

## INVERTEBASE: REACHING BACK TO SEE THE FUTURE: SPECIES-RICH INVERTEBRATE FAUNAS DOCUMENT CAUSES AND CONSEQUENCES OF BIODIVERSITY SHIFTS

Report submitted by: psierwald@fieldmuseum.org  
Report Submitted on: 06/29/2015 - 20:37

### Progress in Digitization Efforts

Data entry ongoing in all participating collections

In addition to hired digitization staff at all collections, most institutions have recruited additional volunteers for data entry

Total records entered: Insects: 86,700, Mollusks: 27,220 (FMNH 11,000 Insects, 5,350 mollusks; AUMNH 60,000 insects; CMNH 5,000 insects, re-mobilizing 6,000 mollusks records; UMMZ 15,870 mollusks, 10,700 insects cataloged and bar-coded records), PSUC, due to different collection type (Odonata in envelopes): imaging preceded data entry

Specimen imaging in first 9 months (largely supported by non-grant funds)

1. Images (many types): 21,600 Insects, 3,147 mollusks, 1,220 invertebrate types imaged (FMNH 11,000 insects, 1,220 invertebrate type images, PSUC: 10,600 Odonata and Hymenoptera; UMMZ: 3,147 mollusks)

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

Taxonomic authority files:

1. North American terrestrial arthropods: Mantodea (CMNH), Odonata (UMMZ, FMNH), Blattodea (FMNH), Coleoptera to subfamily level, Hymenoptera (Apocrita, Symphyta) development in progress
2. Near completion of North American land and freshwater mollusk taxonomic authority list (FMNH, DMNH)
3. Coordination with taxonomic thesaurus in SCAN established
4. North American and regional geographic look-up tables generated, shared and in use for data entry (FMNH, AUMNH)

### Identify Gaps in Digitization Areas and Technology

voice recognition software not yet performing at adequate speed

### Share and Identify Opportunities to Enhance Training Efforts

FMNH digitization staff: georeferencing training workshops for Geolocate completed

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

Collaboration with SCAN regarding the arthropod thesaurus of North American Arthropod taxon names

### Share and Identify Opportunities and Strategies for Sustainability

The awarding of this grant to the Frost Museum has already raised the profile of this arthropod collection within the College of Agricultural Sciences and across Penn State. The College is now investing precious resources into renovating the building (Headhouse III) that houses the Frost Museum. And the University has made the Frost Museum, along with all other museums on campus, part of a billion dollar capital campaign that should result in entirely new, and interconnected (physically and programmatically) facilities.

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

Significant collection improvements due to pre-curation of material to be digitized at all institutions

Updates to the FilteredPush client helper software to support the configuration of multiple Symbiota portals sharing the same Symbiota database and the same FilteredPush node (SCAN and InvertEBase)

### Attachment

N/A

## FOSSIL INSECT COLLABORATIVE: A DEEP-TIME APPROACH TO STUDYING DIVERSIFICATION AND RESPONSE TO ENVIRONMENTAL CHANGE

Report submitted by: adrian.carper@gmail.com  
Report Submitted on: 06/30/2015 - 10:40

### Progress in Digitization Efforts

Since the last report, the University of Colorado Museum of Natural History (UCB) has taken 908 additional fossil insect images, added 1,076 new specimen records to their database, and edited 5,612 records. Digitization was put on hold for several weeks in May while repairs to the UCB imaging station were completed. The Harvard University Museum of Comparative Zoology (MCZ) has taken 1,500 new images, accounting for about 1300 specimens from their collection. In addition, MCZ has assigned 300 new catalog numbers to unnumbered fossils found as they imaged their collection, and corrected image-database mismatches through scripting so all images taken so far are available to be used by the iDigPaleo hub.

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

MCZ is improving the preservation state of their amber collection by placing all specimens into sealed plastic zip-lock bags to protect against moisture and oxygen. This has little impact on the digitization process and it is a first step towards the curatorial improvement of the amber collection. They have also updated the images taken so far and their respective database records to account for a more standardized cataloging process. Each amber inclusion within the same amber piece is now given a syninclusion number (.1, .2, .3...) with the same catalog number, which should help facilitate the extraction of paleoecological data from their amber inclusions.

### Identify Gaps in Digitization Areas and Technology

There is nothing to report.

### Share and Identify Opportunities to Enhance Training Efforts

There is nothing to report.

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

D. Smith (UCB) co-organized the Specimens Full Circle session at the SPNHC Annual Meeting in Gainesville and also participated in the "TCN Coffee Klatch" Special Interest Group Panel. T. Karim, L. Walker, and E. Anderson also presented talks at the SPNHC Annual Meeting in Gainesville in the Specimens Full Circle Session.

MCZ hosted Dr. Conrad Labandeira, from the Smithsonian Institution (National Museum of Natural History, hereafter, NMNH) June 16th – 18th to examine and take a 35-year loan of Mazon Creek fossil insects to be integrated into the new NMNH exhibition, "Deep Time". MCZ also digitized ichneumonid wasp holotypes from Florissant for Dr. Seraina Klopstein, from the Naturhistorisches Museum in Bern, who aims to reconstruct the phylogeny of the group taking into account fossils.

An undergraduate student from the College of William and Mary is utilizing a subset of the Virginia Museum of Natural History's (VMNH) digitized fossil insect images for a senior thesis.

### Share and Identify Opportunities and Strategies for Sustainability

There is nothing to report.

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

Christina Boyd from VMNH attended the iDigBio 3D Digitization for Educators and Citizen Scientists Workshop.

### Attachment

N/A

## MOBILIZING NEW ENGLAND VASCULAR PLANT SPECIMEN DATA TO TRACK ENVIRONMENTAL CHANGE

Report submitted by: sweeneyp0708@gmail.com  
Report Submitted on: 07/02/2015 - 09:42

### Progress in Digitization Efforts

Capture of collection level-information (i.e., "pre-capture") is complete. Approximately 800,000 specimens have been pre-captured -- with at least current identification captured. As part of the primary digitization phase, approximately 428,700 records and 435,700 images have been captured.

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

nothing to report

### Identify Gaps in Digitization Areas and Technology

nothing to report

### Share and Identify Opportunities to Enhance Training Efforts

NEVP held their annual meeting in conjunction with the Consortium of Northeastern Herbaria meeting. The meeting was held at the New York Botanical Garden. There were talks and a workshop on using the FP-Akka workflow for quality control of specimen records. The meeting agenda is here: [http://neherbaria.org/sites/default/files/CNH\\_2015\\_agenda\\_detailed.pdf](http://neherbaria.org/sites/default/files/CNH_2015_agenda_detailed.pdf). The workshop page is here: [http://wiki.datakurator.net/web/CNH\\_Workshop\\_June2015](http://wiki.datakurator.net/web/CNH_Workshop_June2015).

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

We continue to collaborate with, iPlant, the FilteredPush project, the Symbiota team, and iDigBio. We are collaborating with Anne Bashram (U. of AZ), iDigBio, and other TCNs to develop a Augmented Reality tool that will be useful in K-12 education. Sweeney, Neefus, and Allard worked with iDigBio and other ADBC TCNs to produce a Herbarium workflows paper that was submitted for publication in June; this work builds on the Jan 2015 workshop that was hosted by iDigBio and SERNEC.

### Share and Identify Opportunities and Strategies for Sustainability

nothing to report

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

nothing to report

### Attachment

N/A

## THE MACROALGAL HERBARIUM CONSORTIUM: ACCESSING 150 YEARS OF SPECIMEN DATA TO UNDERSTAND CHANGES IN THE MARINE/AQUATIC ENVIRONMENT

Report submitted by: Chris.neefus@unh.edu  
Report Submitted on: 07/02/2015 - 12:29

### Progress in Digitization Efforts

See attached chart.

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

We have begun developing a Best Practices workflow for imaging exsiccati

### 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

Several of the people involved in the Macroalgal TCN are also associated with the NEVP TCN and attended the NEVP TCN annual meeting held at NYBG

### Share and Identify Opportunities and Strategies for Sustainability

We're starting to look at potential projects and funding opportunities that utilize the data and portal to answer research question on biogeography

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

2 of the people associates with the macroalgal TCN participated in the iDigBio API hackathon.

### Attachment

<https://www.idigbio.org/sites/default/files/webform/tcn-reports/digitization%20numbers%206-30-15.pdf>

Digitizing Institution	Start	Collections	Specimens	Percent Complete				
				Records Created	On Portal	Imaged	Transcribed	Geo-referenced
University of New Hampshire	Year 1	10	133,663					
New York Botanical Garden	Year 1	5	163,350					
University of North Carolina	Year 1	7	74,733					
University of Michigan	Year 1	5	95,060					
University of Washington	Year 1	3	37,154					
Duke University	Year 1	1	19,000					
University of Alaska	Year 1	1	9,290					
Bishop Museum	Year 1	1	78,795					
Field Museum	Year 1	1	47,930					
Oregon State University	Year 1	1	9,000					
University of Guam	Year 1	1	13,600					
University of California - Berkeley	Year 2	9	228,861					
University of Hawaii	Year 2	1	2,401					
Harvard University	Year 2	1	150,000					
Academy of Natural Sciences	Year 3	1	37,000					
University of Vermont	Year 3	1	3,500					
Totals		49	1,103,337					

## THE MACROFUNGI COLLECTION CONSORTIUM: UNLOCKING A BIODIVERSITY RESOURCE FOR UNDERSTANDING BIOTEC INTERACTIONS, NUTRIENT CYCLING AND HUMAN AFFAIRS

Report submitted by: bthiers@nybg.org  
Report Submitted on: 07/05/2015 - 13:42

### Progress in Digitization Efforts

So far, approximately 582,517 items have been digitized for this project, including 562,119 specimen labels and 20,398 supplemental images (e.g., specimen photographs, drawings, field notes). The major product of the project is the MycoPortal. To date 1,884,939 specimen records have been added to the portal. 171,874 of these were added in year 3. The Portal contains about 47,000 skeletal records (i.e. locality data yet to be added); 1,837,939 are records with complete text locality information and 214,453 records have geocoordinates. The MycoPortal also has been populated with 582,517 images (approximately 20,398 of these are living fungi, the remainder are images of dried fungi, labels, and field notes), 42 checklists, including a checklist for North American fungi with more than 15,000 entries).

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

Nothing new to report.

### Identify Gaps in Digitization Areas and Technology

We have more than 40,000 specimen records transcribed through the Notes from Nature Zooniverse platform, but so far it has been impossible to retrieve these completed records for reimportation into the institutional databases from which the records derived. This problem results from the lack of an established procedure for re-extraction of transcribed records. Fortunately, a new grant award to the NfN Advisory Committee has been awarded, and will address the technological shortcomings of this site.

### Share and Identify Opportunities to Enhance Training Efforts

A transcription crowdsourcing workshop sponsored by the MaCC project will be held at Botany 2015, led by Dr. George Weiblen of University of Minnesota and Mari Roberts of NYBG.. The workshop is sponsored by the MaCC project.

We were supposed to hold a High School teacher training workshop in conjunction with the Mycological Society of America meeting in Edmonton, but due to lack of enrollment by Canadian high school teachers, the training workshop has been cancelled, and has been re-scheduled for the next MSA meeting in Berkeley, CA in 2016.

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

We have collaborated in sharing expertise with setting up a citizen science/crowdsourcing program for transcribing specimen records. Mari Roberts, who is funded as Volunteer Coordinator for the Lichens, Bryophytes and Climate Change TCN, is stationed at NYBG, and thus interacts closely with other TCN projects here, and has incorporated experiences from all the TCNS that NYBG is involved in into the attached document.

### Share and Identify Opportunities and Strategies for Sustainability

This will be a topic for discussion with Andy Miller of the Microfungi Collection Consortium. We will use the same portal and will have the same sustainability issues,

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

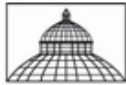
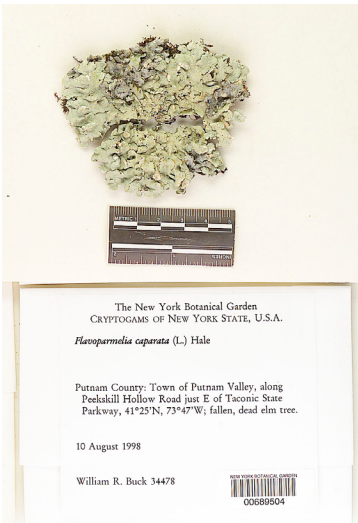
Year 3 Annual report was submitted in June 2015. The project is now on a one year no-cost extension.

### Attachment

<https://www.idigbio.org/sites/default/files/webform/tcn-reports/Crowdsourcing%20Guidelines%20LBCC.pdf>

# Digitization Guidelines

## Instructions for Crowdsourcing



## Digitization Guidelines: Instructions for Crowdsourcing

© 2015 *The New York Botanical Garden*

# Acknowledgements

The William and Lynda Steere Herbarium gratefully acknowledges the support of employees and leaders who helped make this manual possible. Their knowledge and expertise helped make this manual the most comprehensive and up-to-date resource on data entry and management for botanical collections in Symbiota.

## Contributors

### Author

**Mari A. Roberts**

Volunteer Coordinator

Project Manager for TTD-TCN

mroberts@nybg.org

New York Botanical Garden

### Dr. Barbara M. Thiers

Patricia K. Holmgren Director of the William and Lynda Steere Herbarium

Vice President for Administration, Science Division

Editor, *Index Herbariorum*

New York Botanical Garden

### Melissa Tulig

Information Manager for Digitization

New York Botanical Garden

### Kim Watson

Information Manager for Digitization

New York Botanical Garden

## Other Contributors

### Edward Gilbert

Lead Developer for Symbiota

Global Institute of Sustainability (GIOS)

Arizona State University

## Special Thanks to Our Volunteers

Dr. O. R. Anderson

Dr. Maura Flannery

Carl Karp

Joel Siegel



# Contents

---

<b>1</b>	<i>Introduction</i>	1
	Symbiota Software Project . . . . .	1
	Objectives . . . . .	1
	NYBG Projects. . . . .	1
	Overview. . . . .	2
<hr/>		
<b>2</b>	<i>Working in Symbiota</i>	3
	Getting Started . . . . .	3
	Create an Account . . . . .	3
	Crowdsourcing . . . . .	3
	Crowdsource Score Board. . . . .	3
	Record Search Form . . . . .	3
	Setting Up . . . . .	3
	Duplicate Matching . . . . .	3
	Adjusting the Image . . . . .	4
	Going to the Next Record. . . . .	4
<hr/>		
<b>3</b>	<i>Data Entry</i>	5
	Summary Table . . . . .	5
	Fields . . . . .	5
	Collector . . . . .	6
	Associated Collectors. . . . .	6
	Number . . . . .	6
	Date . . . . .	6
	Exsiccati Titles & Numbers . . . . .	6
	Country . . . . .	6
	State/Province . . . . .	6
	County . . . . .	7
	Municipality . . . . .	7
	Locality . . . . .	7
	Cultivated . . . . .	7
	Habitat . . . . .	7

Latitude, Longitude, and Datum . . . . .	7
Elevation and Verbatim Elevation . . . . .	7
UTM . . . . .	7
TRS . . . . .	7
Verbatim Coordinates . . . . .	8
Tips & Additional Details. . . . .	8
Interpreting Unclear Labels. . . . .	8
Other Languages . . . . .	8
<hr/>	
<b>4 Tracking Your Progress</b> . . . . .	<b>9</b>
Volunteers . . . . .	9
User's Status . . . . .	9
Hours . . . . .	9
<hr/>	
<b>5 Examples</b> . . . . .	<b>10</b>
Example 1 . . . . .	10
Example 2 . . . . .	11
<hr/>	
<b>6 References</b> . . . . .	<b>12</b>
Contact Information. . . . .	12
Project-specific Websites. . . . .	12
Digitization Information Websites. . . . .	12
Blogs . . . . .	13
Georeferencing . . . . .	13
Useful Links . . . . .	13
<hr/>	
<i>Appendix A</i> . . . . .	<b>14</b>
Special Characters and Symbols . . . . .	14
Alt Key Code Table . . . . .	14

# 1 Introduction

The purpose of this manual is to provide detailed instructions on entering and editing specimen records in a Symbiota web portal. Although this manual focuses on the New York Botanical Garden's (NYBG) specimen database, the instructions can be used across institutions using the Symbiota software platform.

## Symbiota Software Project

Symbiota is a software platform for creating biodiversity information web portals and communities. Herbaria and other institutions utilize Symbiota's library of webtools to aid biologists in establishing specimen-based virtual floras and faunas in order to understand our world-wide biota.

For more information on the Symbiota Software Project, visit <https://symbiota.org>.

## The William and Lynda Steere Herbarium

The William and Lynda Steere Herbarium of the NYBG is the foundation of the Garden's botanical research program. The Herbarium holds a vast collection of preserved plant specimens filed according to a standardized system of classification. All plant groups -- flowering plants, conifers, ferns, mosses, fungi, lichens, liverworts and algae -- are represented by specimens collected in all parts of the world, but the greatest strength of the Herbarium is the Americas, where The Garden's research has been focused.

Over the last century, The Garden has created one of the most active, best-curated, and most comprehensive herbaria in the world. Containing more than 7,300,000 specimens, it is the fourth largest in the world, the largest in the Western Hemisphere. The Steere herbarium is one of the most frequently used of the approximately 3000 herbaria in the world. The Herbarium is consulted the nearly 200 research projects of Garden scientists and students and an average of 150 visitors travel to New York each year to use the collection, spending on average 1200 person-days per year. Additionally, 30,000 to 50,000 specimens are loaned annually to scientists at other institutions.

## Objectives

The goals of The Herbarium are to make specimen data available electronically for use in biodiversity research projects; to reduce shipping of actual specimens for study; and to reunite data elements (photographs, drawings, manuscripts, gene sequences) derived from each specimen. The digital collections of our virtual herbarium is updated daily as the NYBG pursues the goal of digitizing all of its 7,300,000 plant and fungi specimens. With the help of staff, volunteers, and interns, we plan to accomplish this goal and contribute our efforts to benefit a range of practical applications and research questions.

## NYBG Projects

[The Consortium of North American Bryophyte Herbaria \(CNABH\)](#)

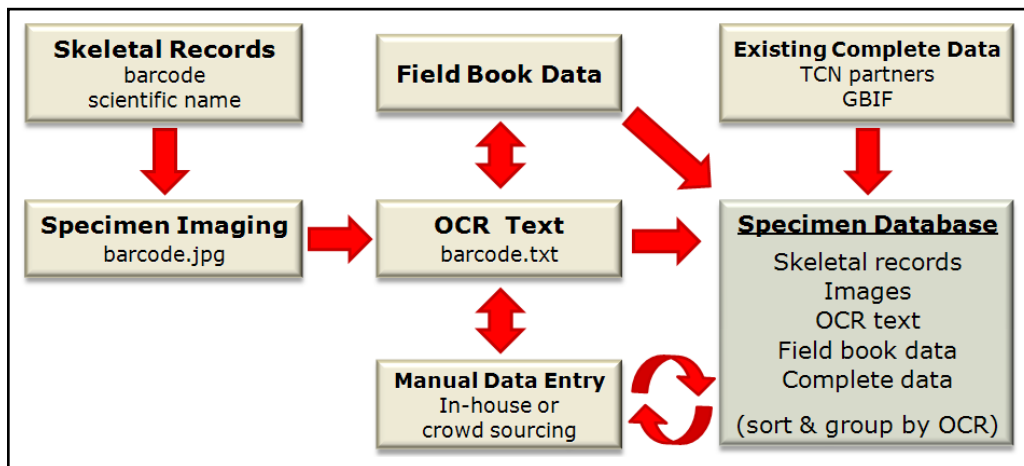
[The Consortium of North American Lichen Herbaria \(CNALH\)](#)

[The Macroalgal Herbarium Consortium](#)

[The Macrofungi Collections Consortium \(MaCC\)](#)

[The Tri Trophic Thematic Collection Network \(TTD-TCN\)](#)

## Overview



*Diagrammatic view of the digitization at the New York Botanical Garden.*

### Digitization Overview

1. Specimens are barcoded, given a scientific name, and filed in the herbarium.
2. Specimens are imaged with DSLR cameras and images are matched with the barcode number.
3. The imaged specimen is uploaded to Symbiota where in-house or crowdsourcing manual data entry occurs.
4. Existing complete data from other institutions and completed data at NYBG are incorporated into Symbiota.
5. Participants can share data with any other project they wish (iDigBio, GBIF).

### Outreach and Training

- Project managers create opportunities for students to learn about botanical collections.
- Coordinate with Volunteer Services to promote opportunities in the herbarium.
- Inform the public about digitization projects by holding tours or workshops, or giving presentations at career fairs and other institutions.
- Interact with the botanical community to ensure that collections data are used for the advantage of research and conservation.

### Our main goal is to:

1. Provide project management.
2. Provide training and documentation.
3. Complete database records.

# 2 Working in Symbiota

The primary use of the Symbiota portal for digitizing institutions will be entering and editing specimen data by capturing label information from an imaged specimen.

Symbiota's data entry tool provides user-friendly functions including auto-completion, duplicate matching, internal and external georeferencing support (Google Maps and GEOLocate), and error checking.

## Getting Started

The minimum requirement for you to start data entry is a Chrome or Mozilla browser.

General guidelines:

- **If a record is too tricky, skip it or ask for assistance.**
- Disable pop-up blocker to allow pop-ups from Symbiota.
- Do not use the back button.
- Use the tab on your keyboard to move from field to field.
- Be sure to save each record before moving forward.
- Once records are saved and you move on to the next record, you may not go back to make edits. Take your time or make notes of any errors made.

## Create an Account

Go to the web portal you wish to work in:

1. Create an account by clicking on **New Account**.
2. Once you have an account, you can get started.

## Crowdsourcing

If you are a volunteer, Symbiota contains crowdsourcing modules per institution to make incomplete records accessible for data entry by the user. Click on the [Crowdsource](#) tab on the [Bryophyte Portal](#) home page.

## Crowdsource Score Board

In the Crowdsource Score Board, you will see 1) Top Scores

of other Users, 2) Current User's Status, and 3) User Stats by Collections.

The score represents the number of records completed by each user.

After completing a record, it will be a Pending point until it is approved by the administrator or your supervisor. The number of records you complete will be tracked by the software.

Under User Stats by Collections, click on the number of **Open Records** for your institution. Then click on the number under **Symbiota ID** to get started.

## Record Search Form

The record search form at the top of the table is where you can specify the type of records to work on.

For example, change the drop down next to **Custom Field 1** if instructed to capture data on a specific genus or region.

## Setting Up

Be sure to use these settings every time:

1. Check the **Auto search** box under **Dupes?**
2. Set the **Status-Auto Set** to **Pending Review**. When saving a record, each completed specimen will automatically save to Pending Review. These will be reviewed by your supervisor.

## Dupes (Duplicate Matching)

When set to auto search, the Symbiota software searches for existing records with the current Collector and collector Number at all institutions with data in Symbiota. This can save you a lot of typing if a duplicate is found at another institution.

The result of the Dupes search will be one of the following:

1. A single collection number match is found. Double-check to make sure all the collection information is indeed the same, and if so "**Transfer to Empty Fields Only**" in order to copy the information from the dupli-

cate record into your new record, and save. Do not click “Transfer All Fields.” This would change the Scientific Name, Catalog Number, and other information that we need to remain the same.

2. Near-duplicate records are found. Such near-duplicates may be previously entered field book records, which would have matching locality data but not the same collection number, or they may be near duplicates from other institutional records (e.g., with collection numbers in a close range when the collector’s number is entered or for the same day if just the date is entered). The user can select a near-duplicate to copy the locality data, update the collection number, and save the record.
3. No duplicate or near duplicate record is found.

## Adjusting the Image

Sometimes, you will need to zoom-in on the collection label, usually in the lower right corner of the herbarium sheet. Hold the Shift key while left clicking the mouse on the area where the label is located (move forward for zoom-in and backward for zoom-out).

You may also change the size of the image box by clicking on the edge of the image frame and dragging outwards or inwards.

## Going to the Next Record

After completing a record, hit the Enter key or press Save Edits. A successful save will generate a message by Action Status, “SUCCESS: edits submitted and activated”. To go to the next record, click the Next Record >> button at the upper right corner of the page (Note: Not to be confused with the Last Record button).

# 3 Data Entry

## Summary Table

Collector	- [Unspecified], [Illegible]
Number	- s.n.
Date	- [s.d.] in the Verbatim Date field
Associated Collectors	- Separate names by semicolon
Exsiccati Title & Number	- Enter <a href="#">Exsiccati title</a> and number.
Country	- Enter if available
State/Province	- Enter if available (Provinces are in Canada)
County	- Enter if available
Municipality	- Only for Mexico. Will specify by “municipio” or “mpio”
Locality	- Enter verbatim. [No loc], [No precise loc]
Elevation (meters)	- Enter the elevation (feet) into the Verbatim Elevation field to convert to meters
Latitude, Longitude	- Enter in decimal degrees or use “Tools” to insert values
Datum	- Enter if provided on label (WGS 84; NAD 27)
UTM	- Use “tools” to insert values and convert to decimal degrees. (UTM 18T 294747 E 4787442N)
TRS	- Enter in the Verbatim Coordinates field (TRS: T1S R3W Sec 13 SE1/4)
Cultivated specimens	- Cultivated check box, Origin:

*A summary table of the data fields in Symbiota. Each field is explained in more detail in this chapter.*

## Fields

A field is the basic element in data entry into which you enter data.

General Guidelines:

- **Do not** change the information in the Scientific Name fields.
- Enter everything **verbatim**. Do not make assumptions. If something is illegible, ask your supervisor or use **square brackets** around the illegible word.
- Do not spend time capturing plant description or habitat description, unless the habitat description is helpful in determining the locality.
- Use the tab key to move from field to field.
- Once data entry is complete, you may hit the Enter key or hit the Save Edits button.

## Collector

The label will indicate who collected the specimen by “Collected by” or “leg”. If this is not shown, a name followed by a number usually indicates the collector and the number assigned to the specimen.

If there is no collector name on the label, enter **[Unspecified]** in the Collector field.

If the Collector name is illegible, enter **[Illegible]** in the Collector field.

If you believe the handwritten name states “Stanley” but you are unsure, put [Stanley] in the Collector field. Square brackets tell us that you made an inference, and these can be looked at later.

Do not worry about the “determined by:” or “det” which identifies who determined the species of the specimen.

Data accuracy is vital to our project so if you are unsure, ask questions or skip the record. For example, if we need to query for all specimens collected by A. A. Rogers, but the determiner was entered (W. A. Murrill) instead, we will not know that this specimen was actually collected by A. A. Rogers.

## Associated Collectors

If there is a list of names, the first name is usually the primary collector and the following names will be put in the Associated Collectors field and separated by a semicolon.

If no associated collectors are shown on the label, leave this field blank.

## Number

If there is no number, enter **s.n.** in the Number field (sine numero, = without a number).

Sometimes numbers have a prefix (e.g., VE-12345) or suffix (e.g., 1839a). If you are cataloging a specimen that has non-numeric characters in the collector number, enter it verbatim.

## Date

Date field: yyyy-mm-dd, yyyy-mm-00 and yyyy-00-00.

Verbatim Date field: [s.d.], mm-dd or range of dates.

If there is no date, enter [s.d.] in the Verbatim Date field.

If only the year is shown, enter yyyy-00-00.

If only the month and day are shown, enter this verbatim in the Verbatim Date field. The Date field requires a year, so you will not be able to type 0000-12-02.

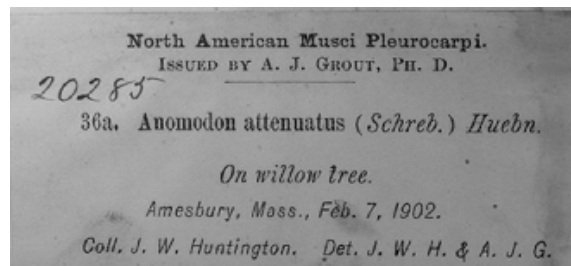
If the date reads 2/3/50, we cannot be certain whether or not the specimen was collected in 1850 or 1950, so enter 2/3/50 into the Verbatim Date field.

For Canadian specimens, sometimes the day and month are switched (i.e. 21-4-1990; 3-4-1998). If you are unsure of which is the day or month, enter it 3-4-1998 in the Verbatim Date field.

If there is more than one collection date, enter the first date into the Date field and all of the dates indicated in the Verbatim Date field.

## Exsiccati Title & Number

An Exsiccati specimen is part of a series or collection of herbarium specimens that are systematically arranged for reference. Many identical copies were made and shipped to different herbaria.



The Exsiccati number is in front of the scientific name (“36a”). Start typing the Exsiccati title and a drop down list will appear (“North American Musci Pleurocarpi”) and make sure the title and the editors match.

For a full list of Exsiccati titles, visit the [Exsiccati Index](#) under the Explore tab in the Bryophyte Portal.

## Country

If the country is not indicated on the label, leave this field blank.

## State/Province

Provinces are for Canada. If the state is not indicated on the



label, leave this field blank.

## County

If the county is not clearly indicated on the label (Co., County, etc.), leave this field blank.

Note: Louisiana has parishes, not counties.

## Municipality

Municipalities are a secondary political subdivision for Mexico. If it is not clearly labeled as a municipality (mpio.), leave this field blank.

## Locality

[No loc]
[No precise loc]

Enter the locality verbatim, aside from the Country, State, County, and Municipality.

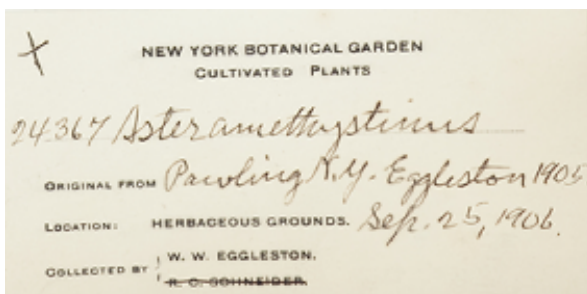
If there is no locality at all, enter **[No loc]** in the locality field.

If there is only information on the Country, State, and County, but no other specific information, enter **[No precise loc]** in the locality field.

If the locality is illegible, enter **[Illegible]** in the locality field. However, you may always do research in Google to decipher difficult handwriting or the spelling of a location.

## Cultivated

If a specimen was cultivated, enter the locality from which the specimen was collected in the locality field. After this, enter "Origin:" followed by where the propagation material originated from.



Country	State/Province	County	Municipality
United States	New York	Bronx	
Locality			
New York Botanical Garden; Herbaceous Grounds; Origin: Pawling, N.Y., Eggleston 1906			

## Habitat

Capture the habitat description (above S-slope of canyon). Habitat information is different from locality information.

## Latitude, Longitude, and Datum

Use these fields to georeference the locality using decimal degree format (e.g., 41.7667, -111.6833).

If the latitude and longitude values are not in degrees, click on the Tools button to convert from formats other than decimal degree.

The geodetic Datum is a coordinate system, and a set of reference points, used to locate places on Earth (i.e. NAD83, NAD27, WGS84).

## Elevation and Verbatim Elevation

Use the former field to record an altitude range in meters (Note: record only the lower range when a single elevation value is cited)

Input altitude in feet into the Verbatim Elevation field and hit the Tab key to convert to meters.

## UTM

Universal Transverse Mercator: Click Insert UTM Values to get the latitude and longitude degrees.

UTM 18T 294658E, 4787442N

Zone:	<input type="text" value="18"/>
East:	<input type="text" value="294658"/>
North:	<input type="text" value="478744"/>
Hemisphere:	<input type="text" value="North"/>
<input type="button" value="Insert UTM Values"/>	

## TRS

Township Range Section: Enter the TRS values verbatim in the Verbatim Coordinates.

Location and Remarks: RIVERSIDE TWP. (SE 1/4 Sec. 2, T. 13, R. 9W). Along roadside, in small stream bottoms. Back gray

## Verbatim Coordinates

If coordinates are confusing or do not fit into the latitude and longitude, UTM, or TRS categories, you may enter the coordinates verbatim in the Verbatim Coordinates.

## Tips & Additional Details

### Interpreting Unclear Labels

Frequently, handwritten labels may be difficult to read. You may use Google to search for tricky localities or ask a supervisor for assistance. If you are unable to interpret a label, skip it.

### Other Languages

Enter the locality information verbatim. Do not translate the label. Use the appropriate Alt key codes in Appendix A.

# 4 Tracking Your Progress

## Volunteers

### Current User's Status

Current Standing
Specimens processed: 3
Pending points: 0
Approved points: 6 ( <a href="#">view records</a> )
Total Possible Score: 6

### Top Scores

User	Score
Karp, Carl	742
Siegel, Joel	536
Logan, Carolyn	142

Under Current User's Stats, you will find the Specimens processed, Pending points, Approved points, and Total Possible Score.

Specimens processed are the number of records you completed. Each successfully completed record is worth 2 points.

Pending points are the points you have pending before they are approved.

Approved points are the points approved.

The Total possible score is the maximum score you may achieve while crowdsourcing each record. If a record is incomplete, you will only be rewarded 1 point instead of 2 points.

## Hours

Keep track of the number of hours you work on records. For example, at NYBG, volunteers fill out a Google doc that can be accessed from any computer with an internet connection. This is useful when volunteers work from home.

Each month, a summary of your hours will be submitted to Volunteer Services to go towards your benefits.

Tracking volunteer, intern, and employee progress is important for making digitization as efficient as possible. We need to know what has been done, who did the work, and the outcomes for the overall projects.

# 5 Examples

## Example 1

IDAHO

Brachythecium rivulare  
forna

Marsh, Yuba River. Atlanta District.

Elmore County. Boise National Forest.

Fay A. MacFadden 19200      1 Oct. 1942  
det. A. J. Grout

Long Form <<

Collector ?      Number ?      Date ?      Dupes?       Auto search

Fay A. MacFadden      19200      1942-10-01

Associated Collectors ?      Verbatim Date ?

Exsiccati Title      Number

Scientific Name ?  
Brachythecium rivulare

Country      State/Province      County

United States      Idaho      Elmore

Locality  
Boise National Forest; Atlanta District; Yuba River.

Latitude      Longitude      Uncertainty ?      Verbatim Coordinates      Tools

Elevation in Meters      Verbatim Elevation

Habitat  
Marsh

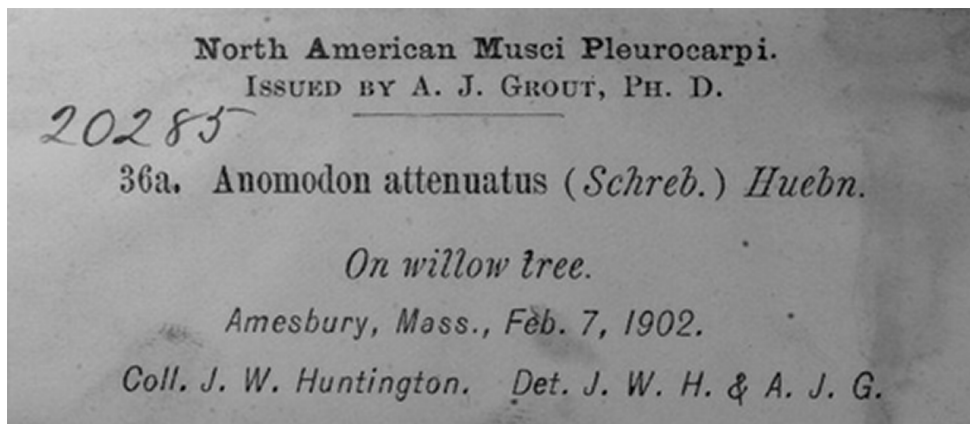
Substrate

Notes

Save Edits

Status Auto-Set: Pending Review ▼

## Example 2



Long Form <<

Collector ?  Number ?  Date ?  Dupes?  Auto search

Associated Collectors ?  Verbatim Date ?

Exsiccati Title  Number

Scientific Name ?

Country  State/Province  County

Locality

Latitude  Longitude  Uncertainty ?  << Verbatim Coordinates  Tools

Elevation in Meters  -  Verbatim Elevation  <<

Habitat

Substrate

Notes

Status Auto-Set:  ▼

# 6 References

## Contact Information

For questions about digitization, equipment, procedures, or public outreach: Contact Mari Roberts.

**Mari A. Roberts**

Volunteer Coordinator

Project Manager for TTD-TCN

mroberts@nybg.org

(718) 817-8878

The New York Botanical Garden

## Project-specific Websites

**The Tri-Trophic Thematic Collection Network (TTD-TCN)**

<http://tcn.amnh.org/>

**The Macrofungi Collections Consortium (MaCC)**

<https://sites.google.com/site/macrongicollectionconsortium/>

**The Lichens, Bryophytes and Climate Change (LBCC)**

<http://lbcc1.acis.ufl.edu/>

**The Consortium of North American Lichen Herbaria (CNALH)**

<http://lichenportal.org/portal/>

**The Consortium of North American Bryophyte Herbaria (CNABH)**

<http://bryophyteportal.org/portal>

**The Macroalgal Herbarium Consortium**

<http://macroalgae.org/portal>

## Digitization Information Websites

### Symbiota

A library of webtools to aid biologists in establishing specimen-based virtual floras and faunas.

<http://symbiota.org>

### Symbiota Help Pages

Tutorials describing features and management tools.

<http://symbiota.org/docs/symbiota-introduction/symbiota-help-pages/>

### KE Software - EMu

A collections management system for all museums.

<http://www.kesoftware.com/>

### Integrated Digitized Biollections (iDigBio)

The National Resource for Advancing Digitization of Biodiversity Collections (ADBC) funded by the National Science Foundation (NSF)

<https://www.idigbio.org/>

### Global Biodiversity Information Facility (GBIF)

Internationally funded open data infrastructure that focuses on making scientific data on biodiversity available online.

<http://www.gbif.org/>

### Discover Life

Information about taxonomy, natural history, distribution, abundance, and ecology of many species worldwide.

<http://www.discoverlife.org/>

## Blogs

### iDigBio Blog

Check here for blogs about techniques, participants, and upcoming workshops and other events.

<http://idigbio.org/tag/blogs>

### NYBG Science Talk Blog

A blog exploring the science of plants, from the field to the lab.

<http://blogs.nybg.org/science-talk>

## Georeferencing

### GEOLocate

A platform of georeferencing natural history collections data.

<http://www.museum.tulane.edu/geolocate/>

### BioGeomancer

A project developing online tools, web services, and desktop applications for georeferencing natural history specimens.

<https://sites.google.com/site/biogeomancerworkbench/home>

### GBIF's Guide to Best Practices for Georeferencing

<http://www.gbif.org/communications/resources/print-and-online-resources/download-publications/bookelets>

### MaNIS Georeferencing Guidelines

<http://manisnet.org/GeorefGuide.html>

## Useful Links

### Exsiccati Index

List of Exsiccati Titles.

<http://bryophyteportal.org/portal/collections/exsiccati/index.php>

### Statoids

Worldwide reference for states, provinces, governorates, counties, districts, and other subdivisions of countries.

<http://www.statoids.com/>

### Tropicos

Online botanical database containing taxonomic information on plants.

<http://www.tropicos.org/>

# Appendix A

## Special Characters and Symbols

Special characters or symbols can be typed by pressing and holding the Alt key while typing the number identifying the character with the keyboard's numeric keypad.

### Alt Key Code Table

0128	€	0160		0192	À	0224	à
0129		0161	ı	0193	Á	0225	á
0130	,	0162	¢	0194	Â	0226	â
0131	f	0163	£	0195	Ã	0227	ã
0132	„	0164	¤	0196	Ä	0228	ä
0133	…	0165	¥	0197	Å	0229	å
0134	†	0166	¦	0198	Æ	0230	æ
0135	‡	0167	§	0199	Ç	0231	ç
0136	^	0168	¨	0200	È	0232	è
0137	‰	0169	©	0201	É	0233	é
0138	Š	0170	*	0202	Ê	0234	ê
0139	<	0171	«	0203	Ë	0235	ë
0140	Œ	0172	¬	0204	Ì	0236	ì
0141		0173	-	0205	Í	0237	í
0142		0174	®	0206	Î	0238	î
0143		0175	¯	0207	Ï	0239	ï
0144		0176	°	0208	Ð	0240	ð
0145	´	0177	±	0209	Ñ	0241	ñ
0146	¸	0178	²	0210	Ò	0242	ò
0147	“	0179	³	0211	Ó	0243	ó
0148	”	0180	´	0212	Ô	0244	ô
0149		0181	µ	0213	Õ	0245	õ
0150	–	0182	¶	0214	Ö	0246	ö
0151	—	0183	·	0215	×	0247	÷
0152	˘	0184	¸	0216	Ø	0248	ø
0153	™	0185	ı	0217	Ù	0249	ù
0154	š	0186	°	0218	Ú	0250	ú
0155	›	0187	»	0219	Û	0251	û
0156	œ	0188	¼	0220	Ü	0252	ü
0157		0189	½	0221	Ý	0253	ý
0158		0190	¾	0222	Þ	0254	þ
0159	ÿ	0191	¿	0223	ß	0255	ÿ



## **DIGITIZING FOSSILS TO ENABLE NEW SYNTHESSES IN BIOGEOGRAPHY- CREATING A PALEONICHES**

Report submitted by: blieber@ku.edu  
Report Submitted on: 07/06/2015 - 13:47

### **Progress in Digitization Efforts**

Paleoniches Update, July 2015

Regarding the University of Kansas portion of the project, led by PI Bruce S. Lieberman and co-PI Una Farrell, we now have a total of 185,021 specimens databased. Further, we now have a total of 144,692 specimens that are georeferenced. In addition, a total of 7,773 localities have been georeferenced (as mentioned previously, thus the georeferencing component of our proposed work is completed). Since the last update to iDigBio, we have now completed databasing all of our targeted gastropods and are almost finished databasing our bivalves. We have also completed imaging of all nautiloid and gastropod species, and are making good progress on bivalves.

Regarding the Ohio University portion of the project, led by PI Alycia Stigall

The primary objective at Ohio University this quarter is the preparation of our data for deployment within the Digital Atlas app—which is under development by the KU group. At present data for 183 species have been prepared for inclusion within the iphone/ipad application. All species will have photographs, descriptions, and georeferenced locality maps present in the app. Furthermore, additional 3D models of select species were added to the Digital Atlas of Ordovician Life website and digitization efforts have continued for the newly acquired Stocker Collection.

Cincinnati Museum Center

Georeferencing and specimen photography work has continued at a slower pace during the summer session as our academic year interns are away from Cincinnati. We are progressing in the objective of preparing data for ingestion with iDigBio, and expect to achieve that objective during the next quarter. At this point, over 29,000 specimens have been georeferenced, which represents more than 42% of the total cataloged invertebrate paleontology collection.

Miami University

Progress continues with georeferencing as well as relocating specimens due to the renovations currently ongoing at Miami University.

Regarding the San José State University portion of the project, led by PI Jon Hendricks:

Since the last update, the Neogene and Pennsylvanian Atlases of Ancient Life have continued to grow and more content has become available online.

Neogene Atlas: Five families of turrid gastropods have been recently added to the Neogene Atlas. At the current time, species-level pages for 360 taxa are now accessible online. In early June, Florida Museum of Natural History (FLMNH) staff sent PI Hendricks the remaining files needed to put the rest of the planned 500 Neogene species online. These will be gradually added to the Neogene Atlas over the course of this summer. A second focus of the Neogene Atlas work this summer will be the construction of occurrence maps for the backlog of Neogene species that currently lack maps. A SJSU undergraduate has been trained in using GIS to make maps and she will begin to address this backlog in the near future.

Pennsylvanian Atlas: Substantial progress has been made on adding species-level pages to the Pennsylvanian Atlas. Since the last update, cephalopods and bivalves have been added and pages for 254 species are now online. Additionally, occurrence maps have now been generated and put online for many of these species. The final group that needs to be added to the Pennsylvanian Atlas is the Gastropoda. We expect to be finished adding species pages to the Pennsylvanian Atlas by the time of the next update.

Finally, for our PEN partners. First, Texas, PI: Ann Molineux, Co-PI: James Sprinkle

They have been continuing their work in this area but received a serious shock in June pertaining to the fact that Ann has just been diagnosed with some very serious health issues. She is heading down to MD Anderson to deal with a potentially life threatening diagnosis. We all wish her the very best and I did not want to bother her with a request for an update at this time.

And at Yale: From PI Susan Butts:

They have recently been approved for a no cost extension. Progress is continuing apace.

### **Share and Identify Best Practices and Standards (including Lessons Learned)**

N/A

### **Identify Gaps in Digitization Areas and Technology**

N/A

### **Share and Identify Opportunities to Enhance Training Efforts**

The University of Kansas has recruited a new graduate student who will begin working on this project in August 2015.

### **Share and Identify Collaborations with other TCNs, Institutions, and Organizations**

N/A

### **Share and Identify Opportunities and Strategies for Sustainability**

N/A

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

A manuscript by PI's Hendricks, Stigall, and Lieberman—titled "The Digital Atlas of Ancient Life: delivering information on paleontology and biogeography via the web" was accepted at Palaeontologia Electronica (an open-access online journal) in June of 2015 and is now in press. This manuscript provides an overview of our Digital Atlas project and goals for the paleontological community.

### **Attachment**

N/A

## **GREAT LAKES INVASIVES: DOCUMENTING THE OCCURRENCE THROUGH SPACE AND TIME OF AQUATIC NON-INDIGENOUS FISH, MOLLUSKS, ALGAE, AND PLANTS THREATENING NORTH AMERICA'S GREAT LAKES**

Report submitted by: kmcameron@wisc.edu  
Report Submitted on: 07/06/2015 - 16:39

### **Progress in Digitization Efforts**

See attached pdf

### **Share and Identify Best Practices and Standards (including Lessons Learned)**

See attached pdf

### **Identify Gaps in Digitization Areas and Technology**

See attached pdf

### **Share and Identify Opportunities to Enhance Training Efforts**

See attached pdf

### **Share and Identify Collaborations with other TCNs, Institutions, and Organizations**

See attached pdf

### **Share and Identify Opportunities and Strategies for Sustainability**

See attached pdf

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

See attached pdf

### **Attachment**

[https://www.idigbio.org/sites/default/files/webform/tcn-reports/GLITCN\\_progress%20report\\_July2015.pdf](https://www.idigbio.org/sites/default/files/webform/tcn-reports/GLITCN_progress%20report_July2015.pdf)

## GREAT LAKES INVASIVES TCN – Bi-monthly report **May 1 – June 30, 2015**

*Fourth GLI TCN report, representing ca. ten months' of effort to date & first full year of grant period.*

Our four regional data processing centers (NY Botanical Garden, Field Museum, Univ of Michigan, and Univ of Wisconsin-Madison) report the following from their constituents:

### 1) Progress in Digitization Efforts TO DATE

#### PLANTS:

- Specimens Barcoded Only (not photographed yet): 600 (WIS) + 33,100 (MIN) + 24,462 (NY) + 18,189 (ILLS) + 1424 (UWM) = **77,775**
- Imaged only AND image uploaded to a portal (i.e., no data record yet): 14,053 (MIN) + 6608 (OSU) = **20,661**
- Imaged only but image not yet uploaded to a portal: 6,198 (MIN) + 59,959 (NY) + 18,332 (F) + 10,000 (MICH) + 4,880 (MU) = **99,369**
- Databased only (skeletal or complete record) AND data uploaded to a portal (i.e., but not imaged yet): 6613 (MOR) = **6,613**
- Databased only but not yet uploaded to a portal: 53,351 (MIN) + 111,081 (NY) + 23,565 (F) + 27,000 (ILLS) + 9624 (MOR) = **224,621**
- Both Image AND a Data Record Uploaded to iDigBio, to the GLI Portal directly or to another Symbiota Portal for editing before transfer to GLI Portal: 62444 (WIS) + 6342 (MSC) + 520 (ILLS) + 3617 (MOR) + 1474 (UWM) + 14,900 (MICH) = **89,297**

**PLANT IMAGING SUMMARY: At least 209,327 images taken. Target stated in grant proposal is 637,000. Imaging goal is 33% complete.**

## **FISH:**

- Specimens Barcoded Only (not photographed yet): **605** (MIN)
- Imaged only AND image uploaded to a portal (i.e., no data record yet): 0
- Imaged only but image not yet uploaded to a portal: 400 (MIN) + 610 lots (F) + 117 (MICH: UMMZ) = **1,190**
- Databased only (skeletal or complete record) AND data uploaded to a portal (i.e., but not imaged yet): **24,000** (ILLS)
- Databased only but not yet uploaded to a portal: 1365 lots (MIN) + 4709 lots/81,324 (F) + The 200,000 UMMZ fish database is complete, but data are currently only being uploaded to the portal when the corresponding images are ready (MICH: UMMZ) = **>6074** lots
- Both Image AND a Data Record Uploaded to iDigBio, to the GLI Portal directly or to another Symbiota Portal for editing before transfer to GLI Portal: 505 (MIN) + 63 (UMMZ)+ 2284 (OSU) = **2,789**

**FISH IMAGING SUMMARY: Five institutions making progress so far. At least 3,979 images have been taken. Target stated in grant proposal is 102,000 lots. Imaging goal is 4% complete.**

## **MOLLUSKS:**

- Specimens Barcoded Only (not photographed yet): 0
- Imaged only AND image uploaded to a portal (i.e., no data record yet): 0
- Imaged only but image not yet uploaded to a portal: **2,645 lots** (UMMZ)
- Databased only (skeletal or complete record) AND data uploaded to a portal (i.e., but not imaged yet): **2,000** (ILLS)
- Databased only but not yet uploaded to a portal: **15,668** (UMMZ)
- Both Image AND a Data Record Uploaded to iDigBio, to the GLI Portal directly or to another Symbiota Portal for editing before transfer to GLI Portal: **855** (UMMZ)

**MOLLUSK IMAGING SUMMARY: Two institutions making progress so far. At least 4,355 images taken. Target stated in grant proposal is 44,000 lots. Imaging goal is 10% complete.**

## **2) Share and Identify Best Practices and Standards (including Lessons Learned)**

From UWM: After experiencing some trouble with autofocus (i.e., autofocus failed to find the focal plane, which was only detectible after opening the image; this lead to time wasted as we deleted the out-of-focus photo, used live view to get an in-focus image...), we found the fastest way for us to image using EOS Utility and Great Lakes TCN Workflow:

1. Open EOS Utility (and Great Lakes TCN Workflow, new session, etc.), open **Live View Shoot** window.
2. Move focus frame to part of specimen, turn **Live Mode ON**, take photo when frame turns green (no need to check each photo for focus now!).
3. In Great Lakes TCN Workflow, enter barcode
4. Enter taxon name...
5. Press Enter
6. Back to step 2.

From Univ of Michigan: "As suggested by OSU, we have been working in pairs when imaging fish, and have found the throughput to be much faster than that of a single technician working alone."

From several herbarium partners: The rate of processing that was stated in the grant proposal (60 sheets per hour) is too high, especially if skeletal records are being created at the same time as imaging. A rate of ca. 40 sheets per hour has been more realistic. As a result, funds are being depleted before all specimens can be digitized.

## **3) Identify Gaps in Digitization Areas and Technology**

Nothing to report.

## **4) Share and Identify Opportunities to Enhance Training Efforts**

Nothing to report.

## **5) Share and Identify Collaborations with other TCNs, Institutions, and Organizations**

The New York State Museum now has its portable imaging station in order to image its respective specimens. They should be contributing data soon.

## **6) Share and Identify Opportunities and Strategies for Sustainability**

Nothing to report.

## **7) Other Progress (that doesn't fit into the above categories)**

A new imaging station was added to UMMZ's Mollusk lineup. Several participants attend the SPNHC meeting. Papers were delivered. Ken Cameron will deliver a paper at the Botany meeting in Edmonton, Canada later this month.

## PLANTS, HERBIVORES AND PARASITOIDS: A MODEL SYSTEM FOR THE STUDY OF TRI-TROPHIC ASSOCIATIONS

Report submitted by: moon@begoniasociety.org  
Report Submitted on: 07/06/2015 - 19:09

### Progress in Digitization Efforts

#### Overall Summary:

Our TCN project has requested a 6 month extension, with a deadline for our project as January 2016. The majority of specimen digitization is over, with subcontracts finishing the remainder of their funds over the summer. The remaining time in the grant will be spent submitting all records to iDigBio, improving data quality through data cleaning and georeferencing, and publication of our dataset in a data publication journal.

#### Plant Records:

The (28 total) NYBG volunteers have transcribed 19,461 complete records since January 2014.  
Total complete records transcribed from all institutions = 208,693  
Total complete records from all institutions = 1,236,138

Two volunteers have received training by our georeferencing staff and have been practicing on NY localities. We anticipate them working solely in the Symbiota Batch Georeferencing Tool by the end of the summer.

Melissa Tulig created an IPT for all NYBG vascular records, including TTD and GLI for iDigBio. The records are in the queue to be harvested by iDigBio but Joanna told us that her team is on vacation as of last week. All of the records have been indexed by GBIF. This includes 1.4 million records from NYBG, including 489,000 that relate to TTD and/or GLI. We submitted links 930,000 images that correspond to these records, including 332,000 that relate to TTD and/or GLI records.

#### Insect Records:

1,151,424 specimen records transcribed.

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

Our focus is now on making partners independent for supplying data to iDigBio. We are in the process of moving collections to SEINet/Midwest herbaria portal and they can set up an RSS feed to iDigBio directly from there.

### Identify Gaps in Digitization Areas and Technology

Nothing to report.

### Share and Identify Opportunities to Enhance Training Efforts

#### Volunteer Events at NYBG:

On June 9th, we held a transcription event for 6 volunteers from Goldman Sachs. They transcribed 488 complete records and imaged 1012 specimens. Starting on July 7, we will have 1 full-time SYEP participant who will be working on TTD records. Starting on July 15, we will have (4-6) ECAG volunteers working on TTD records once a week.

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

Katja Seltmann assisting in the development of the iDigBio Data Management workshop.

### Share and Identify Opportunities and Strategies for Sustainability

Nothing to report.

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

Nothing to report.

### Attachment

N/A

## **INVERTNET: AN INTEGRATIVE PLATFORM FOR RESEARCH ON ENVIRONMENTAL CHANGE, SPECIES DISCOVERY AND IDENTIFICATION**

Report submitted by: chdietri@illinois.edu  
Report Submitted on: 07/07/2015 - 08:57

### **Progress in Digitization Efforts**

The InvertNet technical team completed the transfer of responsibilities for managing the cyberinfrastructure and image ingest workflows to permanent INHS IT staff (one full-time biodiversity informatician and one student hourly). This should provide stability and sustainability moving forward. Work is nearing completion on installation of a tape archiving system to provide additional backup of raw and stitched images. Personnel at INHS and collaborating institutions continued capturing whole drawer images and submitting these to INHS for ingest.

### **Share and Identify Best Practices and Standards (including Lessons Learned)**

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.

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

Nothing to report.

### **Attachment**

N/A



## SERNEC: THE KEY TO THE CABINETS: BUILDING AND SUSTAINING A RESEARCH DATABASE FOR A GLOBAL BIODIVERSITY HOTSPOT

Report submitted by: michael.denslow@gmail.com

Report Submitted on: 07/07/2015 - 09:53

### Progress in Digitization Efforts

#### All SERNEC:

The SERNEC portal (<http://sernecportal.org/>) currently has 32 institutions hosted. The portal now hosts 657,698 specimen records. There are currently 174,744 skeletal records in the SERNEC portal.

We have continued to add new SERNEC images to Notes From Nature to be transcribed. In May, 15,769 images from Southeast Louisiana State University were added to the system. We post blog content to the Notes From Nature blog in order to communicate information about the new images and issues that deal with their transcription (<http://blog.notesfromnature.org/>).

We have begun development of the GEOLocate Symbiota integration. This integration will make the collaborative georeferencing and data return to the individual collections more efficient. The basic workflow of how the two applications will interact has been mapped out. The GeoLocate API that ingests the specifically formatted Darwin Core Archive files from Symbiota has been completed. In addition, a GeoLocate publishing module that outputs DwC-A files with georeferencing results has also been completed.

#### Arkansas:

APCR: The imaging station setup is complete and up and running. At this point, over 1000 barcodes have been applied, 206 skeletal records have been created, and 206 images have been taken. Most of the time spent has been in pre-digitization curation. This has been done by family, using the Atlas of the Arkansas Flora and "Weakley's Flora" as a guide for taxon recognition/correct nomenclature. When student labor starts on 1 July 2015 we will be ready to go by genus/species in each family. Student labor is currently scheduled for the month of July to assist with image/database/upload images. Barcodes have been ordered and delivered for the following Arkansas herbaria: HXC, UCA, ANHC, HEND. Application of barcodes is just beginning at those collections.

#### Georgia:

GA: 20,500 GA specimens were imaged during this time period (29,900 to date); 19,000 GA images were uploaded to the iPlant servers. Server expansion chassis + 8 2-TB drives ordered (with supplement from Plant Biology department) to expand current back-up system. WGC specimens were returned to the University of West Georgia by GA personnel on 8 June 2015; the 5,009 WGC images were uploaded to iPlant by S. Hughes (GA).

VSC: SERNEC Symbiota profiles have been created for VSC and COLG, and Richard Carter (VSC) has ordered barcode labels for COLG. COLG personnel (curator Kevin Burgess and graduate student Samantha Worthy) visited VSC on 18-19 May 2015; they delivered 1,000 COLG specimens to VSC for imaging and received imaging training. Richard Carter (VSC) also contacted Stephanie Harvey (GSW), and imaging at GSW will be starting soon. The plan now is for VSC to contract her student assistant(s) to take the images and enter minimal skeletal data (binomial on the folder) VSC will be purchasing barcodes for GSW.

GAS: The imaging station has been ordered (Photo-eBox set-up, camera, LaCie 8TB USB 2-bay RAID) with supplement from department and institute.

#### Kentucky:

MDKY: The Photo-eBox is supposed to be delivered on 1 July 2015; then our imaging station will be complete. It will likely take a few days to get everything up and running smoothly after Photo-eBox delivery. Two students have been hired and are ready to work the month of July.

#### Louisiana:

We were able to transfer all of the images (15,769) from Southeastern Louisiana University in the CyberFlora Louisiana collection to Notes from Nature for label digitization. More images will be supplied as needed. There are currently 900,000 in the queue. We are awaiting the return of label data to start some testing of GeoLocate throughput.

#### Mississippi:

Legacy databases for MS herbaria have been loaded into Symbiota. The graduate student working on the SERNEC project has been transcribing labels from the IBE collection and recently began georeferencing records in the MSU collection.

Virginia:

GMUF: Imaging and data processing workflows are now fully operational. Between May 1 and June 31, 5,533 specimens were imaged and their skeletal data uploaded to the SERNEC portal. JPEG and DNG file uploads to the iPlant server are ongoing. Approximately 10,000 additional sheets have been barcoded. On May 26th, four undergraduate students began working ca. 9-12 hrs/wk. All have been trained on the barcoding and imaging workflows and are completing RCR certification. A Ph.D. curatorial assistant, employed on a different grant, began working in early May and has done pre-curation steps across ca. 80% of the Virginia-only collection (ca. 30,000 sheets) ahead of undergraduate imagers. These pre-curation steps include removing newsprint bifolds, repairing specimens, mounting duplicate material, and cross-referencing nomenclature on the folders. On June 19, Erika Gonzalez from FARM visited GMUF and received training on all the workflows. Summer imaging will conclude July 24 and resume September 1.

FARM: Imaging and data processing workflows are partially operational. Two undergraduate students have been hired and are receiving training on pre-curation and imaging workflows.

VPI: The Photo-eBox has been ordered, and the curator, Tom Weiboldt has agreed to begin imaging during the Fall Semester.

West Virginia:

Between February 1 and May 1, 2015, Marshall (MUHW) photographed ~8,400 specimens, transcribed ~400 specimens and georeferenced ~150 specimens. We hired two student workers during Spring 2015 and had another eight working via the Federal Work Study Program at Marshall University. All equipment has been purchased for Marshall. West Virginia University (WVA) now has their subcontract in place and is in the process of hiring students and purchasing equipment.

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

All SERNEC:

We have completed several pieces of documentation for the project. These are currently shared publicly here: <http://sernec.appstate.edu/resources> for anyone to access. These documents include the equipment list for the project, as well as protocols for barcoding, image archive, image processing, image capture, image station set up and skeletal data entry. We continue to refine and edit these documents as they get utilized by our partner institutions.

Some collections have written or modified specific protocols to suite their local needs (e.g., Gillespie at MUHW). These are being collected and posted in the SERNEC website as well.

The states of Arkansas, Mississippi, North Carolina, South Carolina, and Virginia have begun to collaborate with local stakeholders to develop a method for selection and a list of sensitive species that will be automatically redacted from public view.

Arkansas:

APCR: It is suggested that barcodes be applied at the bottom of the herbarium sheet if at all possible. This makes scanning the code easier when creating a skeletal record, especially when you are repeating data from the previous specimen (except for the barcode). It is also suggested that when you image specimens you move specimens from the right stack to the left stack and leave them in that order. Then when you rename the files, you reverse stack the specimens back into the order they were in to begin with. It is also suggested that when you rename the files, start with the last image in the window on the screen and progress backward. As files are renamed an image jumps to the front of the line (and out of the way), and the specimens and the images are in the same sequence and the next image to be renamed corresponds to the top sheet in the left stack.

Mississippi:

For anyone dealing with a legacy database, it is strongly encouraged that you link the image with the database record at the time of imaging. For converting label databases to a new format (e.g., Symbiota), make sure to clean the records before uploading and make sure that records are as up-to-date as possible before doing an upload.

Virginia:

Andrea Weeks (GMUF) has an Adobe Illustrator template for a 6 cm ruler with mm markings that can be shared freely - thanks to Travis Marsico. With this template, I have made a ruler with the GMUF logo on it and include it as part of the specimen image scale-bar. STAR

and UARK also have similar. Also, I have made a customized platen for the Photo-eBox to precisely position the herbarium specimen for imaging, modeled on the one used at UFI. This consists of a foam-core base that fits tightly within the Photo-eBox mounted with two strips of foam-core, one along the back wall and one along right wall to form a raised "L". The foam core platen is covered in black vinyl contact paper. The ruler and color chip are mounted permanently on the right-wall foam core strip. Students slide specimens into the Photo-eBox with the label towards the corner of the "L".

## Identify Gaps in Digitization Areas and Technology

All SERNEC:

Nothing to report

Mississippi:

Linking image files with legacy database records is a slow process. It would have been much faster to do this during imaging.

Virginia:

At GMUF, Andrea Weeks has been in correspondence with Herrick Brown about improvements to the Symbiota system regarding: 1) clarifying to users that the first row of data during Excel upload is expected to be field headers, otherwise it is lost; 2) adding customizable field search criteria to return genus-only database entries; 3) providing authorities in database records for varietal autonyms; and 4) improving time-stamp reliability during use of the skeletal data portal to track student productivity. I have also been in contact the GMU College of Science Director of Information Technology to obtain ca. 3-5 Tb permanent archival storage for all GMUF images (CR2, DNG, and JPG) on University servers. Preliminary conversations suggest I will be allotted space soon.

## Share and Identify Opportunities to Enhance Training Efforts

All SERNEC:

A special workshop on Symbiota software was held at the SPNHC meeting in Gainesville. It was attended by members of the SERNEC – TCN as well as iDigBio and other active TCNs.

During year 1, the project has provided mentoring for 5 graduate student, 43 undergraduates, 2 postbaccalaureates, and 1 technician (collections manager) and 1 postdoctoral student -- all trained in museum protocols. In addition, several individuals have gained valuable supervisory experience and exposure to other herbaria and researchers within their states. Responsible Conduct of Research (RCR) training has been offered at each institution. In addition, the lead institution (Appalachian State University) offers this training as a centralized resource for those that do not have access to it from their institutions.

Arkansas:

APCR: I have invited the curators of the other collections that I am responsible for to come to APCR during July, bring their students, and bring a folder of specimens. They can work with the imaging station and get somewhat comfortable with its operation. This would also apply to any colleagues in other "ologies" they have at their institutions.

Virginia:

At GMUF, my students appreciate the specimen-per-hour metric on the SERNEC skeletal data entry site and have built a friendly rivalry over who can top 100/hr. Also, I have printed pictures of how students should arrange windows across the two monitors and posted these by the workstation. This is in addition to the customized imaging protocol I have based on E. Gillespie's document.

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

All SERNEC:

A special workshop on Symbiota software was held at the SPNHC meeting in Gainesville. It was attended by member of the SERNEC – TCN as well as iDigBio and other active TCNs.

The states of Arkansas, Mississippi, North Carolina, South Carolina, and Virginia have begun to collaborate with local stakeholders to develop a method and list of sensitive species that will be automatically redacted from public view.

Mississippi:

Digitization of MS herbaria is a collaboration supported by a CSBR grant. The data from this project are provided to the SERNEC project.

Virginia:

At GMUF during June 2015, I have been in contact with the State Botanist at Virginia Natural Heritage, Johnny Townsend, to augment the SERNEC-wide specimen redacted list. Both he and the State Ecologist, Gary Fleming, are of the opinion to err on the side of sharing data rather than obscuring taxa in need of study/discovery. I have received the Excel spreadsheet of state rare and threatened vascular plant species. Of these species, 11 are listed as State Threatened and will be added to the SERNEC-list, which already lists Federally-listed species as well as species typically over-collected in Virginia, such as ginseng and goldenseal.

### Share and Identify Opportunities and Strategies for Sustainability

All SERNEC:

Nothing to report.

Louisiana:

Tom Sasek was able to persuade the ULM administration to release some previously promised overhead funds to replace the CyberFlora Louisiana central server and RAID backup at LSU. This will protect the database of images for several more years until it is ready to be uploaded to other databases.

Virginia:

At GMUF, I will hire the same summer undergraduate student workers in the fall (Sept 1) and have them train my Independent Study students, who will complete imaging for credit during the Fall semester.

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

All SERNEC:

Nothing to report.

### Attachment

N/A

## NORTH AMERICAN LICHENS AND BRYOPHYTES: SENSITIVE INDICATORS OF ENVIRONMENTAL QUALITY AND CHANGE

Report submitted by: cgries@wisc.edu  
Report Submitted on: 07/07/2015 - 11:17

### Progress in Digitization Efforts

As of June 2015 the number for the LBCC are as follows:

Lichens:

<http://lichenportal.org>

Herbaria actively submitting images or key stroked records to the portal: 75

Specimen records in portal: 1,865,622 (up by 89284 since May 2015)

Specimen records with images: 689,930 (25,555 labels have been imaged since May 2015)

Records with locality information: 1,605,365 (83,929 locality information where added since May 2015)

Bryophytes:

<http://bryophyteportal.org>

Herbaria actively submitting images or key stroked records to the portal: 65

Specimen records in portal: 2,248,833 (up by 156,560 since May 2015)

Specimen records with images: 1,080,391 (92,269 labels have been imaged since May 2015)

Records with locality information: 1,430,069 (92,940 locality information where added since May 2015)

The two consortia together now have 86 members. This includes collections that submitted their data without help from the project from the US (including two private collections and a botany club) and from other countries (8 collections from Canada, Norway, Sweden, and Denmark), and 10 collections that were part of one of the PEN grants. 53 collections have both lichens and bryophytes, 12 only Bryophytes and 22 only lichens. Some changes from the original proposal have occurred as some collections were not able to participate while others that were not listed in the proposal have joined the effort.

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

none to report

### Identify Gaps in Digitization Areas and Technology

Overall automation of transcription needs to be further improved, e.g., OCR, NLP, and Geo-referencing still need more effort and development. Data quality improvements need to be returned to the original collections.

### Share and Identify Opportunities to Enhance Training Efforts

a lot is being done in this area.

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

### Share and Identify Opportunities and Strategies for Sustainability

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

### Attachment

<https://www.idigbio.org/sites/default/files/webform/tcn-reports/LBCCsummaryReport%202015.pdf>

## Digitization TCN Collaborative Research: North American Lichens and Bryophytes: Sensitive Indicators of Environmental Quality and Change

### Overall Project Summary report 2014 - 2015:

This marks the final year of the project for several collaborating institution, however, most have applied for and received a no-cost extension to continue digitizing, transcribing and geo-referencing. Not all final reports were available at the time of this writing; hence, we'll concentrate on summarizing the success of the two consortia and their online portals (CNALH <http://lichenportal.org> and CNABH <http://bryophyteportal.org>).

The two consortia together now have 86 members. This includes collections that submitted their data without help from the project from the US (including two private collections and a botany club) and from other countries (8 collections from Canada, Norway, Sweden, and Denmark), and 10 collections that were part of one of the PEN grants. 53 collections have both lichens and bryophytes, 12 only Bryophytes and 22 only lichens. Some changes from the original proposal have occurred as some collections were not able to participate while others that were not listed in the proposal have joined the effort.

The two portals together provide access to well over 4.1 million records (~1.86 million lichens and 2.25 million bryophytes). As of now this project has added over 1 million records each to the bryophyte and lichen portal. 64% of the bryophyte records and 86% of the lichen records are fully transcribed and 25% and 43% of bryophyte and lichen records, respectively, are geo-referenced. Specifically, during this past year 345,000 bryophyte and 727,000 lichen labels were transcribed. Detailed numbers as of June 2015 are in table 1. Please note that total numbers reported in table 1 include records that were already in the database before the project started, plus records that were added without direct support from this project but were added because the owners see value in publishing their data in a central place (e.g., international and private collections). The table does not include collections that are not (yet) appearing on the either portal.

Relative to the original projections, the following institutions appear to have completed (or nearly so) databases at this time: CUP, ILLS, LSU, MAINE, MONT, NCU, NEBK, NHA, NYS, OSC, TLE (separated from NDA), UMFK, and YPM (Yale, now as part of a PEN). In addition, of the Florida PEN, FLAS, FTG, FTU, and USF are completed. Several other herbaria have nearly completed databases for either lichens or bryophytes but not both (e.g. BRY bryophytes, DUKE lichens, MICH lichens, UC lichens, UNAF bryophytes, URV bryophytes, WIS lichens, etc.)

Approximately 100 people were involved in the project this year. This includes senior personnel, undergraduate (>50) and graduate students, volunteers, and interns specifically mentioned in the reports from collaborating institutions.

Several research projects that are not funded by this project but are using the digitized data, e.g., Duke reports two projects, preparing a manuscript that uses the collections data to study patterns of moss

species richness and turnover in North Carolina and using geo-referenced CNALH records to generate estimates of lichen species richness at various spatial scales across North America.

The Harvard University Herbaria recently established a new Postdoctoral Fellowship, which includes the analysis of data generated by TCN projects. The presently hired candidate is currently not using lichen or bryophyte data but data from vascular plants, for which they are also part of another TCN. But the plan is to discuss adding lichens and bryophytes to his portfolio.

The BRY TCN team has been evaluating the effects of global climate change on the distribution of the alpine tundra lichen *Flavocetraria nivalis*. TCN data have allowed them to document historical distribution patterns in western North America for this species – followed by field work in selected areas to document current distribution patterns. They are also researching the effects of increasing levels of atmospheric nitrogen pollution on the distribution of sensitive indicator species (lichens) in the western United States – particularly at sites in the Intermountain Area with extensive fossil fuel extraction and processing activity (oil and natural gas).

The Field Museum reports two projects focusing on remote assessment of accuracy of taxonomic identifications in portal data, a PhD thesis that will use portal data to analyze lichen biodiversity gradients, and a taxonomic revision of the *Pseudocyphellaria crocata* complex in the Americas, using portal data to detect and map specimens and occurrences.

Both portals have several active species inventory projects. The checklists are linked to voucher specimens in the collections:

- Lichen portal: 60 inventories with 44,610 voucher specimens linked
- Bryophyte portal: 15 inventories with 11,832 voucher specimens linked.

As previously reported many collaborators mention that the increased visibility of their specimens online has led to more requests for information and loans. However, digitizing the collections has also made it easier for the staff to respond to these requests. Some report that they are using the portals on an almost daily basis for collections management and research.

Progress has been made with georeferencing. In cooperation with iDigBio and Tulane, Robert Anglin (WIS) has written code to query the Geolocate service at Tulane University, parse the XML that it produces and convert those results that are acceptable into update SQL that enters these results into the Symbiota database. The data that is submitted consists of the contents of the locality, verbatimCoordinate, country, stateProvince, and county fields for those records that have non-null counties and TRS-values in either the verbatimCoordinate or locality fields.

There are several criteria for acceptability. First, Nelson Rios, the programmer for Geolocate, specified that only those elements in the result set that contain identifiable TRS coordinates in the “ParsePattern” subelement should be considered reliable. Those “ParsePattern” subelements which contain a question mark in addition to the TRS coordinates are rejected as this is an indication of a potentially unreliable result. Those elements that contain more than one “ParsePattern” subelement are averaged if the latitudes and longitudes are within 0.015 degrees of one another. Otherwise they are rejected. Since the

Geolocate results are the same whether or not the TRS values contain subsections (e.g. SW ¼) it is assumed that the results are accurate to within one mile. Applying this to WIS lichen data has increased the georeferenced collections by over 4000.

General quality control of georeferencing results is done by querying the Symbiota database for those records with coordinates in a specified state or province, for instance, and sorting by latitude and longitude in ascending and descending order. Records that are at the top of the record set represent extremes of latitude or longitude and are examined further to ensure that the coordinates given are within the entered state or province. In cases of mismatch either the state or province or the coordinates are corrected. So far, this has only been applied to WIS records, but is available upon request to other collections.

## Outreach

A volunteer coordinator, Mari Robertson, was hired by the New York botanical garden. She has developed a network of volunteers and over 20,000 records have been transcribed and based on her experiences provided documents to guide other institutions in developing such volunteer programs.

Julie Smith (WIS) cohosted four Symbiota webinars with Ed Gilbert: Exsiccati Index and Management, Oct., 2014. Crowdsourcing Data Entry, Sept., 2014. Introduction to Specimen Management, Aug., 2014. Introduction to Specimen Data Entry, Aug., 2014.

Other outreach efforts involved fielding questions from participating institutions and providing data management instructions through Skype and email to participants and volunteers. Visiting participants have also been guided through the transcribing process.

We reported last year on a project to develop lichen and bryophyte flashcards to be used in a 3D smart phone or tablet application. D. Allard (VT) has joined J. Kluse (LSU) and A.M. Basham (ASU) in this effort and one flashcard is being produced as part of a larger iDigBio project. However, lichens and bryophytes may prove too delicate and intricate to capture enough detail with currently available 3D technology for a virtual reality and we are exploring other means to convey the same information.



**Table 1:** Numbers extracted from the Symbiota database for both portals in June 2015. Total numbers of records are including records that were already present before this project started, international and private collections, as well as otherwise contributed already databased collections records.

Acronym	Institution	Lichens			Bryophytes		
		LBCC proposal estimate	total # of records including pre-LBCC	total # of transcribed records including pre-LBCC	LBCC proposal estimate	total # of records including pre-LBCC	total # of transcribed records including pre-LBCC
ALA	University of Alaska Museum	23000	18207	18207	39955	34246	34246
ARIZ	University of Arizona				5000	3738	
ASU	Arizona State University		115288	114434			
BALT	Towson University	6000	4008	2	1000	2	2
BG	University Museum of Bergen		68260	66868			
BING	University of New York Binghamton				8000	20270	1377
BRIT	Botanical Research Institute of Texas	1925	984		11468	14861	2206
BRY	Brigham Young University	78000	40348	39372	2000	3705	3572
C	Natural History Museum of Denmark: The Herbarium of Lichens		8088	8035			
CANL	Canadian Museum of Nature		31102	30412			
CAS	California Academy of Science	17500	558		55000	30694	451
CHRB	Rutgers University	1800	2667	3	5000	6838	
CINC	University of Cincinnati		6930	4731		11797	642
CMN	Canadian Museum of Nature					11227	11129
COLO	University of Colorado	48000	49535	33878	55000	62895	8854
CONN	University of Connecticut	2356	2489	45	7421	6884	90
CUP	Cornell University	6500	5877	5199			
DUKE	Duke University	55632	61952	58105	60396	211845	196924
EVE	Evergreen State College	390	390	388	304	304	304
F	Field Museum of Natural History	34000	62384	29882	34000	119672	49623
FH	Harvard University	93745	60438	59851	130359	43388	42748
FLAS	University of Florida	30000	10687	10685		22737	22733
FTG	Fairchild Tropical Botanic Garden, Florida		139	139		425	425
FTU	University of Central Florida		750	750		791	791
HAW	University of Hawaii	9000	6652	1701	600		
IDS	Idaho State University	1750	1641		126		
ILL	University of Illinois, Urbana	11526	6714	36	27000	25273	11912
ILLS	Illinois Natural History Survey		10453	10416		2743	2738
IND	Indiana University	144	128	56	1584	1560	875
ISC	Iowa State University		5081	4344			
KANU	University of Kansas		18185	18185			
KSP	Pittsburg State University (Kansas)	62			5659	5689	12

<b>LEDLIE</b>	Patricia Ledlie Herbarium (private collection)					1402	1402
<b>LSU</b>	Louisiana State University	32000	30065	30064	4500	3905	3873
<b>MAINE</b>	University of Maine	6500	8302	8008	2000	2687	2675
<b>MCTC</b>	Michigan Technological University	2000	2961		14000	10964	
<b>MICH</b>	University of Michigan	36700	44810	43767	78000	127715	45300
<b>MIL</b>	Milwaukee Public Museum		3240	2892			
<b>MIN</b>	University of Minnesota		144182	140783		42202	4490
<b>MO</b>	Missouri Botanical Garden				242062	323187	218141
<b>MONT</b>	Montana State University	8000	8496	8450			
<b>MOR</b>	Morton Arboretum	19000	16661	7578	1000	489	
<b>MSC</b>	Michigan State University	0	105476	105470	16000	22243	1709
<b>MU</b>	Miami University (Ohio)		8404	1204		3709	2
<b>NBM</b>	New Brunswick Museum (Canada)		10123	10026			
<b>NCU</b>	University of North Carolina Chapel Hill	936	5191	5106	4500	2961	2863
<b>NDA</b>	North Dakota State University	5256	5158	4975	3313	3258	0
<b>NEB</b>	University of Nebraska	16000	16838	10726	11900	10709	6301
<b>NEBC</b>	New England Botanical Club	numbers are included with FH					
<b>NEBK</b>	University of Nebraska at Kearny	178	198	183	0	3827	3708
<b>NHA</b>	University of New Hampshire	40	440	416	500	1167	1093
<b>NY</b>	New York Botanical Garden	135000	121421	63839	50000	389196	319152
<b>NYS</b>	New York State Museum		8935	8927		32452	31756
<b>O</b>	Botanical Museum, Oslo, Norway		168316	166254			
<b>OMA</b>	University of Nebraska at Omaha		19034	18976			
<b>OS</b>	Ohio State University	13180	14360	5	11300		
<b>OSC</b>	Oregon State University	3000	21110	20865	3000	16663	16663
<b>PH</b>	Philadelphia Academy of Science	24763	17829	9611	48401	46581	38272
<b>RM</b>	University of Wyoming		563	3		5613	0
<b>SBBG</b>	Santa Barbara Botanic Garden	5000	33009	32768	0		
<b>SIU</b>	Southern Illinois University					19104	0
<b>SQB</b>	Société québécoise de bryologie (Canada)					3253	3226
<b>SRP</b>	Boise State University	5000	23109	10876	0	817	0
<b>TENN</b>	University of Tennessee	8000	7986	2557	62000	71266	35365
<b>TLE</b>	Ted Esslinger private collection	25000	25924	25413			
<b>UBC</b>	University of British Columbia					167840	156603
<b>UC</b>	University of California, Berkeley	28000	46940	45331	114088	99961	23233
<b>UCR</b>	University of California, Riverside		16041	16036			
<b>UMFK</b>	University of Maine at Ft. Kent	387	46970	46970			
<b>UNAF</b>	University of Northern Alabama	25			3000	413	413
<b>UNLV</b>	University of Nevada Las Vegas	280	169		3500	3154	
<b>UPS</b>	Uppsala University Exsiccati (Sweden)		26226	24102			
<b>URV</b>	University of Richmond	1450	1895	29	450	441	440

<b>US</b>	Smithsonian Institution		75669	68834		5153	4575
<b>USF</b>	University of South Florida		1366	1358		2856	2735
<b>UTC</b>	Utah State University		465	463		2200	2193
<b>UWEC</b>	University of Wisconsin Eau Claire	10000	1180	797		500	
<b>UWFP</b>	University of West Florida					304	304
<b>UWSP</b>	University of Wisconsin Stevens Point	1500	1749	1522	2187	4067	1
<b>VPI</b>	Virginia Polytechnic Institute and State University		759	758			
<b>VSC</b>	Valdosta State University					4069	3771
<b>VT</b>	University of Vermont	73	5100	2745	17637	16969	8652
<b>WIS</b>	University of Wisconsin	112780	111677	100054	44150	34674	16184
<b>WTU</b>	University of Washington		29203	26398	30000	74141	51421
<b>WVA</b>	West Virginia University	15000	8220		20000	3447	2
<b>WWB</b>	Western Washington University		5336			5213	
<b>YPM</b>	Yale University, Peabody Museum of Natural History	4500	5041	4972	48000	36956	31888
<b>TOTAL</b>		<b>922907</b>	<b>1856082</b>	<b>1595835</b>	<b>1344786</b>	<b>2248823</b>	<b>1430060</b>

## **SOUTHWEST COLLECTIONS OF ARTHROPODS NETWORK (SCAN): A MODEL FOR COLLECTIONS DIGITIZATION TO PROMOTE TAXONOMIC AND ECOLOGICAL RESEARCH**

Report submitted by: neilscobb@gmail.com  
Report Submitted on: 07/07/2015 - 22:46

### **Progress in Digitization Efforts**

see attached document

### **Share and Identify Best Practices and Standards (including Lessons Learned)**

see attached document

### **Identify Gaps in Digitization Areas and Technology**

see attached document

### **Share and Identify Opportunities to Enhance Training Efforts**

see attached document

### **Share and Identify Collaborations with other TCNs, Institutions, and Organizations**

see attached document

### **Share and Identify Opportunities and Strategies for Sustainability**

see attached document

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

see attached document

### **Attachment**

[https://www.idigbio.org/sites/default/files/webform/tcn-reports/SCAN\\_June\\_2015.docx](https://www.idigbio.org/sites/default/files/webform/tcn-reports/SCAN_June_2015.docx)

**Southwest Collections of Arthropods Network Update**  
**July 7, 2015**  
**Neil Cobb**

**Progress in Digitization Efforts:**

We are on target to exceed our quota for digitizing labels from pinned specimens, 736,736 records from the original 10 institutions and 878,736 records included the two PEN projects.

**Table 1** presents four sets of statistics derived from our data portal as of March 31, 2015. These include the following data: 1) institutions that are funded by the NSF-ADBC program, including the 2 PEN grants; 2) institutions that have entered data into the SCAN portal but not funded by the NSF-ADBC program; 3) the total of these first two categories; and 4) the total records in the SCAN portal. The fourth column includes records from the first three columns as well as all records we have ingested from aggregators GBIF and iDigBio. The purpose of serving aggregator data is to provide as complete as information as possible to persons that are considering research projects.

Table 1. Number of specimen records digitized and associated summary statistics. From <http://symbiota4.acis.ufl.edu/scan/portal/index.php> . SCAN-funded numbers refer to the 12 museums receiving ADBC funding. SCAN non-funded numbers include nine museums contributing cataloged specimen data and non-cataloged moth specimen data from 22 collections (5 private collections and 17 public museums). Total Served includes all SCAN data and other datasets with North American arthropod records (e.g., GBIF, Tri-Trophic TCN).

	SCAN funded	SCAN non-funded	TOTAL SCAN	Total Served
<b># Specimen Records</b>	925,336	497,257	1,422,593	4,016,683
<b># Georeferenced</b>	727,796	290,428	1,018,224	3,030,446
<b># Identified to species</b>	574,739	170,940	745,679	1,814,198
<b># Families</b>	1,214	1,086	1,695	1,836
<b># Genera</b>	5,912	5,488	8,633	11,968
<b># Species</b>	18,305	16,019	30,938	55,986
<b>Total taxa (including subsp. and var.)</b>	19,079	16,326	31,871	58,657
<b>Type specimens</b>	3,993	37,071	41,064	135,159
<b>% Georeferenced</b>	79%	58%	71%	75%
<b>% Identified to Species</b>	62%	34%	52%	45%

Although we have technically reached our goal for the 12 SCAN museums, we have not thoroughly reviewed all records that SCAN-funded museums have produced to determine how many of those strictly ground-dwelling arthropods, but we expect that 80% or those are target taxa and that we will need over 125,000 more records to meet our project goal. Ten museums will request one-year no-cost extensions and the one PEN museum (BYU) is in their second year. We estimate that we will digitize at least 800,000 ground-dwelling arthropod specimens by

---

the end of the project and over one million total specimens for the original 10 museums. The two additional PEN grants (Harvard and BYU) are on track to meet their quotas.

A subset of SCAN museums are creating high-resolution images and three museums are creating low resolution images that include the specimen and labels in the same image. **Table 2** lists the number of images posted on SCAN by participating museums. Our goal was to produce 15,125 high-resolution images suites. An image suite consists of 1-3 images representing different aspects of a specimen. This will translate into approximately 40,000 images. Three museums are producing low-resolution images (University of Hawaii, University of Arizona, and Texas Tech University).

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

We are identifying best practices on a weekly basis and sharing those with respective people within SCAN.

**Identify Gaps in Digitization Areas and Technology:**

We need to harvest additional data (i.e. beyond SCAN) to better understand the biogeography of arthropod taxa. We are partially meeting this need by incorporating GBIF into the SCAN database.

**Share and Identify Opportunities to Enhance Training Efforts:** Nothing new to report, we are working on activities already described in previous reports

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

We are primarily working with Tri-Trophic TCN in order to develop questions for analyzing ADBC data.

**Share and Identify Opportunities and Strategies for Sustainability:**

We have a sustainability plan for Colorado State University, they are finished using their NSF funding <http://scan1.acis.ufl.edu/content/sustainability> .

Table 2. Number of images posted on SCAN portal from SCAN museums that are focused on producing high-resolution images of specimens. Data are recorded from <http://symbiota1.acis.ufl.edu/scan/portal/imagelib/photographers.php>

Institution	# Images
<a href="#">Arizona State University Hasbrouck Insect Collection (ASU-ASUHIC)</a>	2,622
<a href="#">C.P. Gillette Museum of Arthropod Diversity (CSU-CSUC)</a>	49
<a href="#">Colorado Plateau Museum of Arthropod Biodiversity (NAUF-CPMAB)</a>	1,716
<a href="#">Denver Museum of Nature &amp; Science (DMNS-DMNS)</a>	625
<a href="#">Hymenoptera Institute Collection (UKY-HIC-HIC)</a>	2,085
<a href="#">Museum of Comparative Zoology, Harvard University (MCZ)</a>	12,725
<a href="#">Museum of Southwestern Biology, Division of Arthropods (UNM-MSBA)</a>	193
<a href="#">National Park Collections at Northern Arizona University (NAUF-NPS)</a>	673
<a href="#">New Mexico State Collection of Arthropods (NMSU-NMSU)</a>	1,380
<a href="#">Ohio State C.A. Triplehorn Insect Collection (OSU-OSU)</a>	2,655
<a href="#">SDSU Terrestrial Arthropods Collection (SDSU-TAC)</a>	68
<a href="#">Texas Tech University - Invertebrate Zoology (TTU-TTU-Z)</a>	26,027
<a href="#">UAM Entomology Collection (UAM-UAM ENT)</a>	2,881
<a href="#">University of Arizona Insect Collection (UA-UAIC)</a>	56,952
<a href="#">University of Colorado Museum of Natural History Entomology Collection (UCB-UCMC)</a>	636
<a href="#">University of Hawaii Insect Museum (UHIM-UHIM)</a>	38,587
<a href="#">University of Kansas Natural History Museum Entomology Division (KU-SEMC)</a>	4,445
<a href="#">University of Tennessee at Chattanooga (UTC-UTCI)</a>	1
<a href="#">Yale Peabody Museum, Entomology Division (YPM-ENT)</a>	9,319
<b>SCAN Museums (All Images)</b>	<b>102,973</b>
<b>SCAN Museums (High-Resolution Images)</b>	<b>20,994</b>
<b>All Images served on SCAN</b>	<b>163,014</b>

**Other Progress (that doesn't fit into the above categories):** We continue to provide North American data we have obtained from other sources to increase the quantity of data available to SCAN users. We have grown from serving 10 collection datasets to serving 39 data sets through SCAN (**Table 3**) and four institutions have committed to serving data. These will greatly increase the usability of the existing SCAN data, especially understanding species distributions and more complete species lists. We are re-building our data harvested from North American data from GBIF and are in the process of hosting data from other non-TCN arthropod data sets that have been harvested by iDigBio.

Table 3. List of Collection data sets being served on SCAN data portal.

<b><u>SCAN Funded Collections</u></b>	
1	Arizona State University Hasbrouck Insect Collection (ASU-ASUHC)
2	Brigham Young University Arthropod Museum (BYU-BYUC)
3	C.P. Gillette Museum of Arthropod Diversity (CSU-CSUC)
4	Colorado Plateau Museum of Arthropod Biodiversity (NAUF-CPMAB)
5	Denver Museum of Nature & Science (DMNS-DMNS)
6	Museum of Comparative Zoology, Harvard University (MCZ)
7	Museum of Southwestern Biology, Division of Arthropods (UNM-MSBA)
8	New Mexico State Collection of Arthropods (NMSU-NMSU)
9	Texas A&M University Insect Collection (TAMU-TAMUIC)
10	Texas Tech University - Invertebrate Zoology (TTU-TTU-Z)
11	University of Arizona Insect Collection (UA-UAIC)
12	University of Colorado Museum of Natural History Entomology Collection (UCB-UCMC)
<b><u>SCAN Non-Funded Collections</u></b>	
1	California Academy of Sciences Entomology (CAS-ENT)
2	Denver Botanic Gardens Collection of Arthropods (DBG-DBGA)
3	Dugway Proving Ground Natural History Collection (DUGWAY-DUG-ENT)
4	Entomology Collection at the Natural History Museum of Utah (UMNH-ENT)
5	Gregory P. Setliff Collection - Kutztown University (GPSC)
6	Hymenoptera Institute Collection (UKY-HIC-HIC)
7	National Park Collections at Colorado State University (CSU-CSUNPS)
8	National Park Collections at Northern Arizona University (NAUF-NPS)
9	SDSU Terrestrial Arthropods Collection (SDSU-TAC)
10	The Albert J. Cook Arthropod Research Collection (MSU-MSUC)
11	The Purdue Entomological Research Collection (PU-PERC)
12	United States National Museum, Entomology Collections (USNM-USNMENT)
13	University of Delaware Insect Research Collection (UD-UDCC)
14	University of Georgia Collection of Arthropods (UGCA-GMNH)
15	University of Hawaii Insect Museum (UHIM-UHIM)
16	University of Tennessee at Chattanooga (UTC-UTCI)
17	Utah Department of Agriculture and Food Entomology Collection (UDAF-UDAFE)
18	Western Washington University Insect Collection (WWU-WWUC)
19	Scarab Central (Observational, not served to iDigBio)
20	Moth Observation Database (Observational, not served to iDigBio)





## DEVELOPING A CENTRALIZED DIGITAL ARCHIVE OF VOUCHERED ANIMAL COMMUNICATION SIGNALS

Report submitted by: es269@cornell.edu  
Report Submitted on: 07/21/2015 - 12:31

### Progress in Digitization Efforts

During the reporting period, our TCN digitized over 1650 audio recordings from multiple TCN partners. These recordings (“media specimens”) are now available through, and playable at, the Macaulay Library website (MacaulayLibrary.org), and data are being pushed to iDigBio and VertNet. The list below details the major bodies of material digitized during the latest reporting period:

**Anurans:** We have continued digitizing anuran recordings associated with specimens from several TCN partners. During this reporting period we completed 349 recordings associated with specimens from the Smithsonian Institution, including recordings from Crombie (133) and McDiarmid (118). Another 144 anuran recordings associated with specimens from the Texas Natural History Collection were digitized and archived.

**Orthopterans:** Digitization continued on the David Weissman orthopteran collection with another 450 recordings processed in May and June 2015, bringing total digitized and archived recordings to over 1400.

**Birds:** We continued digitization on the LSU bird collection by archiving 241 recordings from Dan Lane during the period. We also archived 483 recordings by Mark Robbins from the KU birds collection, many with associated specimens.

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

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

During this reporting period, our TCN organized and hosted a major workshop and meeting on digitization of vertebrate specimens, co-organized and supported by iDigBio. This workshop was a major success, with participation by over 70 researchers/staff from museums and other institutions from across the country. Key foci and themes from the workshop included digital media (audio and video recordings associated with specimens, etc), CT and microCT scans of specimens, and resources/strategies to support digitization efforts at smaller institutions. The workshop also included tours of the Cornell Museum of Vertebrates, tours of state-of-the-art facilities for bio-imaging and digital biodiversity media, and field workshops on collecting biodiversity media. Additional details can be found in the workshop reports.

### Share and Identify Opportunities and Strategies for Sustainability

Nothing to report.

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

Nothing to report.

### Attachment

N/A