

# Strategies for Integrating Source Materials and Paleobiology Collections



Patricia Holroyd



# Key Source Material Metadata

## Common Types of Source Material Metadata

- Who Made It?
- When Was It Made?
- What Is It?
- How Much of It Is There?

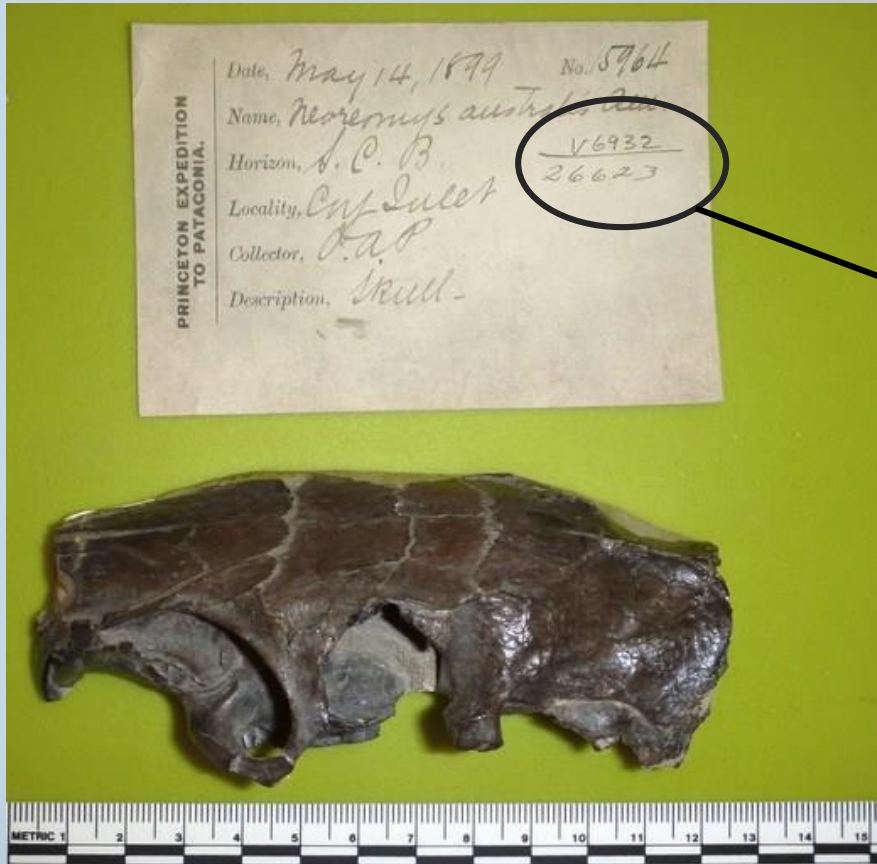
## Key Biodiversity Metadata

- How does it relate to a specimen and to a particular time and place in earth history (i.e., a georeferenced point and a time of occurrence)?
- For paleontology, time of occurrence is in specialized vocabulary (in extended Darwin Core)

# Discipline-Specific Metadata Issues

Collecting info is not always explicitly related to the organism's age or environment

Use locality numbers to relate multiple specimens to age/environmental data

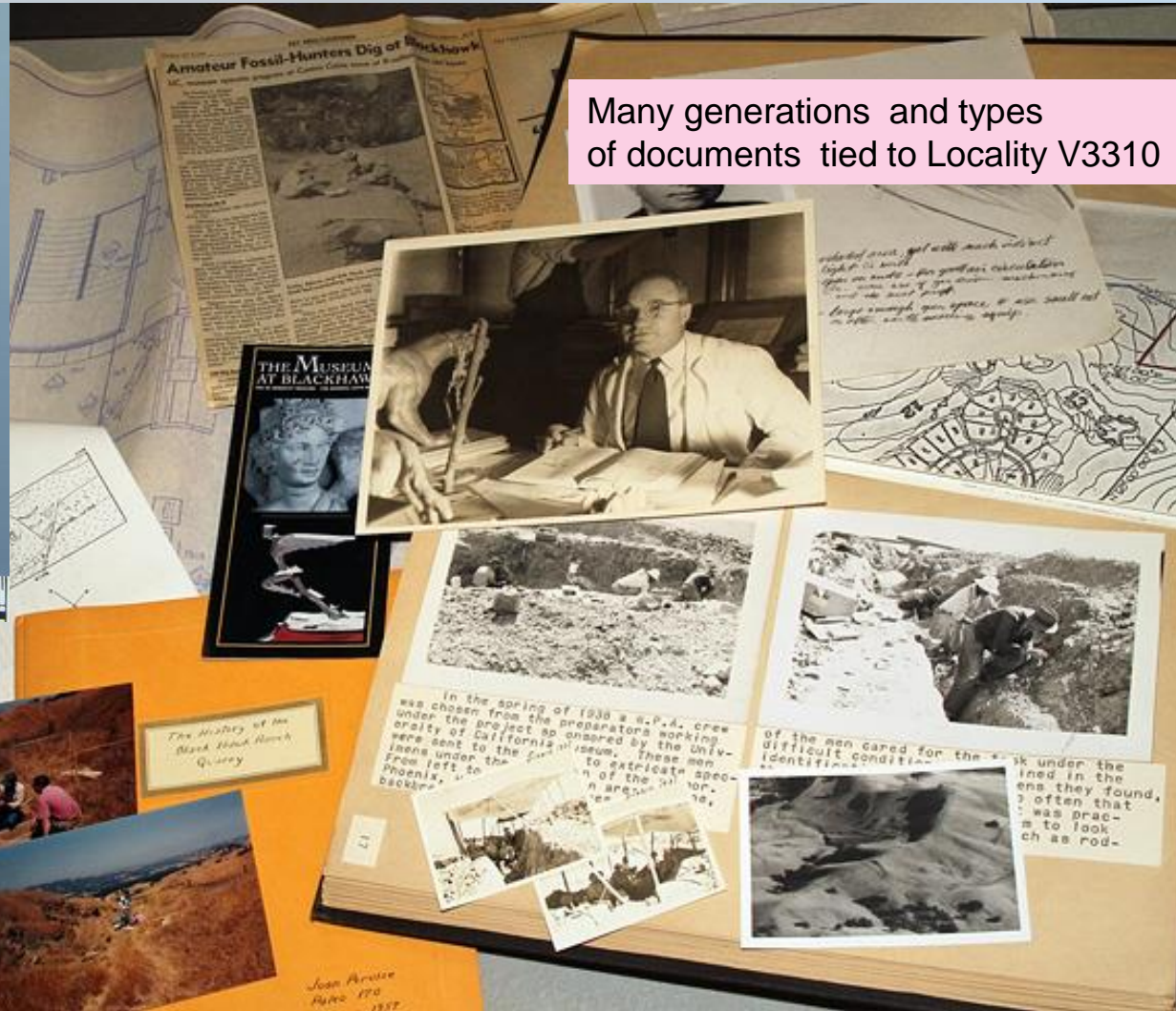




# Organizing Data Tied To Occurrences



Locality V3310  
Camel jaw  
No collector listed



Many generations and types  
of documents tied to Locality V3310

# Common Source Materials



Many Kinds of Media

Different Storage Needs

Differing ways to organize

Field Notes

Maps

Museum Records

Scientific Artwork

Stratigraphic sections

“Gray literature” reports

Personal papers &

correspondence

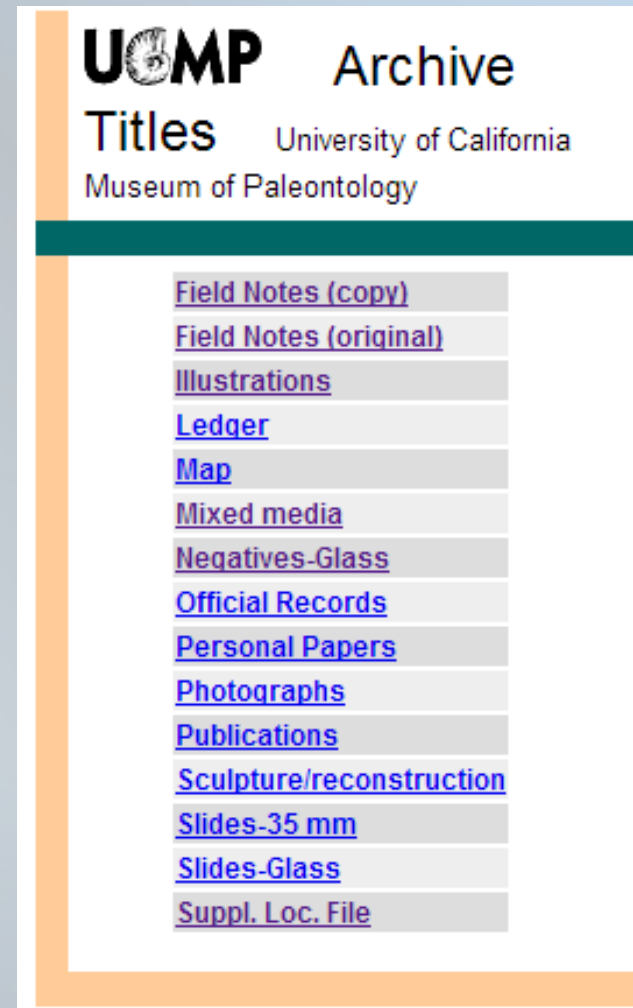
All types of photographic  
media

And increasingly “born  
digital” data (e.g., CT  
data, GPS logs, email)

# Organizing Source Materials in “Functional Units”

Groupings based on

- common uses or points of reference in existing catalogs and curatorial work flows
  - Check-box for supplemental locality files on locality cards
- Media type/storage requirements  
(i.e., maps in a map case)





# Organizing by Occurrence

## Supplemental Locality Files

### California-Santa Cruz County Supplementary Locality Files, 1917-2003 🌐 | University of California Museum of Paleontology

**Title:** California-Santa Cruz County Supplementary Locality Files, 1917-2003 🌐

**Extent:**

➤ [More Extent Information](#)

4 BW polaroids, 9 BW photos, 2 colored photos, 1 plate, 16 letters, 6 locality cards, 9 maps, 3 reports, 2 sketches, 2 strat. sections, 3 notes, 3 copies of field notes, 4 specimen identification cards

➤ [Arrangement](#)

➤ [Subjects](#) (links to similar collections)

➤ [Languages of Materials](#)

➤ [Administrative Information](#)

🖨️ [Printer-friendly](#) | ✉️ [Email Us](#)

**Scope and Contents:** Photos, letters, maps, and documents related to collections made in Santa Cruz County.

[Detailed Description](#)

[Folder 1: Purisima Fm., Miocene: V2701 Santa Cruz 1, V76109 Soquel Cliff, Misc., 1925-1926, 1933, 1971, 1972, 1975, 1976, 1980](#)

[Folder 2: Santa Cruz Mudstone Fm., Miocene: V99496 Borrow Pit Hill, 2003](#)

[Folder 3: Santa Margarita Fm., Miocene: V6857 Taylor Quarry, V6200 De Laveaga Park 1, V6376 Kaiser Olympia Quarry 2, V5244 Felton 2., 1952, 1973, 1990](#)

[Folder 4: Poison Rock, Scaroni Rock, Mussel Point, 1917](#)

[Folder 5: Miscellaneous \(1 of 2\), 1959, 1968, 1970, 1992](#)

[Folder 6: Miscellaneous \(2 of 2\), 1978](#)

The geographic location of an occurrence should be the most stable piece of information, so it is a practical way to associate related data that may have few other apparent commonalities.

Most entries tied to specific localities or geologic units

Most “Misc.” is correspondence related to finds in other museums or still in field

# Source Material Challenges + Decision-making

- Secure and conserve source materials
  - Do you limit access to the source itself?
  - Do you capture metadata and/or the source itself?
- Generate and organize useful data on source materials
  - How do you document the relationship between archival objects and collections (e.g., finding aids, databases)?
  - How do you organize the source material to reflect those relationships?
- Facilitate discovery
  - Draw on user information to guide data capture and generation



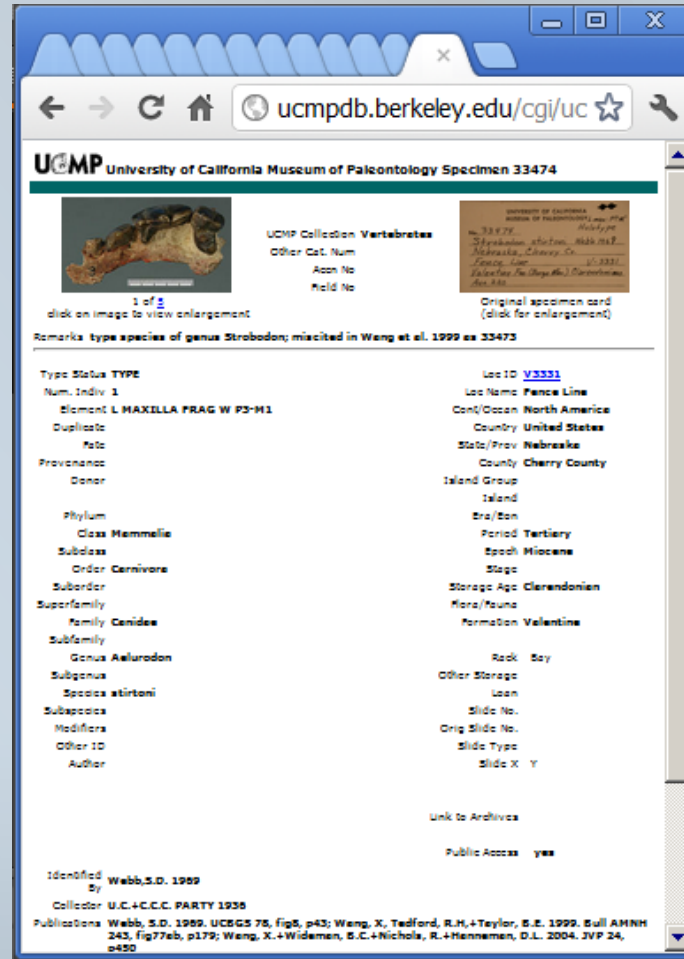


# Ways To Relate These Data: UCMP example

Collection data  
stored in  
two tables:

Specimen &  
Locality

318,056 specimens  
& 90,585 localities  
Online



Images stored in  
CalPhotos, an external  
image database  
Approx.  
6000 labels  
1700 archives images  
3800 specimen images

mySQL tables with  
Perl-scripted web  
interface

# Integrating source metadata + existing database structures

UCMP University of California Museum of Paleontology Specimen 33474

UCMP Collection Vertebrates

Other Cat. Num  
Acn No  
Field No

1 of 2  
click on image to view enlargement

Remarks: type species of genus *Strebodon* mislabeled in Weng et al. 1999 as 33473

Type Status	TYPE	Loc ID	V3331
Num. Indiv	1	Loc Name	Fence Line
Element	L MAXILLA PRAG W P3-M1	Cont/Ocean	North America
Duplicate		Country	United States
State		State/Prov	Nebraska
Provenance		County	Cherry County
Donor		Island Group	
Phylum		Island	
Class	Mammalia	Era/Eon	
Subclass		Period	Tertiary
Order	Carnivora	Epoch	Miocene
Suborder		Stage	
Superfamily		Storage Age	Clerendonian
Family	Canidae	Flora/Fauna	
Subfamily		Formation	Valentine
Genus	Aelurodon	Rock	Say
Subgenus		Other Storage	
Species	stirtoni	Loan	
Subspecies		Slide No.	
Modifiers		Orig Slide No.	
Other ID		Slide Type	
Author		Slide X	Y

Link to Archives **226**

Public Access yes

Identified By Webb, S.D. 1989

Collector U.C.+C.C.C. PARTY 1936

Publications Webb, S.D. 1989. UCGS 78, fig 8, p43; Weng, X., Tedford, R.H., & Taylor, S.E. 1999. Bull AMNH 243, fig 77eb, p179; Weng, X. & Wideman, S.C. & Nicholas, R. & Henneman, D.L. 2004. JVP 24, p450

Problem:  
Often impractical to develop new programming

One solution:  
We just extended data tables and added Archon record number hyperlink

# Source Metadata Record

University of  
California  
Museum of  
Paleontology

[View Cart \(0 items\)](#)

Browse:

[Titles](#)[Subjects](#)[Names](#)[Images](#)[Record Groups](#)

Location: UCMP → J. David Archibald Field Notes

Use to identify materials,  
key disciplinary data,  
standard subject heading

## J. David Archibald Field Notes, 1973-1977 | University of California Museum of Paleontology

**Title:** J. David Archibald Field Notes, 1973-1977

**Extent:** 6.0 Folders

### [More Extent Information](#)

33 polaroid photographs, two geologic maps, one stratigraphic map and 291 pages of field notes

### [Arrangement](#)

Original order

### [Subjects](#) (links to similar collections)

[Cenozoic - Miocene](#)  
[Cenozoic - Paleocene](#)  
[Geology](#)  
[Mesozoic - Cretaceous](#)  
[Paleontology - Vertebrate Paleontology](#)

### [Languages of Materials](#)

[English](#)

### [Administrative Information](#)

**Collection Material Type:** Field Notes (original)

[Printer-friendly](#) | [Email Us](#)

### Scope and Contents:

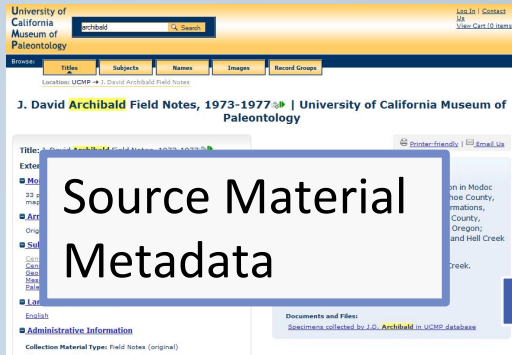
Field notes (1973-1979) from Alturas Formation in Modoc County, California; Massacre Lake Beds, Washoe County, Nevada; Thousand Creek and Virgin Valley Formations, Humboldt County, Nevada; Piute Creek, Lake County, Oregon; Beatty's Butte Beds, Harney County, Oregon; Tullock Formation, McCone County, Montana; and Hell Creek Formation, Garfield County, Montana.

One geologic map of the headwaters of Hell Creek.

### [On-line Images/Records](#)

#### Documents and Files:

[Specimens collected by J.D. Archibald in UCMP database](#)



Records can be associated with multiple links to other resources

Link that generates live SQL query on specimen database

## UCMP UC Museum of Paleontology Specimens

Number of matches: 3176

[next 100](#)

Live query result

- [Download your results](#) (tab-delimited text file with .xls file extension, 3176 lines, file size =427.3 K)
- [Map specimens with a US county](#)

QUERY: SELECT FROM ucmp WHERE spec\_archives like "%46%" ORDER BY cast(specno as unsigned integer)

Click on the Spec # to see the full specimen record

Spec #	Collection	Class	Genus species ssp	Other Name	Type Status	Period	Epoch	Loc ID#	Locality Name	County	State/Prov	Country
<a href="#">76901</a>	Vertebrates	Mammalia	Merychys elegans			Tertiary	Miocene	<a href="#">V6465</a>	Boron	Kern County	California	United States
<a href="#">76902</a>	Vertebrates	Mammalia	Merychys elegans			Tertiary	Miocene	<a href="#">V6465</a>	Boron	Kern County	California	United States
<a href="#">76903</a>	Vertebrates	Mammalia	Merychys elegans			Tertiary	Miocene	<a href="#">V6465</a>	Boron	Kern County	California	United States
<a href="#">76904</a>	Vertebrates	Mammalia	Merychys elegans			Tertiary	Miocene	<a href="#">V6465</a>	Boron	Kern County	California	United States
<a href="#">76905</a>	Vertebrates	Mammalia	Merychys elegans			Tertiary	Miocene	<a href="#">V6465</a>	Boron	Kern County	California	United States
<a href="#">76906</a>	Vertebrates	Mammalia	Merychys elegans			Tertiary	Miocene	<a href="#">V6465</a>	Boron	Kern County	California	United States
<a href="#">76907</a>	Vertebrates	Mammalia	Oxydactylus			Tertiary	Miocene	<a href="#">V6465</a>	Boron	Kern County	California	United States
<a href="#">76908</a>	Vertebrates	Mammalia	Aletomeryx occidentalis			Tertiary	Miocene	<a href="#">V6465</a>	Boron	Kern County	California	United States



University of California Museum of Paleontology

Search:

Location: UCMF # 1 David Archibald Field Notes

**J. David Archibald Field Notes, 1973-1977** | University of California Museum of Paleontology

Title: J. David Archibald Field Notes, 1973-1977

Extents: 33 pages, map and

Administrative Information

Collection Material Type: Field Notes (original)

Documents and Files

Specimens collected by J.D. Archibald in UCMF database

## Source Material Metadata

UCMP UC Museum of Paleontology Specimens

Number of matches: 3176

Download your results (tab-delimited text file with .xls file extension, 3176 lines, file size =427.3 K)

Map specimens with a US county

QUERY: SELECT FROM ucmp WHERE spec\_archives like "%46%" ORDER BY cast(specno as unsigned integer)

Click on the Spec # to see the full specimen record

Spec #	Collection	Class	Genus species ssp	Other Name	Type Status	Period	Epoch	Loc ID#	Locality Name	County	State/Prov	Country
76901	Vertebrates	Mammalia	Merychys elegans			Tertiary	Miocene	V8453	Boron	Kern County	California	United States
76902	Vertebrates	Mammalia	Merychys elegans			Tertiary	Miocene	V8453	Boron	Kern County	California	United States
76903	Vertebrates	Mammalia	Merychys elegans			Tertiary	Miocene	V8453	Boron	Kern County	California	United States
76904	Vertebrates	Mammalia	Merychys elegans			Tertiary	Miocene	V8453	Boron	Kern County	California	United States
76905	Vertebrates	Mammalia	Merychys elegans			Tertiary	Miocene	V8453	Boron	Kern County	California	United States
76906	Vertebrates	Mammalia	Merychys elegans			Tertiary	Miocene	V8453	Boron	Kern County	California	United States
76907	Vertebrates	Mammalia	Oryzodactylus			Tertiary	Miocene	V8453	Boron	Kern County	California	United States
76908	Vertebrates	Mammalia	Alouatta palliata			Tertiary	Miocene	V8453	Boron	Kern County	California	United States

## Live query result

Locality data sent to map display

UCMP University of California Museum of Paleontology Specimen 130335

UCMP Collection **Vertebrates**

Other Cat. Num

Accn No **2961**

Field No **JDA-8-74**

Remarks

Type Status **PUBL**

Num. Indiv **1**

Element **R P4/**

Duplicate

Fate

Provenience

Donor

Loc ID **V84100**

Loc Name **Breakfast Square Butte #2**

Cont/Ocean **North America**

Country **United States**

State/Prov **Oregon**

County **Harney County**

Island Group

Island

Superfamily

Family **Mylagaulidae**

Subfamily

Genus **Hesperogaulus**

Subgenus

Species **gazini**

Subspecies

Modifiers

Other ID

Author

Flora/Fauna

Formation **Beattys Butte**

Rack Bay

Other Storage

Loan

Slide No.

Orig Slide No.

Slide Type

Slide X Y

Link to Archives [46](#)

Public Access **yes**

Identified By

Collector **Archibald, J.D. May 6 1974**

Publications **Calede, J.J.-M.+Hopkins, S.S.B. 2012. ZJLS 164, p436**

Author/Date

## Individual specimen record

Refresh

Map Satellite Hybrid Terrain Day

Feature Location

Record 1 of 5

Link Back [View Detail](#)

PUBL

130335

Concurrence

V

Class

Mammalia

Genus

Hesperogaulus

Map of the United States showing collection locations. A red pin is located in Harney County, Oregon. Other pins are shown in California, Nevada, and Idaho.

500 mi 500 km

UCMP specimen search

Clickable Google map of Archibald collecting

# Workflow for UCMP CLIR project

1. Discovery – We knew we had many documents tied to both collectors and localities and then we found more
2. Prioritization – Field Notes 1st; supplemental locality files 2<sup>nd</sup>; personal papers 3<sup>rd</sup>  
based on frequency of use + funding+ staging of work space and people
3. Concurrent conservation (rehousing) and data extraction/data entry
4. Entry of explicit links (data associations) in related in-house databases

## Future Possible Workflow

1. Discovery – Find more
2. Prioritization – materials near end of lifespan
3. Digital Reproduction of prioritized items
4. Sustainability - Integrating source material data extraction and capture into standard cataloging workflows and increasing granularity

# Acknowledgments & Thanks

**CLIR Hidden Collections Program and NSF-BRC**

**Also: Keck Foundation, Bureau of Land Management, US Geological Survey, and National Park Service, CalTrans**

**UCMP Archive Team:** Dave Smith, Joyce Gross, Mark Goodwin, Bill Clemens, Judy Scotchmoor, Susumu Tomiya, Sarah Werning, Elizabeth Ferrer, Lucy Chang, Kaitlin Maguire, Jeff Benca, Jessica Jedvaj, Dana Riess, Elizabeth Childs, Amanda Burtt, Alexandra Schwarz, Carly Tribull, Cindy Kok, Kathryn Rieck, Kevin Phan, Julie Ding, Min-Sun Jun, and Anoop Mannan, Dave Strauss, Doug Clarke, Claire Englander, Erica Clites, Diane Erwin, Ken Finger

Many colleagues throughout the UC-Berkeley Natural History Museums and library system and the iDigBio community