpm.edu). In addition to establishing the external server and software, MPM was also able to construct and test the export process between the museum's internal database (Axiell EMu) and this new Symbiota instance. This web portal will now benefit all of MPM's natural history collections (not just those involved in TPT) and is served through the museum's Information Services department ensuring long term sustainability in data delivery.

Katja Seltmann is working with Neil Cobb and Evin Dunn at SCAN to add biotic interaction functionality to Symbiota. This builds off of prior work done by Ed Gilbert and Seltmann in 2017-18.

Share and Identify Education and Outreach (E&O) Activities:

- New TCN intro by Jen Zaspel (PU/MPM) at the iDigBio summit 10/2
- iDigBio IAC quarterly report given by Jen Zaspel (PU/MPM) on 11/6
- Nov. 2019 Sullivan, K., Caywood, A., Colby, J., Tyrrell, C., Zaspel, J., et al. (40 co-authors). Digitization TCN: Terrestrial Parasite Tracker: Digitizing collections to trace parasite-host associations and predict the spread of vector-borne disease. Entomological Collections Network Annual Meeting, St. Louis, MO.
- Jen Zaspel (PU/MPM) gave lecture during Milwaukee Museums week on Parasite Tracker 1/24
- Preliminary planning meeting with Brain Scoop team to film at MPM and FMNH
- Zaspel participated in interview for WUWM Lake Effect (Bonnie Noth) *received "Best of 2019" distinction

- Zaspel participated in interview with FOX6 (Ben Handleman)
- Zaspel provided content for numerous press releases for the Terrestrial Parasite Tracker project
- Abstract for TPT talks submitted for SPNHC and Digital Data meetings in 2020.
- Planning for Biology week programming in Milwaukee area schools and at MPM (JZ)
- Planning for MPM “Inside Out” members night (May 1, 2020 JZ)
- Planning for MPM behind the scenes tours for March adult sleepover “Unhealthy relationships: What is a parasite and how do you become one?” (JZ)
- PM Sullivan and MPM PI Tyrrell have developed background content for iDigBio wiki and parasite tracker website (parasitetracker.org)
- (Thursday, 24 October 2019) Seltmann co-organized with Deborah Paul and Vincent Smith a symposium at Biodiversity Next conference, Leiden. The symposium titled "SP39 Increasing Opportunities to Align Data Initiatives for Bio/Geo Collections" was focused on connecting global biodiversity informatics initiatives so that we understand our collective visions and can work together. This symposium included discussion on the Extended Specimen, which prioritizes enhances specimen data, such as biotic interactions being captured in the TPT project.
- An outreach event for the TPT project was a musical performance on October 3, 2019 in Gainesville, Florida. The performance, titled "Our Pestest Pals," included songs about insect behavior and talked about the arthropods that live with us in our homes. Scientifically Speaking with Irene Moon is a pseudonym of Seltmann to bring entomology into rock clubs and art spaces. (https://en.wikipedia.org/wiki/Irene_Moon)
- PR article in UW-Madison College of Agriculture and Life Sciences (CALS) newsletter Article for the Newsletter of the Wisconsin Entomological Society used as the cover piece for the October 2019 Newsletter (UWM).
- Sarah Orlofske provided a public presentation related to ectoparasites for the UWSP Natural History Museum (UWSP)
- Orlofske developed Scientific Method lab for undergraduate students in the Introductory Animal Biology Course. Lab involved concepts of ecology and evolution of ectoparasites and included microscope examination of louse specimens.

Other:
PI Zaspel and Branden Apitz designed a logo for TPT which has been shared across the network and branded on parasitetracker.org and the Parasite Tracker page on GloBI (https://www.globalbioticinteractions.org/parasitetracker)

Published on *iDigBio* (https://www.idigbio.org)

Submission #1608

Submission information

Form: TCN Quarterly Progress Report to iDigBio
Submitted by akasameyer
Friday, February 7, 2020 - 13:11
192.31.105.152

TCN Name:
The Pteridological Collections Consortium: An integrative Approach to Pteridophyte Diversity Over the Last 420 Million Years

Person completing the report:
akasameyer@berkeley.edu

Progress in Digitization Efforts:
For extant specimen progress during this reporting period, Pteridophyte Collections Consortium members created skeletal records for 41,375 specimens, fully transcribed 41,579 specimens, imaged 90,240 specimens, and geo-referenced 7,042 specimen records. The total pteridophyte extant specimen progress including work done prior to the start of the grant is 533,094 (32% of goal) skeletal records created, 988,609 (60% of goal) extant specimens imaged, 699,039 (42% of goal) extant specimens fully transcribed, and 133,658 (8% of goal) extant specimens geo-referenced.

In our Pteridoportal we currently have the following extant specimens:
1,354,306 occurrence records
- 303,423 (22%) georeferenced
- 969,931 (72%) occurrences imaged
- 665,512 (49%) identified to species

For fossil specimen progress during this reporting period, Pteridophyte Collections Consortium members databased 1,862 specimens, imaged 2,050 specimens, and geo-referenced 925 specimen records. The total pteridophyte fossil specimen progress including work done prior to the start of the grant is 28,062 (32% of goal) specimens databased, 26,558 (30% of goal) specimens imaged, and 10,054 (11.5% of goal) specimen records geo-referenced.

At the University of Florida a new student employee is post-processing their remaining 6000 images in Adobe Lightroom.

New York Botanical Garden processed all new images with OCR software and uploaded

text to specimen records to expedite rapid data entry.

Share and Identify Best Practices and Standards (including Lessons Learned):
The Sam Noble Museum is still finalizing revised detailed written museum-specific workflows for various processes of the project and getting approval to share them on their website once complete. Lesson Learned: 1. Changing from a specimen number with multiple identifications to specimens containing multiple specimens requires not just a process and standards but also recording the rationale especially for more complex specimens and/or special cases. 2. Remember to include specific instructions/procedures about how a team left off and need to pick back up so that it is faster to get back into the workflow, especially if it will be for an extended period for all teams (e.g., holidays/breaks). 3. When standards change, a statement/document that explains the change and when it occurred should be created and shared in a publicly accessible manner. As an example, the museum's change to "overall slab that contains multiple fossil" specimens and "identified taxa on an overall slab of other fossils" specimens in addition to "single taxa" specimens (which only have one fossil on overall slab).

Identify Gaps in Digitization Areas and Technology:
The paleo module is still under development; we are expecting an update on the module the week of February 3rd. The unavailability of the module has been an issue for participants.

The Sam Noble Museum is entering what they need for their collection-specific database and working towards ensuring that it is shareable via their IPT server, with the hope that their IPT server data can be used in the PCC Portal.

Another issue at the Sam Noble Museum, while not preventing progress, is that existing "pteridophyte fossil specimens" that already existed in their database as "pteridophyte fossils" and were shared via IPT with UUIDs are requiring careful consideration and slowing complete data record completion. This is because these existing "pteridophyte fossil specimens" need to be treated more like skeletal records that require collection staff to ensure specimens cataloged as "single taxa pteridophyte fossil" and "a pteridophyte taxa fossil on an overall slab that now contain multiple fossils" maintain their appropriate UUID identification in the database and eventual IPT data and do not end up as "overall slabs containing fossils." Staff came to the realization that as they have already shared specimens online and via an IPT server with UUIDs, they should create a statement that indicates their museum's change to overall slab containing fossils specimens and individualized taxa IDed fossil specimens on those overall slab specimens with their own specimen numbers, and that they kept the UUIDs with the "identification" as initially shared/published.

Share and Identify Opportunities to Enhance Training Efforts:
PCC TCN Partners trained one graduate student, two undergraduate students, four interns, and more than ten volunteers on imaging, data entry, and georeferencing.

The University of Texas continued to offer students informal enrichment through periodic discussions of fern morphology and taxonomic classification.

The Sam Noble Museum held an in-house collection-specific training class on paleobotany fossils, photography, collection processes, and workflows for seven new volunteers to assist with the project. Existing volunteers and students were provided with more advanced training on collection photo standards and lighting. They now have a team of 20

(mainly volunteers) assisting with various aspects of the project.

At the University of Michigan a new program was developed in which experienced undergraduate technicians specialize in specific geographical regions and peer review transcription of their selected geographical units. The new protocol allows students to pass partially completed datasets of specimen data and images among each other in order to improve data quality and reduce research time needed for accurate transcriptions. Especially in the Philippines, where many old province names have changed, this effort has reduced the number of records with skeletal-only entry for geography and locality fields.

Share and Identify Collaborations with other TCNs, Institutions, and Organizations:
Participants are coordinating with the following TCNs: Cretaceous World, Endless Forms, Capturing California's Phenology, and the Texas Oklahoma Regional Consortium of Herbaria "American Crossroads."

Share and Identify Opportunities and Strategies for Sustainability:
The Pteridoportal was moved from the IDigBio server to ASU.

The Sam Noble Museum is continuing to take the opportunity of working collaboratively with other collection staff, volunteers, and their museum's IT department to improve in-house technical and collection knowledge as well as establishing in-house processes to appropriately add additional data fields to datasets being shared via the museum's IPT server, ensure data quality, and handle updates as needed.

Share and Identify Education and Outreach (E&O) Activities:
The New York Botanical Garden Herbarium hosted several public tours of the collections which feature the Digital Imaging Center and highlight the TCNs, their digitization staff, and the value of preserving and digitizing natural history collections.

Google Analytics

Other Progress (that doesn't fit into the above categories):
For the paleo taxonomy in the Pteridoportal, we attempted to obtain access to a foundational nomenclatural list (i.e., a list of taxon names with their authorities, showing which are synonyms and what their higher level classification is) for fossil plants through the "Index of Fossil Plant Names." However, the coordinator of that list ultimately declined to share that resource, so we are instead focusing on building up the fossil portion of our taxonomic thesaurus in a more piecemeal way. We have downloaded the names attributed to fossil plant records in GBIF, and are in the process of curating that list--removing non-pteridophytes, correcting errors, removing names that have extant representatives, and applying a preliminary higher classification. In parallel, we have developed a protocol for handling missing higher ranks (fossils are often, for example, named at the level of genus but without placement in a particular family or order), and for indicating which names apply to fossils (versus extant taxa).

Attachment 1

Attachment 2

Source URL: https://www.idigbio.org/node/564/submission/1608

Published on *iDigBio* (https://www.idigbio.org)

Home > Collaborators > TCN Quarterly Progress Report to iDigBio > Webform results > TCN Quarterly Progress Report to iDigBio

Submission #1600

Submission information

Form: TCN Quarterly Progress Report to iDigBio
Submitted by nyeung
Tuesday, February 4, 2020 - 15:47
76.81.71.178

TCN Name:
PILSBRY

Person completing the report:
nyeung@hawaii.edu

Progress in Digitization Efforts:
About 2,300 lots (~27,000) specimens have been entered in the database and about 800 records have been georeferenced. All ledgers from ANSP and BPBM have now been digitized. Several collectors' field notebooks and ledgers have also been digitized in MCZ and BPBM.

Share and Identify Best Practices and Standards (including Lessons Learned):
We have started developing workflows to to batch clean up of taxonomic names, locality details, and inputting references.

Identify Gaps in Digitization Areas and Technology:
Current gaps are personnel issues. Postdoctoral researcher was to start January 1, 2020 but died in research dive in November. PI for MCZ also passed away in December. Recruitment of technicians in some institutions have been slower than expected.

Share and Identify Opportunities to Enhance Training Efforts:
In total, 3 volunteers, 2 technicians, 2 high school students, 7 undergraduates, and 2 post baccalaureates are being trained in databasing, imaging, and georeferencing specimen records. More than half are females and 5 are Pacific Islanders. Additional technicians and students to be recruited during by Spring 2020.

Workflow workshop being organized in Hawaii Feb 14-15th.

Share and Identify Collaborations with other TCNs, Institutions, and Organizations:
Working with WoRMS and MolluscaBase on taxonomic upload protocols

Share and Identify Opportunities and Strategies for Sustainability:
PILSBRY TCN workflow workshop being organized in Hawaii on Feb 14-15 to finalize

methods for data cleanup re: names and localities. Nelson Rios from GeoLocate attending to assist in finalizing Pacific gazetteer to allow consistent locality information that can be used by all collections.

Additionally, techniques being developed for removing tipped-in materials from the ledgers. Investigation of the adhesives originally used is currently under way in collaboration with antique book specialists.

Share and Identify Education and Outreach (E&O) Activities:
Exhibit booths are being organized for several science fairs in the spring and summer. BPBM high school students and post-baccalaureate student submitting poster abstracts for Hawaii Conservation Conferences presenting their georeferencing projects.

Google Analytics

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

Attachment 1

Attachment 2

Source URL: https://www.idigbio.org/node/564/submission/1600