

CHALLENGES TO MAKING PALEONTOLOGICAL DATA USABLE TO A BROADER AUDIENCE

Patricia
Holroyd

University of
California
Museum of
Paleontology

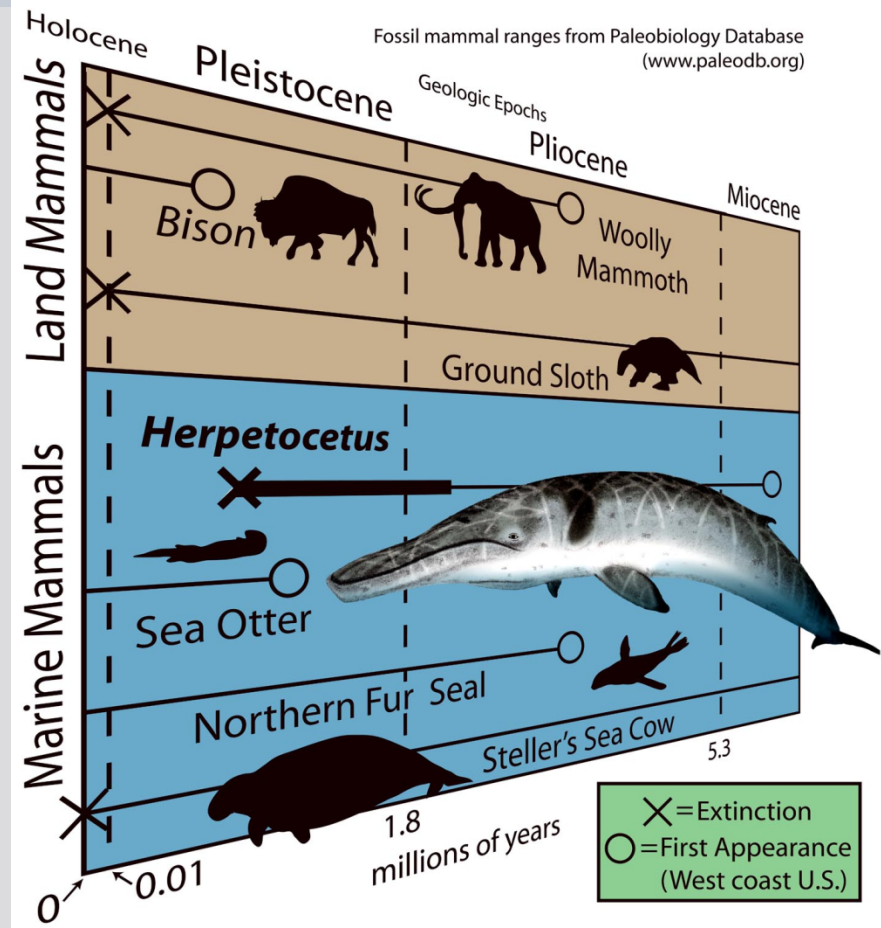


UCMP 53720, a Pleistocene bettong
from Lake Menindee, Australia

GOALS

- incorporate data from the fossil record to complement studies of the modern biota
 - Climate change studies
 - Past distributions
 - Phylogeny & divergence times
 - Functional morphology

UCMP 153132
La Brea Bison
sampled for carbon
& oxygen isotopes



GOALS

- Do good research with good data
 - Need to assure that people know source of data and what types of studies and methods are appropriate



WEB-MOBILIZING SPECIMEN IMAGES IN UC MUSEUM OF PALEONTOLOGY

- Calphotos linked to online specimen data

UCMP Paleontology Collections Photos
University of California Museum of Paleontology

Back to: [UCMP Collection Search](#)

SEARCH FOR: [Specimens](#) • [Specimens \(advanced\)](#) • [Localities](#) • [Photos](#)

BROWSE PHOTOS: [Field Sites](#) • [Modern Cleared Leaves](#) • [Vertebrate Fossils](#) • [Plant Fossils](#) • [Microorganism Fossils](#)

The collections catalog currently contains 9,789 images.

Search for UCMP images using this form. If you choose more than one option, the search will look for records that meet both conditions (e.g. Country AND Taxon).

Type of Photo

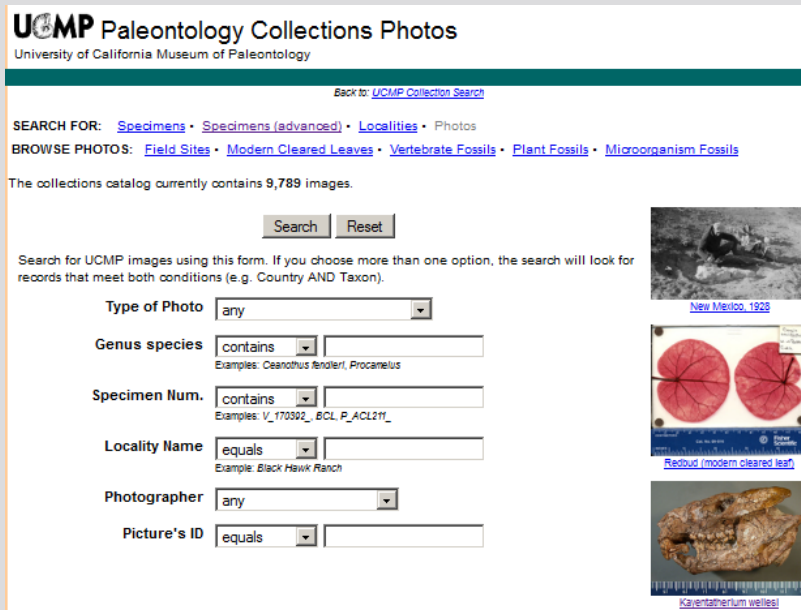
Genus species
Examples: *Ceanothus fendleri*, *Procamelus*

Specimen Num.
Examples: V_170302_BCL_P_ACL21f_

Locality Name
Example: Black Hawk Ranch

Photographer

Picture's ID



Example images shown in the search interface:
- [New Mexico, 1928](#)
- [Redoubt \(modern cleared leaf\)](#)
- [Kaventarium wellies](#)

~10k of >390,301 images in CalPhotos

- NSF-BRC Rehousing & Digitization of former USGS collection at UCMP




Est.
70,000
images

Intended uses: documentation, management, course support, outreach

CHALLENGE 1: MOST ONLINE DATA ARE NOT LINKED TO COLLECTIONS

- Most data are curated datasets of occurrences, images, that may be based on but not tied to collections data

<p>Morphobank media number M133655</p> <p>Taxonomic name †<i>Purgatorius janisae</i></p> <p>Specimen †<i>Purgatorius janisae</i> (UCMPNP:107406)</p> <p>View lower tooththrow occlusal</p> <p>Media loaded by Mary Silcox</p> <p>Copyright holder Silcox</p> <p>Copyright information Person loading media owns copyright and grants permission for use of media on MorphoBank</p> <p>Bibliographic Reference Silcox, M. T. 2001. <i>A Phylogenetic Analysis of Plesiadapiformes and their Relationship to Euprimates and Other Archontans (unpublished Doctoral dissertation)</i>.</p> <p>Media notes Fig. 3.10A from Silcox, 2001 Ip5-m1</p>	<p>Typical Morphobank media record</p>	 <p>» Zoom » Download</p>
---	--	---

PALEO DATA PROVIDERS

- Curated data set providers
 - Occurrence data from published sources



Strong publication bias



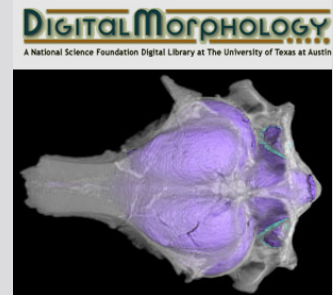
Reference to collections uneven

Data may be very targeted



PALEO DATA PROVIDERS

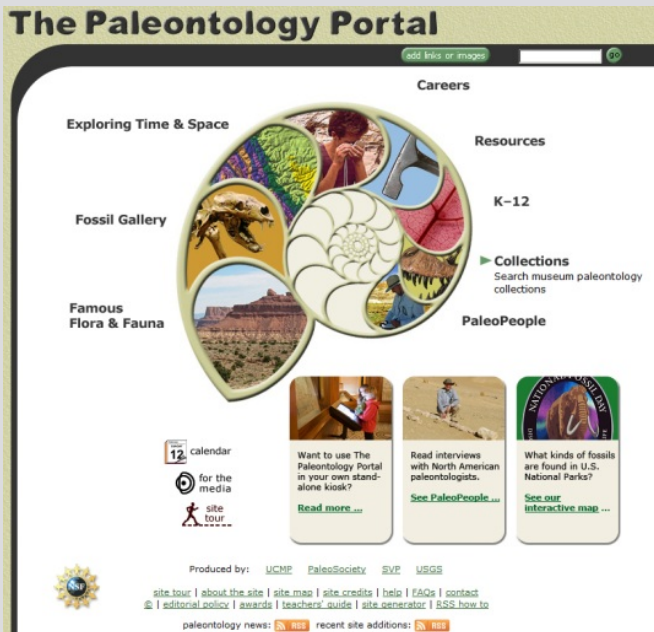
- Curated data set providers
 - Morphological or other derivative data



Data acquisition is purpose-driven (e.g., publication or coursework)
Data standards evolving at cutting edge of technology
New repositories being funded and created

PALEO DATA PROVIDERS

- Specimen data providers



Limited number of data sets

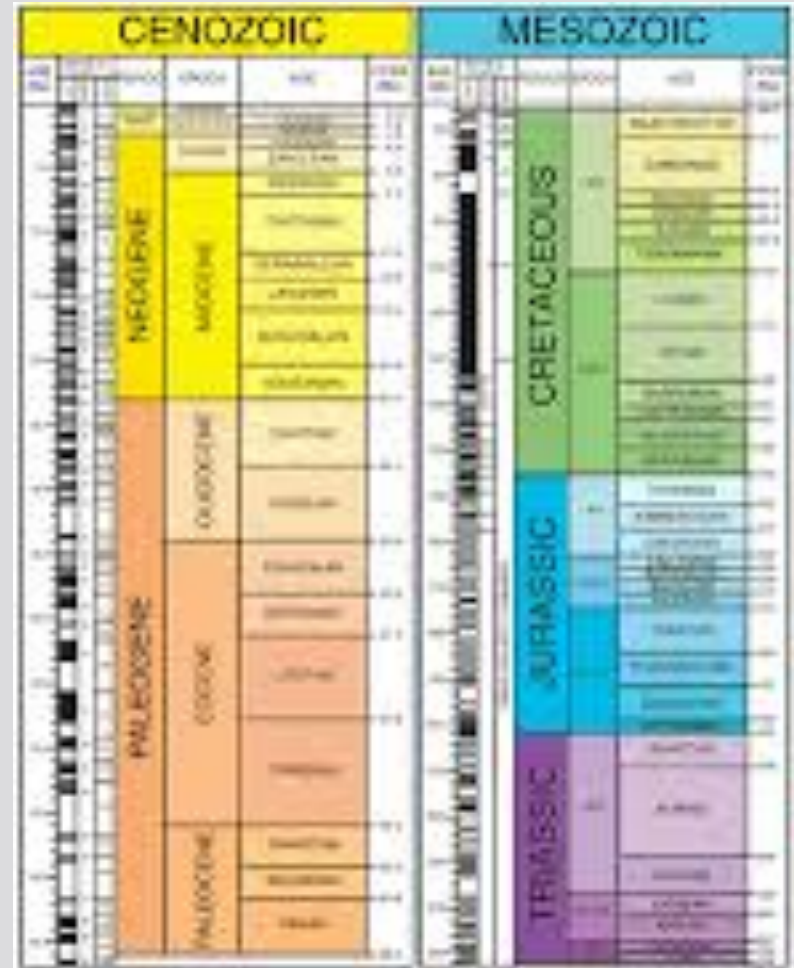
Primarily GIS based analytic tools

CHALLENGE 2: “SOURCE OF DATA” UNIVERSALLY UNCLEAR

- GBIF and BISON tag paleo data as “fossil”
- Like extant data, paleo occurrences may be based on
 - Vouchered specimens in a museum
 - Museum-held duplicates of specimens in private hands
 - Unvouchered specimens in a museum
 - Field observations of taxa
 - Taxonomic lists
 - Utterly baseless speculation

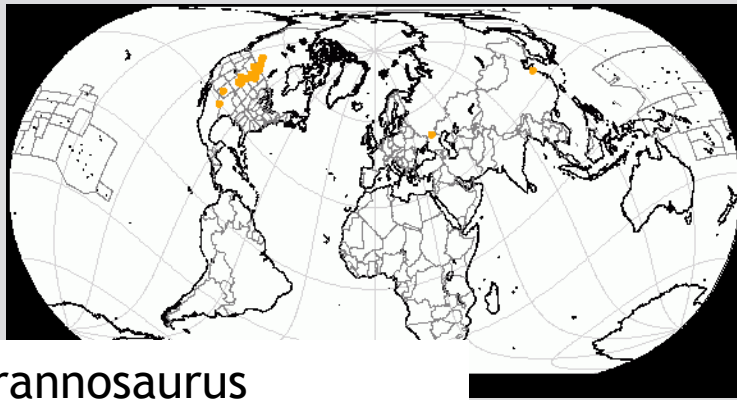
CHALLENGE 3: GEOLOGIC TIME CONFUSES PEOPLE

- Most date info is stored as names
- Darwin Core has these fields
- Data aggregators need to serve at least minimal data on the age and/or geology (often an age proxy)

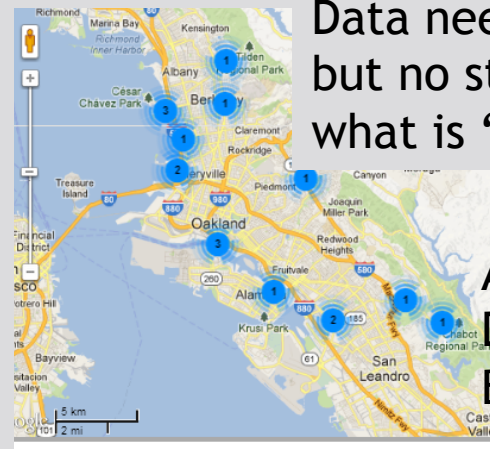


CHALLENGE 4: GEOGRAPHIC DATA MAY BE MISLEADING

Encumbered Data:
Public Law 111-011, Title VI,
Subtitle D, aka Paleontological
Resources Preservation Act (PRPA,
the Act, 16 U.S.C.)
passed in March 2009.



Tyrannosaurus
Paleo map from PBDB



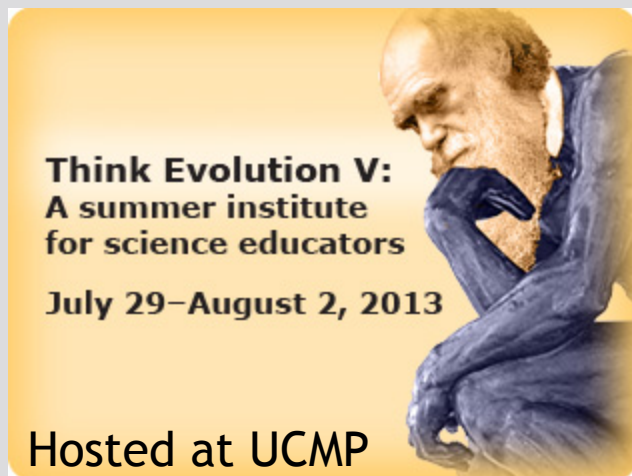
Data need to be fuzzed
but no standards yet for
what is “safe”

Auto-aggregated
Data clusters
Berkeleymapper

...or isn't what you think
because the earth moves

WHERE ARE WE NOW & WHAT CAN WE DO BETTER?

- Many of the informatics solutions are being met or in development
 - we need better concept mapping
 - Improved data interactivity
- Many K-12 resources available and in development

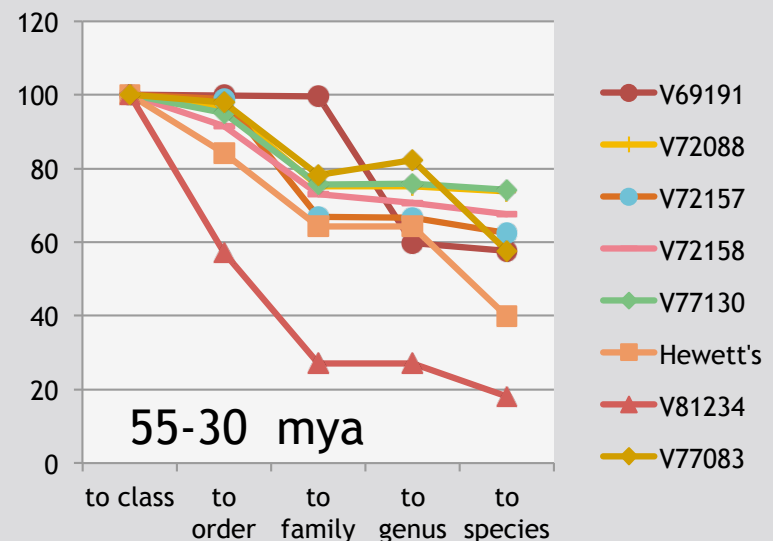
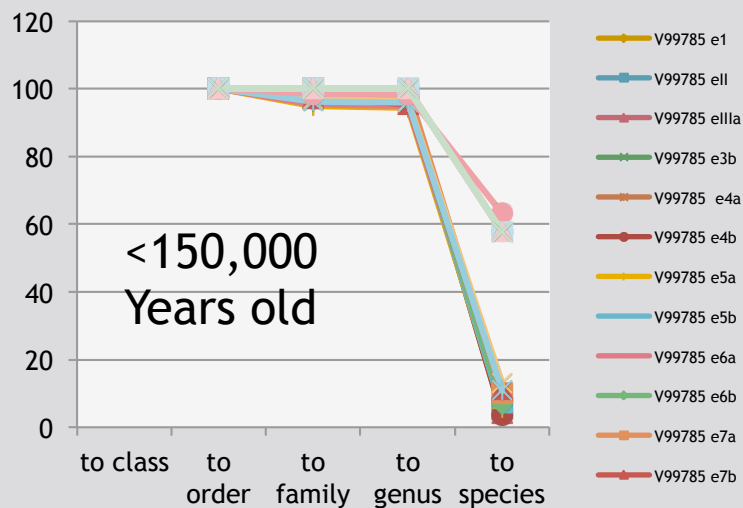


WHERE ARE WE & WHAT CAN WE DO BETTER?

- Need to focus on expanding and improving the research / reuse of data
 - Further develop professional development tools for young and old scientists
 - Collaborate to develop best practices for data use
 - Parham et al. 2012. Systematic Biology - for fossil calibrations

WHERE ARE WE & WHAT CAN WE DO BETTER?

- Need to focus on expanding and improving the research / reuse of data
 - Need data publications to explain collections data
 - Important across taxonomic “silos”



Specimen level taxonomic resolution in Quaternary vs. Tertiary faunas

THANKS

- Pam Soltis and iDigBio
- BIGCB funders: Moore and Keck Foundations
- The many people of Vertnet, CalPhotos, and UCMP who collaborate in our data dissemination efforts