

Representing HathiTrust Resources in your Local ILS

The inclusion of a resource in the HathiTrust repository results from the deposit of bibliographic metadata provided by the contributing institution. HathiTrust resources are then discoverable in the HathiTrust online catalog and via systems and services employed by UC campuses (including Melvyl/WorldCat) and through various knowledge bases (e.g. Primo). It is possible to encourage end-user discovery beyond these environments by adding descriptions of and/or links to HathiTrust resources in your local ILS and thus to your local online catalog.

This document serves as a general guide to assist with the planning and implementation processes for UC technical services and/or systems librarians interested in including links to HathiTrust resources locally, i.e. in the ILS. Public services librarians may also refer to this guide in considering how users encounter links to HathiTrust resources in local discovery interfaces.

There are a number of decisions a campus must make in order to determine the appropriate workflow for including HathiTrust resources their ILS: the scope of resources to include (local holdings only vs. additional HathiTrust volumes); the ILS record type preferred per local practice (single print or electronic records; both print and electronic); and linked resource access status (open access volumes only, or open access volumes plus search-only volumes). These decisions will largely be driven by local cataloging and record management policies and considerations for each are described in more detail below.

Included is a decision tree flow chart (page 3) intended to outline the workflow steps and processes based on the decisions described above. Also provided are descriptions of each workflow step. Sections preceding the flow chart are intended to provide the necessary information and context needed for decision making.

The relationship between HathiTrust resources, OCLC and WorldCat Local

OCLC routinely updates WorldCat with records describing electronic resources in the HathiTrust repository by referring to data published by HathiTrust¹ to initiate or update the description of HathiTrust resources in WorldCat. When a volume is added to the HathiTrust repository, OCLC creates or updates a record describing the electronic resource and indicates that it is held at HathiTrust by adding the holding symbol "HATHI". Per UC decision,² digitized content from UC campuses contributed to HathiTrust are also given the holding symbol "CDLER" in order that they appear in each campus's WorldCat Local implementation.

Data Sources

There are a number of data sources that may be used to add links to HathiTrust resources into local systems. They are described here to provide context for the ILS Workflow Decisions section below.

¹ OCLC creates an electronic record - derived from the print record - after a digitized item is ingested into the HathiTrust repository. They download the Hathifiles (www.hathitrust.org/hathifiles) nightly and use this data to update a table that links the print and electronic OCLC numbers, as well as the HathiTrust ID and other metadata. The print OCLC number is inserted into the 776\$w of the electronic record, as is the electronic OCLC number inserted into the 776\$w of the print record.

² This dates back to a 2011 CoUL decision that attempted to indicate that the book is an electronic resource that is available UC-wide. CoUL had decided that no matter which campus had contributed a specific item to be scanned for inclusion in HathiTrust that the CDLER holdings symbol would display for ALL campuses. So, all campuses appear to hold an electronic version of the item.

HathiTrust Sources

Hathifiles HathiTrust [provides files of tab-delimited bibliographic metadata elements](#) (called “Hathifiles”) for volumes in the repository, including values that can be used as match points to data in local systems (e.g., digital volume identifiers, local bibliographic system numbers, master OCLC numbers for the print resource). This metadata is provided daily in incremental files and monthly in summary files. A description of the file contents and formatting [can be found here](#). These files are intended to be used to aid institutions in loading bibliographic data for volumes in HathiTrust and linking to the volumes from local catalog systems.

HathiTrust Bib API and OAI feed The [HathiTrust Bib API](#) returns bibliographic, rights, and volume information when given a single or multiple standard identifiers (ISBN, LCCN, OCLC, etc.). It is intended to retrieve information about small numbers of items at a time. HathiTrust also makes available an OAI feed of MARC21 and unqualified Dublin Core records for public domain materials. Bulk data retrieval should be done using OAI or the Hathifiles, as described at <http://www.hathitrust.org/data>.

OCLC Sources

OCLC Connexion Connexion is an OCLC service used to create, manage and download high-quality bibliographic and authority records. More information [can be found here](#).

OCLC z39.50 Server This OCLC service is available to libraries who have a cataloging subscription. OCLC maintains a z39.50 server used to access WorldCat to search and retrieve MARC records for cataloging; to edit records in local systems; and to set holding information on WorldCat. More information on the service [can be found here](#).

OCLC WorldCat Search API OCLC maintains an API to allow developer level access to WorldCat, which allows for the search and retrieval of bibliographic records (as MARCXML or Dublin Core). More information on the service [can be found here](#).

ILS Workflow Decisions

In order to determine the appropriate workflow for your campus to enable HathiTrust resource discovery in your catalog, you must make decisions about the following:

Scope of Resources

You may choose to include any or all of the following:

- Links for volumes contributed to HathiTrust by your campus so that end users are aware of resources that your library holds, has digitized and has contributed to HathiTrust.
- Links for volumes held by, but not contributed to HathiTrust, by your campus so that end users are aware of digitized versions of resources in HathiTrust and aware that there are alternate means of accessing the content.
- Links for volumes contributed by other campuses or HathiTrust partners but not held by your library so that your users are able to access resources that may not be physically available in the library.

Access Status of Linked Resources

[The Hathifiles](#) contain the rights status and viewability (open or closed access) for each volume. See [Access Determination for HathiTrust Objects](#) and [HathiTrust Rights Database - Database Layout](#) for more information about the processes and metadata that control the access and use of resources.

You may choose to include:

- Links for volumes that are openly accessible to your users, whether those are public domain or are open for other reasons.
- Links for all volumes regardless of access status (i.e. including links for search-only volumes). Be aware of the implications for your users of including links to objects that are not available in full text.

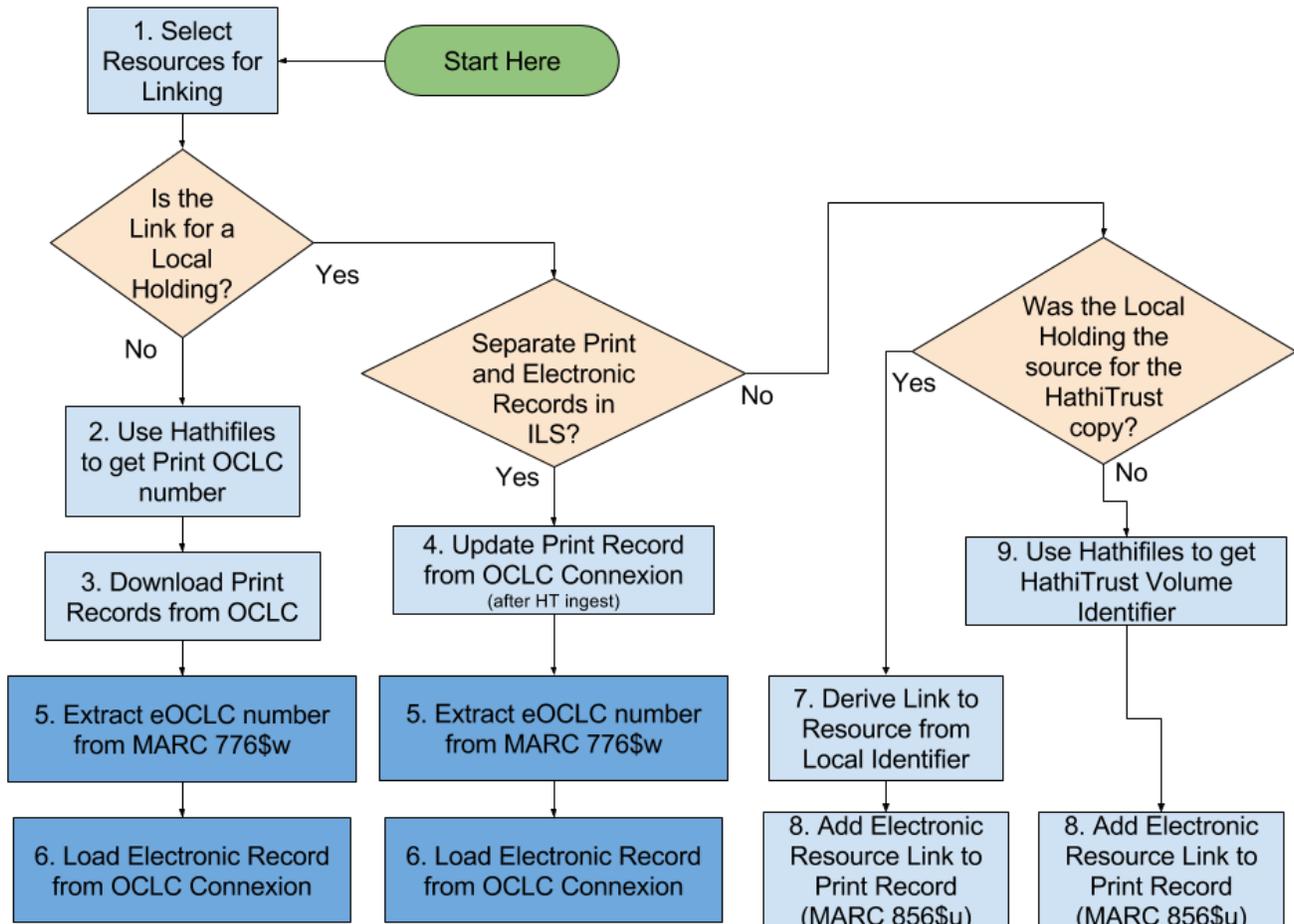
ILS Record Types

You may choose to include one or both of the following record types in the ILS:

- **Print Record**
 - If you hold a print version of the resource locally, you should select this option.
 - You may use this record to link directly to the electronic resource via the MARC 856\$u.
 - You may use this record to link to the OCLC electronic resource record via the MARC 776\$w.
- **Electronic Resource Record**
 - Campus cataloging policy will likely determine whether you include an electronic resource record.
 - If you are adding links for materials you do not hold in print, it is recommended to include the electronic resource record, but not the print record.
 - You may use this record to link to the print record via the MARC 776\$w

Decision Tree Flow Chart

Given the decisions made above, use the Decision Tree below to determine the workflow you should follow. A description of the steps for each action (blue boxes) follows the graphic.



Workflow Steps

1. Select Resources for Linking Determine what resources you wish to link to. Many factors could influence the scope, as described above. You will ultimately need a list of HathiTrust identifiers and/or OCLC numbers (for the print resource) to proceed. You may want to use the Hathifiles to identify targets based on viewability, contributor (using the “source” value), etc. Alternately, should the collection of HathiTrust identifiers prove too difficult or present unique challenges, HathiTrust administrative staff or CDL staff may be able to provide support for the generation of custom lists of volumes.

2. Use Hathifiles to get Print OCLC Number The Hathifiles contain the OCLC number(s) for the print version of the resource in the 8th column of the file. If you wish to constrain the HathiTrust resources that are linked to from your catalog to just those that are in the public domain (and thus openly accessible) be sure to also grab the Access value (i.e. viewability - allow or deny; column 2) and the Rights Attribute (the rights status used to derive the access value; column 3).

3. Download Print Records from OCLC Use OCLC's Connexion, z39.50 server or the WorldCat Search API to download the records into a file for the resources you want to link to. If taking this step, the resources are not local holdings, so you should not load the print records into your ILS; they are only intended for the extraction of the eOCLC number. These print records can be loaded into the ILS, but only if this makes extracting the eOCLC number easier. If these records are loaded into the ILS, it is recommended that they be deleted after the eOCLC numbers are acquired.

4. Update Print Record from OCLC Connexion As described above, once a volume has been ingested into the HathiTrust repository, OCLC will create an electronic resource record and add the OCLC number to the 776\$w of the print resource record. Update the print record in the ILS using OCLC Connexion to gain the OCLC number for the electronic resource. Note that OCLC Connexion allows a maximum of 150 items in batch searches.

5. Extract eOCLC number from MARC 776\$w Using the print records (for a. downloaded into a file from OCLC [for resources not held locally], or b. updated in the ILS [for locally held resources]) extract the OCLC number for the electronic resource from the 776\$w field. If a, use a utility (MarcEdit, UltraEdit, NotePad++, etc.) or script to retrieve the eOCLC number.

6. Load Electronic Record from OCLC Connexion Use the eOCLC number you have extracted from the print record's 776\$w to load the electronic resource record from OCLC Connexion. Note that OCLC Connexion allows a maximum of 150 items in batch searches.

7. Derive Link to Resource from Local Identifier For locally held resources that have been digitized and deposited into HathiTrust, deriving the link to the electronic resource is reasonably straightforward. The HathiTrust volume identifier, upon which the URI is based, uses the identifier employed when the object was digitized (typically the item's local identifier, often a barcode or an ARK - consult your campus' digitization staff for this information). This HathiTrust volume identifier is included among the data values in the hathifiles. HathiTrust combines local identifiers with a "namespace" prefix to create a unique HathiTrust volume identifier (HTID). The form of the persistent links to HathiTrust resources uses the HTID and is described in the Notes section below.

Here are two examples:

<http://hdl.handle.net/2027/uc1.31822010283869>

<http://hdl.handle.net/2027/loc.ark:/13960/t3vt2h87k>

8. Add Electronic Resource Link to Print Record Once you have determined the persistent link to the electronic resource in step 7 above, you can add that link to the MARC field 856\$u. Your campus likely has a policy for how URLs/URIs are cataloged and added to MARC records.

9. Use Hathifiles to get HathiTrust Volume Identifier Using the print OCLC number in your local print record, retrieve the HathiTrust Volume Identifier (HTID) from the Hathifiles. Persistent links to HathiTrust resources use the HTID in a predictable way and are described in the Notes section below.

Notes

- Persistent links to HathiTrust items can be constructed using the HathiTrust Volume Identifier (HTID). The format is: <http://hdl.handle.net/2027/<HTID>>. Example: <http://hdl.handle.net/2027/uc1.b3924435>.
- The HathiTrust Volume Identifier (HTID) can be found in the Hathifiles as the "Volume Identifier". Additionally, the Hathifiles include OCLC numbers and bibliographic "Source institution record numbers" (or Local System Control Numbers) which can be used as match points to records/metadata from local systems.
- Bibliographic data is NOT static and processes should be established to keep the data you include in your ILS current.