YPM-IP workflow
with
voice recognition imaging

Susan Butts
Yale Peabody Museum
Invertebrate Paleontology
Division of Invertebrate Paleontology

- estimated 4.5 M specimens
- 23,000 are type specimens (1,716 holotypes)
- over 317,000 specimen lots are available in the online specimen index including all known type specimens worldwide in scope
- 73% from North America
- over 150 years of curatorial and student collections
Collection strengths:
• Early Pz Maritime Canada and the Appalachians
• Late Pz Mid-Continent
• Western Interior Seaway
• Schuchert Brachiopod Collection
• Ciurca Eurypterid Collection
• St. Jean/Galloway Stromatoporoid Collection
TCN: Fossil Insect Collaborative: A Deep-Time Approach to Studying Diversification and Response to Environmental Change

- Lead: University of Colorado Museum of Natural History, Smith and Karim, 6 other collaborating institutions & 2 participating
- YPM will digitize >10,000 specimens, educational component
PEN: Increasing the Robustness of the Ordovician and Pennsylvanian Dataset of PALEONICHES-TCN

- Lead (TCN): University of Kansas (Lieberman, Beach, and Farrell)
- YPM will digitize
  1. specimens that have been cataloged in EMu, but lack georeferenced data,
  2. material cataloged in ledgers, and
  3. uncataloged material from our stratigraphic collection.
YPM-IP workflow

- Predigitization Curation
- RIPVR - Rapid Image Processing with Voice Recognition
- Data/Image Repository
- Images
- Catalog (KE EMu)
- Distribution
- Digital Specimen Record
predigitization curation

Undigitized material: received from the stratigraphic collection
1. Identification, lexicon check (optional: sites, georef)  
2. Assignment to undergraduate  
3. Conservation assessment (note in database)
4. Retray in acid-free archival trays
5. Clean labels with Groomstick
- KE EMu database with custom data entry form (mandatory fields and authority controlled fields)
- Apply specimen number directly to fossil with Pigma pen
catalog (georeferencing)
USGS GeoLex
Point-radius method (Wieczorek et al., 2004) – using GEO Locate or Google Earth
Imaging station
- copy stand with adjustable lighting
- devoted computer
- consumer-grade camera (Sony H5)
- specimens stabilized with black or white sand (Ward’s Scientific)

New:
- Canon D60
- tethered imaging with Canon Digital Photo Professional
- color calibration

*must occur after catalog number is assigned*
Hands free set up
• Wireless headset (Logitech G930)
• Dragon Naturally Speaking or Windows Speech Recognition
• Voice-driving data entry into Excel

each user needs to “teach” the voice recognition system

simple commands!

look up lists?
Hands free set up

• Wireless headset (Logitech G930)
• Dragon Naturally Speaking or Windows Speech Recognition
• Voice-driving data entry into Excel

*each user needs to “teach” the voice recognition system*

*simple commands!*

*look up lists?*
YPM-IP workflow

Predigitization Curation

RIPVR - Rapid Image Processing with Voice Recognition

Data/Image Repository

Images

Catalog (KE EMu)

Distribution

Digital Specimen Record
Semi-automated upload
• shell script validates camera images, renames with IP naming protocol – specimen number + suffix (orientation)
• ImageMagick resizes and generates composite image
• 2nd shell script prepares CSV files for automated import to EMu multimedia module and automated attachment of images to object record

All developed by Larry Gall, YPM Head of Computer Systems

You NEED a Larry!
Semi-automated upload

- shell script validates camera images, renames with IP naming protocol
  – specimen number + suffix (orientation)
- ImageMagick resizes and generates composite image
- 2nd shell script prepares CSV files for automated import to EMu multimedia module and automated attachment of images to object record

All developed by Larry Gall,
YPM Head of Computer Systems

You NEED a Larry!
Archiving – Yale University Digital Asset Management
• replicates image under multiple file types
• open access jpegs, restricted access tiffs (controllable)
• Administered by Yale Digital Collections Center (YDC2) – museums and libraries at Yale

YPM search
Data aggregators
• iDigBio
• PaleoPortal
• iDigPaleo (TCN:FIC)
• gbif
Archiving – Yale University Digital Asset Management
- replicates image under multiple file types
- open access jpps, restricted access tiffs (controllable)
- Administered by Yale Digital Collections Center (YDC2) – museums and libraries at Yale

YPM search
Data aggregators
- iDigBio
- PaleoPortal
- iDigPaleo (TCN:FIC)
- gbif