GenDB Workflow: Object and Finding Aid Linking

Prepared by the CDL
Digital Special Collections Program
2009

This workflow applies to institutions using GenDB to create digital objects with associated finding aids, for submission to the Online Archive of California (OAC). It is addressed to project managers/technicians at UC campuses that directly use (or coordinate use) of GenDB, and also oversee EAD finding aid encoding.

It assumes the following:

- Your final products will be 1) METS digital objects, and 2) EAD finding aids. Both will be published in the OAC at the end of the workflow.
- You will be using GenDB to create digital objects, and have made arrangements with UCB Digital Publishing Group (UCB DPG -- <http://www.lib.berkeley.edu/cgi-bin/dpg_contact.cgi>) to use the application.
- You will be using a different application or strategy to create EAD finding aids.

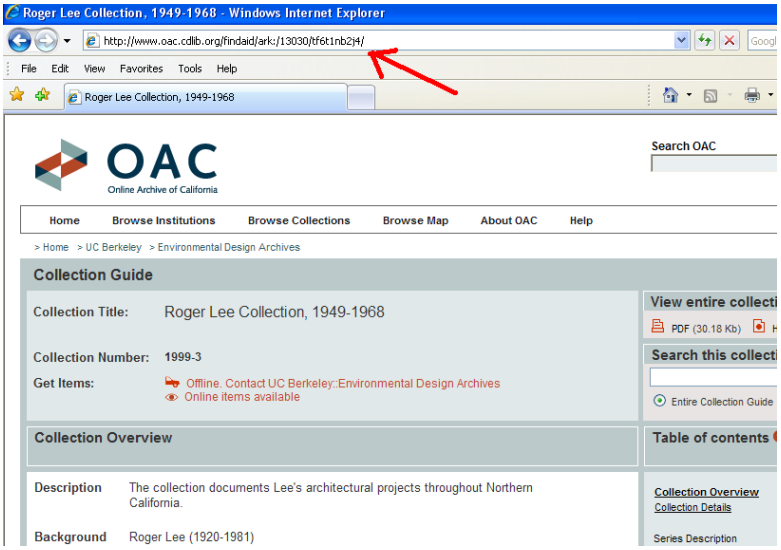
Workflow

Step 1: Determine the ARK URL for an existing finding aid / Create and submit a minimal or collection-level finding aid *without* links to associated objects

Obtain the ARK URL for your finding aid (that will be associated with the objects you're creating in GenDB). The ARK URL is a globally unique, persistent identifier that the CDL uses to track and manage content.

If your finding aid is currently published in the OAC, they will already have an ARK URL. From the list of finding aids on the OAC from your institution at <http://www.oac.cdlib.org/institutions/>, click on a collection title -- the ARK URL will appear in your browser address field.

Example finding aid URL
http://www.oac.cdlib.org/findaid/ark:/13030/tf6t1nb2j4/



If you don't already have a finding aid published in the OAC, you can obtain an ARK URL for it using one of two methods:

- Submit an initial version of the finding aid (e.g., a minimal or collection-level record) to the OAC as a "placeholder" in order to obtain an ARK. *Do not encode any links to associated digital objects at this time* -- you can revise and amplify the finding aid with links to associated digital objects later (see **Step 5**).

- Ask the CDL for an ARK -- send an e-mail to <oacops@cdlib.org>. Create and submit an initial version of the finding aid (e.g., a minimal or collection-level record) to the OAC. *Do not encode any links to associated digital objects at this time* -- you can revise and amplify the finding aid with links to associated digital objects later (see **Step 5**). Before submitting the finding aid to the OAC, encode the end portion of the ARK URL within an <eadid> IDENTIFIER attribute (see bolded example):

  Example:
  <eadid countrycode="us" identifier="**ark:/13030/kt4w10133d**" mainagencycode="CU-SC" publicid="PUBLIC "-//University of California, Santa Cruz::University Library::Special Collections//TEXT (US::CU-SC::MS 74::John Cage Mycology Collection)//EN" "ms74.sgm">ms74.xml</eadid>

  For instructions on submitting finding aids to the OAC, see <http://www.cdlib.org/inside/projects/dsc/ingest/ead_ingest.html>.

Step 2: Create digital objects, and input links back to the finding aid

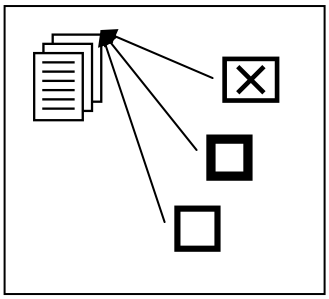
Create your digital objects in GenDB at this time.

In GenDB, the objects need to have the ARK for their associated finding aid in their **Finding Aid** fields. They'll also need their collection title in the **Finding Aid Label** field. If you need to make any edits to a collection title, now is the time to do it. Use the following steps:

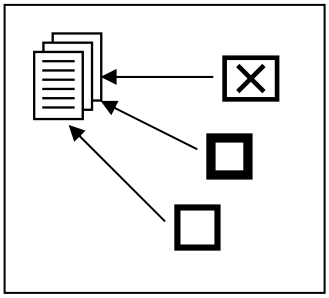
- Logon to GenDB using your project manager logon and password.
- Click on "Go" to the right of **Project Manager Options**.
- Select "Relateds (collection, metacollection)" from the pull-down menu next to **Add Value(s) to**, and click on "Go".
- In the "Create New/Edit Related" window, select the collection title you wish to edit in the **Related Materials** pull-down menu, and click on the "Choose value to edit" radio button.
- Edit the collection title based on *Describing Archives: a Content Standard* guideline #2.3 (see <http://www.archivists.org/catalog/pubDetail.asp?objectID=1279>).
  - Note that the term 'collection' should display in the "Related type" field.
  - The following fields do NOT require data entry:
    - ID/Call#
    - Institution/Creator Name
    - Related URL
- Click on the "Edit value" radio button to save changes.

UCB DPG can batch add this information to your objects, if you'd like: send an e-mail to UCB DPG at <http://www.lib.berkeley.edu/cgi-bin/dpg_contact.cgi> to request this.

Step 2A. If you don't mind the objects linking back to the top level of the finding aid (and not to a specific section in the container list), then you're done with this step. All objects will link back to top level of finding aid by default, if they have the ARK for their associated finding aid. Go directly to **Step 3**.



Step 2B. If you want the objects to link back to a specific point in the finding aid container list, it may involve a bit more work: this will involve encoding a unique identifier in the object, and the same unique identifier at the location in the finding aid to which you want the object to link.



For some context: in the finding aid, the unique identifiers need to be encoded within the ID attribute of the <c0x> Component tags -- see bolded "**7894**" below (instructions covered in **Step 5B** in greater detail).

Example:

```
<c02 id="7894" level="item">
        <did>
                <container type="box" label="Flat file">7</container>
                [etc.]
        </did>
</c02>
```

The *same* unique identifiers need to also be encoded in the **Finding Aid ID** field in GenDB. In other words, the same unique identifier needs to live both in the object, and the specific section in the finding aid you want the object to link back to.

If you already have encoded unique identifiers in your ID attributes of <c0x> Component tags in the finding aid container list, that's great -- you can use those for this step. Otherwise, you'll need to formulate them: just note that each unique identifier must be unique within the finding aid (i.e., do not use duplicate identifiers throughout the finding aid). When formulating a unique identifier, please keep in mind that they can only include letters, digits, periods, hyphens, underscores, and colons.

There are two options for encoding them in your objects in GenDB: asking UCB DPG to automatically batch add them, or manually adding them to individual records in GenDB (by keying the unique identifiers into the **Finding Aid ID** fields yourself).

If you'd like UCB DPG to automatically batch add them, they will utilize the object's **Object ID** as the unique identifer. Send an e-mail to UCB DPG at <http://www.lib.berkeley.edu/cgi-bin/dpg_contact.cgi> to request this.

Step 3: Input rights metadata for digital objects

Step 3A. Send an e-mail to UCB DPG at <http://www.lib.berkeley.edu/cgi-bin/dpg_contact.cgi> and let them know which of the following four rights categories predominate for the majority of your objects:

1) UNKNOWN
2) PUBLIC DOMAIN
3) COPYRIGHTED (copyright owner is "UC Regents")
4) COPYRIGHTED (copyright owner is not UC Regents, but another individual or institution). If selecting this category, please also supply UCB DPG with the following two key pieces of information:

- Copyright owner name. Enter copyright owner name in direct, natural order (e.g., "Jane Doe" not "Doe, Jane")
- Copyright owner contact information -- supply as much publicly-accessible information as possible. If no contact information is provided, this will default to your institution's contact information.

If you specify 1,2, or 4, UCB DPG will change the current default rights record (listed as "UC Regents", with rights category set to COPYRIGHTED) linked to most objects with the new default that you specify. Note that this process will *not* affect rights records and rights categories (with associated links to objects or subobjects) that you have already created in GenDB. Note also that the "UC Regents" rights record will still be available for use.

The following boilerplate text will automatically be supplied in the objects and subobjects (at the point that the object data is compiled into METS by GenDB), based on the rights category you indicate:

1) UNKNOWN: "Copyright status unknown. Some materials in these collections may be protected by the U.S. Copyright Law (Title 17, U.S.C.). In addition, the reproduction of some materials may be restricted by terms of University of California gift or purchase agreements, donor restrictions, privacy and publicity rights, licensing and trademarks. Transmission or reproduction of materials protected by copyright beyond that allowed by fair use requires the written permission of the copyright owners. Works not in the public domain cannot be commercially exploited without permission of the copyright owner. Responsibility for any use rests
exclusively with the user."

2) PUBLIC DOMAIN: "Material in the public domain. No restrictions on use."

3) COPYRIGHTED (Note: copyright owner name -- e.g., UC Regents, or specific individual or institution -- and publicly-accessible contact information will also be displayed with the following rights statement): "Transmission or reproduction of materials protected by copyright beyond that allowed by fair use requires the written permission of the copyright owners. Works not in the public domain cannot be commercially exploited without permission of the copyright owner. Responsibility for any use rests exclusively with the user."

Steb 3B. For specific objects that deviate from the default set in **Step 3A**: you will need to assign the correct rights record with appropriate rights category to each object *and* subobject (if the latter is applicable).

To do this for cases where the rights categories are UNKNOWN, PUBLIC DOMAIN, or COPYRIGHTED by the UC Regents, select the appropriate rights record from the pull-down menu (pull-down menu values are "Unknown", "Public domain", or "UC Regents"). (Note that the values for "UNKNOWN" and "PUBLIC DOMAIN" will be added to the pull-down menu by UCB DPG once you complete Step 3A).

For all cases where the rights category is COPYRIGHTED by another individual or institution: you will need to establish a rights record for the copyright owner, if one does not already exist. Then create a link from the object or subobject to this particular rights record. To establish a new rights record:

- Select the **Create New/Edit Rights** link from the input screen
- In the "Create New/Edit Rights" window, enter the copyright owner name in the **Owner Name** field
- Select "COPYRIGHTED" from the **Rights Category** pull-down menu
- Enter copyright owner contact information (that should be available to the public) in the **Contact Address** field
- Enter any additional comments or notes (e.g., licensing statements, etc.) in the **Comments** field
- Click on the "Create & go back" radio button to save changes

An alternate method for creating Rights values can be implemented using the Project Manager functions:

- Select "Rights" from the pull-down menu next to **Add Value(s) to**, and click on "Go".
- In the "Create New/Edit Rights" window, enter the copyright owner name in the **Owner Name** field
- Select "COPYRIGHTED" from the **Rights Category** pull-down menu
- Enter copyright owner contact information (that should be available to the public) in the **Contact Address** field
- Enter any additional comments or notes (e.g., licensing statements, etc.) in the **Comments** field
- Click on the "Create & go back" radio button to save changes

Step 4: Notify DPG when you've completed your digital objects / Quality control review

When you're finished with encoding the metadata in your objects using GenDB, send an e-mail to UCB DPG at <http://www.lib.berkeley.edu/cgi-bin/dpg_contact.cgi> and ask them to generate METS files for them. In your e-mail, include the finding aid ARK with which the objects are associated.

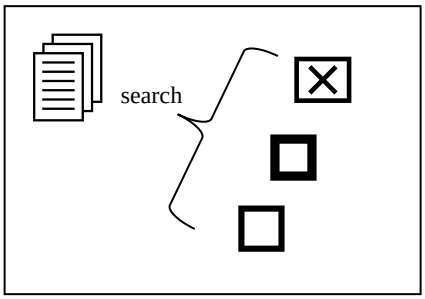
UCB DPG will then provide you with a report listing objects associated with that finding aid. The report will contain some useful metadata in the objects that you may need to refer to later (see **Step 5B**). It may also include some information to allow you to preview and quality control-check the objects.

Step 5: Update the finding aid *with* links to associated objects

Using your local EAD encoding tools, update the finding aid (associated with the objects) with outbound links.

You have two options for creating links from your finding aids back to your objects:

Step 5A (aka "*Easy path/low pain*"): link from a specific level in the finding aid (e.g., collection, series, etc.) to a group of objects, by linking to a search query for the group of objects in the OAC. Note that you will be creating a link to a "canned" search for a group of objects, not to an individual or specific object: for an example of how this works, see the Redwood City Public Library's "Morrish Collection" <http://www.oac.cdlib.org/findaid/ark:/13030/kt70002161>.



To do this, add a <dao> Digital Archival Object tag in <scopecontent> within the <archdesc> level of the finding aid (if linking from the collection level), or add a <dao> to a particular <c0x> Component level in the <dsc> Description of Subordinate Components (container list) section of the finding aid (if linking from major subdivisions or file or item levels). Within the <dao>, encode a link to the search query in an HREF attribute. Use the following specs:

ELEMENTS & ATTRIBUTES	COMMENTS / APPLICATION NOTES
<dao>	Nest the <dao> within <did>.
role= "http://oac.cdlib.org/arcrole/ link/search/"	Encode this snippet.
href= "http://content.cdlib.org/search?style=oac-img&sort=title&relation=**[ark]"**	Encode the search query for a collection of objects in HREF using the following search query syntax. Supply the end portion of the ARK URL for the finding aid in the "relation" parameter of the search query (see bolded "**ark:/13030/kt9290094p**" below): *Example:* http://content.cdlib.org/search?style=oac-img&sort=title&relation= **ark:/13030/kt9290094p**
title=	Use an appropriate label to characterize the nature of the object version or format (e.g., "image," "Online items"). The default label supplied by the OAC is "view

ELEMENTS & ATTRIBUTES	COMMENTS / APPLICATION NOTES
	attached object". Any label supplied in TITLE will override the default label supplied by the OAC.
<daodesc>	Use this optional tag to state information about the contents, usage, or source of the digital objects, when not sufficiently characterized by information in <unittitle> or other descriptive information.

Example (link to a search for a collection of CDL-hosted METS digital objects, from a collection-level description):

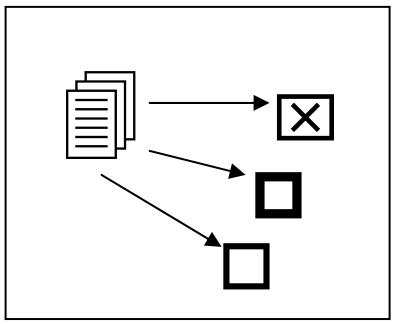
<archdesc level="collection">
<did>
<origination label="Creator">
<corpname source="aacr2">Anaheim Public Library </corpname>
</origination>
<unittitle>Anaheim Public Library photograph collection on Anaheim local history</unittitle>
<unitdate type="inclusive" normal="1860/1970">1860-1970</unitdate>
<unitdate type="bulk" normal="1860/1923">(bulk 1860-1923)</unitdate>
<physdesc>
<extent>1,802 items</extent>
</physdesc>
<abstract>Anaheim Public Library's photograph collection includes images of historical interest of the City of Anaheim and other areas of Orange County from the 1860s to 2002. Images document public, residential and commercial buildings, including businesses, schools, churches, citrus packing houses, fire and police department facilities, theaters (such as the Fox Theater), and the public library; street scenes; neighborhoods; significant individuals, including members of the Los Angeles Vineyard Society which founded Anaheim in 1857 as a wine making colony, early mayors and civic leaders (such as August Langenberger and Charles Pearson); groups and family portraits; annual events, such as the California Valencia Orange Show and the Anaheim Halloween Parade; local geography, including the Santa Ana River, Anaheim Landing and local canyons; rancho families, such as Juan Pacifico Ontiveros and Vicenta Sepulveda Yorba Carrillo; Mission San Juan Capistrano; viticulture and agriculture, including the Anaheim chili peppers; transportation; ethnic communities, including Japanese Americans, Chinese Americans, etc.; natural disasters, such as the 1933 Long Beach earthquake and the 1938 flood; Anaheim Resort area, including Edison International Field (also known as the Big A), Anaheim Convention Center, Arrowhead Pond of Anaheim and Disneyland. Formats include panoramic and aerial photographs. Of particular interest are the large number of photographs which document the development of the Los Angeles Vineyard Society from circa 1860 to 1890. The collection also contains a small number of ambrotypes, daguerreotypes, and tintypes from the 1870s and 1880s.</abstract>
<repository>
<corpname source="lcnaf">Anaheim Public Library.
<subarea>Central Library. History Room.</subarea></corpname>
<address>
<addressline>Anaheim, California 92805</addressline>
</address>
</repository>

```
<unitid repositorycode="CAna" countrycode="US">Consult repository</unitid>
<langmaterial>Collection materials are in <language
langcode="eng">English</language>.</langmaterial>
<dao role="http://oac.cdlib.org/arcrole/link/search/"
href="http://content.cdlib.org/search?style=oac-
img&amp;amp;sort=title&amp;amp;relation=ark:/13030/kt2199p9w7" title="Online items">
<daodesc>
<p>Selected digitized images from this collection.</p>
</daodesc>
</dao>
</did>
...
</archdesc>
```

Example (link to a search for a collection of CDL-hosted METS digital objects, from a series-level description):

```
<dsc>
...
<c01 level="series">
<did>
<container type="box" label="Drawer : Folder ">1 : 56 - 4 : 7</container>
<unitid>Series 2. </unitid>
<unittitle>Correspondence</unittitle>
<unitdate>1840-1947</unitdate>
<physdesc><extent>5.5 cubic ft.</extent></physdesc>
<dao role="http://oac.cdlib.org/arcrole/link/search/"
href="http://content.cdlib.org/search?style=oac-
img&amp;amp;sort=title&amp;amp;relation=ark:/13030/kt9290094p" title="Online items">
<daodesc>
<p>Selected digitized images from this series.</p>
</daodesc>
</dao>
</did>
<scopecontent>
<head>Series Scope and Content Summary</head>
<p>This series contains correspondence between Ellen Browning Scripps and her
family, friends, and acquaintances. Most of the excerpts of correspondence located
at the beginning of this series were compiled by J. C. Harper and are arranged by
subject. Excerpts may also be found in correspondence with Edward Wyllis Scripps
(E. W.), 1868-1926.</p>
<p>This series is arranged alphabetically by correspondent or subject.</p>
</scopecontent>
...
</c01>
...
</dsc>
```

Step 5B (aka, "*Cumbersome path/high pain*"): link from a specific level in the finding aid (e.g., file- or item-level) to a specific, individual object.



To do this, create a <dao> Digital Archival Object tag at the particular <c0x> Component level in the <dsc> Description of Subordinate Components (container list) section of the finding aid where you'd like to make a link to an object. Use the following specs:

ELEMENTS & ATTRIBUTES	COMMENTS / APPLICATION NOTES
<dao>	
role= http://oac.cdlib.org/arcrole/link/"	If you are not linking to an object that is a single image, supply one of the following qualifiers at the end of the URL: **Type of object** / **Qualifier** Multiple images / image+collection Text / text *Example (for a group of images):* http://oac.cdlib.org/arcrole/link/ image+collection
href=	Encode the ARK for the object, in its full URL form. This will be provided by UCB DPG in your list of objects (from **Step 3**). The URL form begins with "http://ark.cdlib.org/...": *Example:* http://ark.cdlib.org/ark:/13030/kt4p3005qx/
title=	Use an appropriate label to characterize the nature of the object version or format (e.g., "image," "Online items"). The default label supplied by the OAC is "view attached object". Any label supplied in TITLE will override the default label supplied by the OAC.

ELEMENTS & ATTRIBUTES	COMMENTS / APPLICATION NOTES
<daodesc>	Use this optional tag to state information about the contents, usage, or source of the digital object, when not sufficiently characterized by information in <unittitle> or other descriptive information.

Example:

```
<c02 level="item">
        <did>
                <container type="box" label="Flat file">7</container>
                <unittitle>Ushiwaka and Benkei duelling on Gojo Bridge or Gojo
                Bridge, an episode from the Life of Yoshitsune, Chronicles of
                Yoshitsune,
                        <unitdate>1881</unitdate>.
                </unittitle>
                <dao role="http://oac.cdlib.org/arcrole/link/"
                href="http://ark.cdlib.org/ark:/13030/kt4p3005qx/" title="View
                image"></dao>
        </did>
</c02>
```

If you created unique identifiers for your objects in **Step 2B**, you'll now need to add them to the appropriate <c0x> Component levels where you want the objects to link back to. Encode the unique identifiers within the ID attribute of the <c0x> Component tags (see bolded "**7894**" below).

Example:

```
<c02 id="7894" level="item">
        <did>
                <container type="box" label="Flat file">7</container>
                [etc.]
        </did>
</c02>
```

Step 6: Submit your digital objects

Notify UCB DPG when you are ready to publish your objects in the OAC; they will confirm with you once they're in the OAC.

Step 7: Submit the updated finding aid

Use voroEAD to upload and process your updated finding aids. Do not upload your updated finding aids until *after* **Step 6**, however -- i.e., after UCB DPG has confirmed that your objects are in OAC.

Step 8: Check in with UCB DPG one last time

Notify UCB DPG and CDL once you're totally done with the above steps.

Step 1: Determine the ARK URL for an existing finding aid / Create and submit a minimal or collection-level finding aid *without* links to associated objects
[] Obtain the ARK URL for your finding aid (that will be associated with the objects you're creating in GenDB). If you already have the finding aid in the OAC, you can determine the ARK URL by viewing the finding aid in OAC.
[] If you don't have an existing finding aid in the OAC, create and submit a minimal or collection-level record to the OAC to obtain an ARK URL. (Alternatively, request an ARK URL from CDL; and input that ARK URL into the finding aid before you submit it).

Step 2: Create digital objects, and input links back to the finding aid
[] If linking from objects to specific sections in the finding aid, add unique identifiers into objects (manually or request UCB DPG to do this). If you don't mind the objects linking back to the top level of the finding aid (and not to a specific section in the container list), then skip this step.

Step 3: Input rights metadata for digital objects
[] Let UCB DPG know which of the three rights categories should be the default for your all of your objects
[] Select appropriate rights category for objects that deviate from the default

Step 4: Notify DPG when you've completed your digital objects / Quality control review
[] Obtain a report of your objects from UCB DPG, and conduct a quality control review.

Step 5: Update the finding aid *with* links to associated objects
[] Update the finding aid so it has outbound links to the associated objects

Step 6: Submit your digital objects
[] Notify UCB DPG to publish your objects in the OAC

Step 7: Submit the updated finding aid
[] Use voroEAD to upload your updated finding aid *with* outbound links to associated objects.

Step 8: Check in with UCB DPG one last time
[] Notify UCB DPG when totally done

Contacts

Accessing and using GenDB
UCB DPG <http://www.lib.berkeley.edu/cgi-bin/dpg_contact.cgi>

Questions about this document
CDL <oacops@cdlib.org>