

Geographical metadata in the California Language Archive

Amy Campbell, Andrew Garrett*, Hannah Haynie, Justin Spence, Ronald Sprouse, and John Sylak (University of California, Berkeley)

General background

The **California Language Archive** integrates the catalogs and the digital content of two archival repositories at Berkeley:

- Berkeley Language Center (blc.berkeley.edu)
- Survey of California and Other Indian Languages (linguistics.berkeley.edu/~survey)

The CLA is in development and will be online in February 2011. In the future, other language archives may also contribute to the CLA (such as the Bancroft Library and the Hearst Museum at Berkeley).

Berkeley Language Center

An archive of **AUDIO RECORDINGS** deposited since about 1950, from languages throughout the Americas with a special focus on California

- About 1,730 hours of audio recordings (cataloged to date)
- Original media: reel-to-reel tapes, cassette tapes, DATs, CDs, video
- 7,842 audio segments in 221 collections, e.g. ‘Catherine A. Callaghan collection of Northern Sierra Miwok field recordings’
- Digitization status: all cataloged material is digitized

Information about geography

- Metadata deposit forms showing consultant linguistic histories

Survey of California and Other Indian Languages

An archive of **PAPER MATERIALS** deposited since about 1950, from languages throughout the Americas with a special focus on California

- About 200 linear feet of material
- Contents include field notebooks, unpublished manuscripts, letters and notes, lexical fileslips, etc.
- 2,451 distinct catalog items in 186 collections
- Digitization status: currently >50,000 scanned images from 200 items in 29 collections

Information about geography

- Available when researchers report information about consultants, or locations of fieldwork sessions
- Occasionally consultant biographical information is accessible through other sources (publications or local communities)

Credits

*Corresponding author (garrett@berkeley.edu). BLC and SCOIL meta-data collaborators and interlocutors include Lisa Conathan, Marianne Garner, Susan Stone, and Maziar Toosarvandani. Photos: Halpern MS image = SCOIL MS Halpern.009.014.0002; Annie Burke photo from *Pomo basketmaking* (1972); Sulphur Creek photo from Bill Tuthill, <http://www.cacreeks.com/sulf-big.htm>

Sources of metadata

Contributing repositories primarily hold legacy material, from which we glean geographical information. (New depositors will be asked to supply geographical metadata.)

Berkeley Language Center

Documentation for LA007: Northern Sierra Miwok
Recorded by Catherine A. Callaghan & Mrs. Aza McCauley, 1960–1961

UNIVERSITY OF CALIFORNIA, BERKELEY LANGUAGE LABORATORY LANGUAGE ARCHIVES DATA SHEET	
Identification No. 7	
Name of language: Northern Sierra Miwok	
Alternative names: Hoquelume	
Dialect: Comanche	
Alternative names:	
Where spoken: Comanche, California	
Family: Miwok	
Name of informant: Mrs. Aza McCauley	
Present residence: Ione, California	
Nationality: USA	Sex: F Age: 85
Occupation: Retired cook	
Education: None	
Period during or conditions under which the informant actually used this language: Native language; used it all her life.	
Other places of residence and approximate duration of stay: Lived in Comanche until age fifteen, then in Pleasanton, California until sometime in adulthood, then moved to Ione, California.	
Other languages spoken and inclusive dates: Learned Plains Miwok at age fifteen as a second language; at some time learned English as an additional language.	
*If more than one informant, supply the information in this block for each on a separate sheet.	

← speaks **Comanche** dialect

← grew up in **Comanche**; has lived in **Ione** and **Pleasanton**

Survey of California and Other Indian Languages

Abraham M. Halpern, field notebook XIV, page 1
Southern Pomo notes with Annie Burke, 7 May 1940

1	
Annie Burke — May 7, 1940	
Cloverdale —	
The name na-ha'no-jamai	
Sulphur Creek	
Others who spoke same language —	
ha-ha'ko-jamai	
Lived at present Cloverdale ca.	
Bana-ko-jamai	
the seat of asti	
at Echo — an extinct bunch	
s you ba-ha'ga — spoke	
more like Hopland	
Betr. Cloverdale and the	
coast were ba-ha'no-jama	
-jamai, who spoke a language	
like him, but a little different	

← currently (1940) lives in **Cloverdale**

← ethnic affiliation is **Sulphur Creek**

- priority assigned to **Sulphur Creek** (for dialectology)

Geographical metadata

Metadata about people

The CLA database includes biographical information about researchers and linguistic consultants:

- Born at [PLACE, DATE]
- Lived at [PLACE, DATE]
- Ethnic affiliation with [PLACE]
- ‘From’ / raised at [PLACE]
- Father ‘from’ [PLACE]
- Mother ‘from’ [PLACE]

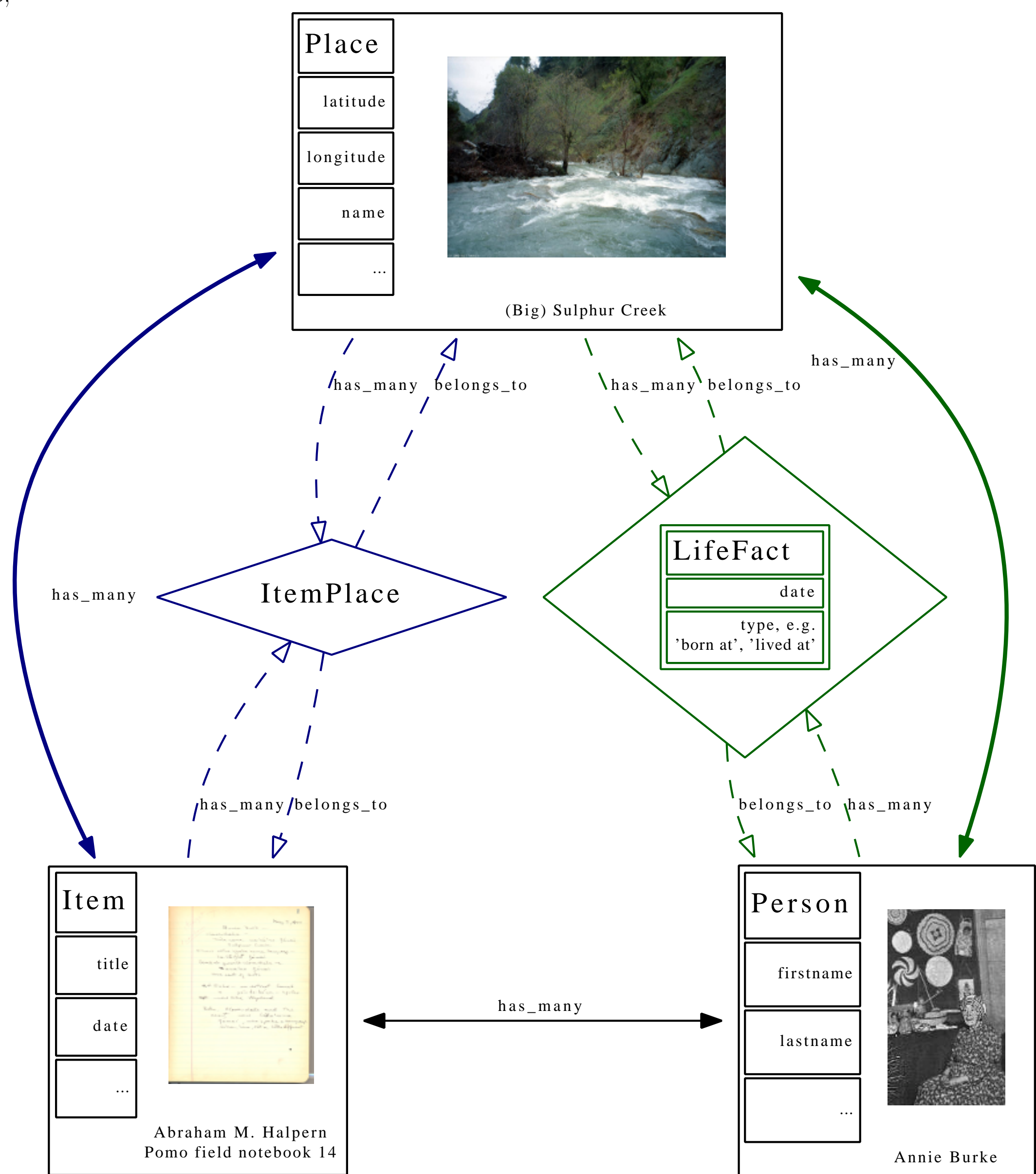
One PLACE per individual can be tagged for dialectological priority.

Metadata about items (field notebooks, etc.)

Legacy material often gives no biographical or geographical information about consultants; sometimes they are not even named. We record:

- Fieldwork PLACE

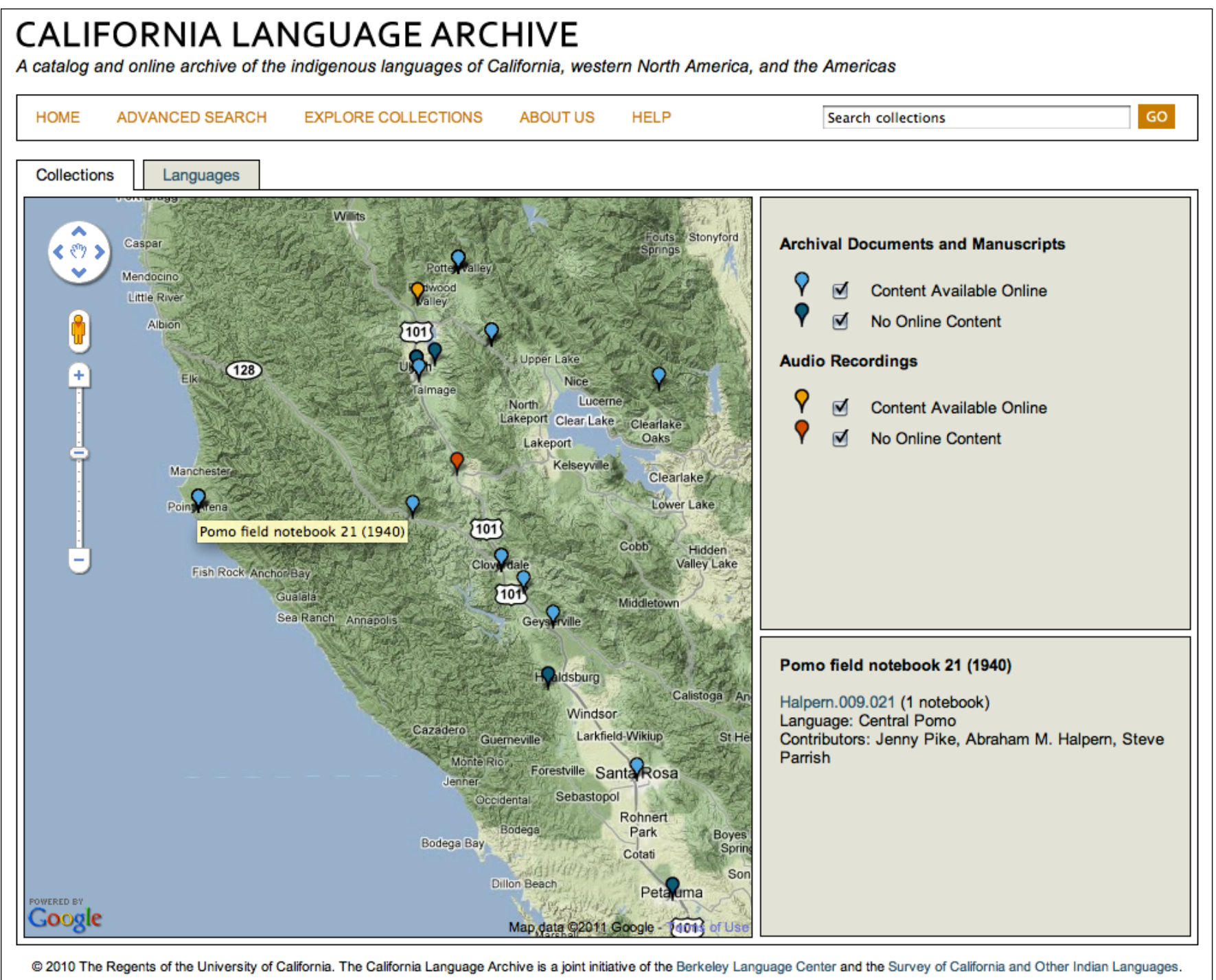
Conceptual database schema



User interface

Map-based searching

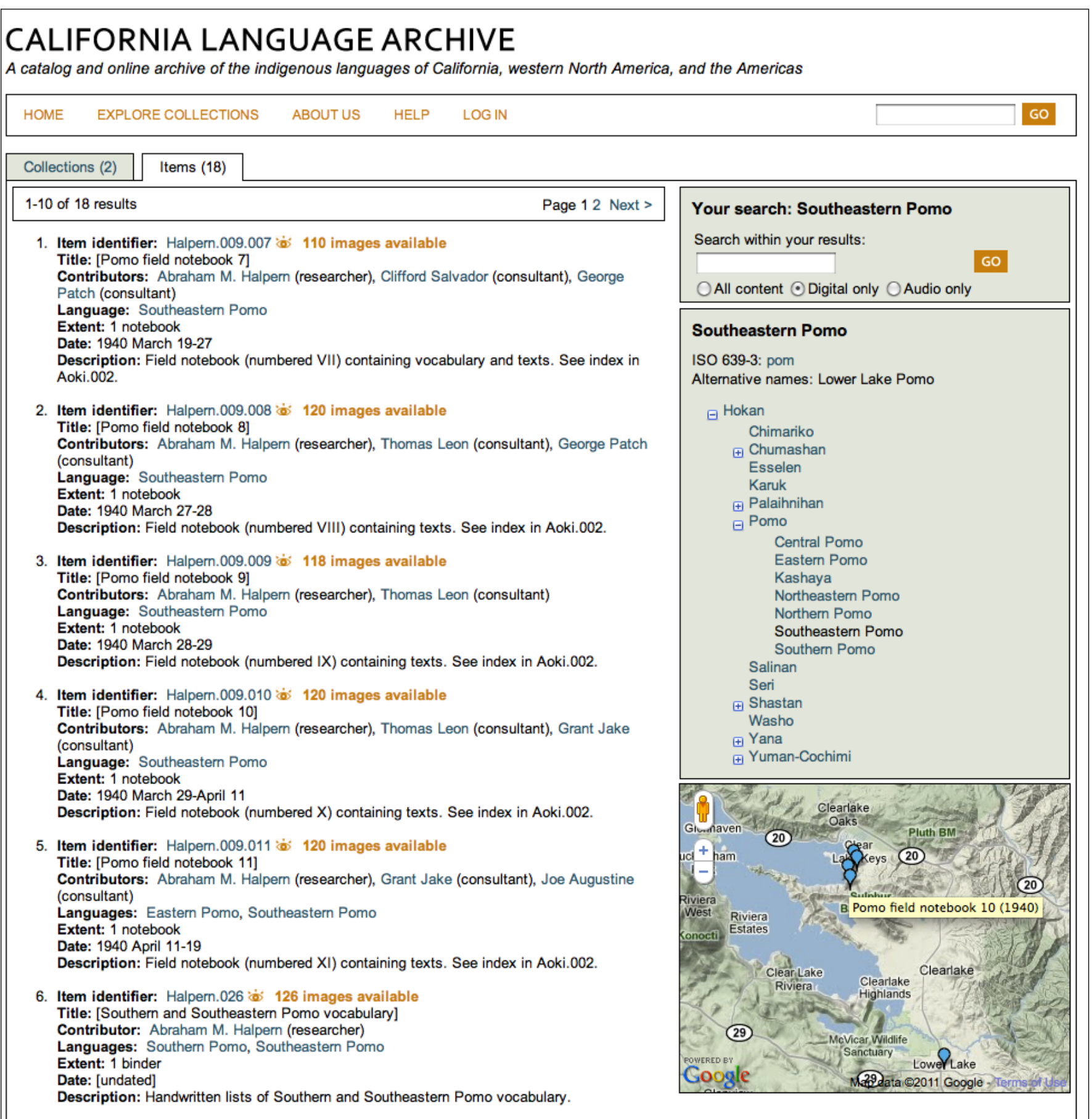
Exploring a map interface shows items with geographical metadata in the CLA database:



The map pins distinguish audio vs. paper material with vs. without digital content.

Maps in search results

A search for Southeastern Pomo material with digital content yields the following partial result:



The map shows places for items with geographical metadata in the CLA database.