CONNECTING NH COLLECTION DATA

Jonathan L. Dunnum

Division of Mammals
Museum of Southwestern Biology
University of New Mexico
Albuquerque, New Mexico USA
Our challenge

- Huge digitization efforts underway
- Need to look outside our own collections
  - What’s the big picture
- Opportunity to do more than just facilitate your collection management of specimens
- Identify who we need to connect to and what information they need
Historically
- NH collections and data limited to subset academia
- Guarded access
- Destructive sampling very rare
- Museums have paid a price

Today - future
- Can’t afford to be as restrictive
- Now creating infrastructure for everyone else
- Renaissance for NH collections
Significant questions are centered on our ability to assess change.

- Climate change
- Habitat conversion
- Pollutants
- Emerging pathogens
- Introduction of exotics
- Loss of biotic diversity

- Baseline or historic information is critical to documenting changing environments.
- Only NH collections have these samples
  - New questions
  - New technology
Critical time for research and education on biodiversity issues.

- Build connections
- Rethink the traditional role of museum’s and databases
- Facilitate use
  - Access allows for creativity, new uses and new ideas
- Connect Big Data
  - Cyberinfrastructure for Informatics
Critical time for education on biodiversity issues

• Data resource will keep growing
• Need to develop human resource

• Train Future Investigators
.......to creatively explore, utilize and integrate these vast resources across disciplines and into critical science initiatives.
Connecting to ...........

- Wide array of academic researchers
- Federal and state agencies
- Public health organizations
- NGOs / Environmental consultants
- Undergrad education
- K-12 educators
- Artists

Varying levels of needs, questions, complexity
Non-traditional users

- Making natural history collections available to non-traditional users
- Artists provide not only new perspectives, but also are advocates for biodiversity issues

Gendron Jensen lithograph
Last brown bear taken from NM Pecos wilderness in 1917

Morphology and Evolution with Brian Conley
- Community and comprehensive collection management system
- Open-access specimen database [http://arctos.database.museum](http://arctos.database.museum)
- 3M records
- 90 collections, 19 institutions
  - Shared instance - TACC
  - Single instance - MCZbase Harvard

---

**Museum of Southwestern Biology**
University of New Mexico

**The Museum of Vertebrate Zoology at Berkeley**

**Museum of the North**
University of Alaska

**Denver Museum of Nature & Science**

**Cornell University Museum of Vertebrates**

**Harvard University**
Museum of Comparative Zoology
Breadth of Data Types in ARCTOS

Audio  
DNA  
Relations  
GIS  
Projects/Pubs

SPECIMENS

Fieldnotes, Maps, and Images
Queriable database

Internal collection management

Connectivity to other databases
Connecting data within a collection

- Diverse parts and associated data
- Stored in many places
- In multiple databases

Holistic voucher

- Traditional skin, skeleton or fluid voucher
- Ecto & endo parasites
- Field notes, ecological data
- Frozen tissues
- Genomic data
- Karyotypes
- Virus isolates
DB needs to keep everything connected.
One piece wraps cryotube.

Other piece affixed to data sheet.

Less transcription work. Less (no?) transcription errors.

Tubes (containers) can be installed immediately. Specimen’s parent container is scanned in at data-entry.

Tissue, parasites, skels, etc.
Every container is within a parent container.

Recorded dimensions and logic prevent placement of a large container into a smaller parent.
Ties projects, specimens, permits, loans, publications together
Conversion of catalog numbers to GUIDs

- URL and Darwin core triplet
- Unique and standardized
- Can go around VPDs
- Allows your data to go out and play with others
Between collections -- within and among Institutions

- Integrated collections
- Host / parasites
- Division of parts between institutions
Symbiotype example

| Symbiotype | GenBank=HQ285047; GenBank=HQ285046; GenBank=HQ285047; GenBank=EU192164; GenBank=EU192165; | MSB: Mamm:96073 | 101580 |

**MSB: Mamm:96073**

**101580**

**Symbiotype**

**GenBank=HQ285047; GenBank=HQ285046; GenBank=HQ285047; GenBank=EU192164; GenBank=EU192165;**

**NK: 101580**

**Oligorzyomyx fultevscens**

**Central America, Panama, Los Santos**

**02-Mar-2000**

**Identification**

**GenBank=HQ285047**

**GenBank=HQ285046**

**GenBank=EU192164**

**GenBank=EU192165**

**Preparator number: 444**

**Linkage to publications**

**Citation**: Symbiotype of *Oligorzyomyx fultevscens* in Vincent et al. 2000, not specifically cited but one of the parasites from the group that was subsequently sequenced in Nelson et al. 2010.

In format that can be accessed and served by other databases.

- IDigBio
- GBIF
- VertNet

Active links back to Arctos
Other Big Data

- GIS
- GenBank
- IsoBank
Linkages to GIS applications

- Geolocate
- Berkeleymapper
- Googlemaps
Linkage to GenBank

- Reciprocal links in Arctos and GenBank
- Darwin Core triplet (MSB:Mamm:123456)
- Cat # in voucher and definition fields
- Tying genes to time and space – only vouchered material
Andalgalomys pearsoni dorbignyi (MSB 55245)

DNA sequence submitted to GenBank AF159285

Genetic Data
10. Alvarado-Serrano and D’Elia. 2013
11. Schenk et al. 2013
12. Ventura et al. 2013
13. Pardinas et al. 2014

22 publications using data from a single specimen collected in 1984

Morphologic Data

Taxonomic References
Google Scholar Profile

MSB Division of Mammals:

<table>
<thead>
<tr>
<th>Citation indices</th>
<th>All</th>
<th>Since 2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>Citations</td>
<td>42723</td>
<td>15296</td>
</tr>
<tr>
<td>h-index</td>
<td>87</td>
<td>48</td>
</tr>
<tr>
<td>i10-index</td>
<td>511</td>
<td>276</td>
</tr>
</tbody>
</table>

Use recognized metrics to demonstrate value
Questions?

Demos?