OHMS (Oral History Metadata Synchronizer)
Getting Started

1.0 OVERVIEW
OHMS is designed to enhance access to online oral histories in an efficient and affordable fashion. It also uses non-proprietary software and file formats in order to be interoperable and sustainable for the future. OHMS is not a repository. It is a system for making repositories better. The primary purpose of OHMS is to empower users to discover information in an oral history interview online by connecting them to moments in an interview that correspond to their search results.

There are two main components of the OHMS system:

**OHMS Application:** The OHMS application is where the work is done. This is the back-end, web-based application where interviews are imported and metadata is created. In the OHMS application, transcripts are time-coded and interviews are indexed. Upon completion, the interview record, which includes the synchronized transcript and/or time-coded index, is exported as an XML file. When this XML file is located on a web server, it interfaces with your content management system through the OHMS Viewer. The Louie B. Nunn Center for Oral History currently hosts the central installation of the OHMS application. Accounts are free.

**OHMS Viewer:** The viewer is OHMS’ user interface. When a user clicks on an interview within the repository, the OHMS viewer loads. The viewer combines select interview level metadata and the intra interview level metadata with your interview media.

The OHMS viewer currently utilizes jPlayer (HTML5) for delivering the audio/video (http://jplayer.org/).

**OHMS Viewer Examples:**
OHMS Viewer: synchronized transcript:
https://nyx.uky.edu/oh/render.php?cachefile=02OH110RPWR05_MalcomX.xml

OHMS Viewer: interview index:

OHMS Viewer: synchronized transcript + interview index:
2.0 OHMS APPLICATION: OVERVIEW
The OHMS application is a user-authenticated web application so individual users must have their own account with a username and password. Once you have an account and have been given a user name and password, log into OHMS. When you first log in, you will be asked to change your password.

Once logged into the OHMS application, you can begin to prepare your oral history interview for access. There are several major components of the OHMS application:

- **Interview Manager**: Central area to initiate an interview. Provides navigation to major functional areas of OHMS and monitoring of workflow.

- **Metadata Manager**: Input or edit collection or interview level metadata.

- **Indexing Module**: Index/annotate an interview.

- **Transcript Synchronization Module**: Place timecode into a transcript at appropriate locations.

- **Thesaurus Manager**: Manage controlled vocabularies used while indexing.

- **Interview Import**: Import interview metadata to begin indexing and/or transcript synchronization.

- **User Management**: Assign user rights and permissions.
3.0 INTERVIEW MANAGER
The Interview Manager is the main hub of the OHMS application. The Interview Manager is where you:

A. Create new interview records
B. Preview completed interviews
C. Update or edit interview-level metadata
D. Initiate the indexing module
E. Upload transcripts
F. Initiate transcript synchronization
G. Flag an interview with a note
H. Workflow management
I. Export the final XML or CSV file for an interview
J. Batch export or delete interviews
4.0 METADATA EDITOR

Interview-level metadata can either be created in OHMS or imported and, therefore, retain metadata associated with the interview throughout the OHMS process. OHMS utilizes metadata fields commonly associated with oral history.

The “Metadata” link for an interview is a navigational link to the interview-level metadata editor for that particular interview. This is where the most critical information for initiating activity in OHMS takes place.

In order to begin work in the OHMS application, three metadata fields must be completed:

- Title
- Media format
- Media URL (and any pertinent streaming server data, depending on your host site).

Once these three elements are established in OHMS, an interview can be indexed, or a transcript can be synchronized as soon as the text is uploaded. Not all of the fields visible in the interview-level metadata records will be visible in the OHMS Viewer (such as interview-level subjects and keywords). However, all of the interview-level fields remain associated with the interview in the XML document accessed by the OHMS Viewer. These can be indexed, made searchable, harvested by a digital library, archive, repository, or CMS.

The Nunn Center has worked with our university's digital library to develop efficient workflows that utilize the resulting OHMS XML document to ingest interviews into the digital library’s CMS, but also to serve as the source for the deposit of an interview into the preservation repository. As a result, the OHMS XML document contains both interview and intra-interview level metadata to maximize future flexibility.

4.1 OHMS METADATA FIELDS (INTERVIEW LEVEL)

The following is a list of interview-level metadata fields that are utilized by the OHMS application and the OHMS viewer:
• Title
• Accession Number
• Interviewee (First Last)
• Interviewer (First Last)
• Interview Date (YYYY-MM-DD)
• Date (Non-preferred format) - This is utilized for partial dates
• Collection ID
• Collection Title
• Series ID
• Series Title
• Summary
• Keywords
• Subjects
• Duration
• Media Format
• Media File Host (Choices vary by selection. See section 4.2 for more information.)
  o Host
    • Media Host Account ID
    • Media Host Player ID
    • Media Host Clip ID
    • Media URL
  o Brightcove
    • Media Host Account ID
    • Media Host Player ID
    • Media Host Clip ID
    • Media URL
• Media Filename
• Alt Sync URL - Used to link the audio version of a video source for transcript syncing. Capability to directly sync a transcript from the video source is pending.
• Usage Statement
• Rights Statement
• Funding
• Format
• Language
• Type
• Thesaurus - Choose relevant thesauri for accessing controlled vocabularies in the indexing module.
• Transcript Sync Data - Metadata created in OHMS after an interview transcript has been synchronized.
• Repository - Automatically generated
• CMS Record ID - Automatically generated
• OHMS XML Filename
• Use Restriction
When editing the Metadata for an interview, you can also change the status of the metadata using a drop-down menu, which is at the bottom or top of an interview's metadata editor page.

The status begins as "In Process" by default and can be changed to "Ready for QC", "Active QC", or "Complete". This changes the color of the metadata indicator in the Interview Manager, which gives you a visual cue as to the status of the interview.

The metadata fields necessary for the OHMS Viewer to fully function are the intra-interview level metadata created within the OHMS application itself, such as transcript synchronization and/or interview index metadata, as well as select interview-level metadata fields, such as title, interviewee, interviewer, date, collection name, series name, accession number, usage statement, and rights statement.

4.2 DIGITAL MEDIA HOST
OHMS is not a digital repository, so you do not upload the audio or video interview into OHMS. You must tell OHMS where your digital media is hosted, either by a direct hyperlink or by adding pertinent streaming server information. The media host section of the metadata editor contains multiple fields, which allows OHMS to effectively interface with the digital object via progressive download or a streaming server.

• If you are using progressive download with a direct link to the media file via URL, you will need to select "Host" when setting up your interview.

• In most cases, direct linking to the media file will only require populating the media URL field. The Nunn Center has been using progressive download for digital audio using mono, low bitrate encoded mp3 and mp4 audio files.
• If you are working with digital video, you can either upload your videos to Youtube or use the streaming server Brightcove.

• If you are working with Brightcove, you will need to place the host account, player, and item ID numbers into the appropriate fields and choose Brightcove as your media host in order to access the media file.

• If you are working with YouTube you will need to publicly share your video there and insert the shared link (not the full link) into the media URL field and select YouTube as your media host.

• The OHMS viewer is customizable but will require some programming if you are working with streaming services other than the ones listed.

4.3 METADATA EDITOR: IMPORTANT NOTES

• **Required fields:** title, media format, and media URL.

• **Media host:** This section determines the delivery source of the media file. If you are using YouTube or a streaming media host such as Brightcove or Kaltura, selecting this option provides the fields necessary for OHMS to work with these streaming services.

• **Media URL:** This is the web location for the digital audio or video. If you are not using a streaming service that requires additional information, the Media URL must be populated with a direct link to the media file in order to function in OHMS.

• **Transcript sync module and video compatibility:** The transcript sync module does not currently work directly with a video source. If you want to sync the transcript for a video interview, use the media URL field for the true URL linking to the video file. However, you must also use the alt sync URL field to link to an audio version of the video. The time codes must match in order to be accurate. The transcript sync module will utilize the audio file but on the user side, the viewer will match the completed sync data with the original video interview. (See section 7.0 of this guide for more information.)

• **Transcript sync data:** This is created only after syncing the transcript and is therefore blank in the beginning. It will be automatically populated upon the completion of a synchronized transcript.

• **Thesaurus (keywords, subjects, titles):** Assignment of a thesaurus for titles, subjects or keywords that will be autosuggested while indexing a particular interview depends on whether the repository has uploaded a thesaurus in these categories. (See section 6.0 of this guide for more information.)
• **Use Restriction:** This field does not impact the OHMS Viewer but has been added as a field. Use of this field can serve to potentially flag restricted content prior to putting an OHMS XML file/interview online. This field is both importable and exportable in the OHMS XML and CSV files.

• **OHMS XML Filename:** The input in this field determines the filename of the exported xml file. This will make it easier to determine each interview’s OHMS Viewer hyperlink.

• **Keywords:** This field functions much like the Subjects field to transport interview level metadata. This field is both importable and exportable in the OHMS XML and CSV files and does not impact the OHMS Viewer.

• **Acceptable Date Formats:** OHMS has been updated to accept a wider variety of date formats. It now accepts dates in the form of mm/dd/yyyy which is the default date format in Microsoft Excel. Now the dates in your OHMS csv should be autocorrected in Excel and import properly.

5.0 INDEXING MODULE

The indexing module is deployed by clicking an interview's “Index” button in the interview manager.

Activating an interview’s indexing module for the first time will bring up a view similar to this one:
**5.1: INDEXING CONTROLS AND WORKFLOW**

In order to begin indexing an oral history interview, you must press the “play” button on the player.

![Play button](image)

The player must be playing in order for you to create an index point. This is created by pressing the “Tag Now” button at the appropriate moment.

When you press “Tag Now” you are presented with the tag data module. This includes player controls and a series of descriptive fields to fill out.
The indexer can control the player within the tagging module. The nature of indexing means that the indexer will, inevitably, tag a segment late. In other words, you don’t know if something is important until after you have heard it. For this reason the player backtracks a few seconds each time the tagging module is activated.

5.2 INDEX MODULE METADATA FIELDS
An OHMS Index segment contains the following metadata fields:

- Time Stamp
- Segment Title (required)
- Partial Transcript
- Keywords (semi-colon delimited)
- Subjects (semi-colon delimited)
- Segment Synopsis
- GPS Coordinates
- GPS Description
- Hyperlink
- Link Description

Indexing is a very subjective process and can vary greatly in breadth and depth. The Nunn Center has worked out several different levels of indexing, which correspond to the level of detail of our desired outcome. For specific guidance on how the Nunn Center uses each field, consult the Nunn Center’s guide, Indexing Interviews in OHMS: An Overview.

Although we have identified a function for each index field, specific policies regarding the particular use of these fields should be discussed by individual institutions.
**Time Stamp**

This is created as soon as you press the “Tag Now” button. However, you will need to adjust your time stamp so that the segment correlates to the appropriate time.

**Segment Title**

Segment titles are required. It is best to use descriptive titles, because these act as chapter titles for the interview.

**Partial Transcript**

The Nunn Center has found that including a brief partial transcript helps corroborate locations for the user. We tend to use the first spoken words of a segment (140 characters or so).

**Keywords**

This field also allows for multiple entries, separated by a semi-colon, and can be used in conjunction with a thesaurus to control or suggest terms. (See section 6.0 of this guide for more information.) The Nunn Center uses this field for local terms and tags, as well as for project specific vocabularies.

**Subjects**

This field allows for multiple entries, separated by a semi-colon, and can be used in conjunction with a thesaurus to control or suggest terms. (See section 6.0 of this guide for more information.) The Nunn Center uses this field for Library of Congress Subject Headings, or another controlled vocabulary.

**Segment Synopsis**

The segment synopsis is designed to contain a descriptive statement about a segment.

**GPS Coordinates**

When clicked, this field connects a user to a location on Google Maps. Of course, the institution can change the map resource if desired. Coordinates are entered in the format "XX.XXX, YY.YYY", where X is latitude (north or south) and Y is longitude (east or west). Only one set of coordinates is allowed per segment at this time. There must be a space following the comma.

**GPS Description**

This field identifies the location corresponding to the GPS coordinates.

**Hyperlink**

This field allows a hyperlink to be present, connecting the oral history segment to an external resource of any type. Only one hyperlink is allowed per segment at this time.

**Link Description**

This identifies or describes the hyperlink.
5.3 USING A THESAURUS IN THE INDEX MODULE METADATA FIELDS

The segment title, subjects and keywords fields can correlate to an uploaded thesaurus of terms. Each field can be assigned a different thesaurus for a particular interview. If a thesaurus has been uploaded and properly assigned in the metadata manager, terms will be suggested based on a partial keying of letters. For example, if the indexer begins typing “segrega” in the subjects field, all of the terms in the assigned thesaurus containing “segrega” will be suggested. The indexer selects the term desired and the term is dropped into the subjects field.

This feature requires a thesaurus to be uploaded first, and then assigned to the interview. (See section 6.0 of this guide for more information.)

5.4 IMPORTANT NOTES ABOUT INDEXING INTERVIEWS

Indexing is an incredibly subjective process and can be as comprehensive as you choose. For models on indexing and the Nunn Center’s approach, we recommend consulting the Nunn Center’s guide, Indexing Interviews in OHMS: An Overview.

Remember, you are creating time-dependent metadata and, therefore, the time code of the media being indexed must match the time code on the media that is accessed by the OHMS viewer.
6.0 THESAURUS MANAGER

OHMS allows you to upload thesauri for the indexing module. These thesauri contain the terms that will be autosuggested in the indexing fields. This helps to manage consistency in titles and descriptive terms. Thesauri can be uploaded for the following fields in the interview index:

- Title
- Subjects
- Keywords

These thesauri must be single-column Excel files saved in .CSV format.

In order to assign a thesaurus
to a field for a particular interview, access the metadata module for the interview. Within the metadata module, you are able to identify which thesaurus to use for titles, keywords and subjects for a particular interview. You are able to change the thesaurus mid-indexing if needed.

6.1 UPLOADING A THESAURUS

When uploading a thesaurus:

• Your .csv file must be a single column.
• If you are on a Mac and using Microsoft Excel, format the .csv as a Windows formatted .csv.
• Give your thesaurus a descriptive title and a version number so that you can upload a replacement file with more terms if needed.

7.0 TRANSCRIPT SYNC MODULE

The purpose of the transcript sync module is to connect a user's search within the text to corresponding moments in the online audio/video interview. In order to do this, OHMS has a user friendly, efficient method to drop time code into a transcript at 1-minute intervals. Currently, the transcript sync module is operational for audio only. If you are syncing a video, strip the audio from the video, upload the audio, and add the audio URL to the alt sync URL metadata field for that interview. This will point the transcript sync module to the audio version. This will only affect the transcript sync process, and will not be exported or associated with the interview on the user side of OHMS.

7.1 UPLOADING A TRANSCRIPT

In order for the “sync” function to be enabled, you must first upload a plain text version (.txt) of the interview transcript.

Once you have uploaded a transcript, you can begin the sync process. The “Sync” column of the interview manager will indicate an interview’s eligibility for the transcript sync process.

• “No T”: No transcript has been uploaded.
• “No AV”: No hyperlink is present in the media URL field of an interview’s metadata.
• “No AV/T”: No transcript has been uploaded and no hyperlink is present in the media URL field of an interview’s metadata.

• “Sync”: If this word is blue, an interview is ready to sync. If red, sync is ready for QC. Purple means Active QC. If green, the transcript has been synced successfully.

7.2 INITIATING TRANSCRIPT SYNCHRONIZATION

OHMS enables the placement of a time sync marker at each minute mark. When beginning a transcript sync, press “Sync” in the “Sync” column of the interview manager. This will open the transcript with a player in the top right corner. The OHMS transcript sync module only allows you to listen for a short period of time at the end of each minute of an interview. The default interval is 10 seconds. You can shorten or lengthen this lead time by changing the time interval next to the clock.

7.3 SYNCHRONIZING A TRANSCRIPT

1) The audio will begin upon opening the transcript sync module for an interview.

2) Click the forward button to be taken to the 00:00:50 mark on the first minute.

3) You will hear two chimes. Your first chime (lower pitched) is a warning, indicating that you have ten seconds to find the corresponding section of the transcript.

4) When you have found the text in the transcript that corresponds to what you hear, follow along.

5) At the minute mark, a second chime (higher pitched) will ring.
6) When the second chime rings, click on the appropriate word in the transcript that is heard during or closest to the second chime. This places a time code marker into the text (marked in green). If you succeed in placing the marker, you are automatically taken to the next interval (00:01:50). If you fail to place the marker in the allotted time, the same audio segment will replay after the post roll completes (default is 10 seconds). It will continue to repeat until the marker is placed.

7) You can adjust the sync point by rewinding back or forwarding to the appropriate segment. This is indicated by the clock and the minute-interval counter. Then you can correct your sync placement when the bell rings on the minute.

8) You must save your work prior to exiting or your sync data will be lost. There are save buttons at the top and bottom of the transcript.

If at any time during the syncing process you wish to correct or edit the transcript, you can click the "Edit Transcript" button at the top or bottom of the transcript. This will automatically save the points you have already synced, and take you to the editing module.
From here, you may make changes to the transcript, but be aware that this may change the placement of any sync points further along in the transcript.

7.4 TRANSCRIPT SYNCHRONIZATION HINTS AND REMINDERS

**Hint:** Using the control+f/command+f keyboard shortcut search for a word heard in the 10-second lead time quickly locates the Syncer in the transcript.

**Hint:** It is recommended that you do not upload the transcript with title or disclaimer pages prior to the beginning of the actual transcript, as this will interfere with your time-code syncing efforts.

**Reminder:** OHMS transcript sync only works on audio files at this time. (See section 7.0 of this guide for information on syncing a transcript for a video interview.)
8.0 NOTES

The notes function in OHMS can serve many purposes. These include flagging interviews with audio or video malfunctions, flagging interviews that need to be reviewed for possible restriction, or other issues that need to be addressed by your repository's administrator. Clicking "Notes" in the interview manager allows you to write a note explaining an issue.

Administrators will be automatically notified when a note is created. They can click on the notes column to read the note and mark it "Resolved" when the issue has been fixed. If a note is unresolved, the "Notes" indicator in the interview manager will be red. When resolved, the button will turn green.

9.0 WORKFLOW MANAGEMENT

The OHMS interview manager has a workflow management component to demonstrate the processing and quality control status as well as an overall status update. The Nunn Center uses this feature to communicate readiness to various participants in the OHMS workflow (indexers, syncers, archivists, digital library representatives). An indexer can use the workflow status to mark that an interview has been started but is not completed, indicate which phase of the process the interview is in, and indicate that the interview has been completed.
Processing

The status component of workflow management is handled individually through the metadata, indexing, and transcript syncing modules. Each page has a drop-down menu that allows you to choose from four options to set the status of each component.

Changing the status of these modules will automatically change the color of that module's indicator in the interview manager, giving you a visual indication of the status.

- "In Process": blue
- "Ready for QC": red
- "Active QC": purple
- "Complete": green

The indexing module also has an additional option: "Not Applicable", which will change the color of the word "Index" to gray in the interview manager.

Status

The status measurement is a way to communicate to other team members where the interview is in the overall OHMS process. Because we often have multiple interviews from...
different projects that are processed all at once, it is helpful to have an “at-a-glance” view of their statuses. Additionally, it is useful to have final confirmation that the OHMS export and the archival ingest of the XML file has been completed so that you can remove the interview from the OHMS application, in preparation for making it publicly accessible via the OHMS viewer. This process is automated based upon the status of each module within the interview manager. An interview automatically begins with a blue "In Process" status. When all status indicators for an interview are green (or gray, indicating that they are not applicable), the status indicator will automatically turn green indicating that the interview is "Complete".

9.1 PREVIEW YOUR WORK

As soon as an interview has been loaded into OHMS and has active links to a media source, a preview option will be available.

Clicking on the preview option activates the OHMS viewer where you can preview the interview index or transcript synchronization as the user would experience it. This function helps optimize the effectiveness of the quality control phase. This is a back-end viewer used for preview purposes, and is only accessible to those given permission to log into the OHMS application.

10.0 EXPORTING AN INTERVIEW

Exporting an XML

When you are satisfied with the interview’s level of processing, the XML file must be exported for use with the OHMS Viewer. Upon clicking the “XML” export link, an XML document will be created and downloaded with a filename derived from the “OHMS XML filename” field in the interview. If no filename is present in the field, it will default to the interview’s accession number. It is recommended that you change the filename to the interview’s unique identifier. Upon completion of your indexed or synchronized interview:
1) You will need to export the OHMS XML file either individually or using the batch export function.

2) Place the downloaded XML file in the “Cachefiles” directory in the OHMS Viewer installation.

3) Construct a hyperlink to the XML file to call up the OHMS viewer. Part 1 of the URL will always remain the same, linking to the location of the OHMS viewer. Part 2 will always be the corresponding filename for the OHMS XML file, placed in the cachefiles subdirectory of the OHMS viewer.

4) Place hyperlink in appropriate location of CMS.

_Exporting a CSV_

This functionality has been added with the purpose of transporting metadata from OHMS to a content management system. Although transcript and index data is included in the CSV export, the CSV export contains no OHMS functionality. It has been primarily added to include a convenient way of transporting metadata from OHMS to a CMS such as Omeka or CONTENTdm. The OHMS XML is still required for the OHMS Viewer to function. Transcript and index data are included for easy inclusion into the global search of a CMS but are not necessarily intended for public viewing. This CSV export differs from the initial metadata import, which requires the provided OHMS CSV import template. **The exported CSV is not designed to be re-imported.**

### 10.1 VIEWING THE INTERVIEW IN OHMS

Assuming you have successfully installed the OHMS viewer (https://github.com/uklibraries/ohms-viewer), it is very simple to make the interview accessible in the OHMS viewer.
11.0 INTERVIEW IMPORT

You can batch import interview metadata in order to initiate the creation of a transcript sync data or an interview index, or to re-import an OHMS XML file to update or edit the record (remember, a .csv file exported from OHMS cannot be re-imported). Initial import of metadata can be in the form of a .csv or XML file structured in accordance with the OHMS import templates provided.

Metadata Fields imported into OHMS as a CSV

At this phase in OHMS development, it is recommended that you utilize the .csv template for initial import of interview records. If you are on a Mac and using Microsoft Excel, format the .csv as a Windows formatted .csv. Do not use a .csv exported from OHMS as a template as it will not import correctly.

- Title (required)
- Accession #
- Interviewee (multiple values separated by a semicolon “;”)
- Interviewer (multiple values separated by a semicolon “;”)
- Date (interview date—must be structured as YYYY-MM-DD or MM/DD/YYYY)
- Collection ID
- Collection Title
- Series ID
- Series Title
• Summary
• Subjects (multiple values separated by a semicolon “;”)
• Duration
• Media URL
• Video ID
• Usage
• Rights
• Funding
• Format
• Language
• Type
• Transcript Sync Data
• CMS Record ID
• Date Non Preferred Format
• XML Filename
• Alt Sync URL
• Use Restrictions
• Keywords (multiple values separated by a semicolon “;”)

It is important that your Date field retains the structure YYYY-MM-DD or MM/DD/YYYY or else the import will fail.

Metadata Fields imported into OHMS as XML

When importing interviews into OHMS as XML, it is best to utilize the OHMS XML as a template. See the sample OHMS import XML (individual and batch import).
12.0 REMINDERS/KNOWN LIMITATIONS/BUGS

• Internet Explorer is not recommended for the OHMS Application at this time. This does not involve the user interface of the OHMS Viewer, just the OHMS Application. Compatibility is best when using Firefox, Chrome, or Safari.

• The OHMS transcript sync and indexing modules create time-dependent metadata. You should not make any changes to the media time-code after processing in OHMS. Any edits that change the duration of the interview will alter the location of the sync markers and index points, and thus disrupt the correspondence between markers and content.

• The OHMS transcript sync module does not currently work with video. If you want to sync the transcript for a video interview, utilize the media URL field for the true URL linking to the video file. However, use the alt sync URL field to link to an audio version of the video you are delivering. The time codes must match in order to be accurate. The OHMS Viewer will then match the sync data created with the audio sync with the video interview for the user. (See section 7.0 of this guide for more information).