InvertNet: 2013-14 Progress and Goals

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.

Rationale

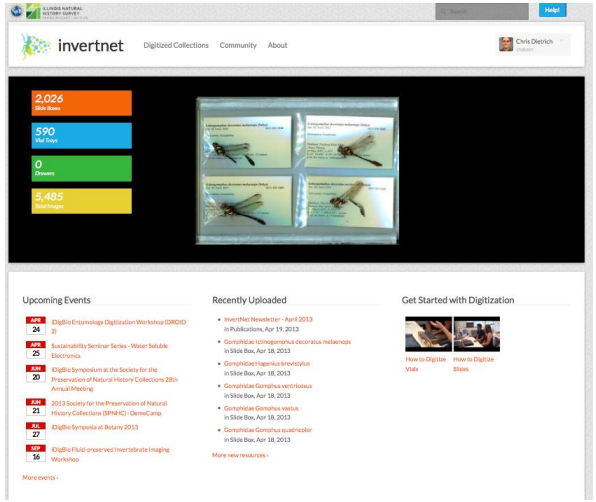
- vast majority of specimens in U.S. collections are invertebrates
  - primarily insects and related arthropods
  - less than 5% available online
  - only label data usually provided
- most invertebrate biodiversity research is specimen-based
  - all knowledge of many species is embodied in collections
- existing digitization methods are inadequate
  - slow and expensive ($1+ per specimen)
  - risk of damage to specimens from handling

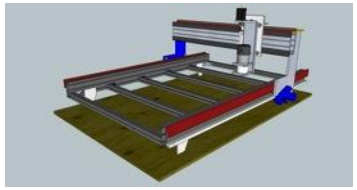
Goals

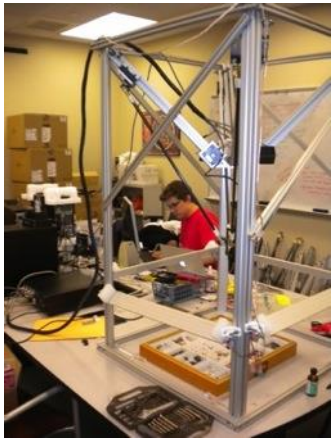
- Digitize all holdings of 22 midwestern arthropod collections (~50 million specimens)
  - Specimen images and metadata (label info)
  - Drawers, vials, slides
  - Advanced imaging (including 3D)
  - Best quality at reasonable cost (~$0.10/specimen)
- Provide access to images and other data via online 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

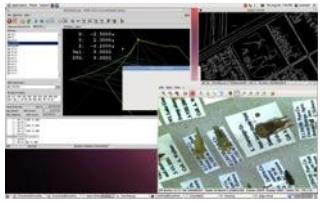
Accomplishments

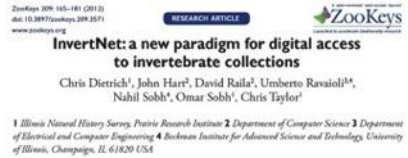
- Created InvertNet cyberinfrastructure platform based on HUBzero (invertnet.org)
- Implemented efficient workflows for slides and vials using 2D scanning technology
- Built and tested several prototype whole-drawer digitization systems
- Built 13 robotic drawer digitization systems for collaborators (deliveries underway)
- Built 180 TB storage system to house InvertNet image library
- Ingested 17,000+ images and metadata from collaborating institutions representing >300,000 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











Year 4 Goals

- Finish delivering drawer digitization systems, train users
- Capture images of ~80,000 drawers from all collaborating institutions and provide access via InvertNet.org
- Crowdsource label data capture from images of slides, vials and drawers
- Ingest existing specimen-level data from collaborating institutions
- Improve 3D reconstruction tools to allow virtual tilting of drawer and specimen images via a web interface