Ichthyology collection digitization

Andrew Bentley
Biodiversity Institute, University of Kansas
Vertebrate/Ichthyology advantages in digitization

- Collections relatively small
- Good historical practices of use of ledgers, card catalogs, field notes
- Early adopters of digitization
- Ichthyology – use of lots rather than individual specimens
Digitization progression

Path of digitization

**Undigitized**
- Card catalogs
- Ledgers
- Field notes
- Labels

**Primary digitization**
- Object records in database

**Secondary digitization**
- Georeferencing
- Field notes
- Specimen images
- Locality images

**Web accessibility**
- Local portal
- Fishnet2
- GBIF
- iDigBio
- Vertnet

**Vertebrates**

**Paleo and entomology**

**Herbaria**

**Vertebrates**
Digitization steps

- Primary digitization – primary data to create specimen records
  - Ledger
  - Card catalog
  - Labels
  - Field notes

- Secondary digitization – augmentation of specimen data with additional information from digitized content
  - Field notes
  - Specimen images (35mm slides and digital), X-Rays
  - Locality images
  - Other media – video, audio
  - PDFs of reference works
Primary digitization

Bound (ledger) or unbound (card catalog), labels and in some cases, field notes can be scanned or photographed. In color to capture all data and amendments, attach to relevant records in a robust, customizable database with the facility to attach media. Attached images can be used to capture primary data during the database selection process.
Bound vs. unbound

- **Unbound**
  - Scanner – preferably document feeder
  - Camera – light box, copy stand
  - Color

- **Bound**
  - More problematic
  - Scan or photograph – not ideal unless item can be opened flat
  - Book scanner – expensive and not readily available
  - Professional company – Scan America
Specimen Imaging

- Most ichthyology material better taken wet – either in water or preservative – to prevent desiccation and improve image quality – morphology rendered in higher quality

- Vast variety of sizes from larval to sharks and rays

- Cleared and stained and skeletal material

- Cost per unit effort – available resources – increases with increasing complexity – 2D field image, 2D high quality image with post processing, 3D image, MRI
Problems:

• Shadow
• Uneven lighting
• Glare
MK Digital Direct Light Box

http://www.mkdigitaldirect.com

Photo e-Box
Photo e

- Box 2032 system
  - $3,700
- Camera - Canon EOS T3i
  - $850
- Copy stand with adjustable rail
  - $450
- Laptop optional
  - ~ $5,000

Internal platform size – 32 inches x 20 inches
Features

- Individually adjustable multiple light sources
- 20 bulbs (6500K), 9 switches
- Camera ports top and front
- Adjustable copy stand head with rail
- Larger size
- Various colored backgrounds
- Black internal upper surface to prevent reflection
- Extractor fans
Wet boxes

Glass box with Plexiglass frame

32 inches x 20 inches

16 inches x 10 inches
Ichthyology

Priorities

• Type material *
• Tissue vouchers
• Rare/threatened/endangered
• Sp. material for identification
• As accumulated or requested – internal and external
In the field...
Utility of Images

- Confirm or make identification
- Morphological examination – landmarking (3D)
- Color morphs*
- Field guides
- Publication
- Other external user groups – K-12, artists, publishers
- 3D, MRI – landmarks, internal structure
Utility of Field Notes

- Confirm collection information – dates, times, locality, collectors etc.

- Augment collection information

- Additional information – associated species, environmental parameters, specimens collected but not kept etc.

- Travelogue

- Next step - Transcription – e.g. Notes from Nature
  - Higher value and integration of data
Other augmented data

- Genbank sequences
  - Reuse of non-renewable resource
  - Metric of collection use
  - CO1 on BCoL and BOLD

- Publications and citations
  - Metric of collection use
  - Repeatable science
  - Copyright infringement?

- Video and audio – behavioral information, environment, interspecies interactions
Infrastructure & Management

- Database able to handle digital content – attachments able to accommodate all kinds of media – images, video, audio, PDFs etc.

- Server infrastructure able to handle volume of material – space requirements

- Ability to serve this content up to the outside world – local portal and webpage, aggregators, research users, other users
Access to data

- Final step to provide access to all data
- IPT portal and publishing through aggregator(s)
  - Vertnet, GBIF, iDigBio, Fishnet2
  - The more the merrier...
- Database web portal
- Data cleanliness – controlled vocabularies
- Exposing data to research community
- Provides assistance with data cleanup and ID
- Large and diverse external user community – big data
- Increase in use (loan traffic) and advocacy for collection
Other developments

- iDigBio
- SPNHC – best practices wiki
- Biodiversity Collections Network (BCON)
- Prezi - [https://prezi.com/h9l4jfwexwsc/bcon-diagram/](https://prezi.com/h9l4jfwexwsc/bcon-diagram/)
- ASIH – Collection Committee and best practices
- TCN...