

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

Report submitted by: [neilscobb@gmail.com](mailto:neilscobb@gmail.com)  
Report Submitted on: 05/14/2014 - 12:55

### **Progress in Digitization Efforts**

see attached file

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

see attached file

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

see attached file

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

see attached file

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

see attached file

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

see attached file

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

see attached file

### **Attachment**

[https://www.idigbio.org/sites/default/files/webform/tcn-reports/SCAN\\_May\\_2014.docx](https://www.idigbio.org/sites/default/files/webform/tcn-reports/SCAN_May_2014.docx)

**Southwest Collections of Arthropods Network Update**  
**July 14, 2014**  
**Neil Cobb**

**Progress in Digitization Efforts:**

We are on target to meet our second-year quota for digitizing labels from pinned specimens. Table 1 presents three sets of statistics as of March 18, 2014. These include data from institutions that are funded by SCAN, institutions that are entering data into the SCAN portal but not funded by SCAN, and the total records in the SCAN portal. We have added

Table 1. Number of specimen records digitized and associated summary statistics. From <http://symbiota1.acis.ufl.edu/scan/portal/collections/misc/collstats.php>

	<b>SCAN funded</b>	<b>SCAN non-funded</b>	<b>TOTAL SCAN</b>
<b># Specimen Records</b>	486,539	96,862	583,401
<b># Georeferenced</b>	361,581	34,517	396,098
<b># Identified to species</b>	345,687	64,627	410,314
<b># Families</b>	675	537	1,212
<b># Genera</b>	6,339	3,650	9,989
<b># Species</b>	13,793	8,130	21,923
<b>% Georeferenced</b>	74	36	69
	71	67	71

We have also started creating high-resolution images taken by a subset of SCAN museums that are committed to producing specimen images. Table 2 lists the number of images posted on SCAN by these participating museums. Our goal is to produce ~16,000 images suites. An image suite consists of 1-5 images representing different aspects of a specimen. This will translate into approximately 40,000 images. We are currently behind on our projections due to unexpected logistical challenges but we expect to greatly increase our productivity over the summer, 2014.

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	# High-Resolution Images
Arizona State University	1296
Colorado State University	48
Northern Arizona University	1020
Denver Museum of Nature and Science	574
University of New Mexico	80
Northern Arizona University - NPS	673
New Mexico State University	853
Texas Tech University	56
University of Colorado at Boulder	0
<b>TOTAL</b>	<b>4600</b>

**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. We presented a joint paper at the 21<sup>st</sup> Century meeting May 5, 2014. We are working with Pam Soltis and Charlotte Germain on collaborative ecological niche modeling and biodiversity issues.

**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> .

**Other Progress (that doesn't fit into the above categories):** We are starting to share North American data from other sources to increase the quantity of data. These will greatly increase the usability of the existing SCAN data, especially understanding species distributions and more complete species lists. We are hosting North American data from GBIF and are in the process of hosting data from Tri-Trophic TCN and other non-TCN arthropod data sets that have been harvested by iDigBio. With these additional records we are currently serving over 3.8 million records for 44,023 species.

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

Report submitted by: [neilscobb@gmail.com](mailto:neilscobb@gmail.com)  
Report Submitted on: 06/02/2014 - 12:38

### **Progress in Digitization Efforts**

see attached

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

see attached

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

see attached

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

see attached

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

see attached

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

see attached

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

see attached

### **Attachment**

[https://www.idigbio.org/sites/default/files/webform/tcn-reports/SCAN\\_June\\_2014.docx](https://www.idigbio.org/sites/default/files/webform/tcn-reports/SCAN_June_2014.docx)

**Southwest Collections of Arthropods Network Update**  
**July 14, 2014**  
**Neil Cobb**

**Progress in Digitization Efforts:**

We are on target to meet our second-year quota for digitizing labels from pinned specimens. Table 1 presents three sets of statistics as of June 1, 2014. These include data from institutions that are funded by SCAN, institutions that are entering data into the SCAN portal but not funded by SCAN, and the total records in the SCAN portal. We have added

Table 1. Number of specimen records digitized and associated summary statistics. From <http://symbiota1.acis.ufl.edu/scan/portal/collections/misc/collstats.php>

	<b>SCAN funded</b>	<b>SCAN non-funded</b>	<b>TOTAL SCAN</b>
<b># Specimen Records</b>	617,317	89,301	706,618
<b># Georeferenced</b>	448,581	38,283	486,864
<b># Identified to species</b>	406,150	66,332	472,482
<b># Families</b>	674	538	1,212
<b># Genera</b>	6,479	3,675	10,154
<b># Species</b>	14,446	7,953	22,399
<b>% Georeferenced</b>	73%	43%	69%
	66%	74%	67%

We have also started creating high-resolution images taken by a subset of SCAN museums that are committed to producing specimen images. Table 2 lists the number of images posted on SCAN by these participating museums. Our goal is to produce ~16,000 images suites. An image suite consists of 1-5 images representing different aspects of a specimen. This will translate into approximately 40,000 images. We are currently behind on our projections due to unexpected logistical challenges but we expect to greatly increase our productivity over the summer, 2014.

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	# High-Resolution Images
Arizona State University	1,338
Colorado State University	48
Northern Arizona University	1,020
Denver Museum of Nature and Science	574
University of New Mexico	80
Northern Arizona University - NPS	673
New Mexico State University	910
Texas Tech University	325
University of Colorado at Boulder	0
<b>TOTAL</b>	<b>4968</b>

**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. We presented a joint paper at the 21<sup>st</sup> Century meeting May 5, 2014. We are working with Pam Soltis and Charlotte Germain on collaborative ecological niche modeling and biodiversity issues.

**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> .

**Other Progress (that doesn't fit into the above categories):** We are starting to share North American data from other sources to increase the quantity of data. These will greatly increase the usability of the existing SCAN data, especially understanding species distributions and more complete species lists. We are hosting North American data from GBIF and are in the process of hosting data from Tri-Trophic TCN and other non-TCN arthropod data sets that have been harvested by iDigBio. With these additional records we are currently serving over 4.8 million records for 56,441 species.



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

Report submitted by: [mws244@cornell.edu](mailto:mws244@cornell.edu)  
Report Submitted on: 06/09/2014 - 10:13

### Progress in Digitization Efforts

We are now nearing completion of the first year of activities for this project. The key outcome from our Year 1 activities has been the digitization of many hours of audio recordings to create over 10,100 new media specimens, as detailed below. These recordings ("media specimens") are now available through, and playable at, the Macaulay Library website ([MacaulayLibrary.org](http://MacaulayLibrary.org)), and we are developing plans to push the data to iDigBio and VertNet. The list below details the major bodies of material digitized during this period:

- (1) Bird audio recordings from University of Kansas, Natural History Museum. KU is a major partner in this project, and during the first year we have nearly completed digitizing and making accessible ALL of the bird audio material from this institution. This material includes: a total of 8,192 audio recordings from researcher Mark Robbins; 219 recordings collected by Peter Hosner in Peru (2008, 2009, 2012); and 134 recordings from researcher Mike Andersen, including recordings from Congo (2012) and Fiji (2010, 2011).
- (2) Bird audio recordings from Louisiana State University. Digitization of over 100 hours of analog audio tape recordings, in particular material from recordist Daniel Lane collected in Peru and Ecuador during the late 1990's. Of this total digitized material, 690 recordings have been processed to become audio media specimens with complete data records in the Macaulay archive.
- (3) Bird audio recordings from the Yale Peabody Museum. We have now digitized and created media specimens for 877 audio recordings from Yale researcher Krzysztof Zyskowski.
- (4) Fish EOD recordings from Cornell University Museum of Vertebrates. Macaulay Library staff worked with CUMV research technician John Sullivan (supported by this award) and researcher Carl Hopkins to develop and test protocols for the digitization, archival, web-presentation, and delivery of electric organ discharge (EOD) signal data from mormyrid and gymnotiform fishes (a signal modality that does not lend itself to current audio standards).

With these media now digitized and archived at the Macaulay Library, particularly the large body of material from KU, the stage is set to create the links across databases between physical specimen and media specimen.

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

The Macaulay Library uses an audio archival standard of 96kHz 24-bit, the audio standard recommended by Sound Directions: Best Practices for Audio Preservation <<http://www.dlib.indiana.edu/projects/sounddirections/papersPresent/index.shtml>> and a standard adopted by leading audio archival institutions such as the Library of Congress and The British Library.

### Identify Gaps in Digitization Areas and Technology

There are no accepted standards for the preservation and subsequent presentation of electric organ discharges produced by e-fish. During the past year, Macaulay Library audio archival staff worked with staff at CUMV to develop archival and web-proxy presentation protocols in collaboration with e-fish researchers that will serve as a model format for EODs.

### Share and Identify Opportunities to Enhance Training Efforts

Personnel from this TCN project visited partner institutions and participated in meetings/summits to facilitate the work undertaken and for exchange of information. For example, LSU PI Daniel Lane visited the Macaulay Library in May 2014 to assist with metadata entry and digitization of specimens, PI's Mike Webster, Greg Budney and Rafe Brown all attended the iDigBio Summit in October 2013, and Co-PI Kim Bostwick attended the iDigBio Education & Outreach Workshop in January 2014. Importantly, we have submitted a proposal to the iDigBio Steering Committee to conduct, in partnership with others, a workshop on digitization of vertebrate specimens. If approved, this workshop will be held in early May 2015 at Cornell.

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

No collaborations with other TCNs at this time, but we are exploring data-cleaning and geo-referencing capabilities developed by other TCNs.

### Share and Identify Opportunities and Strategies for Sustainability

National-level reporting of iDigBio achievements, e.g. Heretofore resources now available to the public.

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

In addition to the digitization outcomes detailed above, several of the partner institutions have made significant progress on data cleaning and data migration in preparation for linking their specimens to digitized audio recordings. For example, CUMV recently completed the migration of its specimen data from Specify to Arctos, making all specimen records accessible online to anyone at [http://arctos.database.museum/cumv\\_all](http://arctos.database.museum/cumv_all). Individual records now have stable GUIDs that can be linked to Macaulay Library media records, and these data are now being pushed to FishNet 2, VertNet and GBIF on a regular basis. Incorporation of these data into the iDigBio portal is planned for the future. Similarly, CUMV staff are continuing to complete clean up of specimen metadata (e.g, georeferencing locality data, taxonomy, etc.), and have begun adding links to online resources tied to their voucher specimens (i.e., GenBank records, DOIs and citations for publications citing e-fish vouchers, etc.), and have starting adding images of electric fish voucher specimens (e.g., see <http://arctos.database.museum/guid/CUMV:Fish:96774>).

### Attachment

N/A

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

Report submitted by: [moon@begoniasociety.org](mailto:moon@begoniasociety.org)  
Report Submitted on: 06/19/2014 - 15:22

### Progress in Digitization Efforts

Mari Roberts has trained and supervised employees and volunteers at the Herbarium of The New York Botanical Garden. In May and June, she oversaw 3 volunteers who completed an additional 2,300 specimen records. Currently, there are 3 volunteers and 3 NYBG employees completing specimen data entry in Symbiota, the project web portal.

Robert Naczi curated grass specimens in the Herbarium of The New York Botanical Garden, to ensure the most accurate identifications possible in advance of posting specimen data online. In May and June, he identified and annotated 285 specimens of Poaceae at NYBG. Nicole Tarnowsky, NYBG Administrative Curator, assisted by entering the identifications in the NYBG Herbarium database.

University of Illinois, University of Kansas, and Texas A & M submitted insect records to iDigBio (Texas submitted via SCAN).

Our recent annual report to NSF, for plant and insect numbers, is attached.

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

Nothing to report.

### Identify Gaps in Digitization Areas and Technology

We need to get serious about attribution fields for specimen records. Both funding source and institution minimally should be attributed for each record. We have datasets we would like to provide, but cannot until this issue is sorted out.

IPT is limiting. Data providers are forcing information into Darwin Core fields that are pushing the limits of the definitions of these fields. Complex data, such as interactions, require fields outside of IPT.

### Share and Identify Opportunities to Enhance Training Efforts

We have many volunteers helping over the summer. Our limitation tends to be space, computers, and advisory staff at this point.

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

Finished a 3 day workshop in Riverside California on Data Mining and Distribution modeling with collection data. <http://tcn.amnh.org/home/data-mining-workshop>. Members of multiple TCNs, iDigBio, and outside institutions were in attendance. The presentations from the workshop were recorded and will be made available.

### 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

<https://www.idigbio.org/sites/default/files/webform/tcn-reports/TTD-TCNAnnualReportNumbers.pdf>



Lyctocoridae	0	0	0	0	0	0	0	0	0	0	70	0	0	0	4	0	0	1	75
Lygaeidae	242	1461	0	0	85	0	0	0	1248	589	600	1129	0	2906	205	0	0	68	8533
Machaerotidae	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	3
Malcidae	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	8
Margarodidae	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
Meenoplidae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Membracidae	951	0	0	0	454	0	0	0	4251	0	517	1591	0	0	7098	0	0	2	14864
Mesoveliidae	0	43	0	0	0	0	0	0	121	161	0	0	0	0	0	0	0	0	325
Microphysidae	0	0	0	0	0	0	0	0	0	0	13	0	0	0	0	0	0	0	13
Miridae	38403	8746	5636	0	3	0	0	0	6076	2029	1563	30660	19403	92	999	0	0	17	113627
Monophlebidae	0	0	0	0	0	0	0	0	0	17	0	0	0	0	0	0	0	1	18
Myerslopiidae	0	0	0	0	0	0	0	0	13	0	0	0	0	0	0	0	0	0	13
Mymaridae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Nabidae	0	958	0	0	0	0	0	0	2967	0	0	0	0	0	134	0	0	134	4193
Naucoridae	0	0	0	0	0	0	0	0	37	0	0	0	0	0	0	0	0	0	37
Nepidae	0	0	0	0	0	0	0	0	127	0	0	0	0	0	0	0	0	0	127
Ninidae	0	0	0	0	0	0	0	0	3	7	11	7	0	0	5	0	0	0	33
Nogodinidae	0	0	0	0	0	0	0	0	2	0	0	0	0	0	4	0	0	0	6
Notonectidae	0	34	0	0	52	0	0	0	522	0	0	0	0	0	8317	0	0	1	8926
Ochteridae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ortheziidae	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	44	45
Oxycarenidae	0	0	0	0	0	0	0	0	6	1	3	108	0	23	0	0	0	0	141
Pachygronthidae	0	0	0	0	0	0	0	0	216	272	659	0	0	80	74	0	0	65	1366
Pentatomidae	461	258	0	0	895	0	0	0	438	0	177	0	30	2101	1185	0	0	43	5588
Phylloxeridae	0	0	0	74	0	0	0	0	0	0	0	0	0	0	0	0	0	0	74
Piesmatidae	0	0	0	0	0	0	0	0	235	0	0	0	0	97	62	0	0	25	419
Plataspidae	0	45	0	0	0	0	0	0	25	0	0	24	0	0	44	0	0	0	138
Pleidae	0	11	0	0	0	0	0	0	128	0	0	0	0	0	1	0	0	0	140
Pseudococcidae	0	0	0	0	0	0	0	0	0	685	0	0	0	0	0	0	0	0	685
Psyllidae	0	0	0	0	0	0	0	0	1387	0	0	0	0	0	0	0	0	95	1482
Putoidae	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
Pyrrhocoridae	0	0	0	0	0	0	0	0	198	51	0	0	0	0	0	0	0	2	251
Reduviidae	11	0	0	0	0	0	0	0	3554	1	741	0	0	990	201	0	0	218	5716
Rhopalidae	44	175	0	0	0	0	0	0	381	826	1282	1148	0	936	3	0	0	55	4850
Rhyparochromidae	460	0	0	0	303	0	0	0	713	1557	2685	526	0	735	184	0	0	106	7269
Ricaniidae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Saldidae	2235	333	0	0	0	0	0	0	340	184	0	0	0	0	5	0	0	14	3111
Schizopteridae	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
Scutelleridae	0	0	0	0	74	0	0	0	591	0	423	0	0	0	55	0	0	22	1165
Stenocephalidae	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	3
Tessaratomidae	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0	0	0	12
Tettigometridae	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	4
Thyreocoridae	0	0	0	0	261	0	0	0	1024	0	864	0	9	320	93	0	0	9	2571
Tingidae	2753	206	0	0	0	0	0	0	1691	0	2552	0	0	328	364	0	0	135	8029
Triozidae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	97	97
Tropiduchidae	0	0	0	0	0	0	0	0	1	0	0	0	0	0	3	0	0	0	4
Veliidae	0	273	0	0	0	0	0	0	255	369	0	0	0	0	0	0	0	11	908
<b>Total specimens databased (2013-2014)</b>	<b>62046</b>	<b>12993</b>	<b>5636</b>	<b>6860</b>	<b>5481</b>	<b>2</b>	<b>11428</b>	<b>17556</b>	<b>88637</b>	<b>13248</b>	<b>22522</b>	<b>85592</b>	<b>26658</b>	<b>18615</b>	<b>21587</b>	<b>45842</b>	<b>10620</b>	<b>1934</b>	<b>457257</b>

Funds expended (\$) (2013-2014)	\$83,462.24	\$48,770.55	\$27,248.69	\$47,761.23	\$19,128.82	\$0.00	\$25,668.66	\$78,021.15	\$70,743.35	\$10,931.53	\$27,813.77	\$26,468.12	\$22,392.68	\$30,460.89	\$43,566.34	\$15,256.86	\$21,067.72	\$16,736.94	\$615,499.54
---------------------------------	-------------	-------------	-------------	-------------	-------------	--------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	--------------

<b>Total specimens databased (2011-2014)</b>	<b>110,324</b>	<b>17,465</b>	<b>7,792</b>	<b>13,736</b>	<b>11,853</b>	<b>15,350</b>	<b>23,732</b>	<b>23,813</b>	<b>149,130</b>	<b>26,638</b>	<b>62,863</b>	<b>89,838</b>	<b>86,568</b>	<b>49,647</b>	<b>43,093</b>	<b>66,710</b>	<b>20,581</b>	<b>6,282</b>	<b>825,415</b>
--	----------------	---------------	--------------	---------------	---------------	---------------	---------------	---------------	----------------	---------------	---------------	---------------	---------------	---------------	---------------	---------------	---------------	--------------	----------------

<b>Goal (Specimen No.) Total</b>	<b>333000</b>	<b>70000</b>	<b>40000</b>	<b>75000</b>	<b>15000</b>	<b>15000</b>	<b>30000</b>	<b>50000</b>	<b>73000</b>	<b>50000</b>	<b>75000</b>	<b>40000</b>	<b>150000</b>	<b>75000</b>	<b>20000</b>	<b>50000</b>	<b>35000</b>	<b>15000</b>	<b>1211000</b>
----------------------------------	---------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	---------------	--------------	--------------	--------------	--------------	--------------	----------------

<b>Percent Complete (in total)</b>	<b>33.1%</b>	<b>25.0%</b>	<b>19.5%</b>	<b>18.3%</b>	<b>79.0%</b>	<b>102.3%</b>	<b>79.1%</b>	<b>47.6%</b>	<b>204.3%</b>	<b>53.3%</b>	<b>83.8%</b>	<b>224.6%</b>	<b>57.7%</b>	<b>66.2%</b>	<b>215.5%</b>	<b>133.4%</b>	<b>58.8%</b>	<b>41.9%</b>	<b>68.2%</b>
------------------------------------	--------------	--------------	--------------	--------------	--------------	---------------	--------------	--------------	---------------	--------------	--------------	---------------	--------------	--------------	---------------	---------------	--------------	--------------	--------------

<b>Percent Complete (2014)</b>	<b>18.6%</b>	<b>18.6%</b>	<b>14.1%</b>	<b>9.1%</b>	<b>36.5%</b>	<b>0.0%</b>	<b>38.1%</b>	<b>35.1%</b>	<b>121.4%</b>	<b>26.5%</b>	<b>30.0%</b>	<b>214.0%</b>	<b>17.8%</b>	<b>24.8%</b>	<b>107.9%</b>	<b>91.7%</b>	<b>30.3%</b>	<b>12.9%</b>	<b>37.8%</b>
--------------------------------	--------------	--------------	--------------	-------------	--------------	-------------	--------------	--------------	---------------	--------------	--------------	---------------	--------------	--------------	---------------	--------------	--------------	--------------	--------------

**Table 2.** Numbers of botanical specimens imaged or databased in April 1 2013 - March 30 2014 reporting period. Institutions in bold are lead institutions; other institutions are subcontracted partners.

Botanical Specimens	Institution													Totals
	Eastern Michigan U.	Iowa State U	Miami U	Missouri Botanical Garden	New York Botanical Garden	U of Colorado	U of Illinois, Illinois Natural History Survey	U of Kansas	U of Maine	U of Michigan	U of Minnesota	U of Texas	U of Wisconsin	
Imaged	6614	21065	25238	4334	103172	25733	55788	10000	12,000	52672	52,400	40548	83,790	<b>486740</b>
Databased but not imaged	0	228	0	10226	0	0	0	31903	0	0	0	0	0	<b>42357</b>
<b>Total specimens digitized (2013-2014)</b>	<b>5032</b>	<b>21293</b>	<b>25238</b>	<b>14560</b>	<b>103172</b>	<b>25733</b>	<b>55788</b>	<b>41903</b>	<b>12000</b>	<b>52672</b>	<b>52,400</b>	<b>40548</b>	<b>83,790</b>	<b>529097</b>
<b>Total specimens digitized (2011-2014)</b>	<b>5032</b>	<b>50857</b>	<b>33175</b>	<b>67744</b>	<b>241290</b>	<b>81400</b>	<b>68123</b>	<b>54403</b>	<b>17500</b>	<b>113558</b>	<b>68064</b>	<b>62706</b>	<b>144876</b>	<b>1003696</b>
<b>Goal (No. of specimens) Total</b>	<b>10000</b>	<b>102000</b>	<b>35000</b>	<b>101000</b>	<b>274000</b>	<b>67000</b>	<b>124000</b>	<b>97000</b>	<b>34000</b>	<b>115000</b>	<b>70000</b>	<b>105000</b>	<b>90000</b>	<b>1224000</b>
<b>Percent completed (Total)</b>	<b>50.3%</b>	<b>49.9%</b>	<b>94.8%</b>	<b>67.1%</b>	<b>88.1%</b>	<b>121.5%</b>	<b>54.9%</b>	<b>56.1%</b>	<b>51.5%</b>	<b>98.7%</b>	<b>97.2%</b>	<b>59.7%</b>	<b>161.0%</b>	<b>82.0%</b>
<b>Percent completed in 2013 - 2014</b>	<b>50.3%</b>	<b>20.9%</b>	<b>72.1%</b>	<b>14.4%</b>	<b>37.7%</b>	<b>38.4%</b>	<b>45.0%</b>	<b>43.2%</b>	<b>35.3%</b>	<b>45.8%</b>	<b>74.9%</b>	<b>38.6%</b>	<b>93.1%</b>	<b>43.2%</b>

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

Report submitted by: [blieber@ku.edu](mailto:blieber@ku.edu)  
Report Submitted on: 06/23/2014 - 11:21

### Progress in Digitization Efforts

Paleoniches Update, June 2014

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 129,067 specimens databased. Of these, there are a total of 125,337 specimens databased that have clean, proofed localities. Further, we now have a total of 104,821 specimens that are georeferenced. In addition, a total of ~6,813 localities have been georeferenced.

Since the last update, PI Hendricks (San Jose State University; SJSU) and his undergraduate and graduate student assistants have continued to generate content for the Neogene and Pennsylvanian components of the "Digital Atlas of Ancient Life" and have put it online.

PI Hendricks has been migrating Neogene content from the old Digital Atlas page (<http://www.geosun.sjsu.edu/~jhendricks/AtlasTemp/>) to the new WordPress-based page at [www.neogeneatlas.org](http://www.neogeneatlas.org), which is now accessible; due to changes in programming, previously generated species content (information and images) must be entered into the new website formatting on a species-by-species basis (Hendricks is doing this work and is correcting errors along the way). Hendricks intends to have all previously and newly created Neogene content transferred to the new domain and format by the time of the next report (at which time the old website will likely be taken offline).

Progress continues to be made on developing new content for both the Neogene and Pennsylvanian Digital Atlases. Neogene content (taxon information, images, and distributional maps) has been completed for the bivalve families Arcidae and Glycymeridae and content construction is underway for the bivalve families Mytilidae and Ostreidae, as well as additional gastropod genera belonging to the family Muricidae. We intend to have corresponding webpages for all of these taxa online by the time of the next report. Substantial progress has also been made on generating content (especially for crinoids) for the [www.pennsylvanianatlas.org](http://www.pennsylvanianatlas.org) page, though it has not yet been put online (the structure of the Pennsylvanian site will soon mirror that of the new Neogene site). Our goal is to have the new Pennsylvanian page online by the time of the next report.

For the Ordovician part of the Paleoniches project led by Ohio University:

Ohio University:

Work continues on construction of the Ordovician Atlas of Ancient Life ([www.ordovicianatlas.org](http://www.ordovicianatlas.org)). There are now 550 published webpages within the site including 77 species and associated higher taxa.

Miami:

The Limper Museum collections at the Miami University are currently in transition. Construction is beginning on a new geology building in August, so most collections were in the process of being packed for storage during the past 2 months. Kendall is currently working with the type collection and expects major digitizing efforts to resume when classes begin again in August.

Cincinnati:

A new exhibit, "Cincinnati: Life under the sea" premiered June 14th in the Ruthven Gallery at the Union Terminal Exhibits Building. The exhibit features an interface to the PaleoNiches developed website: The Ordovician Atlas of Ancient Life ([www.ordovicianatlas.org](http://www.ordovicianatlas.org))

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

The last few months have been a continuation of the previous. Slow accumulation and verification of the required data continues and the first batch of sample data has moved to VertNet on its way to iDigBio. Once the data is approved through VertNet we shall send more. We upgraded to Specify 6.05.4 and continue to attach images to records within the database. Students had been trained and are completing much of the georeferencing and imaging. The summer graduate student is digesting a large collection of Paleogene type and figured specimens along with images. A second has scanned the field locality images of Helen Plummer. These will be attached to relevant Paleozoic records.

And finally for Yale:

Our final goal for the project was to catalog our ledgers. We have finished cataloging from the ledgers now, but the job was not as expected, since a high proportion of ledger records had already been cataloged. However, the YPM-IP collections have a large amount of uncatalogued material in our systematic collections from the focus areas in this grant. We have selected the top taxa from the Ordovician and the Pennsylvanian (50 most abundant genera from each time period) and are proceeding to digitize those. There is no shortage of material and the progress is nearly identical (greater, if anything) to what was expected from the ledger project. Since the previous (March 2014) report, we have cataloged 980 objects and have photographed 2,034 specimens (each includes images from 2-4 orientations). We have added 105 new localities and georeferenced 3,012 localities that were previously entered in the database.

### 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

Regarding KU, in other relevant news, Una Farrell will be presenting a poster at the SPNCH meeting in Wales on label imaging workflows. One of the graduate students supported by the grant, Erin Saupe, has just received her Ph.D. with honors and is going on to do a post-doc at the Yale Institute for Biospheric Studies.

Regarding Ohio University: Three new undergraduate employees are in training to continue developing the atlas over the summer term. A poster was presented at Annual IGCP 591(Early and Middle Paleozoic Revolution) meeting in Tartu, Estonia that introduced the website and digitizing project to an international community.

Regarding Texas, they co-hosted the iDigBio imaging workshop and found it to be very productive.

Regarding Yale, two EVOLutions students (Peabody Museum afterschool program for STEM education for under-served populations) are working with Butts (senior personnel) and have produced videos and are working on storyboards for the Paleoniches multimedia kiosk.

### 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)

Regarding KU, PI Bruce Lieberman Erin Saupe, along with the PI from San José State Jon Hendricks, (and a few other authors) have several papers in review describing our ecological niche modeling (ENM) relying on georeferenced and environmental data from mollusks housed in the museum of one of our partners on the grant (FLMNH).

### Attachment

N/A



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

Report submitted by: [chdietri@illinois.edu](mailto:chdietri@illinois.edu)  
Report Submitted on: 06/24/2014 - 09:47

### **Progress in Digitization Efforts**

An additional 2,304 images, mostly of slide trays from collaborating institutions, were uploaded to [invertnet.org](http://invertnet.org) and tagged with metadata, bringing the total to 16,737. Among these are the first test images of whole drawers of pinned insects. Work is continuing on building and testing 14 robotic drawer digitization systems and our technical team anticipates completing this work within the next few weeks, after which we will schedule deliveries to collaborating institutions. We successfully tested stitching algorithms on perpendicular and angled views of drawers and are now working to make these more computationally efficient and incorporate them into the drawer digitization workflow. We are also implementing a new image viewer that allows users of the website to switch between static perpendicular and angled views of the drawer.

### **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

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

Report submitted by: [p\\_sweeney@att.net](mailto:p_sweeney@att.net)  
Report Submitted on: 06/25/2014 - 12:15

### **Progress in Digitization Efforts**

Capture of collection level-information (i.e., "pre-capture") is almost complete. At this stage approximately 800,000 specimens have been pre-captured -- with at least current identification captured. The project has transitioned to the primary digitization phase. A second conveyor unit was installed at Harvard in May and light-boxes are deployed at three institutions, and primary digitization has begun at those institutions. As part of the primary digitization phase, approximately 115,000 records and 60,000 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**

nothing to report

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

We continue to collaborate with, iPlant, the FilteredPush project, the Symbiota team, and iDigBio.

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

nothing to report

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

A web-based data entry application was developed for use by institutions utilizing a light-box.

### **Attachment**

N/A

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

Report submitted by: [cgries@wisc.edu](mailto:cgries@wisc.edu)  
Report Submitted on: 06/26/2014 - 17:04

### Progress in Digitization Efforts

As of June 2014 the number for the LBCC are as follows:

Lichens: <http://lichenportal.org>

Herbaria actively submitting images or key stroked records to the portal: 57

Specimen records in portal: 1,277,112 (up by 77197 since March 2014)

Specimen records with label images: 576,065 (73474 labels have been imaged since March 2014)

Bryophytes <http://bryophyteportal.org>

Herbaria actively submitting images or key stroked records to the portal: 53

Specimen records in portal: 1,857,964 (up by 76561 since March 2014)

Specimen records with label images: 757,864 (76750 labels have been imaged since March 2014, plus about 200 have been added to existing records)

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

nothing to share

### Identify Gaps in Digitization Areas and Technology

nothing report

### Share and Identify Opportunities to Enhance Training Efforts

nothing to report

### 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)

I am attaching the annual project report summary that we just submitted to NSF.

### Attachment

<https://www.idigbio.org/sites/default/files/webform/tcn-reports/summary%20-%202014%20annual.pdf>

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

### Overall Project Summary report 2013 - 2014:

In this third year of funding several digitization centers have finished their imaging commitment to the project. Most small collections have either been completely imaged, are in the process, or were not able to participate for various reasons. Most collections are now beginning to transcribe label information and we have begun hiring independent contractors to help with this task. During the month of May, 2014 over 90 people have transcribed more than 25,000 labels.

Although the overall digitization process is the same for all institutions and well documented on our website, it is interesting to note that almost every institution seems to be using different software to accomplish the tasks of imaging, creating the skeletal records, and managing the internal imaging process (e.g., Excel, Specify, Arctos, FileMaker Pro, custom application provided by the project, KE-EMu, SilverBiology, Tropicos).

This reporting period saw the addition of over 700,000 records to the two portals (~ 290,000 lichen records in <http://lichenportal.org> and 425,000 bryophyte records in <http://bryophyteportal.org>), bringing the total number of available records to 3.1 million. Considering that there are large numbers of already digital records at collaborating institutions waiting to be loaded into Symbiota, it appears that the overall goal of the project will be reached by the end of the funding.

Approximately 140 people were involved in the project, with about half of those not being paid directly by the project. This includes senior personnel, undergraduate (>60) and graduate students, volunteers, and interns specifically mentioned in the reports from collaborating institutions.

Authority information has been improved in many areas, e.g., collector names, taxonomic nomenclature (index fungorum, International Plant Names Index), geographic names.

As the emphasis is shifting to transcriptions development has also focused on supporting this aspect of the project with functionality to rapidly incorporate exsiccata information into multiple herbaria databases through the duplicate tab searches. Major improvements to crowdsourcing have been implemented. Content is continuously being added to the volunteer website (<http://lbcc.lichenportal.org> and <http://lbcc.bryophyteportal.org>) and detailed accounting of transcriber activities is now possible to support payment of independent contractors. The Natural Language Processor is fully integrated and has proven very successful while it is continuously improved as more label formats are becoming available.

Data that were mobilized through this project have been used in 2 scientific studies (PhD research) by collaborating institutions. Aspects of the project were presented locally, and at national and international conferences, the ADBC summit, and several guest lectures at partner organizations. This year several demonstrations of the software were conducted by project participants other than the WIS core team. The annual LBCC meeting was held in Tempe, AZ in February 2014 and attended by

representatives of 10 collaborating collections. In addition representatives from 7 non-imaging collections were invited and received training in the use of Symbiota for transcription. Training in transcribing is increasingly conducted remotely by Julianne Smith at WIS and John Brinda at MO via teleconferences and online.

In addition to the already reported increase in visibility leading to more loan requests, collaborators are noting the importance of biodiversity informatics training/awareness students are receiving through this effort. As the data are becoming more accessible, the motivation to improve the collections and the data are other important outcomes of this project, as reported by several collaborators.

### **Outreach:**

Again the digitizing effort and its implications have been introduced to the public in many open houses, 'meet a scientist', and 'behind the scene' tours at several herbaria and museums. Furthermore, data have been used in formal education (e.g., ALA BIOL488 Artic Vegetation Ecology: Geobotany). Five collections are working with volunteers, including high school students and the general public, while two others have integrated the project into teacher training efforts and one provides extensive internship opportunities as part of the project.

J. Kluse (LSU) has started to collaborate with M. Basham (ASU) to develop lichen and bryophyte flashcards to be used in a 3D smart phone or tablet application. The application was developed and demonstrated for insects (<http://junaio.wordpress.com/2014/02/13/insectarium-beyond-a-bug-in-a-box/>) and provides amazing opportunities for outreach and education. Integrating lichens and bryophytes into the application poses new challenges for the developers and creative solutions will be needed.

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

Report submitted by: [bthiers@nybg.org](mailto:bthiers@nybg.org)  
Report Submitted on: 07/01/2014 - 18:34

### Progress in Digitization Efforts

We have just completed our annual report. Here is a summary of work accomplished to date:

#### Specimen and Ancillary Item Digitization

Between 1 Apr 2013 and 31 Mar 2014, 285,003 items have been digitized (includes specimens, specimen labels, photographs, field notes, field book records) at a cost of \$251,219, or an average of \$0.88 per digitized item. The cost of digitization does not include time spent training for the project, administering it, or performing outreach activities related to the project, nor does it include indirect costs, equipment or supplies.

The major product of the project is the MycoPortal. To date 1,653,361 specimen records have been added to the portal (265,762 records have been added from 35 different institutions this year). The Portal contains 97,114 skeletal records (i.e. locality data yet to be added), 1,518,472 are records with complete text locality information, and 156,821 records have complete locality data with geocoordinates. The MycoPortal also has been populated with 364,753 images (approximately 10,000 of these are living fungi, the remainder are images of dried fungi, labels, and field notes), 41 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 add

### Identify Gaps in Digitization Areas and Technology

Record completion continues to be our greatest challenge. We have proposed some changes to the Symbiota software that would help to increase the efficiency of this process. The main change requested is the ability to sort records in table view to identify blocks of records with similar data, to make full use of the carry forward function.

### Share and Identify Opportunities to Enhance Training Efforts

nothing to report

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

We are collaborating with the other TCNS represented in our institution (3, addition to Macrofungi) to engage high school students in data transcription. We are planning a day of orientation and training, and from this introduction, hope to develop a longer term relationship with 4-10 high school students who might be willing to make a commitment to enter data for an entire semester. These high school students are enrolled in summer programs at the Garden.

### Share and Identify Opportunities and Strategies for Sustainability

nothing new to report.

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

A project meeting was held at the Mycological Society of America meeting on 7 June 2014 at Michigan State University in East Lansing for 14 project participants.

### 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](mailto:Chris.neefus@unh.edu)  
Report Submitted on: 07/07/2014 - 16:05

### Progress in Digitization Efforts

Digitization is proceeding at all 6 of the project's primary digitization centers and at all of the other digitizing institutions that were scheduled to state in Year 1. Currently 152,478 digitized specimens are available via the project portal ([macroalgae.org](http://macroalgae.org)).

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

Presentations on the project and workflow have been recently made at SPNHC in Cardiff (June 22-27, 2014) and at a Consortium of Northeast Herbaria (CNH) meeting held in conjunction with the Canadian Botanical Association annual meeting in Montreal (June 14-18, 2014). All three presentations are available on our project's documentation website ([macroalgae.unh.edu](http://macroalgae.unh.edu))

### Identify Gaps in Digitization Areas and Technology

No gaps, but we discovered a couple of bugs in the georeferencing module of Symbiota that we reported to Ed Gilbert. 1. When a locality string for an occurrence record contains a # or a !, the locality string that Symbiota passes to GeoLocate is truncated at that point. 2. In the Batch Georeferencing tool, the Datum field doesn't get filled in automatically. 3. In the Batch Georeferencing tool if you use the option to copy coordinates from a previously georeferenced record, it enters the error radius from the previous record in the datum field of the new records.

### Share and Identify Opportunities to Enhance Training Efforts

We have been talking about doing a short training video on imaging exsiccatae.

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

A number of the PIs and other personnel in the Macroalgae project are also involved in the other herbarium based TCNs (Lichens and Bryophytes, Macrofungi, and NE Vascular Plants) so we share experiences across projects.

### 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

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

Report submitted by: [Dena@colorado.edu](mailto:Dena@colorado.edu)  
Report Submitted on: No report provided.

### **Progress in Digitization Efforts**

No report provided.

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

No report provided.

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

No report provided.

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

No report provided.

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

No report provided.

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

No report provided.

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

No report provided.

### **Attachment**

N/A