Developing a Centralized Digital Archive of Vouchered Animal Communication Signals

Report submitted by: msw244@cornell.edu
Report Submitted on: 11/25/2014 - 10:37

Progress in Digitization Efforts

Our TCN project has now digitized over 11,500 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:

(1) Kansas University Biodiversity Institute (ornithology). We have recently digitized 228 recordings of analog audio tape recordings from the Philippines and Mognola, collected by KU researcher Peter Hosner. With this addition, all of the analog recordings from the KU Ornithology group have been digitized, including material from recordists/collectors Hosner, Robbins, and Andersen.

(2) Anuran audio recordings. We have made substantial progress on digitizing analog audio recordings from a number of different TCN partners. First, we have made substantial progress on digitizing the very large body of neotropical frog recordings collected by William Duellman (University of Kansas); to date 826 of these have been digitized and archived in the Macaulay Collection (550 newly archived since the last report). We have also digitized 125 hours of anuran material from the Smithsonian Institution, and this material is in the process of being archived at the Macaulay Library. Finally, we have received and accessioned 1,040 analog tapes from the Texas Natural History Collections, to be digitized in the near future.

(3) Insect recordings. Macaulay Library staff have now received the first 20 open-reel tapes of orthopteran recordings from researcher David Weissman. This material will be digitized in the near future, and will be associated with specimens deposited at the California Academy of Sciences.

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 formats 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

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.

iDigBio is funded by a grant from the National Science Foundation's Advancing Digitization of Biodiversity Collections Program (#EF1115210)

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

Nothing to report.

Attachment

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

Report submitted by: blieber@ku.edu
Report Submitted on: 11/25/2014 - 13:01

Progress in Digitization Efforts

Paleoniches Update, November 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 143,294 specimens databased. Of these, there are a total of 138,393 specimens databased that have clean, proofed localities. Further, we now have a total of 115,418 specimens that are georeferenced. We are now close to completing all of the cephalopods and bivalves we aimed to database. In addition, a total of 8,007 localities have been georeferenced, meaning that we have effectively completed the entire georeferencing component of our proposed work.

Further, PI's B. Lieberman, J. Hendricks, and co-PI J. Beach have continued to work with the developer of the Paleoniches iPad Atlas application (Rod Spears Consulting) and they have designed a prototype. A set of screen shots of the prototype are attached as a pdf.

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

Progress was made on two major initiatives during this quarter: georeferenced data were ingested into iDigBio and GBIF and geoferenced species locality maps were deployed within the Ordovician Atlas.

We are excited to have to successfully exported the Ohio University Invertebrate Paleontology Kallmeyer collection data from within Specify to a Darwin Core Archive file that was made available via a newly established IPT via VertNet. These data were then made available to GBIF (www.gbif.org/dataset/3c001217-eea8-4f59-8b28-885699f8cd6c) and iDigBio (https://www.idigbio.org/portal/recordsets/0d05a365-36e8-4150-a350-23ed33f79b17).

Work on the Ordovician Atlas website continues by a group of one grad and five undergraduate students. Currently, the Ordovician Atlas contains 84 genus and 117 species pages that are live. The Arthropods and Graptolites are completed, with the Edrioasteroids, Porifera, and brachiopods close to completion.

The major innovation on the Atlas website since the last report, is that species and genus locality maps have been incorporated into the pages that display data dynamically queried from iDigBio. These maps were developed using the iDigBio api to genus and species pages (example here: http://www.ordovicianatlas.org/atlas/arthropoda/trilobita/asaphida/asaphidae/isotelus/isotelus-maximus/). This new feature was featured in the iDigBio newsletter for November (http://us4.campaign-archive1.com/?u=5c564b4cf1e8157b450723e1c&id=bef6df12dd&e=b17da2d5a2)

Miami University
Over the last two months, Hauer had one undergraduate student, Maggie Perme, working on georeferencing. In that time, she has georeferenced approximately 115 locations, which correspond to about 300 specimens. The localities are all non-Shideler localities in Ohio.

Cincinnati Museum Center

In terms of specimen digitiziation, since the beginning of September, the new UC student intern, Ian MacAdam, has been focusing on georeferencing Indiana locality records. He has worked a total of 126 hours and has georeferenced 533 locality records for a total of 2,745 georeferenced catalogue records. In total, we now have 1,874 sites georeferenced and 15,568 catalogue records in Emu with georeferencing data.

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) and his students have continued to develop and add content to the Neogene Atlas of Ancient Life, with the assistance of Invertebrate Paleontology staff at the Florida Museum of Natural History. In particular, the bivalve family Pectinidae has now been added to the website. Species-level page are now online for 302 species (out of 500 planned pages).

The student at SJSU responsible for developing the Wordpress sites for the Neogene and Pennsylvanian Atlases graduated and attained a full time job elsewhere. This student was also responsible for adding content to the Pennsylvanian Atlas. PI Hendricks is currently in the process of developing a "user guide" for adding content to the Neogene and Pennsylvanian Atlases, which will assist the next individual who takes on this part of the project. We expect that the addition of new content to the Pennsylvanian Atlas (much of which has already been generated) will resume by the time of the next report. PI Hendricks is currently on sabbatical leave in Ohio; he plans to hire a new student assistant upon returning to San José in January.

(Also see Jon's activities mentioned above under KU pertaining to the development of the portable device app.)

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

The following is progress made thus far:

1. For the Ordovician 2,283 records and 6,900 specimens are databased with 1,133 specimens georeferenced.
2. For the Carboniferous 14,149 records and 43,000 specimens are databased with 3,845 specimens georeferenced.
3. For the Paleogene/Neogene 23,272 records and 70,000 specimens are databased with 7,406 specimens georeferenced.
4. For the Quaternary 12,310 records and 37,000 specimens are databased with 3,913 specimens georeferenced

Further, 3,200 images have been attached in Specify with 10,000 more imaged and awaiting attachment. There are also 20,000 type and figured specimens attached in PaleoCentral.org . In addition, PaleoCentral.org with deep time mapping is in beta testing. Finally, the data are currently with VertNet as they get their migration system to adapt to handle invertebrate specimens. The first pass reports have been examined and they should be in iDigBio soon.

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. Since the previous report, we have modified or inserted 5,600 records in KE EMu. Each of these records has 1-3 images (depending on the preservation of the fossil) and georeferenced.

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)

Finally, in other relevant news, Una Farrell attended the most recent iDigBio Summit IV, and presented a poster and a lightning talk describing our work on the TCN. In addition, post-doc Michelle Casey and co-PI Jim Beach each gave talks in the session "Advancing the Digitization of Paleontology and Geoscience Collections: Projects, Programs, and Practices I" at the 2014 Annual Meeting of the Geological Society of America in Vancouver.

Co-PI Farrell traveled to New Haven to help Yale PEN personnel PI Briggs, Senior Personnel Butts, and Museum Assistant Utrup. This occurred over a two day meeting on November 8 and 9 abd progress, practices and protocols for digitization, and future directions were discussed.

The Ohio University group has also been very active this quarter in promoting the Ordovician Atlas project in scientific community. Posters on the Ordovician project were presented at the International Paleontological Society Meeting in Mendoza, Argentina (http://www.ipc4mendoza2014.org.ar/abstracts/) and as part of the Digitization symposium at the Geological Society of America meeting in Vancouver, Canada (https://gsa.confex.com/gsa/2014AM/webprogram/Paper244473.html). Additionally, a paper on the Ordovician project has been accepted for publication (due Dec 22, 2014) in the Estonian Journal for Earth Sciences.

In October, PI Hendricks from SJSU gave a presentation on the Neogene Atlas of Ancient Life in the session "Advancing the Digitization of Paleontology and Geoscience Collections: Projects, Programs, and Practices I" at the 2014 Annual Meeting of the Geological Society of America in Vancouver. The presentation PowerPoint file was uploaded to the GSA website for anyone to view.

PEN partners at Yale Briggs and Butts presented on the PaleoNICHES PEN at the Geological Society of America Annual Meeting (Vancouver, BC – October, 2014). Butts and Co-PI at KU Farrell created a protocol (written procedure and scripts) for batch automated re-numbering and attachment of records from source files, specifically targeted at Specify users. This protocol will be made available to other Specify users.

Finally, PI's Hendricks, Stigall, and Lieberman are finalizing a manuscript that provides an overview of the Digital Atlas project. We plan that this will be submitted for review sometime in December and we will provide an update on its progress towards publication in the next report.

Attachment

https://www.idigbio.org/sites/default/files/webform/tcn-reports/POC_Screenshotsoptimized.pdf

Neogene Altas iPad App

Proof of Concept
Screen shots

Digital Atlas of Ancient Life Electronic Field Guide

Funding for development and construction of this app was provided by the National Science Foundation. The Neogene Atlas of Ancient Life is one component of the overarching Digital Atlas of Ancient Life project.

START

BROWSE

Phylum:
Mollusca

Class:
Bivalvia

Order:
Bivalvia

Order: **Bivalvia**

Family:
Ostreidae

Family:
Ostreidae

Family:
Ostreidae

Family:
Ostreidae

Family:
Ostreidae

Family:
Ostreidae

Phylum:
Mollusca

Class:
Bivalvia

Order:
Bivalvia

Family:
Ostreidae

Genus:
Undulostrea

Species:
Undulostrea locklini (Gardner, 1945)

Information

Geological Range

Late Pliocene; Extinct.

Paleogeographic Distribution

Southern Florida to Georgia.

Stratigraphic Occurrences

Late Pliocene

- Duplin / Raysor formations (GA)
- Raysor Formation (GA)
- Tamiami Formation (S. FL)
- Tamiami Formation (Ochopee Lime stone) (S. FL)
- Tamiami Formation (Pinecrest Beds)

Phylum:
Mollusca

Class:
Bivalvia

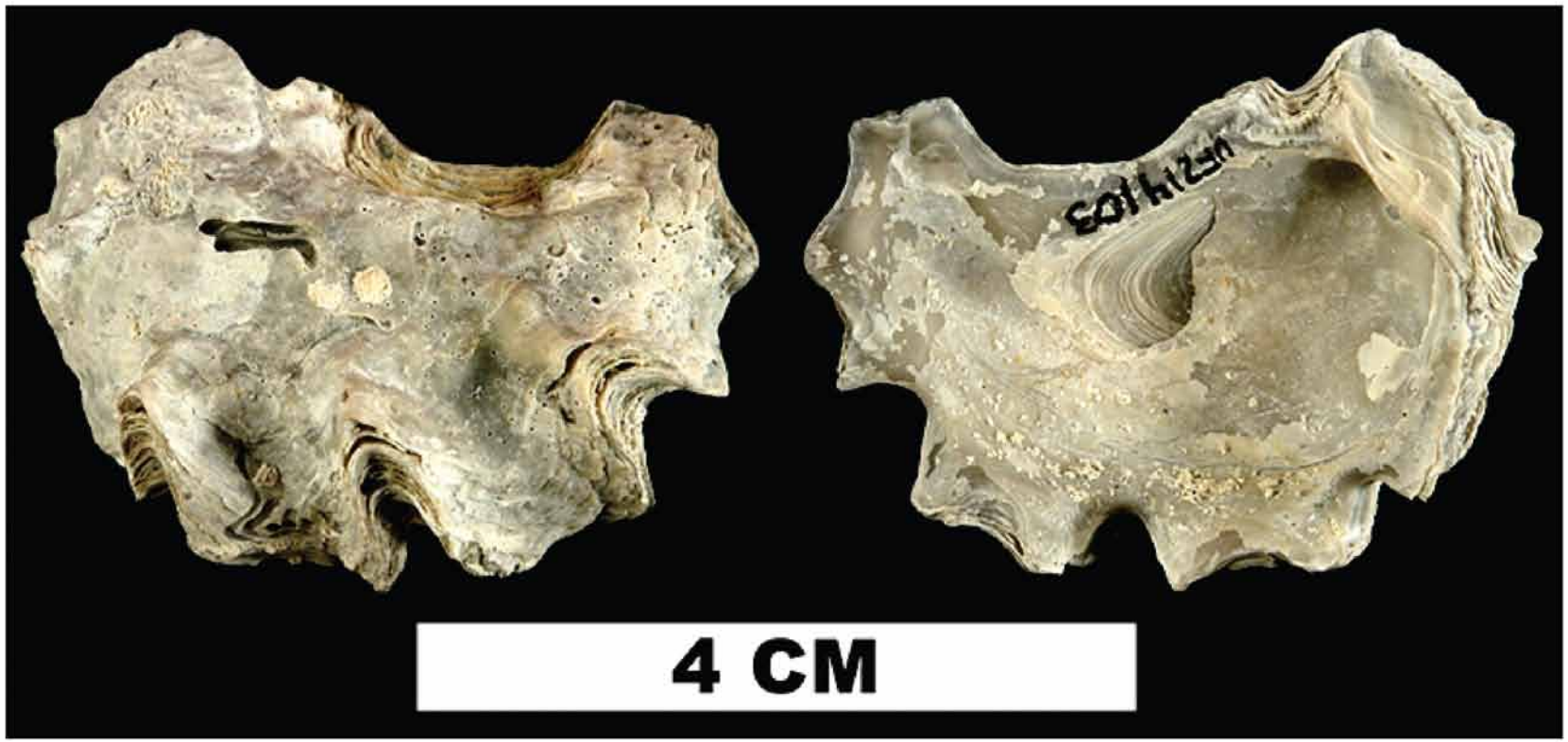
Order:
Bivalvia

Family:
Ostreidae

Genus:
Undulostrea

Species:
Undulostrea locklini (Gardner, 1945)

Images



Phylum:
Mollusca

Class:
Bivalvia

Order:
Bivalvia

Family:
Ostreidae

Genus:
Undulostrea

Species:
Undulostrea locklini (Gardner, 1945)

Maps

Late Pliocene Map

Late Pliocene Map

Late Pliocene Map

Bursidae


Orthaulax gabbi


Orthaulax gabbi

Conidae


Orthaulax gabbi


Orthaulax gabbi


Orthaulax gabbi


Orthaulax gabbi


Orthaulax gabbi


Orthaulax gabbi


Orthaulax gabbi


Orthaulax gabbi


Orthaulax gabbi


Orthaulax gabbi

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: 11/25/2014 - 16:24

Progress in Digitization Efforts

First GLI TCN report, representing three months' of effort to date.

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

Plants:
Specimens Barcoded Only: 47,198 (NY) + 1,325 (IL) = 48,523
Barcoded and Imaged to Date: 4,136 (MI) + 760 (IL) + 12,324 (NY) + 14,936 (WI) = 32,156
Databased to Date: 15,127 (MICH) + 3,437 (NY) + 45,574 (WI) = 64,138
Uploaded to the GLI Symbiota Portal: = 45,574 (WI)

Mollusks:
Only Michigan has made progress so far:
884 lots of invertebrates have been imaged, representing 2 genera and 4 species.
4363 records of invertebrates have been completed, representing 4 genera and 126 species.

Fish:
Nothing yet to report

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

We are in the progress of reviewing our traditional workflows (i.e., those used by our participants under other TCNs such as 'tri-trophic') and experimenting with new workflows for the fish collections. Best practices should emerge after our Dec 15, 2014 meeting (see below) and during year 1.

Identify Gaps in Digitization Areas and Technology

Nothing yet to report.

Share and Identify Opportunities to Enhance Training Efforts

GLI TCN participants from five states will be meeting in Chicago, IL (at their own expense) for a scheduled meeting on December 15th. This will allow for discussion and training on extant methodologies being used by those who have started imaging, data processing, etc. We also plan to discuss suggested workflows proposed by iDigBio for the three different collection types we represent (3D wet things in jars, 3D dry things in boxes, and 2D dry things on paper).

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

A subset of herbarium TCN members has begun formation of a Symbiota portal for Great Lakes Herbaria under the SEINET umbrella (separate from this GLI TCN). This will provide an editing platform for our TCN records from the Great Lakes area in order to benefit from duplicate-matching/georeferencing that may have already been entered via other Symbiota portals when processing OCR from imaged specimens currently lack data records.

The University of Michigan participants working on mollusks are collaborating with the Invert EBase TCN to make sure that data flows to both projects.

Share and Identify Opportunities and Strategies for Sustainability

Nothing to report

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

After December 1 most key senior personnel will have been hired. NYBG hired their first project intern recently . Michigan has hired its Regional Project Manager; Field Museum and UW-Madison have new Project Managers scheduled to begin on Dec 1. The IT Specialist associated with the Lichen/Bryophyte TCN has begun working with our GLI TCN as he transitions from the older to the newer project.

Project Webpage established- http://herbarium.wisc.edu/GreatLakes.htm
TCN Symbiota Portal established - http://greatlakesinvasives.org/portal/index.php

Attachment

Fossil Insect Collaborative: A Deep-Time Approach to Studying Diversification and Response to Environmental Change

Report submitted by: talia.karim@colorado.edu
Report Submitted on: 11/26/2014 - 15:14

Progress in Digitization Efforts

AMNH: We will probably finish the entire collection of Dominican and Mexican amber by the end of December (images and databasing). This will be about 6,000 pieces alone, approximately 8,000 inclusions. Then it will be on to the Baltic collection (smaller, but with lots of rarities), then Cretaceous.

CU, Boulder: We are continuing to database, image, and enhance records in our Specify database. We currently have 47,994 insect and spider records in Specify, and 9,553 images attached to those records.

Harvard-MCZ:
- Since our last report (September) we have taken about 1600 images more, accounting for about 1400 specimens from the collection.
- Hired two digitization assistants, Patrick McCormack and John Mewherter, working part time. Date of start: beginning of October.
- Developed and started to use a script to 1) update the specimen's determinations in MCZbase as they are imaged and 2) check for inconsistencies, like overlooked specimens.
- Added about 400 entries to the database of F.M. Carpenter's bibliographic collection on fossil insects.

Yale-Peabody: We have completed our electronic cataloging as outlined in the grant proposal and have nearly half of our specimens imaged. As a result of improved visibility of our fossil insect collections we have received, and continue to receive, large quantities (1000s) of fossil insects from an avocational collector and are incorporating this material into our digitization efforts (2957 cataloged, 1903 imaged, since the start of the FIC).

VMNH: Fossil insect digitization paused - exception: periodic digitization of newly excavated Solite matieral (insects, vertebrates, plants). Due to the departure of Dooley (PI) that occurred at the museum a few months ago, priorities were shifted and grant funding was put on hold until a new curator can be hired (expected date July 2015). Byrd is currently acting as collections manager of the paleontology department.

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

Nothing to report.

Identify Gaps in Digitization Areas and Technology

Yale-Peabody: The balance between the desire for high resolution photography and a rapid imaging rate is, as always, difficult to maintain.

CU, Boulder: Need for better and easier access to archival storage for large files. At present, these files are stored on the PETA Library (CU research computing project funded by NSF) are not accessible to the public or easily sharable. Botany is working with our research computing group on campus to try and work through some of these issues.

Share and Identify Opportunities to Enhance Training Efforts

Nothing to report.

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

AMNH: Planning an amber preparation and imaging workshop for February 2015. The workshop will be open to all TCN members as well as a select few additional invitees.

Share and Identify Opportunities and Strategies for Sustainability

Nothing to report.

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

Yale-Peabody:

- Butts and Norris continue to have biweekly Skype and quarterly face-to-face meetings with Seth Kaufman (Whirl-i-Gig) on the development of functionality for iDigPaleo and incorporation of datasets from the FIC PIs. A prototype of the database is available at: http://idigpaleo.whirl-i-gig.com/.
- A prototype of the database is available at: http://idigpaleo.whirl-i-gig.com/. The database currently has datasets from UC Boulder, Yale Peabody Museum and the Harvard Museum of Comparative Zoology. With these datasets registered users can browse, curate collections, filter by institution, location, age and/or taxonomy, and comment and tag images. Curated collections (galleries) can be used to auto-generate field guides and classroom handouts. The latest meeting with Whirl-i-Gig discussed the next functions to add: mapping (including paleomapping utilizing GPlates), incorporation of common names, and the reporting of comments to institutions.

CU, Boulder:
- Smith presented an overview of the TCN at the GSA annual meeting, "The Fossil Insect Collaborative: An NSF-Funded Paleontology Collections Digitization Project"
- Karim presented a talk on compression fossil digitization workflows at the GSA annual meeting, "Fossil Insect Digitization Workflow at the University of Colorado." Two graduate students funded on the TCN, Walker and Levy, were co-authors on the talk.
- Smith and Heads (PI INHS) organized a fossil insect symposium for the annual ESA meeting, "How the Fossil Record Can Contribute to Our Understanding of Insect Ecology and Evolution"
- Smith presented "Macroevolutionary history of the Coleoptera: A quantitative analysis of fossil occurrences" at the ESA session.
- Karim presented an update of the TCN at the ECN meeting in Portland, "The Fossil Insect Collaborative Year 2: data acquisition, publication, and use"

Harvard-MCZ:
- Perez de la Fuente presented “Digitization of the fossil insect collection from the Museum of Comparative Zoology” at the 4th International Paleontological Congress in Mendoza (Argentina), that took place from Sept 28th to Oct 3rd. Authors: Ricardo Pérez-de la Fuente and Brian D. Farrell.
- Provided support and technical means to Prof. Alexander Rasnitsyn, from the Paleontological Institute of Moscow, during a visit to the MCZ to study Paleozoic insects, from Sept 15th to Sept 17th.
- Collaborated in a side project with Dr. Eric S. Chivian, founder and Director of the Center for Health and the Global Environment at Harvard Medical School, on dragonfly wing topography related to flight kinematics in insects. Obtained images and 3D models from extant dragonflies with the equipment and software used to digitize the fossil insect collection.
- Shared fossil insect images and expertise with Rosie Powell-Tuck, from Colossus Productions, UK, working on a new 3DAttenborough documentary on animal flight.

INHS:
- Heads presented at the ESA annual meeting, "Fossil insects from the Lower Cretaceous Crato Formation of Brazil"
- J. Thomas (Heads lab manager, undergraduate student) presented at the ESA annual meeting, "Rediscovery of the Milton Sanderson Dominican amber collection"

Attachment

North American Lichens and Bryophytes: Sensitive Indicators of Environmental Quality and Change

Report submitted by: cgries@wisc.edu
Report Submitted on: 11/26/2014 - 18:48

Progress in Digitization Efforts

As of November 2014 the number for the LBCC are as follows:
Lichens:
http://lichenportal.org
Herbaria actively submitting images or key stroked records to the portal: 64
Specimen records in portal: 1,594,005 (up by 299,158 since August 2014)
Specimen records with images: 620,085 (20,161 labels have been imaged since August 2014)
Rrecords with locality information: 1,274,404 (329,795 locality information where added since August 2014)

Bryohpytes:
http://bryophyteportal.org
Herbaria actively submitting images or key stroked records to the portal: 58
Specimen records in portal: 2,039,717 (up by 139,461 since August 2014)
Specimen records with images: 925,662 (122,528 labels have been imaged since August 2014)
Records with locality information: 1,214,101 (80572 locality information where added since August 2014)

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

Regular Symbiota training sessions are being held remotely.

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

Share and Identify Opportunities and Strategies for Sustainability

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

Attachment

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

Report submitted by: tmarsico@astate.edu
Report Submitted on: 11/30/2014 - 17:25

Progress in Digitization Efforts

The grant has been set up at the lead institution in Arkansas (STAR), and the subaward at APCR also has been executed. One collection in Arkansas, STAR, has all the flowering plant vouchers from within the state imaged and databased. This represents 16,791 collections. These were databased in Specify. All of these collections are available for viewing and searching online at herbarium.astate.edu. In the coming months we have planned to deposit these digital accessions in Symbiota and iDigBio. APCR is obtaining details about purchasing barcodes before imaging. STAR is working with UARK to begin imaging set up with them. Marsico will be attending the SERNEC training meeting in Valdosta, Georgia in January 2015.

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

We will keep good notes in Arkansas about what we learn as we set up mobile imaging stations. Currently we have nothing to report here.

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

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

Report submitted by: moon@begoniasociety.org
Report Submitted on: 12/01/2014 - 08:04

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 (as of October 20, 2014):

1,081,834 specimens data captured
593,251 specimens georeferenced
785,320 data records provided to iDigBio

Our project plans to continue digitization with the current funding through July 2015.

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

Continue to work with EOL and iDigBio to export associations data as a DwC-A extension and improved attribution for our dataset.

Identify Gaps in Digitization Areas and Technology

We need guidelines from iDigBio that are explicit about the statistics we are collecting regarding digitization rates. Our project adds in curation to time for digitization, others do not. Because the digitization would not happen without this level of curation, we feel we should include it. Ultimately, these differences will make it difficult to sum across efforts.

Share and Identify Opportunities to Enhance Training Efforts

Nothing to report.

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

Katja Seltmann is actively working with iDigBio to organize a Data Carpentry workshop series. Ongoing collaboration with SCAN to summarize insect digitization efforts to date.

Joint presentations at the Entomological Society of America and Entomological Collections Network.
N. Cobb, K.C. Seltmann and N. Franz. 2014. Title: The current state of arthropod biodiversity data: Addressing impacts of global change. Entomological Collections Network. November 2014. Portland, OR.

K.C. Seltmann and D. Paul. Title: Recreating biomes one label at a time. Entomological Society of America. November 2014. Portland, OR.

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

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

Report submitted by: chdietri@illinois.edu
Report Submitted on: 12/01/2014 - 09:54

Progress in Digitization Efforts

University of Illinois personnel delivered and set up twelve robotic drawer digitizing systems at collaborating institutions (Purdue, Michigan State U., Carnegie Museum, South Dakota State U., North Dakota State U., Minnesota, U. Wisconsin-Madison, Milwaukee Public Museum, U. Missouri, U. Kansas, Kansas State U., and Iowa State U.) and collaborators were trained in their use. A user's manual and troubleshooting guide was also developed and delivered to collaborators. Collaborators have been capturing images and sending image sets to the U. of Illinois for processing (stitching and posting on InvertNet.org). The InvertNet technical team is continuing to test and optimize image capture workflows to improve speed and image quality. They are also optimizing the image ingest workflow to facilitate more rapid uploading of image sets to the cyberinfrastructure platform.

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

Nothing to report.

Identify Gaps in Digitization Areas and Technology

A major gap in existing technology is in stitching software. Available stitchers are too slow to work in real time, so we have not yet installed stitching software on the local computers used to operate the drawer digitizing system but, rather, are having collaborators save and send raw image sets to the U of Illinois for processing. We are working with open-source software to create a real-time stitcher by parallelizing some of the processes, but need to overcome some technical hurdles first.

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)

Attachment

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

Report submitted by: psierwald@fieldmuseum.org
Report Submitted on: 12/01/2014 - 10:50

Progress in Digitization Efforts

Invertebase TCN-wide activities: Preparation for digitization, capture of collection level-level information at all participating collections (Field Museum Nat Hist; Cleveland Museum Nat Hist; Auburn University Museum; University of Michigan Museum; Delaware Museum Nat Hist; Frost Ent Museum). Development of digitization workflow protocols
Field Museum Nat Hist: digitization of arthropod and invertebrate type material (including specimen imaging) through other funding ongoing, digitization staff search for Invertebrate TCN ongoing.
Cleveland Museum Nat Hist:
Auburn University Museum: Personnel digitizing specimens: 2 undergraduates (10 hours/week each); 4 graduate students (2 at 15 hours/week and 2 at 4 hours/week). Purchased 2 laptop workstations and 3 barcode readers. Established initial protocols for digitization and adding barcodes to pinned specimens (2 dimensional format). Specimen digitization started – Ephemeroptera, Odonata, Orthoptera, Phasmida, Dermaptera – 12,525 specimens digitized and barcoded (21 Nov 2014).
University of Michigan Museum of Zoology: 1. Three undergraduate students were hired for mollusk data entry. These new students plus two returning students (one Ph.D. and one undergraduate curatorial assistant) and two new undergraduate students hired for the Great Lakes Invasives TCN will continue data entry into the UMMZ Mollusk Division Specify database. 2. Since September 2014, total 4886 records, representing three freshwater snail families (Lymnaeidae: 717; Pysidae: 311; Valvatidae: 1035) and one bivalve family Sphaeriidae (2823), were entered.
Delaware Museum Nat Hist: We have begun updating the taxonomy of the unionid bivalves in the DMNH collection according to Musselp website.
Frost Ent Museum:

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

Field Museum Nat Hist: nothing yet to report
Cleveland Museum Nat Hist:
Auburn University Museum: nothing yet to report other than rate of digitization increasing; logistics of organizing workers and volunteers problematic due to limited workstations.
University of Michigan Museum of Zoology: nothing yet to report
Delaware Museum Nat Hist: DMNH has meetings set up on 2 December with senior staff from the Academy of Natural Sciences of Philadelphia to discuss grant administration and financial reporting best practices.
Frost Ent Museum:

Identify Gaps in Digitization Areas and Technology

Field Museum Nat Hist: Nothing yet to report
Cleveland Museum Nat Hist:
Auburn University Museum: Tremendous variation in label quality – unlikely something that can be automated.
University of Michigan Museum of Zoology: nothing yet to report
Delaware Museum Nat Hist: Sourced a new server and laid out a timeline for installation with our IT partners.
Frost Ent Museum:

Share and Identify Opportunities to Enhance Training Efforts

Field Museum Nat Hist: nothing yet to report
Cleveland Museum Nat Hist:
Auburn University Museum: Nothing yet to report
University of Michigan Museum of Zoology: nothing yet to report
Delaware Museum Nat Hist: Using the Specify Friday afternoon HelpCast webinars to get familiar with database operation.
Frost Ent Museum:

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

Invertebase TCN (P. Sierwald) established collaboration with SCAN TCN (Neil Cobb) for collaborative portal development and collaborative FilteredPus development, collaborative taxonomic thesaurus development with SCAN
Field Museum Nat Hist: nothing yet to report
Cleveland Museum Nat Hist:
Auburn University Museum: Collections acquisition for digitization (marine mollusks) - East Carolina University
University of Michigan Museum of Zoology: The University of Michigan participants are collaborating with the Great Lakes Invasives TCN to make sure that data flows to both projects.
Delaware Museum Nat Hist: nothing yet to report
Frost Ent Museum:

Share and Identify Opportunities and Strategies for Sustainability

Field Museum Nat Hist: nothing yet to report
Cleveland Museum Nat Hist:
Auburn University Museum: Nothing yet to report
University of Michigan Museum of Zoology: nothing yet to report
Delaware Museum Nat Hist: nothing yet to report
Frost Ent Museum:

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

Field Museum Nat Hist: Invertebase TCN (PS) established Invertebase Symbiota Portal in collaboration with Ed Gilbert and Nico Franz. PIs P. Sierwald and R. Bieler attended TCN summit IV in Gainesville. Conducted three Invertebase TCN members meetings at the TCN summit IV in Gainesville. Developed management plan for taxonomic authority files (Arthropods: currently Deans, Svenson, Lee and Sierwald; mollusks: Bieler, Lee, and Shea)) and georeferencing (leader R. Bieler).
Cleveland Museum Nat Hist: taxonomic thesaurus development Hymenoptera. PI G. Svenson attended TCN summit IV in Gainesville.
Auburn University Museum: Nothing yet to report
University of Michigan Museum of Zoology: taxonomic thesaurus development Odonota. T. Lee attended TCN summit IV in Gainesville.
Delaware Museum Nat Hist: Taxonomic thesaurus development for North American land and fresh water mollusks with FMNH, especially freshwater unionid bivalves. Developing protocols and timeline for hiring temporary staff for data manipulation and entry. PI E. Shea attended TCN summit IV in Gainesville.
Frost Ent Museum: PI A. Deans attended TCN summit IV in Gainesville.
FilteredPush development: PI Hanken and TCN member David Lowery attended TCN summit IV in Gainesville.

Attachment

MOBILIZING NEW ENGLAND VASCULAR PLANT SPECIMEN DATA TO TRACK ENVIRONMENTAL CHANGE

Report submitted by: p_sweeney@att.net
Report Submitted on: 12/01/2014 - 12:24

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 231,000 records and 160,500 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.

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

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: 12/08/2014 - 14:05

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/SCAN_November_2014.docx

Southwest Collections of Arthropods Network Update

December 10, 2014

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 720,254 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 ~20,000 more records to meet our project goal by July 1, 2015. The biggest challenges will be to increase the value

Table 1. Number of specimen records digitized and associated summary statistics. From http://symbiota4.acis.ufl.edu/scan/portal/index.php

	SCAN funded	SCAN non-funded	TOTAL SCAN	Total Served
# Specimen Records	720,254	127,193	847,992	4,972,417
# Georeferenced	527,748	49,255	578,289	3,970,404
# Identified to species	462,765	62,933	526,232	2,997,916
# Families	713	358	751	1,448
# Genera	7,069	3,356	8,588	19,626
# Species	16,183	6,836	20,780	57,851
% Georeferenced	73%	39%	68%	80%
% Identified to Species	64%	49%	62%	60%

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 2,000 high-resolution images as part of their 25,238

images uploaded. We currently have 7,869 high-resolution 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	# High-Resolution Images
Arizona State University	1,984
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)	48,798
University of Hawaii (low res images)	10,533
University of Colorado	0
TOTAL	90,438

\

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 are continuing 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 4.9 million records for 57,851 species.

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

Report submitted by: emilylgillespie@gmail.com
Report Submitted on: 12/08/2014 - 17:54

Progress in Digitization Efforts

We received final word of funding in August 2014. I received account numbers, enabling purchasing and hiring, in Oct 2014. As of today (8 Dec 2014) I have received all equipment but a laptop, which is in-house but being 'imaged' by the IT specialist, and I have been able to get a student hire through HR as of this week.

We have been preparing the herbarium this semester for digitization, particularly clearing space, reading cabinets for maneuverability, reorganizing according to APG, training students in navigating the collection, and making purchases and payroll requests.

As of Dec. 8, 2014, we have begun selecting specimens and barcoding. We've barcoded about 500 sheets in the last 10 days, and we will begin imaging the moment our laptop is connected to our camera and we troubleshoot the workflow. I anticipate that we will move fairly quickly as soon as spring semester begins.

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

I've discovered that preparation is a huge undertaking, if you're dealing with a collection that isn't completely up to date curation-wise (I am fairly new at my institution). This is a step that should not be underestimated by curators. However, I was able to fill the time it took my grants office to process my grant for this purpose, so it worked out well.

Identify Gaps in Digitization Areas and Technology

nothing to report so far.

Share and Identify Opportunities to Enhance Training Efforts

I have about six federal work study students who are going to ultimately be helping with this effort; one of them is officially hired by the grant so far. This one student is proofreading and troubleshooting protocols as I write them, before we train the other students. This will hopefully minimize stops/restarts as well as student frustration.

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

nothing to report so far

Share and Identify Opportunities and Strategies for Sustainability

nothing to report so far

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

nothing to report so far

Attachment

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: 12/09/2014 - 15:42

Progress in Digitization Efforts

There are currently 1,779,399 records available for searching the MycoPortal. My best guess is that this includes at least 75% of all records that have been created so far -- there are a few institutions that have rather convoluted systems for getting data from their own systems in to the MycoPortal, and so they only do uploads every few months. As best I can tell, there are still between 100-200k specimens to be digitized for this project

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

In recent months we haven't been following up with all participants the way we did at the beginning of the project, and this is a mistake. I feel that we have lost touch a bit with some of the participants who do not seem to be making particularly good progress. So, I intend to have the project coordinator do a year end outreach to all participants, and then re-establish more regular contact after that.

Identify Gaps in Digitization Areas and Technology

We still struggle with record completion but with funding from the MaCC project, we have commissioned some changes to Symbiota that should make record completion go much faster. These involve the ability to sort records in table view, and also improved data parsing through the incorporation of Salix data parsing techniques.

Share and Identify Opportunities to Enhance Training Efforts

Nothing to report at this time.

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

We anticipate that the Symbiota improvements described above will help all of the TCNs that use Symbiota.

Share and Identify Opportunities and Strategies for Sustainability

Nothing new to report

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

We are in a transitional situation with the MaCC project at the moment, since the Portal Manager, Scott Bates, has left the project. I am in the process of figuring out how to manage his subcontract going forward, and how to maintain the level of service that we have provided up until now.

Attachment