

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: 01/14/2015 - 18:26

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_January_2015.docx

Southwest Collections of Arthropods Network Update
March 3, 2015
Neil Cobb

Progress in Digitization Efforts:

We are on target to meet our third-year quota for digitizing labels from pinned specimens (525,349 specimens of ground-dwelling arthropod taxa). Table 1 presents three sets of statistics as of December 1, 2014 from our data portal. These include data 1) from institutions that are funded by the NSF-ADBC program; 2) institutions that are entering 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 complete information to persons that are considering research projects. Although we have already surpassed our goal of 525,349 specimen records, we have not thoroughly reviewed the 746,410 records that SCAN-funded museums have produced to determine how many of those strictly ground-dwelling arthropods, but we expect that 70% or those are target taxa and that we only need ~4,000 more records to meet our project goal by July 1, 2015.

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
# Specimen Records	746,410	134,578	880,988	5,371,762
# Georeferenced	544,874	52,686	597,560	4,342,496
# Identified to species	470,785	66,356	537,141	3,098,473
# Families	718	362	776	1,450
# Genera	7,037	3,463	8,923	20,014
# Species	16,569	7,016	21,558	58,709
% Georeferenced	73%	39%	67%	81%
% Identified to Species	63%	49%	61%	58%

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 is to produce 15,125

high-resolution 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 continue to greatly increase our productivity over the fall, 2014. Three museums are producing low-resolution images (University of Hawaii, University of Arizona, and Texas Tech University). Texas Tech University has produced about 3,000 high-resolution images as part of their 25,238 images uploaded. We currently have 9,527 high-resolution images (out of 95,805 total images) and we will continue to focus resources towards the continued imaging of exemplar specimens.

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
Arizona State University	2,006
Colorado State University	49
Northern Arizona University	1,020
Denver Museum of Nature and Science	624
University of New Mexico	135
Northern Arizona University - NPS	673
New Mexico State University	1,384
Texas Tech University (mostly low-res images)	25,238
University of Arizona (low res images)	48,98
University of Hawaii (low res images)	636
University of Colorado	15,242
TOTAL	95,805

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 Entomological Collections Network meeting November 15, 2014.

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 continue to provide North American data we have obtained from other sources to increase the quantity of data available to SCAN users. 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 4.9 million records for 58,709 species.

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: 01/30/2015 - 10:22

Progress in Digitization Efforts

The AMNH Paleontological Collection (PEC) initially entered and further modified over 5000 collection objects in amber from three major localities (the Burmese, New Jersey and Baltic deposits). Additionally, over 5,500 new records of fossil arthropods from the Dominican Republic have been databased, with 500 entries since Dec 1st. All of these are being exported to the recently updated Specify database for consistent data entry from volunteers and students. Editing will continue to be done by PI, D. Grimaldi, and by Curatorial Specialist, P. Nascimbene, with taxonomic names/terms/fields revised where necessary. PEC has now accumulated more than 4000 images of individual inclusions in amber (spanning 75 drawers of specimens), with two volunteers assisting in specimen preparation and photomicrography.

CU-Boulder Invertebrate Paleontology added 302 new specimen records to its Specify database since Dec 1st and is working to convert collection and determination dates from co-opted text fields into a standard ISO format within Specify. The initial conversion should be completed by the end of Jan 2015, though manual revision of partial dates will subsequently be needed. Similar updates and data validation are also being performed on the "Type Status" field to match current standard database formats. An additional 1,037 images have also been acquired.

The Harvard Museum of Comparative Zoology (MCZ) has taken 1350 additional images from 1200 collection specimens since Dec 1st. An additional 200 new catalog numbers have been assigned to unnumbered fossils found as the collection is imaged. MCZ has also completed databasing of F.M. Carpenter's alphabetically organized bibliographic collection on fossil insects by adding the final 200 entries.

The Virginia Museum of Natural History (VMNH) has taken a hiatus from insect digitization over the last few months due to a salvage operation of the Triassic insect-bearing beds at the Solite Quarry, which will ultimately result in the collection of a large number of additional insects. The salvage operation was funded by National Geographic Society, and VMNH has provided additional funding to extend the Paleontology Technician position for an additional year, to ensure completion of the insect digitization project.

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

PEC has updated record dating and georeferencing. The most accurate date for individual amber deposits, based on radiometric dating where possible, is being shared across institutions to standardize specimen dating. Also, utilizing the GeoLocate function in Specify, information is being entered for entire deposits, giving access to an interactive map, precise coordinates, and range/margin of error, which can be applied to future collection objects entered into the database. This also allows georeferencing information to be standardized and shared across institutions.

Identify Gaps in Digitization Areas and Technology

PEC found it necessary to utilize an updated version of Specify that allows associations between collection objects and localities (called the Swedish model), given that typically, all specimens from individual amber localities share the same geological and geographical information.

Share and Identify Opportunities to Enhance Training Efforts

The Division of Invertebrate Zoology at AMNH will host the next TCN Workshop/Meeting on February 26 and 27, 2015. There will be presentations and demonstrations / workshops on the digitization of fossils in amber, including preparation, imaging, etc.

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

CU-Boulder PI, D. Smith, is co-organizing a session with D. Paul and V. Blagoderov on digitization for the 2015 Society for the Preservation of Natural History Collections (SPNHC) annual meeting in Gainesville, FL. Co-PI T. Karim and digitization specialist, G. Nelson, have also submitted a proposal to iDigBio and the Geological Society of America (GSA) for a digitization session at the 2015 GSA annual meeting in Baltimore, MD.

Share and Identify Opportunities and Strategies for Sustainability

There is nothing new to report.

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

The fossil insect collection and digitization workspace at MCZ was moved at the end of December to the Northwest Building, in which the rest of MCZ's paleontology invertebrate collections are now kept. This move is included in a general restructuring that the MCZ is suffering in order to accommodate future changes. The new space is better equipped in terms of space and fossil preparation, providing, for instance, a polisher that can be used to remove fractured surfaces in old amber samples to improve the visibility of their inclusions prior to imaging them. During the move, MCZ transferred and pre-cleaned "rediscovered" fossil insect material that had been stored apart from the rest of the collection, originally inside wooden drawers found in an industrial unit in Cambridge. The current MCZ entomology curators were unaware of the material. Rough accounts estimate that this new material represents about 20,000 fossil insect specimens, namely rock fossils from the Permian of Kansas and Elmo and the Eocene Green River Fm, which had been excavated and/or curated by Prof. F.M. Carpenter, former curator of fossil insects at the MCZ and who passed away two decades ago. Preliminary inspections of this material show that most of the samples are in excellent condition and have a great scientific potential. Although this material was not in our current records and lacks modern catalog numbers, it shows some degree of processing, as most of the specimens are marked on the rock surfaces and identification numbers were assigned to them. Some specimens are even taxonomically determined. Most of the samples were kept (even stacked up) inside the wooden drawers without being boxed or multiple of them were stored in dusty cardboard boxes, so often samples were exposed to suffer damage by friction. Samples were thus transferred to plastic boxes or their cardboard boxes were cleaned. Mechanical protection was ensured preventing direct contact between fossiliferous surfaces, often using plastic or paper layers. Boxes were stored as in the rest of collection, i.e., in cardboard trays that are kept inside metal drawers placed into metal cabinets. Nine cabinets were required, which joined the official MCZ collection.

Attachment

Nothing to report.

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

Report submitted by: eshea@delmnh.org
Report Submitted on: 02/02/2015 - 13:58

Progress in Digitization Efforts

Field: Generated data entry spreadsheet with North American geographic file, drop-down menus to county level. Insects: designed and tested complete workflow and taxonomic authority file (spread sheet with drop down menus) for papered and pinned FMNH Odonata data entry, uploaded workflow and Odonata taxonomic authority file on InvertEBase google drive. Odonata digitization in progress, close to 1,000 Odonata digitized, barcodes are added, through pre-curation major improvement of FMNH Odonata collection. Invertebrates: taxonomic authority file under development. Identified mollusk collection units for digitization.

Cleveland: We consulted with staff at C.A Triplehorn Insect Collection, Ohio State University who maintain database and server, a well-developed 2-step digitization protocol and full-time Biodiversity Informatics Manager. We have since transcribed label data from more than 1,500 mantid specimens which approximately half have been entered in to the OSU DEA2 database. We are starting digitization efforts on Hymenoptera with a newly installed thermal printer generating barcode labels.

Auburn: Additional two undergraduate students hired. Total of 26,109 specimens digitized; currently working through Hemiptera.

Michigan: Three undergraduate students were newly hired in January making total six undergraduates working on data entry. Total 2917 records representing three freshwater snail families (Lymnaeidae: 403; Planorbidae: 632; Pysidae: 953; Valvatidae: 292) and one bivalve family Sphaeriidae (637) were entered during 22 Nov. 2014- 22 Jan. 2015.

Delaware: We have reached out to Specify personnel to discuss timeline for DMNH Mollusk database creation and how the two new major upgrade releases (Specify 6.6 and Specify 7) will impact our conversion.

Frost: We finished scanning the sucking louse collection (Insecta: Anoplura), just over 15,000 slides, using InvertNet's slide-scanning protocol. The images are stored locally on 4 redundant hard drives, pending transfer to InvertNet. We also finalized our dragonfly envelope imaging strategy and have begun imaging our synoptic collection (stored separately from the rest). More than 1,800 Odonata have been imaged.

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

Field: purchased and/or tested all equipment: barcode cutter, stage for pinned insect for quick label imaging, photographic station for label imaging tested

Cleveland: nothing to report

Auburn: nothing to report

Michigan: nothing to report

Delaware: nothing to report

Frost: After experimenting with a 16-cell dragonfly template and a single specimen template we've determined to image the specimens one at a time (less error prone, same time per specimen). An image of the template is available online, and we are running tests that use scripts to separate fields (identifier, specimen, metadata, color standard). Results will inform our future efforts.

Identify Gaps in Digitization Areas and Technology

Field: label imaging techniques to be demonstrated via video conference in March to resolve label imaging issues

Cleveland: We have been working to resolve label imaging procedure, but have not settled on an optimal protocol.

Auburn: nothing to report

Michigan: nothing to report

Delaware: We have received an updated quote for a server to house collections database and identified technology deficits in DMNH computing architecture that will need to be resolved prior to installation.

Frost: Lighting for our Odonata imaging is sufficient for digitization, but needs to be brighter to make the images more useful for research beyond label transcription. We are using four 100w-equivalent daylight spectrum compact fluorescent bulbs. We are likely to move to a flash-based workflow starting late January 2015.

Share and Identify Opportunities to Enhance Training Efforts

Field: Insects digitization person hired, prospective student summer intern interviewed

Cleveland: Three work-study students from Case Western Reserve University were hired and are assisting in digitization.

Auburn: Additional 2 undergraduates hired and trained.

Michigan: Nothing to report.

Delaware: We continue to seek out opportunities for new hire and CM to attend web meetings regarding Specify.

Frost: We are advertising within the Penn State system to bring in undergraduate researchers. Selected participants will assist in digitizing the collection, while also addressing basic research questions. A curriculum is being developed for this activity (we expect it to take awhile to develop).

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

Field: Insects and Invertebrates: Sierwald and Bieler further developing InvertEBase Symbiota portal

Cleveland: We are working with the Biodiversity Informatics Manager, Joe Cora, at OSU reporting issues experienced to help develop their database further.

Auburn: Nothing to report

Michigan: The University of Michigan participants are collaborating with the Great Lakes Invasives TCN to make sure that data flows to both projects.

Delaware: Nothing to report.

Frost: We are working with the Speciesfile group to parse data from the Odonata images. Scripts have already been written, and tests are ongoing. The slide scans will also be part of this process, in collaboration with Julie Allen (INHS).

Share and Identify Opportunities and Strategies for Sustainability

Field: Nothing to report

Cleveland: Nothing to report.

Auburn: Nothing to report

Michigan: Nothing to report.

Delaware: Nothing to report.

Frost: Nothing to report.

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

Field: Nothing to report

Cleveland: A thermal printer was purchased for printing barcodes.

Auburn: Nothing to report

Michigan: Nothing to report.

Delaware: We have a new grant-supported staff member starting on/about Feb. 1

Frost: Nothing to report.

Attachment

Nothing to report.

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

Report submitted by: cgries@wisc.edu
Report Submitted on: 02/05/2015 - 16:57

Progress in Digitization Efforts

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

Lichens:

<http://lichenportal.org>

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

Specimen records in portal: 1,722,502 (up by 128,497 since November 2014)

Specimen records with images: 660,324 (40,239 labels have been imaged since November 2014)

Records with locality information: 1,398,530 (124,126 locality information where added since November 2014)

Bryophytes:

<http://bryophyteportal.org>

Herbaria actively submitting images or key stroked records to the portal: 59

Specimen records in portal: 2,075,953 (up by 36,236 since November 2014)

Specimen records with images: 952,375 (26,713 labels have been imaged since November 2014)

Records with locality information: 1,264,019 (49,918 locality information where added since November 2014)

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

Our digitization coordinator participated in the workshop entitled 'Workflows for Herbarium Digitization' and contributed to a composite set of workflow documents that capture the best of what has been learned in our community over the past few years of ADBC activity.

Identify Gaps in Digitization Areas and Technology

none 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

Nothing to report.

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: 02/06/2015 - 17:00

Progress in Digitization Efforts

see attachment

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

see attachment

Identify Gaps in Digitization Areas and Technology

see attachment

Share and Identify Opportunities to Enhance Training Efforts

see attachment

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

see attachment

Share and Identify Opportunities and Strategies for Sustainability

see attachment

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

see attachment

Attachment

https://www.idigbio.org/sites/default/files/webform/tcn-reports/GLITCN_progress%20report_Jan2015_0.pdf

GREAT LAKES INVASIVES TCN – Bi-monthly report Dec 1, 2014 – Jan 30, 2015

Second GLI TCN report, representing five months' of effort to date.

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: 3,923 (NY) + 4,000 (ILLS) + 264 (MINN) = **8,187**

Barcoded and Imaged to Date: 27,536 (WIS) + 12,243 (NY) + 12,452 (OSU) + 163 (MINN) + 8,674 (MICH) + 159 (ILLS) + 2,695 (F) + 4880 (MU) = **68,802**

Databased to Date: 45,401 (WIS) + 8,110 (NY) + 27,500 (MINN) + 15,127 (MICH) + 27,000 (ILLS) + 9,478 (F) = **132,616**

Uploaded to the GLI Portal directly or to another Symbiota Portal for editing before transfer to GLI Portal: 45,401 (WIS) + 7,849 (MICH) + 5,783 (F) + 42 (MINN) + 9,624 (MOR) + 9,804 (ALBC) = **78,503**

MOLLUSKS:

Barcoded and Imaged to Date: **1,617** lots (MICH) have been imaged, representing 3 genera and 5 species.

Databased to Date: **6,594** records added by MICH, representing 6 genera and 92 species.

Uploaded to the GLI Portal or another Symbiota Portal: 2,341 images from MICH have been processed and, of these, **1,404** have been uploaded to the GLI portal.

FISH:

Databased to Date: 27,145 (ILLS) + 81,324 specimens [in 4,709 lots (F)] = **108,469**

Georeferenced: 25,000 (ILLS)

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

We continue to experiment with alcohol resistant barcode options. Researchers at Ohio State have learned that after several months, the adhesive backing on standard vinyl barcodes used by the herbarium has not dissolved.

Our plant workflow offers efficiencies that take advantage of duplicates from institutions that are not funded through our TCN. Records (complete or skeletal) and images are uploaded mostly to the larger Midwest Consortium of Herbaria Symbiota Portal, a subset of the larger SEINET portal. Transcription and Georeferencing takes place there prior to the completed record being migrated into the GLI Portal. This offers a higher probability of encountering a duplicate in the nationwide portal.

Lesson learned: there is no substitute for a face-to-face meeting, especially at the start of a collaborative project that crosses disciplines. At their own personal expense, 20 participants from MN, WI, IL, MI, & OH traveled to the Field Museum in Chicago on December 15 for a TCN workshop. This was a critically important meeting that brought the participating botanists and zoologists together – many for the very first time, and possibly last. A request for minimal financial support for the workshop to iDigBio was rejected. To their great credit the Field Museum covered the cost of parking and lunch for the participants, instead; the TCN is extremely grateful for this.

3) Identify Gaps in Digitization Areas and Technology

Still trying to resolve best practices for physical application to and use of barcodes on liquid preserved fish and 3D mollusk specimens. Their use is not routine among these collection managers.

4) Share and Identify Opportunities to Enhance Training Efforts

Having current employees train new hires is efficient, and also beneficial for constantly revising and updating the workflow with potential improvements.

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

Project managers have discussed mollusk imaging with Paul Callomon at Philadelphia Academy of Natural Sciences.

We have loaned a digitization workstation to participants at UW-Milwaukee, thereby starting to bring in some of the smaller but important institutions.

6) Share and Identify Opportunities and Strategies for Sustainability

Nothing to report

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

We have constructed and implemented a custom 'app' that allows for the creation of skeletal records of imaged data before uploading into Symbiota in order to accelerate the transcription process. It is available to downloading on our project website together with installation and use instructions.

University of Michigan PI Rabeler represented the project at an iDigBio Herbarium Workflows workshop at Valdosta State University, 26-30 January 2015.

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

Report submitted by: blieber@ku.edu
Report Submitted on: 02/07/2015 - 13:18

Progress in Digitization Efforts

Paleoniches Update, February 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 151,873 specimens databased. Of these, there are a total of 147,550 specimens databased that have clean, proofed localities. Further, we now have a total of 122,714 specimens that are georeferenced. We have now completed databasing all of the cephalopods we aimed to database. We are almost finished databasing our bivalves and have begun databasing our gastropods. In addition, a total of 8,114 localities have been georeferenced. Thus we have effectively completed the entire georeferencing component of our proposed work. We have also since the last update to iDigBio significantly expanded the number of species that have been imaged, especially focusing on adding crinoids and both articulate and inarticulate brachiopods.

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

In the last two months, several new genus and species pages have been released on the Ordovician Atlas website. The addition of maps to live pages is complete, and we continue to put maps on pages that are not currently live. In addition, members of the Stigall lab visited the Cincinnati Museum Center in December to photograph ~30 specimens for the website. Currently, the website has 130 species pages live, and all the cephalopods, arthropods, brachiopods, crinoids, graptolites, and corals are now live. We are currently focusing our efforts to finish and make live the gastropod, rhombiferan, and bivalve species pages.

Miami University

During December 2014 and January 2015, there has been no activity due to the fact that no students were available to work during the holiday break and Miami's J-term (January term).

Cincinnati Museum Center

Since the middle of November, Brenda Hunda's University of Cincinnati student intern Ian MacAdam, has been focusing on georeferencing locality records. He has worked a total of 76.5 hours and has georeferenced 727 locality records, with his cumulative total of 1260 sites. His work to date has resulted in 6621 catalogue records with georeferenced site data. For the project as a whole to date 2,602 site records have been georeferenced resulting in 19,533 georeferenced catalogue records in Ke Emu. This constitutes georeferencing of 28.75% of our Invertebrate Paleontology collection.

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

Since the last update, PI Hendricks (San José State University; SJSU)—in collaboration with Invertebrate Paleontology staff at the Florida Museum of Natural History—has continued to develop and add content to the Neogene Atlas of Ancient Life. In particular, seven additional families of gastropods have been added since the last updated (distributional maps for these will be published online soon). Species-level pages are now online for 332 species (out of 500 planned pages). Hendricks is currently recruiting a new undergraduate student assistant to help with the development of content for the Pennsylvanian Atlas. The new student assistant will be hired, trained, and actively generating new web content for the Pennsylvanian Atlas by the time of the next update.

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

1. We have a data set with VertNet awaiting their re-processing prior to iDigBio access. Data was run through the migrator and we are now expecting a rerun with relevant tweaks to the migrator to accommodate our data set. Ironically the dataset currently with VertNet is now somewhat behind the times so we'll need to send an update!
2. Georeferencing is at 32%
3. About 3300 image attachments have been made. These include whole drawer images, specimen images, labels, field images, and notebooks. We have to pass through two updates of the database before we can batch upload the bulk of the images files. We expect this to happen in the next month.
4. The deep time version of PaleoCentral is ready, Tomislav Urban is bullet proofing at the moment.

And at Yale: From PI Susan Butts:

We are working on digitizing the most abundant taxa from the Ordovician and the Pennsylvanian (50 most abundant genera from each time period) and are proceeding to digitize that material from our systematic collection. We have done at least one drawer of the top 50 most abundant Ordovician fossils and at least 18 of the top 50 Pennsylvanian drawers (5 are in progress). At this rate, we will certainly achieve our digitization goals by the end of this project (June 2015). We have roughly 550 specimen images (most with three views) awaiting upload to the database.

Since the previous report, we have modified or inserted 1,202 records in KE EMu. We have attached photos to 1,113 additional records (1-3 photos per specimen). All items digitized via our PEN are now available to iDigBio via the Peabody IPT.

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

N/A

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 submitted for peer review to *Palaeontologia Electronica* (an open-access online journal) on December 18, 2014. This manuscript provides an overview of our Digital Atlas project and goals for the paleontological community.

Attachment

Nothing to report.

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

Report submitted by: barbara.thiers@gmail.com
Report Submitted on: 02/08/2015 - 11:23

Progress in Digitization Efforts

Since our last report (December 2014), 90,530 specimen records have been added to the Portal. This represents about 80% of the records actually created during this period. The total number of records available for searching in the MycoPortal stands at 1,803,604, approximately 400,000 more than promised in the original proposal. Fourteen of the participating institutions have completed their digitization work. We estimate that there remain to be digitized about 250,000 images across the 21 institutions still participating institutions.

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

Nothing new to report.

Identify Gaps in Digitization Areas and Technology

We have identified as many gaps as we can reasonably address in this proposal.

Share and Identify Opportunities to Enhance Training Efforts

Nothing new to report.

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

Using funds from the Macrofungi grant, we have paid for some additional work that will benefit all Symbiota portals. It is now possible to sort records in edit view on any indexed field, which will greatly expedite record completion/

Share and Identify Opportunities and Strategies for Sustainability

The most important step toward sustainability that we can make is to continue the digitization to include all fungi. A proposal to ADBC is pending.

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

Plans are in development for a large presence for the MaCC project at Botany 2015. There will be a workshop for attendees, and also a high school teacher's workshop and special field trip for high school teachers.

Attachment

Nothing to report.

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

Report submitted by: chdietri@illinois.edu
Report Submitted on: 02/09/2015 - 09:06

Progress in Digitization Efforts

Software for the InvertNet whole drawer digitization system was upgraded to include feature-point-based stitching software written by co-PI John Hart that automatically stitches sets of images to create whole-drawer gigapixel-scale zoomable panoramas in near real time (~3 minutes per image set) on the local computer that also controls the robot and captures and saves the raw images to disk. This upgrade has been installed on the production system at INHS and will be installed remotely on the systems at collaborating institutions within the next few weeks. We are also upgrading camera head hardware to address some camera motion issues and improve the speed of image stitching. Replacement parts are being manufactured and should be ready to send to collaborators with the new software in the next few weeks. Images and metadata for approximately 100 drawers have been captured at INHS and are in the queue for upload to invertnet.org. Collaborators at most institutions have now scanned at least a few of their drawers but some have opted to wait until we can provide the above mentioned upgrades before proceeding further.

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

Nothing to report.

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

Report submitted by: adeans@gmail.com
Report Submitted on: 02/09/2015 - 09:59

Progress in Digitization Efforts

Field: Generated data entry spreadsheet with North American geographic file, drop-down menus to county level. Insects: designed and tested complete workflow and taxonomic authority file (spread sheet with drop down menus) for papered and pinned FMNH Odonata data entry, uploaded workflow and Odonata taxonomic authority file on InverteBase google drive. Odonata digitization in progress, close to 1,000 Odonata digitized, barcodes are added, through pre-curation major improvement of FMNH Odonata collection. Invertebrates: taxonomic authority file under development. Identified mollusk collection units for digitization
Cleveland: We consulted with staff at C.A Triplehorn Insect Collection, Ohio State University who maintain database and server, a well-developed 2-step digitization protocol and full-time Biodiversity Informatics Manager. We have since transcribed label data from more than 1,500 mantid specimens which approximately half have been entered in to the OSU DEA2 database. We are starting digitization efforts on Hymenoptera with a newly installed thermal printer generating barcode labels.
Auburn: Additional two undergraduate students hired. Total of 26,109 specimens digitized; currently working through Hemiptera.
Michigan: Three undergraduate students were newly hired in January making total six undergraduates working on data entry. Total 2917 records representing three freshwater snail families (Lymnaeidae: 403; Planorbidae: 632; Pysidae: 953; Valvatidae: 292) and one bivalve family Sphaeriidae (637) were entered during 22 Nov. 2014- 22 Jan. 2015.
Delaware: We have reached out to Specify personnel to discuss timeline for DMNH Mollusk database creation and how the two new major upgrade releases (Specify 6.6 and Specify 7) will impact our conversion.
Frost: We finished scanning the sucking louse collection (Insecta: Anoplura), just over 15,000 slides, using InvertNet's slide-scanning protocol. The images are stored locally on 4 redundant hard drives, pending transfer to InvertNet. We also finalized our dragonfly envelope imaging strategy and have begun imaging our synoptic collection (stored separately from the rest). More than 1,800 Odonata have been imaged.

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

Field: purchased and/or tested all equipment: barcode cutter, stage for pinned insect for quick label imaging, photographic station for label imaging tested
Cleveland: nothing to report
Auburn: nothing to report
Michigan: nothing to report
Delaware: nothing to report
Frost: After experimenting with a 16-cell dragonfly template and a single specimen template we've determined to image the specimens one at a time (less error prone, same time per specimen). An image of the template is available online, and we are running tests that use scripts to separate fields (identifier, specimen, metadata, color standard). Results will inform our future efforts.

Identify Gaps in Digitization Areas and Technology

Field: label imaging techniques to be demonstrated via video conference in March to resolve label imaging issues
Cleveland: We have been working to resolve label imaging procedure, but have not settled on an optimal protocol.
Auburn: nothing to report
Michigan: nothing to report
Delaware: We have received an updated quote for a server to house collections database and identified technology deficits in DMNH computing architecture that will need to be resolved prior to installation.
Frost: Lighting for our Odonata imaging is sufficient for digitization, but needs to be brighter to make the images more useful for research beyond label transcription. We are using four 100w-equivalent daylight spectrum compact fluorescent bulbs. We are likely to move to a flash-based workflow starting late January 2015.

Share and Identify Opportunities to Enhance Training Efforts

Field: Insects digitization person hired, prospective student summer intern interviewed
Cleveland: Three work-study students from Case Western Reserve University were hired and are assisting in digitization.
Auburn: Additional 2 undergraduates hired and trained.
Michigan: Nothing to report.

Delaware: We continue to seek out opportunities for new hire and CM to attend web meetings regarding Specify.

Frost: We are advertising within the Penn State system to bring in undergraduate researchers. Selected participants will assist in digitizing the collection, while also addressing basic research questions. A curriculum is being developed for this activity (we expect it to take awhile to develop).

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

Field: Insects and Invertebrates: Sierwald and Bieler further developing InvertEBase Symbiota portal

Cleveland: We are working with the Biodiversity Informatics Manager, Joe Cora, at OSU reporting issues experienced to help develop their database further.

Auburn: Nothing to report

Michigan: The University of Michigan participants are collaborating with the Great Lakes Invasives TCN to make sure that data flows to both projects.

Delaware: Nothing to report.

Frost: We are working with the Speciesfile group to parse data from the Odonata images. Scripts have already been written, and tests are ongoing. The slide scans will also be part of this process, in collaboration with Julie Allen (INHS).

Share and Identify Opportunities and Strategies for Sustainability

Field: Nothing to report

Cleveland: Nothing to report.

Auburn: Nothing to report

Michigan: Nothing to report.

Delaware: Nothing to report.

Frost: Nothing to report.

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

Field: Nothing to report

Cleveland: A thermal printer was purchased for printing barcodes.

Auburn: Nothing to report

Michigan: Nothing to report.

Delaware: We have a new grant-supported staff member starting on/about Feb. 1

Frost: Nothing to report.

Attachment

Nothing to report.

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

Report submitted by: moon@begoniasociety.org
Report Submitted on: 02/09/2015 - 10:05

Progress in Digitization Efforts

Plant Digitization Numbers:

Total Skeletal records completed in Symbiota during the course of the project: 81,734 (NY = 50,702)

Total Complete Records = 1,248,123

Total Specimens Imaged = 976,658 (NY = 237,604)

Total Images uploaded to iDigBio = 435,265

Insect Digitization Numbers:

Total Complete Records = 1,211,000

The TTD-TCN expects to continue digitizing at a high rate until summer 2015. Afterward, we will focus on data cleaning, data sharing and georeferencing until the final day of the project (December 31, 2015).

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

The TTD-TCN has been using the iDigBio portal in several research efforts. From these we produced a list of improvements to share with the group. We suggest a general call for IAC members to provide detailed feedback.

Download format and term definitions

The columns after download are not in logical order. All columns that are identifiers should be clustered together, locality information clustered together, collecting event clustered etc. Within the clusters the data elements can be in a loose order, but the elements should be together.

Several terms are included in the download that represent the same information, but are named only slightly different (ex. VerbatimEventDate, verbatimEventDate). These should be merged in the download file or at least returned next to each other in the download file.

There is no document that defines the terms. One should be provided. Further, those definitions should have URI identifiers so that individuals can reuse them with confidence (including them in a meta.xml).

Portal behavior

When searching the portal, certain fields should not be exact match. These include Collector and Locality fields. There are others, but these were the most limiting.

Higher taxonomy should be included to improve the search. Family name being the most important. If it is not in the dataset from the provider, it should automatically be added upon ingestion to iDigBio. Without the higher taxonomy, a user will miss specimen records they are likely looking for.

Minor issues

Terms should be evaluated for continuity. The term "row number" contains a space.

Ideally would like a tsv as well as a csv download.

Identify Gaps in Digitization Areas and Technology

Nothing to report this period.

Share and Identify Opportunities to Enhance Training Efforts

Mari Roberts (NYBG) is pushing forward with using volunteers for transcribing records. Right now she has total of 4 volunteers coming to the Garden to transcribe labels and 2 volunteers who transcribe from home. She is attending local career fairs to increase this number.

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

Katja Seltmann (AMNH) has been working with iDigBio to organize the Field to Database and Data Management workshops.

Share and Identify Opportunities and Strategies for Sustainability

I am not sure how this translates into sustainability, but on a number of occasions now we have been asked the question "Do you have an API?". Our answer has been "No, but iDigBio does". A robust API is software that is difficult to maintain, and one of t

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

Nothing to report this period.

Attachment

Nothing to report.

MOBILIZING NEW ENGLAND VASCULAR PLANT SPECIMEN DATA TO TRACK ENVIRONMENTAL CHANGE

Report submitted by: p_sweeney@att.net
Report Submitted on: 02/09/2015 - 10:39

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 280,000 records and 231,000 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. We are collaborating with Melody Bashram (U. of AZ), iDigBio, and other TCNs to develop a Augmented Reality tool that will be useful in K-12 education. PI Sweeney participated in the Herbarium Digitization Workflows workshop (Jan 2015) that was hosted by iDigBio and SERNEC. NEVP advised SERNEC on their overall data workflow, particularly aspects related to the flow of data to and from iPlant.

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

Nothing to report.

DEVELOPING A CENTRALIZED DIGITAL ARCHIVE OF VOUCHERED ANIMAL COMMUNICATION SIGNALS

Report submitted by: msw244@cornell.edu
Report Submitted on: 02/09/2015 - 12:46

Progress in Digitization Efforts

Our TCN project has now digitized over 13,000 audio recordings from several different 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:

The collection of anuran recordings from famed herpetologist William Duellman is now completely archived: 1,334 total archived recordings, with 957 associated physical specimens; 500+ of these recordings have been archived since November 2014 update.

We have initiated digitization/archival work for recordings from the David Weissman Orthopteran Collection: 526 recordings archived in December and January.

We have also initiated digitization/archival work for the AMNH anuran collection: 350+ Rex Cocroft recordings archived in December and January.

We have now digitized the first 100 tapes (out of 1,040) of anuran recordings from Texas Natural History Collections.

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. In particular, Matthew Medler (Cornell), Rafe Brown (Kansas Univ) and Robin Abraham (also KU) participated in the iDigBio summit in late October 2014.

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

In January, Macaulay Library staff visited partner institution Museu Paraense Emílio Goeldi to retrieve analog (48 cassettes) and digital recordings for digitization and archival.

Share and Identify Opportunities and Strategies for Sustainability

National-level reporting of iDigBio achievements, heretofore resources now available to the public.

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

Nothing to report.

Attachment

Nothing to report.

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: 02/11/2015 - 11:50

Progress in Digitization Efforts

See attached chart

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

Several members of the Macroalgal TCN participated in the recent Herbarium Digitization Workflow workshop at Valdosta State University. The participants are updating the workflows published on the iDigBio website and plan to finish by the end of March.

Identify Gaps in Digitization Areas and Technology

nothing to report

Share and Identify Opportunities to Enhance Training Efforts

See the first entry above regarding the recent Herbarium Digitization workshop.

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

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

Digitizing Institution	Start	Collections	Specimens	Percent Complete				
				Records Created	On Portal	Imaged	Fully Transcribed	Geo-referenced
University of New Hampshire	Year 1	10	131,468					
New York Botanical Garden	Year 1	5	163,350					
University of North Carolina	Year 1	7	74,166					
University of Michigan	Year 1	5	91,683					
University of Washington	Year 1	3	37,154					
Duke University	Year 1	1	19,000					
University of Alaska	Year 1	1	8,300					
Bishop Museum	Year 1	1	78,795					
Field Museum	Year 1	1	37,494					
Oregon State University	Year 1	1	9,000					
University of Guam	Year 1	1	13,600					
University of California - Berkeley	Year 2	9	228,764					
University of Hawaii	Year 2	1	2,000					
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	1,085,274	527,387	422,658	311,399	222,656	134,935