

## California Revealed

### Statement of Work

January 2018

#### Project description

California Revealed is a State Library initiative to help California's public libraries and other local heritage groups to digitize, preserve, and serve online historically significant Californiana. Participating institutions select items from their collections, create discovery metadata, and send materials and metadata to the California Preservation Program (California Revealed), who oversee the digitization process (outsourced to vendors) as well as provide online access and long-term preservation.

Our current goal is to digitize approximately 8,000 objects, primary source materials (e.g., books, documents, photographs, and audiovisual recordings) - drawn from 81 public libraries, historical societies, and archives. California Revealed staff will work with the institutions, vendors, and digital repositories.

To see a list of California Revealed's current partner institutions and browse the *California Light and Sound* collection, please visit: [californialightandsound.org](http://californialightandsound.org).

Collections are still arriving, and we do not yet know the exact item count, dimensions, page numbers, durations or condition of the materials. We expect to receive:

Photographs: ~1,435 negatives and slides; 939 prints; 9 aperture cards

Loose materials: ~1,267 flyers, letters and manuscripts; 1,016 posters; 104 postcards; 17 maps (largest is 24 x 38 in.)

Folder-level clippings and documents: ~252 folders

Bound documents: ~400 newsletters and pamphlets; 360 books; 352 periodicals; 218 yearbooks; 96 brochures and catalogs; 33 scrapbooks; 9 photograph albums

Newspapers: ~5,306 issues; 22 bound volumes; ~19 rolls of microfilm

Video: 212 VHS and VHS-C tapes; 18 ¾" U-Matics; 18 MiniDVs; 7 2" open reels; 6 Betacams; 5 ½" open reels; 3 Hi8s; 2 DVDs; one DVCAM; one 1" open reel.

Audio: ~229 ¼" open reels, 644 audio cassettes, 20 microcassettes; 18 transcription discs; 2 LP records.

Film: ~35 16mm; 21 Regular 8mm; 9 35mm; 1 Super 8mm

We anticipate sending materials for digitization in January 2018. Digitization work must be accomplished within 6-8 weeks of receipt of original materials.

Please direct all correspondence related to this project to: Pamela Jean Vadakan, Preservation Department, 20 Doe Library, University of California, Berkeley, CA 94720, 510.642.4665, [pamelaje@berkeley.edu](mailto:pamelaje@berkeley.edu)

## PROJECT WORK PLAN

### Technical specifications for digitization

Please comply with the following file format and process specifications for each digital file instantiation - preservation master and access copies. Specs for documents assume a 100-megapixel array camera. See special requirements below for handling instructions.

### Bound documents

#### e.g. books, directories, photograph albums and scrapbooks

- Preservation master: TIFF, 400 ppi (up to 20”h x 28”w) / 300 ppi (up to 28”h x 37”w), 24-bit RGB color, uncompressed
- Access copy: PDF with uncorrected raw OCR
- Include foldouts and inserts in position to maintain context. Unfold albums with layers to capture images or text beneath. Maintain resolution across pages and inserts - do not magnify or reduce.  
\*Capture two-up when possible, split images to single page TIFFs at the gutter, minimizing the margin without eliminating original content, and crop excess copy board on the non-gutter edges along with 10 pixels around

### Multiple image, unbound, documents

#### e.g. postcards - front and back, letters of more than one page, pamphlets, and double-sided brochures

- Preservation master: TIFF, 400 ppi (up to 20”h x 28”w) / 300 ppi (up to 28”h x 37”w), 24-bit RGB color, uncompressed
- Access copy: JPG and PDF with uncorrected raw OCR
- Include foldouts and inserts in position to maintain context. Maintain resolution across pages and inserts - do not magnify or reduce.

### Newspapers

- Preservation master: TIFF\*, 400 ppi (up to 20”h x 28”w if a two-page spread) / 300 ppi (up to 28”h x 37”w if a two-page spread), 24-bit RGB color, uncompressed
- Access copy: PDF with uncorrected raw OCR  
\*Capture two-up when possible, split images to single page TIFFs at the gutter, minimizing the margin without eliminating original content, and crop excess copy board on the non-gutter edges along with 10 pixels around

### Single image, unbound, documents

#### e.g. photographs, maps, prints, and posters - front only

- Preservation master: TIFF, 400 ppi (up to 20”h x 28”w) / 300 ppi (up to 28”h x 37”w), 24-bit RGB color, uncompressed
- Access copy: JPG

### Photographic negatives and transparencies

- Preservation master: TIFF, 3,000 ppi (up to 4”h x 5”w) / 2,000 ppi (larger than 4”h x 5”w), 24 bit RGB color, uncompressed
- Access copy: JPG
- Use an overhead array camera and cold light transilluminator light box with glass

## Audio recordings

- Preservation master: Broadcast WAV - 24 bit, 96 kHz, 2,304 Kbps for mono, 4,608 Kbps for stereo. L and R channels interleaved.
- Access copy: MP3, 160 Kbps for mono – 320 Kbps for stereo, 44.1 kHz. L and R channels interleaved.
- Broadcast WAV files should not exceed 3 GB per file. Please split files so they fit within these parameters and name the split files as parts.
- For born-digital sources, maintain original specifications and embedded metadata.
- Approximate size: 47 MB, mono and 94 MB, stereo for a 40 min. recording

## Film and video recordings

### Preservation master

- File format: MOV, 10-bit uncompressed (4:2:2), QuickTime wrapper\*
  - Stream bitrate: -226 Mbps (sound) / -224 Mbps (silent) – preferably fixed
  - Codec: v210
  - Frame width: 720 pixels; Frame height: 486 pixels
  - Frame rate: 29.97 fps (NTSC)
  - Display aspect ratio: Maintain original (often 4:3)
  - Standard: NTSC
  - Color space: YUV
  - Audio: Linear PCM, 2,304 Kbps stream bitrate, 2 Channel mono or stereo track, 48kHz sampling rate, 24-bit depth\*\*
  - For born-digital sources, maintain original specifications and embedded metadata.
  - Approximate size: 102 GB for a 60 min. recording
- \*If partners want a higher resolution scan - high definition or 2k - they will be asked to cover additional charges.  
\*\*If video source has more than two channels of audio, capture all channels.

### Access copy

- File format: MOV, H.264 MPEG-4 Part 10
  - Stream bitrate: 3.5 Mbps minimum – 4 Mbps maximum – preferably fixed at 3.5 Mbps\*
  - Codec: avc1
  - Frame width: 720 pixels; Frame height: 540 pixels (minimum)\*\*
  - Frame rate: 29.97 fps
  - Display aspect ratio: Maintain original (often 4:3)
  - Standard: NTSC
  - Color space: YUV
  - Scan type: progressive
  - Bit depth: 8 bits
  - Audio: AAC, 256 Kbps stream bitrate, 2 Channels, 48 kHz sampling rate
  - Approximate size: 1126 MB for a 60 min. recording
- \*Raise minimum data rates for audio and video to ensure video level remains at 3.5 Mbps or higher and audio data rate remains at 160 Kbps or higher.  
\*\*Frame size is according to the Internet Archive's preferred specs in order to display video as "HD". Dimensions can be increased as long as original aspect ratio is maintained.

## Technical requirements

California Revealed asks partners to prepare their materials so they are scanner ready (see [shipping guidelines](#)). California Revealed will include an inventory for each shipment. Upon arrival at the facility, the vendor will carefully unpack and confirm contents and report back any discrepancies within 30 days of receiving, prior to scheduling digitization. The vendor will review a small sample of the materials from each shipment to evaluate the overall condition.

Materials will be handled with clean hands, gloves when necessary, and in a clean work studio with clean scanning equipment. The vendor will inspect, prepare, and digitize the object according to California Revealed specifications and will include a condition report/technical evaluation about the original source that includes any damage or errors

discovered upon check-in or transfer. Examples include, but are not limited to, the following: excessive dust/dirt, stains, tears or faint image for still image and text documents, and sticky shed, dropout or low levels for audiovisual recordings. If any material received which in the vendor's opinion cannot be digitized without damage to the material, California Revealed will be notified and the material set aside pending the partner institution's decision.

Special handling needs will be noted by the partner and California Revealed in the "Additional Technical Metadata" field. **The vendor will contact California Revealed if the general condition of a batch may impact the quality of the capture results, and if any special handling - such as repair, treatment or disbinding - is necessary, or if handling costs exceed the initial estimate per item, so that California Revealed can ask the partner institution for permission to proceed and possibly cover the additional handling cost if needed. In particular, please be attentive to fragile materials.** The vendor will record any treatments in their digitization notes.

Once the materials are prepared for digitization, the vendor will make every effort to handle them carefully to avoid or minimize damage. The vendor will digitize the collections using non-destructive measures to maintain original condition. The vendor will ensure that the content is captured faithfully and accurately with no information loss or degradation. The vendor will not introduce any image or audiovisual processing or enhancement at any point in the creation of the preservation masters. The vendor will take responsibility for the safe care and handling of materials at all times. When not being scanned, materials will be stored in a locked room.

### **Special requirements for still image and text documents**

The vendor will limit any factors that may pose a risk to the objects being captured. The vendor will not use any sheet-fed scanners or any devices with automatic page turning. The vendor will clean photographs as needed using canned air or a soft Chinese Hake brush. If glass is used to flatten documents, the glass will be cleaned with deionized water between captures to minimize dust and fibers. Materials should be illuminated using low ultraviolet-emitting light.

Capture both color and black-and-white originals as color. No color/tone enhancement, unless such treatment or enhancement is requested by the partner.

Crop images to include the page edge along with 10 pixels around. Unless otherwise specified by the partner, the image area should exclude mounts and matting. By default, include foldouts and loose inserts in position to maintain context. Unfold albums with layers to capture images or text beneath. Maintain resolution across pages and inserts - do not magnify or reduce. Capture the front and back of postcards. The partner and California Revealed will add specific, clear, digitization instructions regarding mounts, inserts and rectos/versos to the metadata record, along with a physical note on the item. After capture, the vendor will return materials to their original arrangement, as they were packed and shipped to the vendor.

California Revealed scans cover to cover but prefers not to scan multiple blank pages in a row. If there are blank or missing pages, partners will insert a 8.5"x 11" "target" page to be scanned to alert the user that blank pages have been skipped or there are missing pages.

### **Special requirements for audiovisual recordings**

The vendor will photograph original audiovisual recordings and containers if possible.

Film and video sources will be transferred as best light with luminance values within "legal" limits of the waveform monitor scope (below 110 IRE or above 7.5 IRE). Luminance, chrominance, and tone should be adjusted to the bars/tone if present on source tapes; if color bars are not present or are clearly inaccurate, levels should be adjusted to the content of the tape using known references (such as blue sky, known blacks and whites, flesh tone, etc.) without any clipping or crushing. For audio sources, minimize headroom by maintaining peaks around -6 db. Boost by 5 db as needed, as long as it does not cause distortion and potential clipping. If a recording is particularly quiet and has no machine or room noise, boost by 10 db. No image/sound processing should be introduced to the signal chain at any point in the creation of preservation masters. This includes, but is not limited to, dropout compensation, noise reduction, audio equalization, limiting and filters.

Fully monitored 1:1 transfers are ideal; but if possible, batch transfers are preferred to help keep project costs low. Please begin all files with 00:00:00;00 timecode. If there's timecode, capture as much of the tape with timecode as possible.

Retain the sound configuration of the original source recording – left and right mono as left and right mono and left and right stereo as left and right stereo. Retain mono audio sources as mono/single channel – do **not** split the sound into left and right channels. For film and video sources with mono sound, **do** split the sound into left and right channels as split mono. If two channel sound is skewed low on one channel or one channel is unlistenable, **do** split the best level channel into left and right channels as split, balanced, mono.

Do not capture the audio track (or delete it) if the original film or video source is silent.

Recordings should be captured from the physical start until the end of the program, along with consistent heads and tails for all files, with ten seconds of silence at heads and tails for audio recordings and ten seconds of black leader at heads and tails for video and film recordings.

For born-digital sources, such as MiniDV tapes, maintain original specifications and embedded metadata for the preservation master. For DVD-formatted disc media, create a disk image capture for the preservation master. Then please derive an access file from the preservation master according to specifications above. For disc media used as file storage, follow instructions for born digital sources.

### **Review before digitization**

The partner is responsible for identifying duplicate content, and California Revealed will attempt to highlight potential duplicates and choose the best source to digitize, if possible. The vendor will contact California Revealed immediately if duplicate content exists or if there are any discrepancies between the descriptive metadata supplied and the content.

For some audiovisual cases, the content of a recording is unknown: there is little or no description available and playback is impossible. We request the option to preview the content before digital capture in the event that we suspect duplicate or unrelated material. We will note such cases in the metadata and shipping letter and physically mark the recording as well.

California Revealed's focus is to preserve archival, non-commercial Californiana. If upon preview the vendor discovers that content does not match these criteria, please do **not** capture the recording and contact California Revealed immediately for further instruction. California Revealed will consult with the partner if there is a possibility the other content may be significant and worth transferring. Otherwise, California Revealed prefers to transfer only California content.

To conserve data-storage space, please do **not** capture or deliver a file if it is not substantive or is without potential research value, such as ambient room noise, silence/blank side of a tape, or blue frames. Contact California Revealed with questions as such cases are discovered upon initial ingest.

### **Review after digitization**

The vendor is expected to review 100% of the digital objects, including metadata, for quality, completeness, and accuracy of the new digital files and metadata. Still image and text documents will be viewed at a full object view sized at 100% using imaging software. The vendor will check the digital file against the original source to ensure a faithful reproduction of the original and confirm that no artifacts were introduced in the digitization process. The vendor will note inherent abnormalities in the "Vendor Quality Control/Transfer Notes" field - these will be particularly helpful for California Revealed staff when they inspect the files.

California Revealed, with partner participation, will review deliverables within 30 days of receipt of the files to ensure that all project specifications are met. Materials will be returned directly to the partner after all parties have confirmed that the files are acceptable and there is no necessary rework.

## Subcontracting

The vendor will perform all digitization on its premises. No subcontracting of this work will be permitted without prior communication and approval from California Revealed.

## XML metadata

The vendor will deliver one metadata record per object. The [Dublin Core](#) schema will be used for still image and text documents. According to the California Digital Newspaper Project, Dublin Core wrapped in METS schema ([NDNP Lite](#)) will be used for newspapers. The [PBCore](#) schema will be used for media objects. These metadata records will combine descriptive, rights, and administrative metadata supplied by the partner and California Revealed, along with technical metadata supplied by the vendor that tracks equipment information used for digitization as well as transfer and quality control notes. Please see California Revealed's [Metadata Fields](#) for a complete list of required fields per object.

## Sample model Dublin Core records

Image object (one page)

[https://archive.org/download/c\\_000167/c\\_000167\\_metadata.xml](https://archive.org/download/c_000167/c_000167_metadata.xml)

Text document (two pages)

[https://archive.org/download/cscrm\\_000380/cscrm\\_000380\\_metadata.xml](https://archive.org/download/cscrm_000380/cscrm_000380_metadata.xml)

Bound document

[https://archive.org/download/cgl\\_000007/cgl\\_000007\\_metadata.xml](https://archive.org/download/cgl_000007/cgl_000007_metadata.xml)

## Sample model PBCore records

Simple Moving Image Recording (one reel or tape)

[https://archive.org/download/car\\_000211/car\\_000211\\_PBCore.xml](https://archive.org/download/car_000211/car_000211_PBCore.xml)

Simple Sound Recording (one tape)

[https://archive.org/download/car\\_000132/car\\_000132\\_PBCore.xml](https://archive.org/download/car_000132/car_000132_PBCore.xml)

Complex Moving Image Recording (multiple reels or tapes)

[https://archive.org/download/car\\_000196/car\\_000196\\_PBCore.xml](https://archive.org/download/car_000196/car_000196_PBCore.xml)

Complex Sound Recording (multiple tapes)

[https://archive.org/download/car\\_000090/car\\_000090\\_PBCore.xml](https://archive.org/download/car_000090/car_000090_PBCore.xml)

## Embedded metadata

For still image and text objects, the vendor will ensure embedded technical metadata are readable for all file instantiations, including file size and type, width, height, color channels, and resolution.

The vendor will embed descriptive metadata supplied by California Revealed\* for each preservation and access file, including:

- Embedded Title (dc:Title) with part number (e.g., Scotia Bridge – part 1 of 1). Vendor will provide part number after digitization when total number of parts are confirmed.
- Embedded Year\*\* (dc:Date)
- Embedded Holding Institution (dc: Source)
- Embedded Comment/Project Note for California Revealed (dc:Description)

- Embedded Copyright Status: Public Domain or Copyrighted (first two sentences of the copyright statement) (dc:Rights)  
\*If unknown, California Revealed will leave blank  
\*\*Formatted as YYYY; YYYY-MM; YYYY-MM-DD

For audiovisual recordings, the vendor will ensure embedded technical metadata are readable by Media Info for all file instantiations, including file size and type, duration, frame size, codec, number of audio channels and total bitrate for video sources and file size and type, duration, number of audio channels and total bitrate for audio sources.

The vendor will embed descriptive metadata supplied by California Revealed for each preservation and access file, including:

- Embedded Title (pbc:Title) with part number of file (e.g. 1955 Flood – part 1 of 2). Vendor will provide part number after digitization when total number of parts are confirmed.
- Embedded Year\*\* (pbc:Date)
- Embedded Holding Institution (pbc: collectionSource)
- Embedded Comment/Project Note for California Revealed (pbc:Extension)
- Embedded Copyright Status: Public Domain or Copyrighted (first two sentences of the copyright statement) (dc:Rights)  
\*If unknown, California Revealed will leave blank  
\*\*Formatted as YYYY; YYYY-MM; YYYY-MM-DD

### **File naming and directory specifications**

File names are based on the Object Identifier number (e.g., clgam\_000001), which includes the owning institution's Marc organization code followed by a unique, sequential number. The source identifier serves as the prefix for all digital file instantiations that represent the media object. Please see metadata spreadsheet supplied by California Revealed per shipment for Object Identifiers.

Label preservation masters ObjectIdentifier\_prsv plus extension (e.g., cusb\_000001\_prsv.wav)

Label access files ObjectIdentifier\_access plus extension (e.g., cusb\_000001\_access.mp3)

Label .md5 checksums in the following manner:

- File name: cusb\_000001\_prsv.mov
- Checksum: cusb\_000001\_prsv.mov.md5

The directory structure for the digital files and the supporting metadata and documents must be "flat" (without subdirectories) for each recording.

Create a folder for each object and label it by the Object Identifier number (i.e., CA-R2017/marc organization code/object identifier). The following items should be within each folder per object:

- preservation file(s)
- preservation file .md5(s)
- access file(s)
- access file .md5(s)
- xml metadata record
- technical evaluation form or notes about the transfer (if not already supplied in the xml record)
- picture file .jpg(s) of the original recording (for a/v only)

### **File naming for still image and text documents**

If a still image or text object consists of multiple pages or parts (e.g., if a photograph is made up of several negatives), add "\_p#" (for "page" or "part") to the file name (e.g., cgl\_000001\_p001\_prsv.tif, cgl\_000001\_p002\_01\_prsv.tif)

If a scrapbook object consists of multiple pages and inserts add "\_p#" (for "page") and "a" or "b" (for each side) to the file name (e.g., cgl\_000002\_p001\_a\_psv.tif, cgl\_000002\_p001\_b\_prsv.tif)

## File naming for audiovisual recordings

Label video access files ObjectIdentifier\_access.HD plus extension (e.g., cusb\_000002\_access.HD.mov)

The '.HD' should **not** be included in the names of audio access files (e.g., ObjectIdentifier\_access.mp3)

If an audio or video object consists of multiple tapes, add “\_t#” to the file name (e.g., cusb\_000001\_t01\_prsv.mov; cusb\_000001\_t02\_prsv.mov)

If a film object consists of multiple reels, add “\_r#” to the file name (e.g., cusb\_000001\_r01\_prsv.mov; cusb\_000001\_r02\_prsv.mov)

If a recording consists of multiple sides, add “a” or “b” to the file name (e.g., cusb\_000001\_a\_prsv.wav; cusb\_000001\_b\_prsv.wav)

If a recording consists of multiple tapes or reels and multiple sides, file names should be labeled (e.g., cusb\_000001\_t01\_a\_prsv.wav; cusb\_000001\_t01\_b\_prsv.wav; cusb\_000001\_t02\_a\_prsv.wav; cusb\_000001\_t02\_b\_prsv.wav).

If there is a need to break up an audiovisual object into multiple parts (e.g., if a recording duration goes beyond the file capacity, or a recording is multi-tracked or speed changes in the middle of a recording), add “\_p#” (for “part”) to the file name (e.g., cusb\_000001\_p01\_prsv.wav, cusb\_000001\_p02\_prsv.wav)

If a recording includes a blank side, or duplicate content is discovered on a tape/reel and later deleted, please drop the reference to the deleted side(s) or tape(s) in the file names and re-label remaining files accordingly.

Label .jpg of the original recording in the following manner: cusb\_00001.jpg

If more than one picture is taken of the recording, label .jpg(s) in the following manner: cusb\_00001-1.jpg; cusb\_00001-2.jpg; cusb\_00001-3.jpg; etc.

## Deliverables

Please see attached vendor workflow/timeline. California Revealed's files will be retained on the vendor's storage system until they are checked and accepted by California Revealed and the partner institution within six months of receiving the files. The vendor will communicate confirmation of the deletion of the files via email within six months or whenever California Revealed has given permission to delete the files, after quality control has been completed and all parties approve.

Initially the vendor will batch-deliver digital files to California Revealed by shipment on borrowed exFAT-formatted hard disk drives (HDDs), preferably with Firewire or USB 3.0 ports. The vendor will produce an .md5 checksum per file state (preservation and access) at the time of ingest to verify that files were not modified in transfer.

California Revealed will perform quality control on the files and may ask the vendor to review original objects for which it is unclear whether or not artifacts are inherent to the source. California Revealed will report quality control feedback from the partner and will ask the vendor for replacement files for any deliverables that do not meet specifications. Replacements should be received within 30 days of request.

After quality control and upon approval of the files by California Revealed and the partner institution, the vendor will deliver all files (including preservation files, access files, checksums, photos, xml metadata, and tech sheets, if included) on a mirror set of two LTFs-formatted LTO7 tapes. Include only the final approved files for each object. Please ensure that all the associated files that represent an object are on the same LTO tape. The vendor will check the checksums on the LTO after writing the tapes to confirm that files were not altered during transfer. California Revealed will return the borrowed HDDs upon receipt of the final LTO deliverable.

The vendor will create a manifest of the LTO tapes in excel form, including the following information, in order:

- LTO\_MediaLabel (including unique identifier, e.g., CA0010L6 and/or barcode)
- Object Identifier (e.g., cusb\_000001)
- Folder Name (e.g., cusb\_000001)

- General File Name (e.g., cusb\_000001\_prsv.mov)
- Path and File Name (e.g., CA-R2018\cusb\cusb\_000001\cusb\_000001\_prsv.mov)

Partner institutions will contact the vendor directly for copies of their approved files and associated metadata – either on a mirror HDD borrowed or bought from the vendor or on one supplied by holding institution – within 60 days upon the vendor’s file delivery to California Revealed. The holding institution is expected to pay only for storage media and shipping for copies of their files. California Revealed will inform the vendor which institutions expect copies at the start of digitization work. When confirming shipment of the backup copies to the partner, the vendor will cc: [dps@calpreservation.org](mailto:dps@calpreservation.org).

## Shipping

All originals will be shipped from California Revealed via the vendor’s FedEx account (please invoice California Revealed for shipping).

Please send deliverables before April 1, 2018:

Pamela Vadakan  
California Preservation Program  
20 Doe Library  
Berkeley, CA 94720  
510.642.3885  
[dps@calpreservation.org](mailto:dps@calpreservation.org)

Please contact Pamela Vadakan for shipping instructions after April 1, 2018.

After final approval of the files and metadata by California Revealed and the holding institution, the vendor will return the materials as they were originally shipped directly to the holding institution via the vendor’s FedEx account (please invoice California Revealed for shipping). A list of shipping addresses and contact information for all participating partner institutions will be provided by California Revealed so that the vendor can send shipping information to both the institution and California Revealed upon return.

Please send all shipping notices to: [dps@calpreservation.org](mailto:dps@calpreservation.org).

## Billing

Califa, California Revealed’s fiscal agent, handles all payments for California Revealed project.

Mail or email the invoice(s) to Pamela Vadakan; approved invoices will be forwarded to Califa for payment.

Ensure invoices are labeled with the following Bill to/Send to information:

### Bill to:

Andrew Yon  
Pacific Library Partnership (PLP)  
2471 Flores Street  
San Mateo, CA 94403  
650.349.5538  
[yon@plsinfo.org](mailto:yon@plsinfo.org)

### Send to:

Pamela Vadakan  
California Preservation Program  
20 Doe Library  
Berkeley, CA 94720  
510.642.4665  
[pamelajeane@berkeley.edu](mailto:pamelajeane@berkeley.edu)

**The California Preservation Program looks forward to working with you!**

Pamela Vadakan  
[California Light and Sound](http://CaliforniaLightandSound.org)  
510.642.4665, [pamelajeane@berkeley.edu](mailto:pamelajeane@berkeley.edu)

## California Preservation Program

### Vendor Workflow/Timeline

| Process  | Time Frame   |
|--|--|
| California Revealed calls for nominations  | 30 days  |
| California Revealed compiles nominations for California Preservation Program Steering Committee  | 2 weeks  |
| California Revealed reviews nominations  | 2 weeks  |
| California Revealed sends award letters to partner institutions  | 2 weeks  |
| Partner ships originals to California Revealed   | 30 days  |
| California Revealed checks inventory against award letter  | Upon receipt of originals  |
| California Revealed processes and ships batches of originals (minimum 3 partners per batch) to vendor in the order in which they were received                       | 30-60 days   |
| <b>Vendor digitizes originals</b>  | <b>4-6 weeks for documents; 6-8 weeks for audiovisual recordings</b>                     |
| California Revealed QC's new files and metadata and sends feedback to vendor per batch/institution   | 30 days upon receipt of files (first round)  |
| California Revealed uploads files and metadata to IA per institution and sends email notification to partner to check files and order backup from vendor             | 30 days upon receipt of files (first round)  |
| <b>Vendor responds and provides redos as needed</b>  | <b>30 days upon CA-R feedback</b>  |
| California Revealed QC's redos   | 15-30 days upon receipt of redos (second round)  |
| <b>Vendor sends California Revealed invoice for digitization</b>   | <b>30 days upon receipt of first round</b>   |
| California Revealed pays invoice   | 30 days upon receipt of first round  |
| Partner QC's, updates descriptive metadata, checks image and sound quality of access files, and sends feedback to California Revealed*<br>*not a project requirement | 30 days upon delivery of partner email (30-60 days upon receipt of first round of files) |
| California Revealed responds and relays technical questions, if any, to vendor   |  |
| California Revealed signs off on files – final approval for LTO creation   |  |
| <b>Vendor returns originals directly back to partner and provides file backups</b>   | <b>1-2 weeks</b>   |
| <b>Vendor provides final LTO deliverable to California Revealed</b>  | <b>30 days upon final file approval</b>  |
| California Revealed QC's LTO   | 1-2 weeks  |
| <b>Vendor sends California Revealed invoice for LTO</b>  | <b>30 days</b>   |
| California Revealed pays invoice   |  |
| California Revealed publishes Islandora records and syncs records in WorldCat, Calisphere and DPLA.  | 2-4 weeks  |

## Metadata Map For the California Preservation Program (CPP)

**Red = Required** | **Green = Preferred** | **Blue = Ideal** | supplied by Partner Archive

Gray = supplied by CAVPP | **Purple = supplied by Digitization Vendor**

| Field   | PBCore element   | Dublin Core element          | Notes  |
|---|--|------------------------------|--|
| <b>[Descriptive Metadata]</b>                   |  |                              |  |
| Institution Name                                | collectionSource; Identifiersource   | provenance                   |  |
| Title   | Title [titleType=Main/Supplied]  | title                        |  |
| Creator (repeatable)                            | creator<br>[creatorRole=Writer/Producer/Director/<br>Interviewer/Performer]        | creator                      | source="Library of Congress Name Authority File" |
| Date (repeatable)                               | AssetDate<br>[dateType=Created/Published]  | date-created;<br>date-issued | YYYY-MM-DD                                       |
| Significance to California history              |  |                              |  |
| Contributor (repeatable)                        | contributor [creatorRole=Camera/<br>Editor/Sound/Interviewee/Music/<br>Cast/Other] | contributor                  | source="Library of Congress Name Authority File" |
| Publisher/Distributor (repeatable)              | Publisher<br>[publisherRole=Publisher/Distributor]                                 | publisher                    | source="Library of Congress Name Authority File" |
| Language of Material (repeatable)               | instantiationLanguage  | language                     | source="ISO 639.2"                               |
| Description (repeatable)                        | description [descriptionType=Content<br>Summary/Abstract/Transcript/Shotlist]      | description                  |  |
| Subject [subjectType=Topic/Entity] (repeatable) | subject  | subject                      | source="Library of Congress Subject Headings"    |
| Genre   | Genre  |                              | source="The Moving Image Genre-form Guide"       |

## Metadata Map For the California Preservation Program (CPP)

|  |  |                         |  |
|--|--|-------------------------|--|
| Coverage (repeatable)                                      | Coverage<br>[coverageType=Spatial/Temporal]                        | coverage                | source="Library of Congress Subject Headings";<br>source="Library of Congress Extended Date/Time Format" |
| Collection Guide Title and URL                             | collectionTitle  | relation-ispartof       | collectionSource=Institution Name  |
| Relationship (repeatable)                                  | [relationType=Has Part/Is Part Of/Has Version/Is Version Of/Other] |                         |  |
| Additional Title (repeatable)                              | [titleType=Alternative/Reel/Tape/Series]                           | title-alternative       |  |
| Call Number (repeatable)                                   | Identifier [identifierType=Call Number/Other]                      | identifier              | source=Institution Name  |
| Additional Descriptive Notes for Overall Work (repeatable) | description [descriptionType=Additional description]               |                         |  |
| <b>[Rights Metadata]</b>                                   |  |                         |  |
| Copyright Statement  | rightsSummary  | rights                  | rights Summary<br>annotation=Copyright Statement   |
| Country of Creation  | [Extension]  |                         | annotation=countryofcreation<br>source="ISO 3166.1"  |
| Copyright Holder   | rightsSummary  |                         | rightsSummary<br>annotation=Copyright Holder   |
| Copyright Holder Info                                      | rightsSummary  |                         | rightsSummary<br>annotation=Copyright Holder Info  |
| Copyright Date   | rightsSummary  |                         | rightsSummary<br>annotation=Copyright Date   |
| Copyright Notice   | rightsSummary  |                         | rightsSummary<br>annotation=Copyright Notice   |
| Institutional Rights Statement (URL)                       | rightsSummary  |                         | rightsLink annotation=Institutional Rights Statement   |
| <b>[Technical Metadata about the original]</b>             |  |                         |  |
| Total Number of Reels/Tapes/Pages Generation               | Extent<br>instantiationGenerations                                 | format-extent<br>format |  |
| Media Type   | [instantiationMediaType=Moving Image/Sound/Still Image/Text]       | type                    |  |

## Metadata Map For the California Preservation Program (CPP)

|  |                                   |               |  |
|--|-----------------------------------|---------------|--|
| Dimensions (still image/text only)                       | instantiationPhysical             | format-extent |  |
| Gauge and Format (a/v only)                              | instantiationPhysical             | format-medium |  |
| Duration (a/v only)                                      | instantiationDuration             | format-extent |  |
| Silent or Sound (a/v only)                               | instantiationTracks               | format        |  |
| Color or Black & White (a/v only)                        | instantiationColors               | format        |  |
| Aspect Ratio (a/v only)                                  | instantiationEssenceTrack         |               | annotation=AspectRatio                             |
| Running Speed (a/v only)                                 | instantiationEssenceTrack         |               |  |
| Track Standard (a/v only)                                | instantiationEssenceTrack         |               | annotation=TrackStandard                           |
| Channel Configuration (a/v only)                         | instantiationChannelConfiguration | format        |  |
| Subtitles/Intertitles/Closed Captions (a/v only)         |                                   |               |  |
| Additional Technical Notes for Overall Work (repeatable) | instantiationAnnotation           |               | annotation=[Additional Technical Notes for Overall |

### [Administrative Metadata]

|                         |  |             |   |
|-------------------------|--|-------------|---|
| Internet Archive URL    | Identifier   | identifier  |   |
| Asset Type              | AssetType  |             | Media Object  |
| Object Identifier       | Identifier [identifierType=Object Identifier]                  |             | source=CAVPP  |
| Project Identifier      | Identifier [identifierType=Project Identifier]                 |             | source=CAVPP  |
| InstitutionURL          | collectionSource   |             |   |
| Project Note            | Extension  | description | California Audiovisual Preservation Project (CAVPP) |
| Quality Assurance Notes | instantiationAnnotation<br>annotationType=CPPQA/<br>Partner QA |             |   |
| Grant Cycle             |  |             |   |
| Cataloger Notes         |  |             |   |

## Metadata Map For the California Preservation Program (CPP)

|   |                           |  |  |
|---|---------------------------|--|--|
| <b>[Technical<br/>Metadata about<br/>the digital files]</b> | Digital File Identifier   | instantiationIdentifier                        | annotation=File Name   |
|   | Creation Date             | instantiationDate                              |  |
|   | Frame Size                | instantiationDimensions                        | unitsOfMeasure="pixels"  |
|   | File Extension            | instantiationDigital                           |  |
|   | Standard and File Wrapper | instantiationStandard                          | source="IANA MIME Media Types" or source="PRONOM Technical Registry" |
|   | File Location             | instantiationLocation                          | annotation=LTO Tape Number and Barcode Number                        |
|   | Media Type                | instantiationMediaType=Moving Image/Sound      | type   |
|   | Generation                | instantiationGenerations                       |  |
|   | Size                      | instantiaionFileSize                           | unitsOfMeasure="GB" or "MB"  |
|   | Duration                  | instantiationDuration                          |  |
|   | Colors                    | instantiationColors                            |  |
|   | Tracks                    | instantiationTracks                            |  |
|   | Channel Configuration     | instantiationChannelConfiguration              |  |
|   | Language                  | instantiationLanguage                          | source="ISO 639.2"   |
|   | Track Type                | essenceTrack<br>[essenceTrackType=Video/Audio] |  |
|   | Encoder                   | essenceTrackEncoding                           |  |
|   | Sampling Rate             | essenceTrackDataRate                           | unitsOfMeasure=Mbps/kHz  |
|   | Frame Rate                | essenceTrackFrameRate                          | unitsOfMeasure=fps   |
|   | Playback Speed            | essenceTrackPlaybackSpeed                      | unitsOfMeasure=fps/ips/rpm   |
| Bit Depth   | essenceTrackBitDepth      |  |  |

## Metadata Map For the California Preservation Program (CPP)

|               |  |   |   |
|---------------|--|---|---|
|               | Aspect Ratio                               | essenceTrackAspectRatio   |   |
|               | Data Compression                           | essenceTrackAnnotation  |   |
|               | Relationship/Transfer Vendor/Creation Date | instantiationRelation<br>[instantiationRelationType=Derived from] or [instantiationRelationType=Is part of] | source="CAVPP" or "Vendor Name YYYY-MM-DD"<br>annotation="Source Item Number" |
| [PREMIS-like] | MD5/MD5 Date                               | objectIdentifierValue<br>[objectIdentifierType=filename]/messageDigest                                      | source="Vendor Name YYYY-MM-DD"<br>annotation="checksum"                      |
|               | Creating Application and Version           | instantiationAnnotation<br>annotationType=creatingApplicationName   |   |
|               | Source Deck Manufacturer/Model             | instantiationAnnotation<br>annotationType=Source Deck   |   |
|               | Digitizer Manufacturer/Model               | instantiationAnnotation<br>annotationType=Digitizer   |   |
|               | Technical Evaluation report                | instantiationAnnotation<br>annotationType=Vendor Technical Evaluation report                                | Transfer; Quality Control   |
|               | Vendor Quality Control Notes               | instantiationAnnotation<br>annotationType=Vendor Quality control  |   |