InvertNet: A New Paradigm for Digitization of Invertebrate Collections

Chris Dietrich
INHS-UIUC
chdietri@illinois.edu
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 since July 2011

- Created robust cyberinfrastructure platform based on HUBzero (invertnet.org)
- Implemented efficient workflows for slides and vials using 2D scanning technology
- Curated and staged collections for digitization
- Ingested 12,000+ images and metadata from 11 collaborating institutions representing >200,000 specimens
- Linked InvertNet data repository to iDigBio portal
- Developed and tested three different drawer digitization prototype systems
- 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
- Added two institutions as InvertNet collaborators (U. Hawaii, U. Iowa)
- Trained 15 grad students and >30 undergrads
Goals for years 3-4

- Capture images of ~80,000 drawers from all collaborating institutions and provide access via InvertNet.org
- Crowdsourced data capture from images of slides, vials and drawers
- Ingest existing specimen-level data from collaborating institutions
- Link to BugGuide
- Develop improved, real-time photo stitcher
- Develop improved 3D reconstruction algorithms that allow virtual tilting of drawer and specimen images via a web interface