



invertnet Year 5 and Beyond

Chris Dietrich
 Illinois Natural History Survey
 University of Illinois
 chdietri@illinois.edu



iDigBio is funded by a grant from the National Science Foundation's Advancing Digitization of Biodiversity Collections Program (Cooperative Agreement EF-1115210). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Objective

- develop and implement an efficient workflow for cost-effective, high-throughput digitization of insect collections



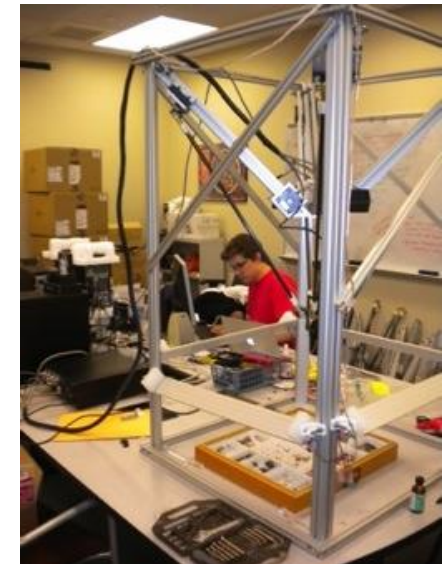
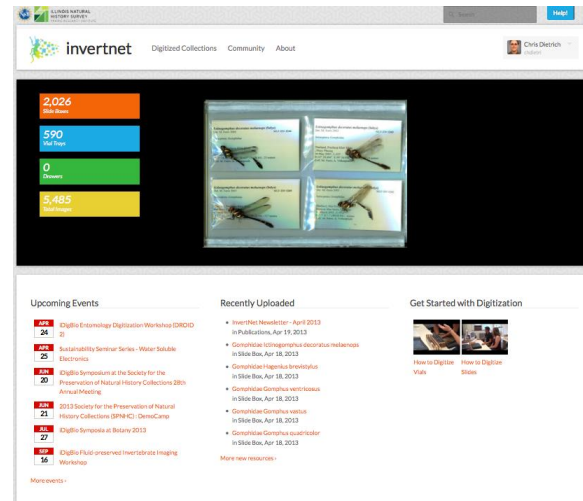
Specific Goals

- Digitize all holdings of 22 midwestern arthropod collections (~50 million specimens)
 - Specimen images and metadata (label info)
 - Drawers, vials, slides
 - Advanced imaging (including 3D)
 - Best quality at reasonable cost (~\$0.10/specimen)
- Provide access to images and other data via online virtual museum
 - browsable/searchable/zoomable web interface
 - link to other data providers (GBIF, iDigBio etc.)
- Provide platform for research and development of additional tools and resources
 - Data mining and analysis
 - Community building, collaboration, and support
 - Education, outreach, and reference



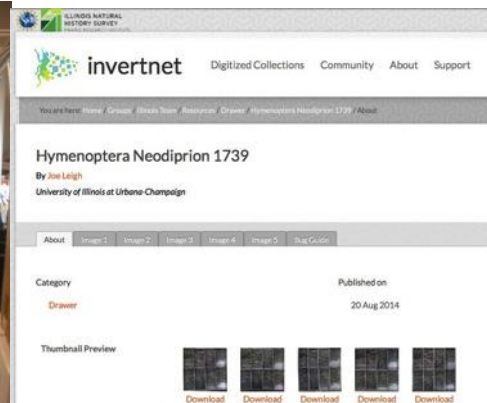
Accomplishments

- Created InvertNet cyberinfrastructure platform based on HUBzero (invertnet.org)
- Built 180 TB storage system to house InvertNet image library
- Implemented efficient workflows for slides and vials using 2D scanning technology
- Built 14 robotic drawer digitization systems & delivered to collaborators
- Ingested >46,000 images and metadata from collaborating institutions representing >2.5 million specimens
- Developed image annotation tool to facilitate specimen-level data capture
- Linked InvertNet data repository to iDigBio portal and BugGuide.net
- Held two training workshops for collaborators (April 2012 and November 2013)
- Participated in numerous workshops, symposia and planning meetings
- Published 2 papers describing our high-throughput digitization approach
- Trained 15 grad students and >30 undergrads



Ongoing Activities

- Capturing whole-drawer images at collaborating institutions
- Seeking additional funding for gameification of label data capture



Illinois Natural History Survey: Insect Collection

Search INHS Insect database (also incl. Arachnida, Myriapoda, Onychophora, & Xiphosurida)

| Search Term | Where | Partial Search | Sort |
|--------------|----------------|----------------|----------|
| mongoljassus | Genus | Whole field | Unsorted |
| and | Species | Whole field | Unsorted |
| and | Catalog Number | Whole field | Unsorted |
| and | Label | Whole field | Unsorted |

Results per page: 25 | Add a line | Clear the form | Search

Found records: 125 (Page: 1 of 5)

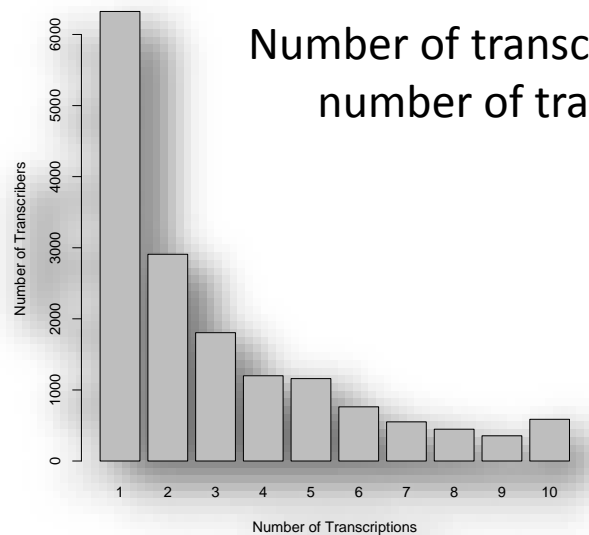
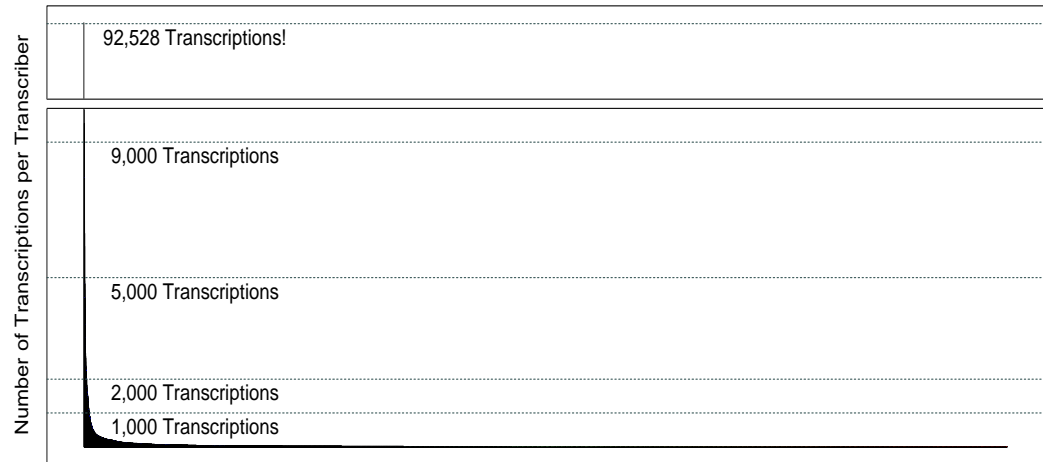
| Catalog Number | Class | Order | Suborder | Superfamily | Family | Subfamily | Tribe | Genus | Subgenus | Species | Subs |
|-----------------|---------|-----------|-----------------|-------------|--------------|-----------------|-------------|--------------|----------|------------|------|
| Homoptera 27772 | Insecta | Homoptera | Auchenorrhyncha | Membracidae | Cicadellidae | Deltocephalinae | Paralimnini | Mongoljassus | | Stanshaica | |
| Homoptera 27773 | Insecta | Homoptera | Auchenorrhyncha | Membracidae | Cicadellidae | Deltocephalinae | Paralimnini | Mongoljassus | | Stanshaica | |
| Homoptera 27774 | Insecta | Homoptera | Auchenorrhyncha | Membracidae | Cicadellidae | Deltocephalinae | Paralimnini | Mongoljassus | | Stanshaica | |
| Homoptera 27775 | Insecta | Homoptera | Auchenorrhyncha | Membracidae | Cicadellidae | Deltocephalinae | Paralimnini | Mongoljassus | | Stanshaica | |
| Homoptera 27776 | Insecta | Homoptera | Auchenorrhyncha | Membracidae | Cicadellidae | Deltocephalinae | Paralimnini | Mongoljassus | | Stanshaica | |
| Homoptera 27777 | Insecta | Homoptera | Auchenorrhyncha | Membracidae | Cicadellidae | Deltocephalinae | Paralimnini | Mongoljassus | | Stanshaica | |
| Homoptera 27832 | Insecta | Homoptera | Auchenorrhyncha | Membracidae | Cicadellidae | Deltocephalinae | Paralimnini | Mongoljassus | | Stanshaica | |



Why Gameify?

- a few citizen scientists now do most of the work
- crowd sourcing has limited appeal to broader community
- >100 million Americans play online games on a regular basis

Number of transcriptions for the ~21,000 transcribers



Number of transcribers for each number of transcriptions

Gameification goals

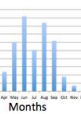
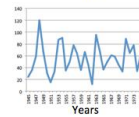
- recruit beyond the core communities usually inclined to participate in citizen-science projects (e.g., online gamers)
- improve user experience
- teach participants about biodiversity and evolution
- Transcribe specimen data rapidly and accurately



Diversity Score: 653
Rank: 14
Orders: 12
Families: 133
Genera: 3,565
Species: 7,251
Point total: 53,478



Jane's Species

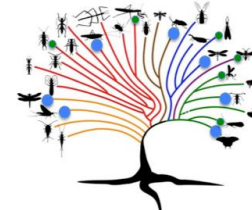


Empire of Insects

JANE'S LAB COMMAND CENTER



Empire Map



Diversity Tree

Empire Building Tasks:

Find Specimens

Find Scientific Names

Find Localities

Find Collecting Dates

Find Collectors

Full Annotations

Search for:



Diversity Score: 653
Rank: 14
Orders: 12
Families: 133
Genera: 3,565
Species: 7,251
Point total: 53,478

Points earned this session: 14

Scoring

Find and segment a specimen (per view): 1 pt
Bonus for segmenting all 5 views of a specimen: 3 pts
Bonus for completing a unit tray with 10 or more specimens: 5 pts
Bonus for completing a drawer with 100 or more specimens: 20 pts

Empire of Insects

Find Specimens



Instructions:
Zoom in on the drawer so the boundaries between individual insects (and labels) are visible. Then draw a box or polygon around each individual specimen and repeat for as many specimens as you can. Be sure to enclose the entire specimen within the box/polygon including legs, wings, antennae and all labels attached to the same pin). If multiple views of the same drawer/specimens are available, you will be prompted to highlight the same specimen on all the different views before moving on to the next specimen.



Diversity Score: 653
Rank: 14
Orders: 12
Families: 133
Genera: 3,565
Species: 7,251
Point total: 53,478

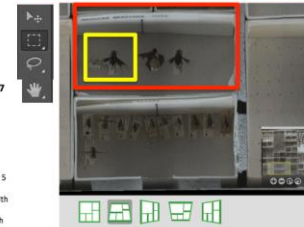
Points earned this session: 27

Scoring

Find and segment a unit tray with visible header label: 1 pt
Segment all 5 views of a specimen: 5 pts
Bonus for completing a unit tray with 10 or more specimens: 5 pts
Bonus for completing a drawer with 100 or more specimens: 20 pts

Empire of Insects

Find Scientific Names



Instructions:
The scientific name for a group of specimens can usually be found in at least one of two places: on a vertical label at the top of the unit tray and/or on the specimen at the top left corner of the tray. The vertical labels are usually only visible in the back-tilted view of the drawer. The ID labels attached to specimens are usually placed below the tilted view (usually tilted to the right) so it may be necessary to look at a tilted view (usually tilted to the right) to see as much of this label as possible. Use the tools to draw a box around the unit tray for which you want to capture the scientific name and then transcribe as much of the text of this label as you can see. If an ID label is visible on the specimen, draw a box around the specimen and transcribe as much label text as you can see into the box provided.

Unit tray header label

Anthrax mexicana Cole

Specimen identification label

Anthrax mexicana
Cole
det. C. D. Michener

InvertNet data management and use

- Data management
 - HUBzero-based cyberinfrastructure
 - 180TB storage system with tape backup and archiving
 - local image capture to duplicated removable hard drives
 - all data and processes are open access
- Research use of data
 - 3D modeling and reconstruction
 - automated identification using computer vision and machine learning
- Management of the network including oversight and processes
 - INHS permanent IT staff have assumed management responsibility for managing website and cyberinfrastructure
 - InvertNet PIs continue to provide oversight