Whole Drawer Digitization

By

Paul Mayer

The Field Museum

Terry Evans, *Field Museum, drawer of Cardinals, various dates*, 2001
From the *Specimen Series*
Museum of Contemporary Photography

Digitizing Invertebrates

- Difficult due to large numbers
- Traditional done one at a time
  - Image individually
  - Enter data from label
  - Advantage is high quality data for specimens
  - Disadvantage
    - Handling delicate specimens
    - Time consuming

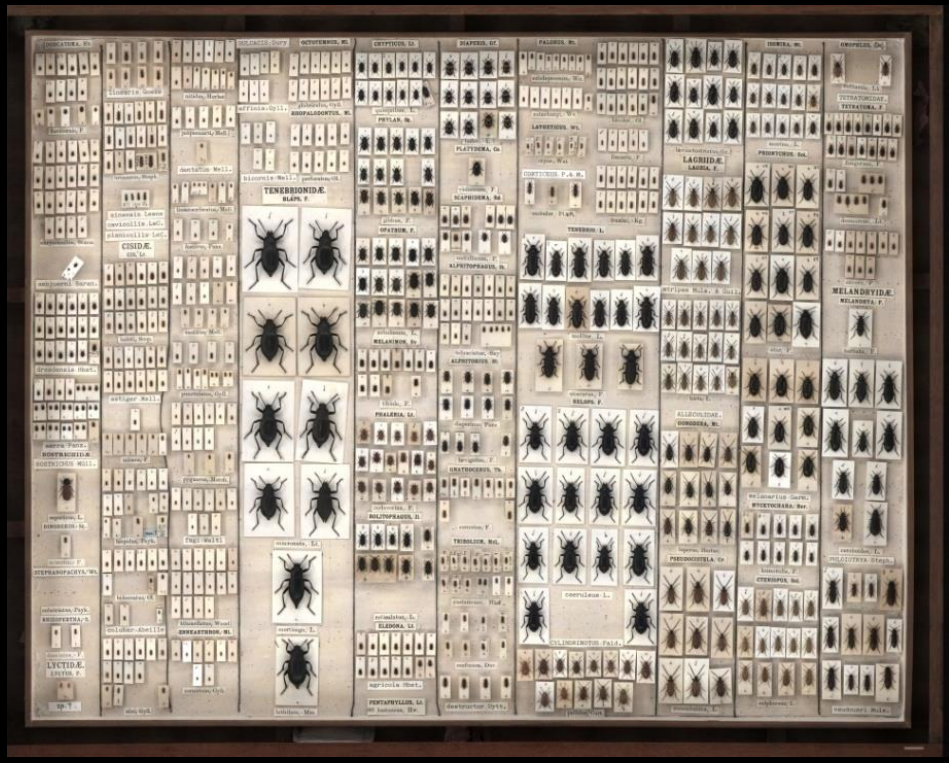
Digitizing Whole Drawers

- Image many specimens/labels at once
- Virtual access to collections
  - Researchers
  - Staff
  - Students/teachers
  - Exhibits
- Inventory
  - Present/Absent
  - Collection space
  - Curation levels
- Security
- Art

Terry Evans, *Field Museum, drawer of Meadowlarks, various dates*, 2001
From the *Specimen Series*
Chicago Art Institute

Advantages to Whole Drawer Digitization

- Quick
- Many specimens at once
- Automated
- Less handling of specimens



Advantages of Whole Drawer Digitization

- Whole-drawer images facilitate more effective collection management, virtual curation, and public engagement
- Promote and encourage remote curation of unsorted specimens
- Assist with loan requests
- Provide a method for auditing the collection
- Permit morphometric analysis of at least some specimens; and
- Encourage public engagement with biological collections.

- Prioritize the collections
  - Calculate a collection health index and tracking as it changes over time.
- Tighten inventory control
  - Develop a maps
  - Pin-point specimens that
    - High monetary value
    - Holotypes
    - Taxa represented by a single specimen
    - Damaged by pests
- Create a visual base-line inventory to serve as a basis for future inventory control.
- Planning tool
  - Space
  - New projects

Disadvantages of Whole Drawer Digitization

- Hard to capture label data
- Not all specimens visible
- Prep time for drawers
- Difficult to keep different size specimens all in focus
- Positioning of specimens
- Can be expensive
- Difficult to link image to specimen
- Loans and reorganizing collections moves specimens

Two Approaches for Imaging Drawers?

- Low resolution
  - Smart phones or Point and shoot cameras
  - Often performed by volunteers or interns
  - Low resolution
  - Fast way to inventory collection
- High resolution
  - Large images lots of megapixels or gigapixels
    - Often automated
    - Cut up image and assigned separately to individual specimens
    - Fast way to image many specimens

The Field Museum
Fossil Invertebrates

- Specimens are stored in 550 cabinets
- 14,000 drawers
- Labels stand upright in back of box
- Some labels are on outside of drawer

The Field Museum
Fossil Invertebrates

- Volunteers photograph whole drawer and outside drawer labels with iphones and point and shoot cameras.
- Drawers are heavy difficult to move.
- Cabinets extend to ceiling, top drawers difficult to photograph

The Field Museum Fossil Invertebrates

- Little prep time is spent to arrange labels or specimens
- Image is used only for an inventory and security purposes.
- Allows research a rough idea of what material is available, but not designed for specimen level examination.

University of Texas Austin
Ann Molineux

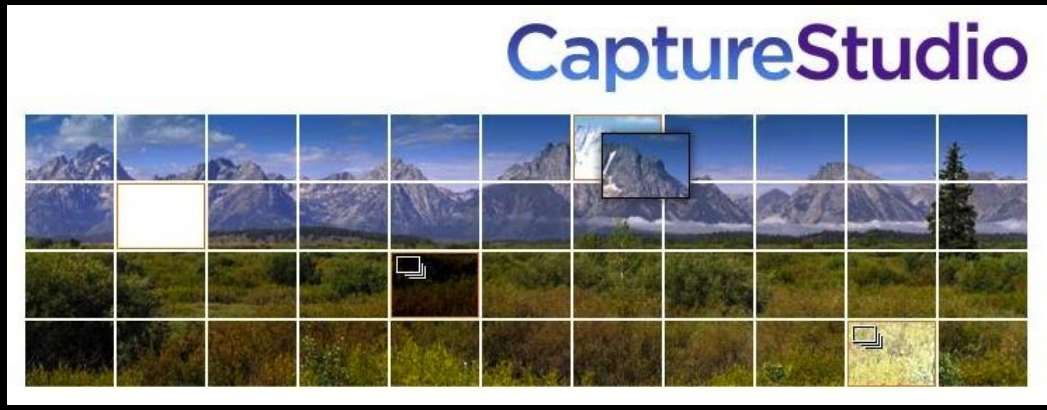
- Create virtual access to collections
- Collection storage area is ~10 miles (16 km) from center building
- Volunteers use ipads and iphones to photograph whole drawers

University of Texas Austin
Ann Molineux

- Combined database and spatial data using ArcGIS Online
- Related a third aspect, drawer and specimen images that are linked to the specimen records and can be viewed alongside the data and spatial records.
- Online resource that allows the user to
  - Virtually examine the collection, query a database file and
  - View the drawer content, individual specimens, and label
  - Images when they exist.

Gigapixel Imaging

- Automatically capture multi-gigapixel panoramas
- Use almost any camera and lens combination
- Accurately take hundreds or thousands of photos
- Create stunningly detailed, zoomable panoramas
- Control unit with an intuitive on-board interface
- Get the only complete solution to capture, stitch, upload, zoom, share, embed, tag, and print ultra-high-resolution images

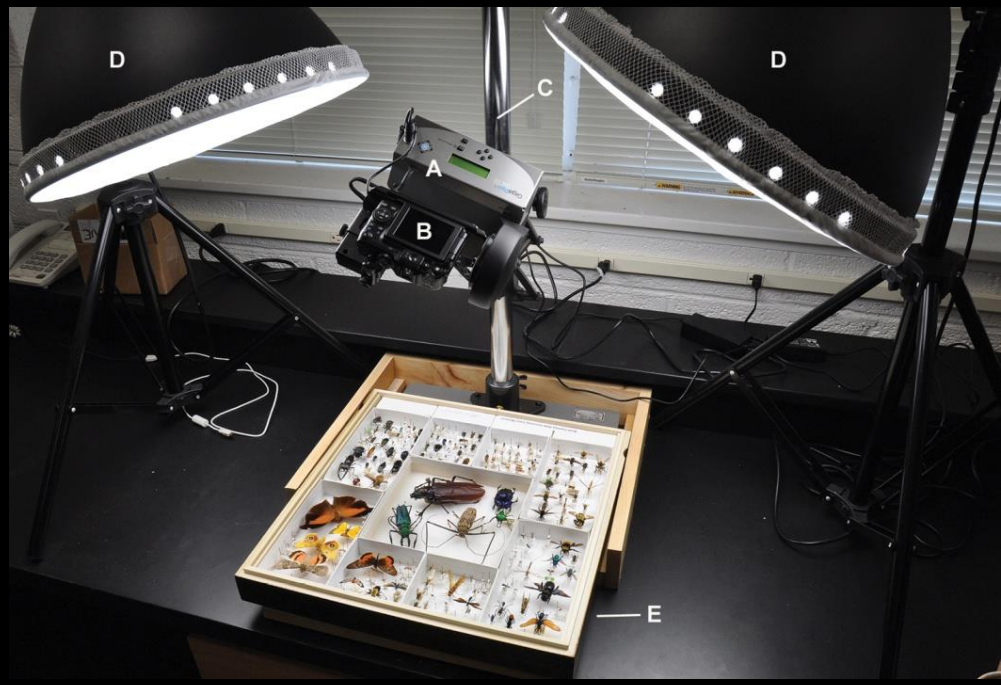


North Carolina State University (Bertone and Deans 2010)

- **Results and insights from the NCSU Insect Museum GigaPan Project**
- ZooKeys 209: 115-132 (20 Jul 2012)
doi: 10.3897/zookeys.209.3083
- Matthew A. Bertone
- Robert L. Blinn
- Tanner M. Stanfield
  Kelly J. Dew
- Katja C. Seltmann
- Andrew R. Deans

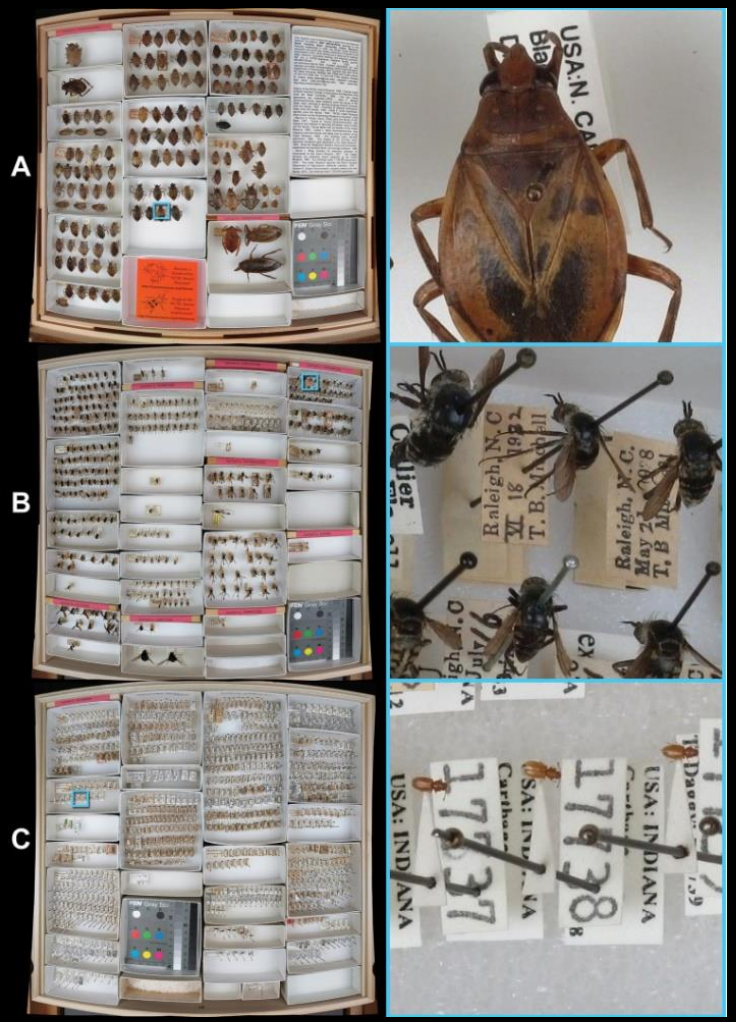
North Carolina State University (Bertone and Deans 2010)

- GigaPan Epic 100 (~$450)
- Canon G11 (~$500)
- Lighting (~$500)
- Copy stand (~$100)
- Other accessories (~$50)



North Carolina State University (Bertone and Deans 2010)

- GigaPan robot
  - Capture images of their entire collection, comprising over 2700 drawers
  - 1.5 million specimens
  - 150MB each drawers panorama data (including original photos, raw tile data, and gigapan panorama file). Thus, for the entire 2, 700 drawer collection, ~405 gigabytes of storage space was needed.
  - Limited view of label data
    - Only image of text
  - Camara mounted over center of drawer large angles at edges of drawers



Australian National Insect Collection

- **Whole-drawer imaging for digital management and curation of a large entomological collection**
- Beth Louise Mantle, John La Salle, and Nicole Fisher

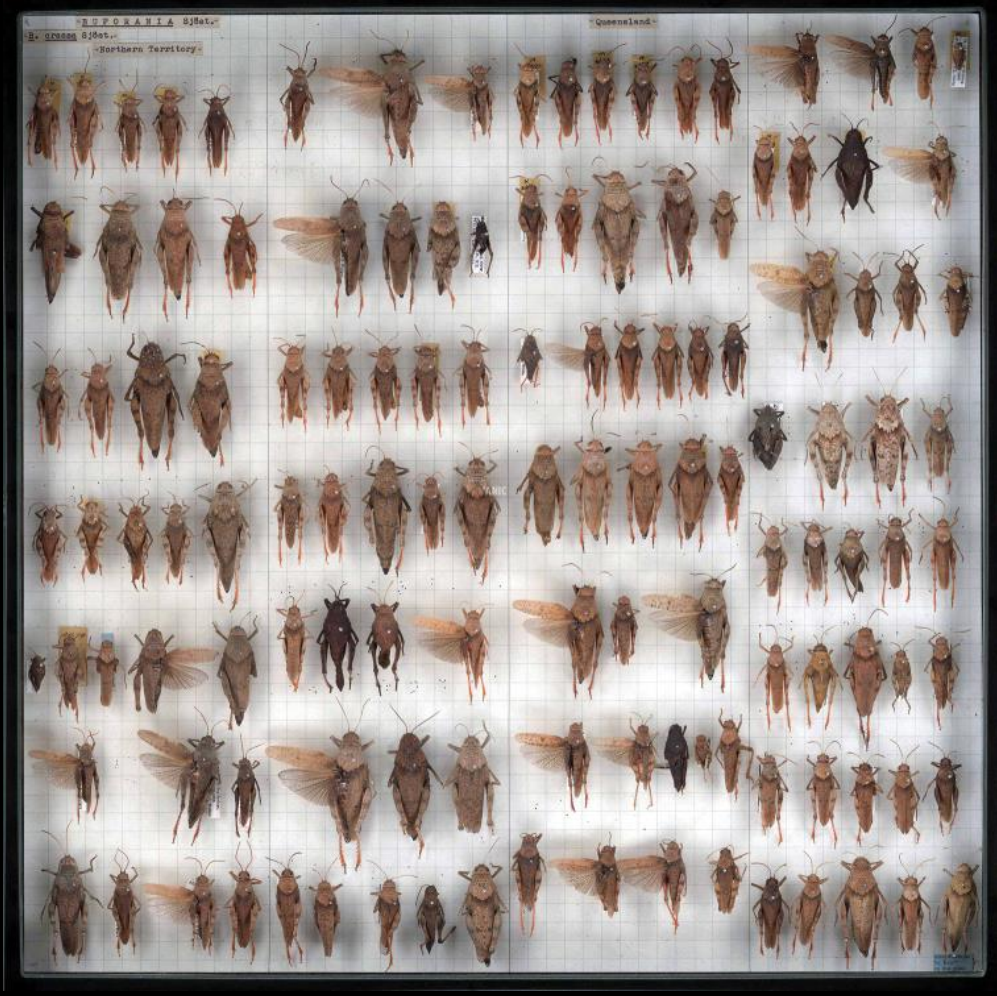
Zookeys. 2012; (209): 147–163.
Published online Jul 20, 2012.
doi: 10.3897/zookeys.209.3169

Australian National Insect Collection

- SatScan captures sequential “tile” images (200 – 400 per drawer
- Edmund Optics 0.16× telecentric lens #NT56-675
- Given an average capture time of 5–7 minutes per drawer, a skilled operator can process up to 60 drawers of specimens each day, and up to 90 final pictures can be stitched in 12 hours (e.g. overnight). These times are typical for a trained operator and bug-free software.

- Each drawer was assigned a unique identifier that also acts as a location code for the drawer within the collection

  ANIC entomology drawer measuring 480 × 500 mm produces a final image of 15000 × 14000 pixels, and file size of ~780MB (BMP) or 340MB (TIFF)
- At the time of publication, more than 1,500 collection drawers (from a current total of 22,000 drawers) have been imaged and uploaded to Morphbank-ALA

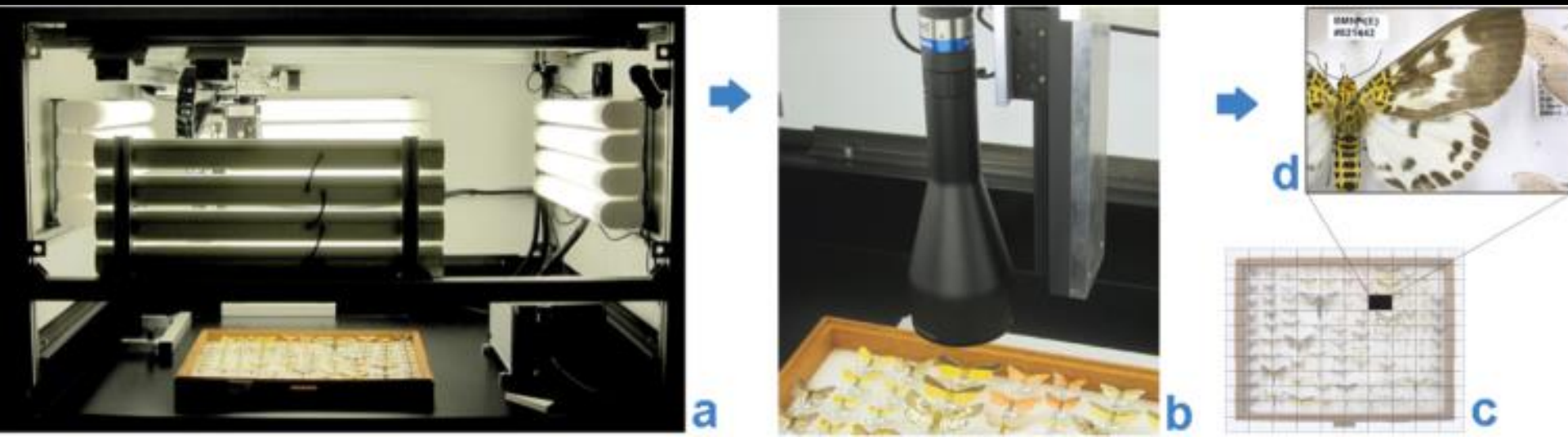


Australian National Insect Collection

- Whole-drawer images comprised of large specimens may facilitate species identification, images of small specimens have a higher probability of revealing useful and extractable label data.

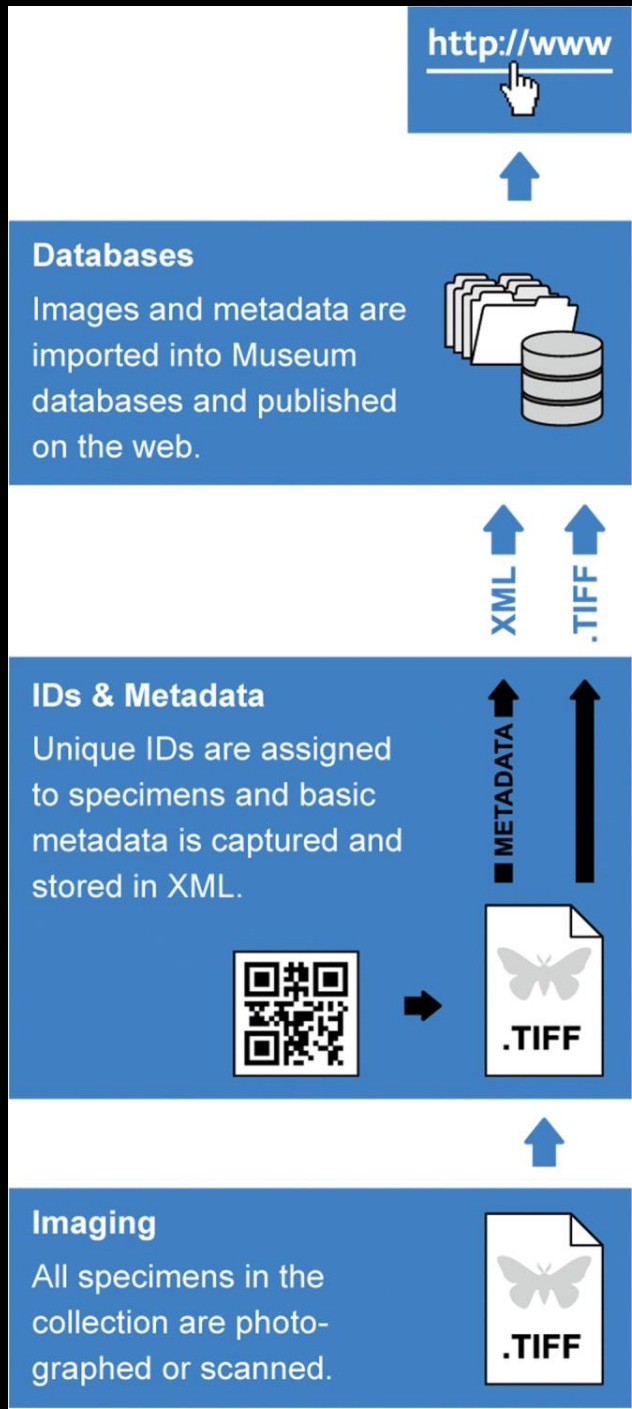
Natural History Museum, London

- **No specimen left behind: industrial scale digitization of natural history collections**
  Zookeys. 2012; (209): 133–146. Published online Jul 20, 2012. doi: 10.3897/zookeys.209.3178
- Vladimir Blagoderov
- Ian J. Kitching
- Laurence Livermore
- Thomas J. Simonsen
- Vincent S. Smith



Natural History Museum, London

- Scatsan
  - The entomology collections of the NHM have about 30 million insect specimens, mostly pinned, housed in 135,000 collections drawers. Assuming that 80% of the collection is appropriate to be imaged using the SatScan® system, rough calculations based on the above figures suggest that the entire collection could be imaged and basic metadata captured in 18 person-years.
  - Average scanning time 8-10minutes
  - Stitching time 13-26 minutes
  - File size 500 to 720 mb
  - (about 1.2%) have so far been databased and assigned a uID.

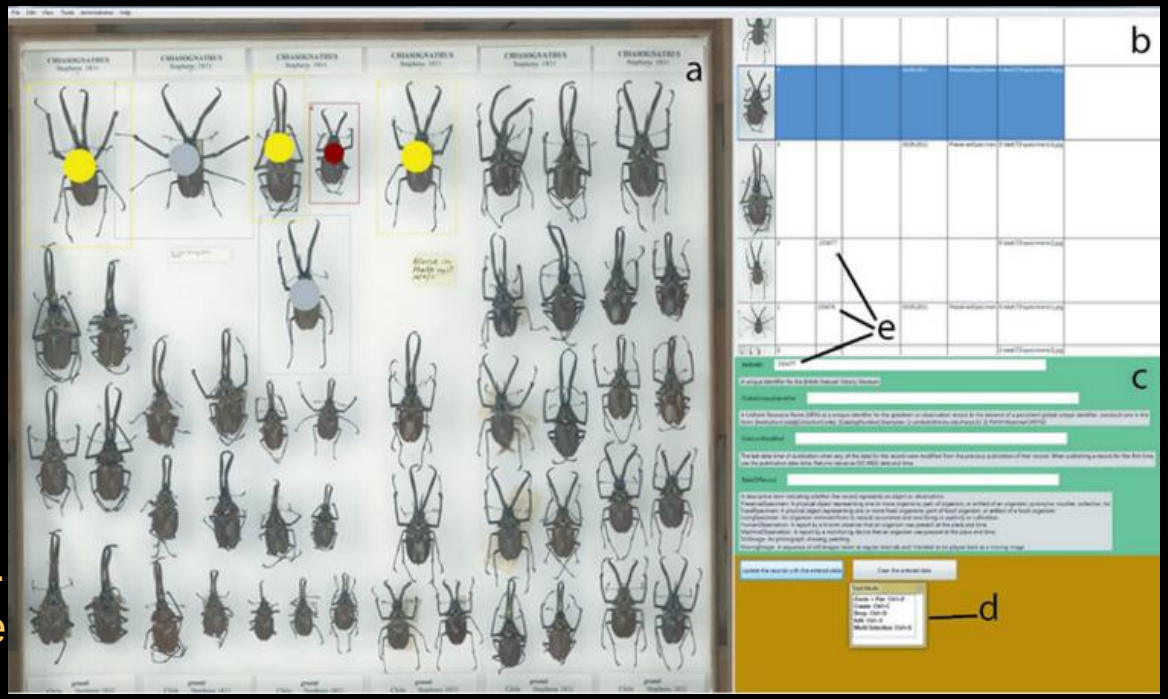


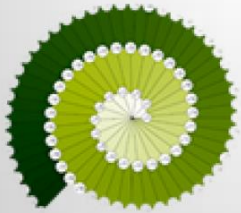
Natural History Museum, London

- The SatScan system solves distortion problem by using a orthographic camera lens.
- Camera does not tilt moves across the drawer
- 250 MB – 3 GB compressed TIFF file

Natural History Museum, London

- Software program, Metadata Creator designed to allow
  - fast capture of specimen data and
  - associating these with the image of the specimen
  - Users can mark individual specimens on the panoramic image by drawing rectangular boxes around them, selecting these areas and annotating them individually or in batches.
  - Label data not capture very we





InvertNet

Advancing Digitization of Biological Collections

InvertNet: A New Paradigm for Digitization of Invertebrate Collections

Imaging Hardware for larger specimens

Chris A. Taylor

Prairie Research Institute

Illinois Natural History Survey

University of Illinois

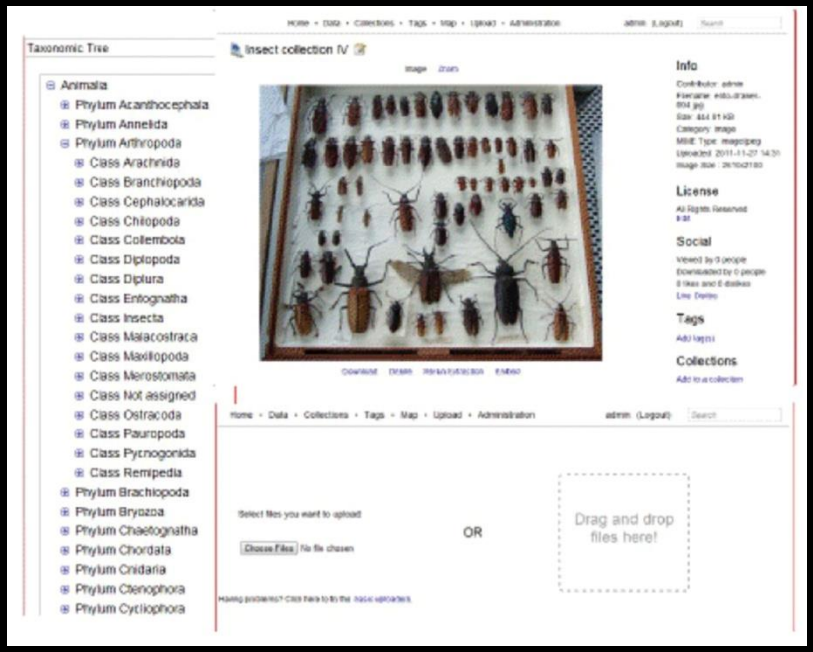
ctaylor@inhs.illinois.edu

InvertNet: a new paradigm for digital access to invertebrate collections

- Goal provide digital access to 60 million specimens (primarily insect from 22 different collections
- Chris Dietrich
- John Hart
- David Raila
- Umberto Ravaioli
- Nahil Sobh
- Omar Sobh
- Chris Taylor

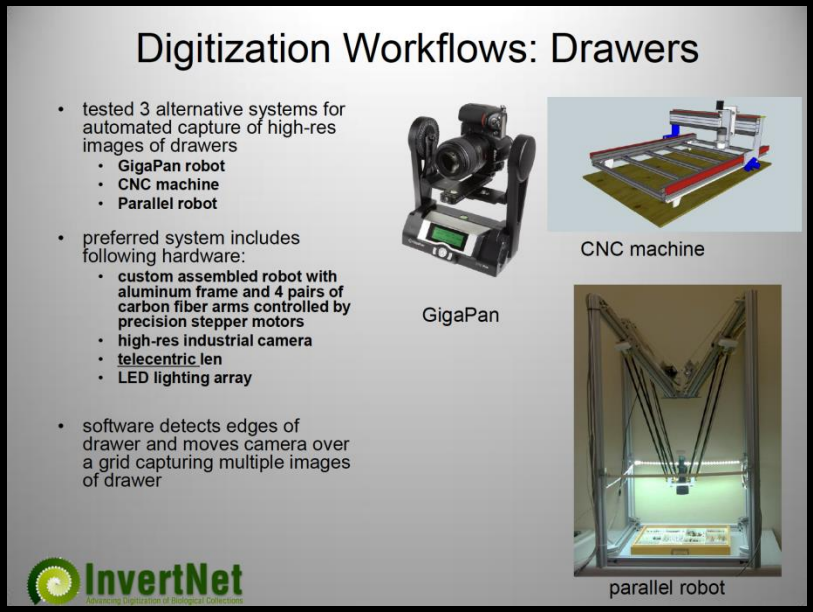
The InvertNet approach

- Primary goals
- automated as possible
- should capture an array of high resolution images from multiple viewpoints
- be inexpensive to purchase, operate, and maintain
- flexible to adapt
- upgradable once deployed



Three Approaches Tried

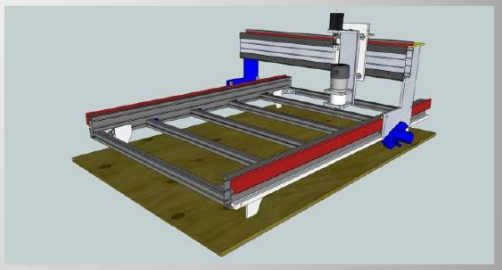
- GigaPan-style panoramas from different viewpoints
  - Increase time, effort, and the probability of human error if the drawer and/or camera must be re-positioned manually
- Computer Numerical Control (CNC) machine
  - Big, heavy difficult to set up and expensive
- Delta Robot
  - Best of three approaches



Digitization Workflows: Drawers

- tested 3 alternative systems for automated capture of high-res images of drawers
  - **GigaPan robot**
  - **CNC machine**
  - **Parallel robot**
- preferred system includes following hardware:
  - **custom assembled robot with aluminum frame and 4 pairs of carbon fiber arms controlled by precision stepper motors**
  - **high-res industrial camera**
  - **telecentric len**
  - **LED lighting array**
- software detects edges of drawer and moves camera over a grid capturing multiple images of drawer

GigaPan



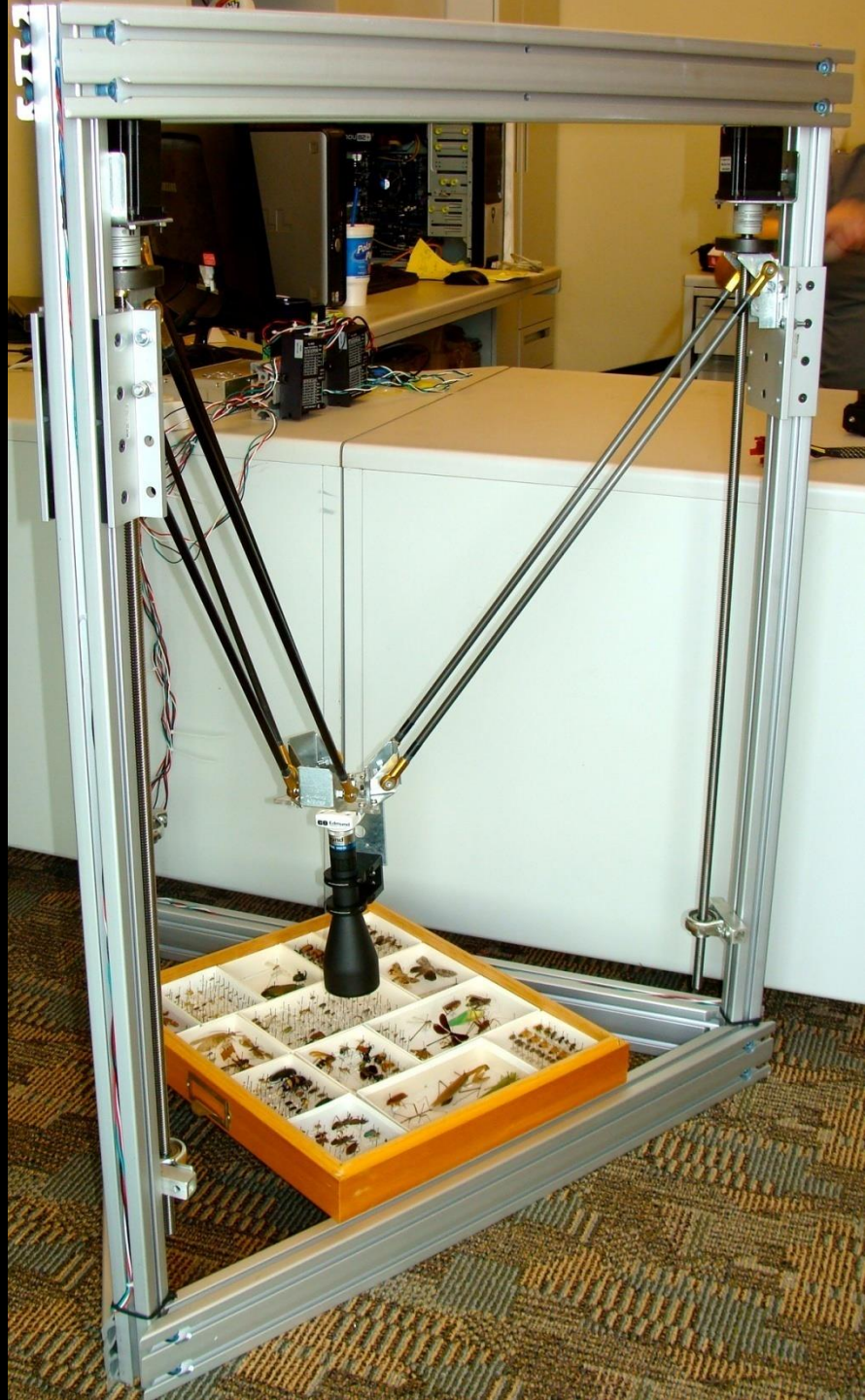
CNC machine

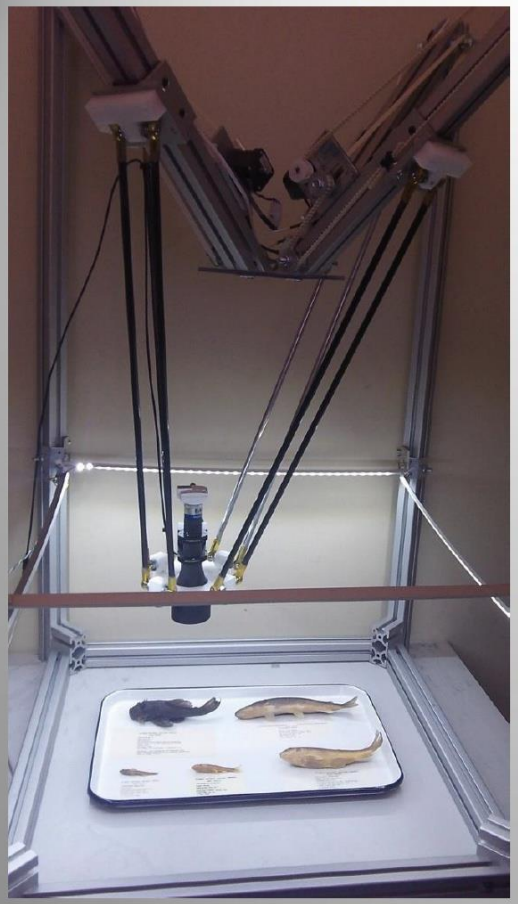
parallel robot

InvertNet
Advancing Digitization of Biological Collections

Stitching software

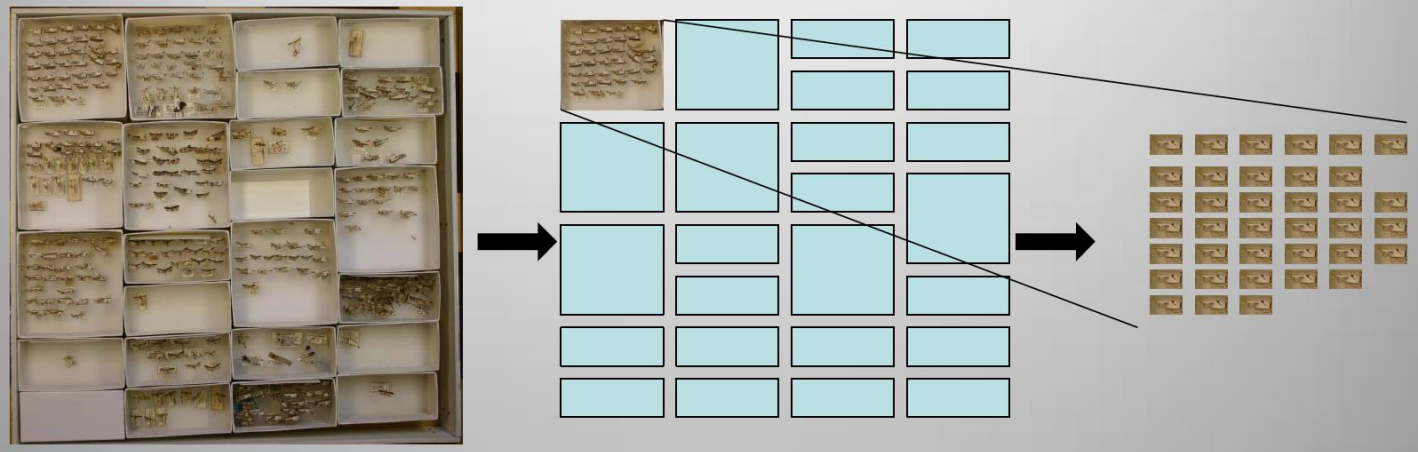
- GigaPan software system
- Open source
  - Hugin, OpenCV
  - Zoomable User Interfaces (ZUIs)
    - Zoomify
  - 3-D Reconstructions





InvertNet
Advancing Digitization of Biological Collections

Stitched images post-processing



1. capture image of drawer + metadata (location, contents)
2. segment unit trays (image analysis software)
3. segment specimens
4. capture label data

InvertNet
Advancing Digitization of Biological Collections

Digitization of other kinds of specimen storage units

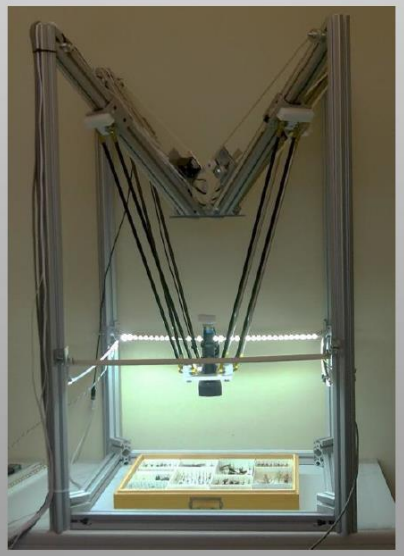
- Fluid-preserved specimens in vials
- Specimens mounted on microscope slides.

Hardware Cost

- Total cost of hardware and software (including computer) is around $5000

- Interface will be very user friendly

- Software needs to be pre-loaded by us



InvertNet
Advancing Digitization of Biological Collections