

A Guide to OHMS

Expanding Metadata and Indexing Capabilities for Oral
History Projects

Sasha Gaylie

University of British Columbia School of Information

Table of Contents

Introduction.....	2
1. The Oral History Metadata Synchronizer	2
1.1 Why OHMS?	2
1.2 Beyond Oral History?.....	5
2. On Metadata	5
2.1 Oral History Metadata: An Overview	5
2.1.1 The Oral History Association Metadata Task Force	6
2.1.2 Further reading	8
2.2 Interview-Level Element Set	9
3. The OHMS Application: A Practical Guide	10
3.1 Metadata Editor	11
3.2 Transcript Synchronization.....	14
3.2.1 Importing transcripts: Formatting considerations.....	15
3.2.2 The Sync module	16
3.3 Indexing	17
3.3.1 Segment description.....	18
3.3.2 Expanding functionality: GPS coordinates, hyperlinks, and images	19
3.4 Bilingual Indexes and Transcripts.....	21
3.5 Web Viewer.....	23
Conclusion	23
Appendix A	24
Appendix B	26
Bibliography.....	27
Resources	29

Introduction

The Oral History Metadata Synchronizer [OHMS],¹ developed by the Louie B. Nunn Center for Oral History at the University of Kentucky Libraries, is a web-based, open source system for: (1) creating oral history metadata records and expanding indexing capabilities for audiovisual resources, and (2) facilitating researcher/user access to media resources through a viewer interface. This guide provides context for the OHMS system, a discussion of current best practices in oral history metadata, and a full walkthrough of the OHMS application. Throughout this guide, I will be using examples and images from an interview I conducted and indexed for an oral history project on the year 2020, hosted by the University of British Columbia School of Information.²

1. The Oral History Metadata Synchronizer

1.1 Why OHMS?

As mentioned in the introduction, OHMS is a **set** of tools—an application for creating metadata records/indexes (hereafter referred to as the “OHMS application”), and a viewer interface (the “web viewer”)—that seeks to improve user access to oral history interview data. It is important to note that OHMS is not a repository and does not store media files on its own. OHMS works with your chosen content management system to improve discoverability of audiovisual resources and expand access to the data within them.

As oral history practitioners at the Louie B. Nunn Center for Oral History, the creators of OHMS identified this need to improve researcher access to oral histories, and in a way that would still fit in with most institutional budgets.³ Imagine for a moment that a research team at a university has managed to record an astounding 10-hour interview with a former university president for a centennial project, and would like to make this data available to potential researchers through their institutional repository. Ignoring the fact that a project of this scale would likely also include interviews with several other narrators, how many hours would need to be budgeted for transcription in order to make all the contents of one 10-hour interview text-

¹ “OHMS Oral History Metadata Synchronizer,” Louie B. Nunn Center for Oral History. Accessed November 20, 2021. <https://www.oralhistoryonline.org/>.

² “The 2020 (Re)Collection,” University of British Columbia, April 15, 2021, <https://blogs.ubc.ca/2020recollection/>.

³ Douglas A. Boyd, “The Oral History Review OHMS: Enhancing Access to Oral History for Free,” *The Oral History Review* 40, no. 1 (2013): 106. <http://dx.doi.org/10.1093/ohr/oht031>.

searchable? How many hours towards manual correction of transcripts run through automated transcription software?

As in this example, the time-intensive nature of transcription, combined with unwieldy media duration times, often results in uncertain futures for these types of resources. Hours of data-rich footage can sit in folders for years, awaiting a budget that allows for thorough indexing. Or, perhaps even more tragic, the resources do make it to an end-user interface but only with bare-bones description, and researchers are unable or unwilling to manually sift through the content to find relevant data within.

The OHMS application and viewer present a few solutions to these common collections problems: transcript synchronization, interview-level metadata input and segment-level indexing. The indexing feature of OHMS, building on traditional analogue indexes that list segment timecodes and descriptions in a text document, allows for timecodes of the actual media file to be tagged with subject headings and/or keywords, described in synopses, and enriched with GPS coordinates or hyperlinks. Figure 1 below shows a snippet of an interview index in the OHMS viewer interface, while Figure 2 on the next page displays some of the indexing options within a segment. Segment length is completely up to the indexer and the extent to which each segment is described can vary according to institutional needs.

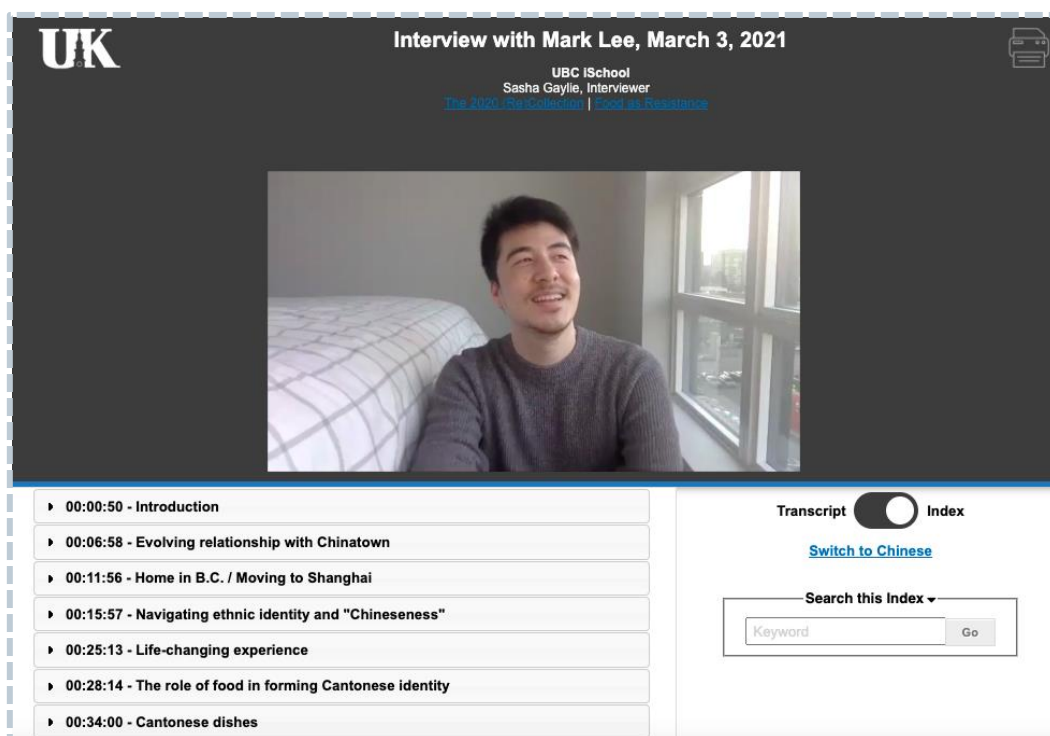


Figure 1: An index as it appears in the OHMS web viewer.

The option to index media segments alone opens up many doors for institutions hoping to make data more readily available. Kristopher Turner describes the efforts of the University of Wisconsin Law School Library, which sought to build a new institutional repository to house its department's scholarship in 2015.⁴ Oral histories that had been collected over several decades included interviews in excess of 20 hours. Turner notes that transcription for these longer

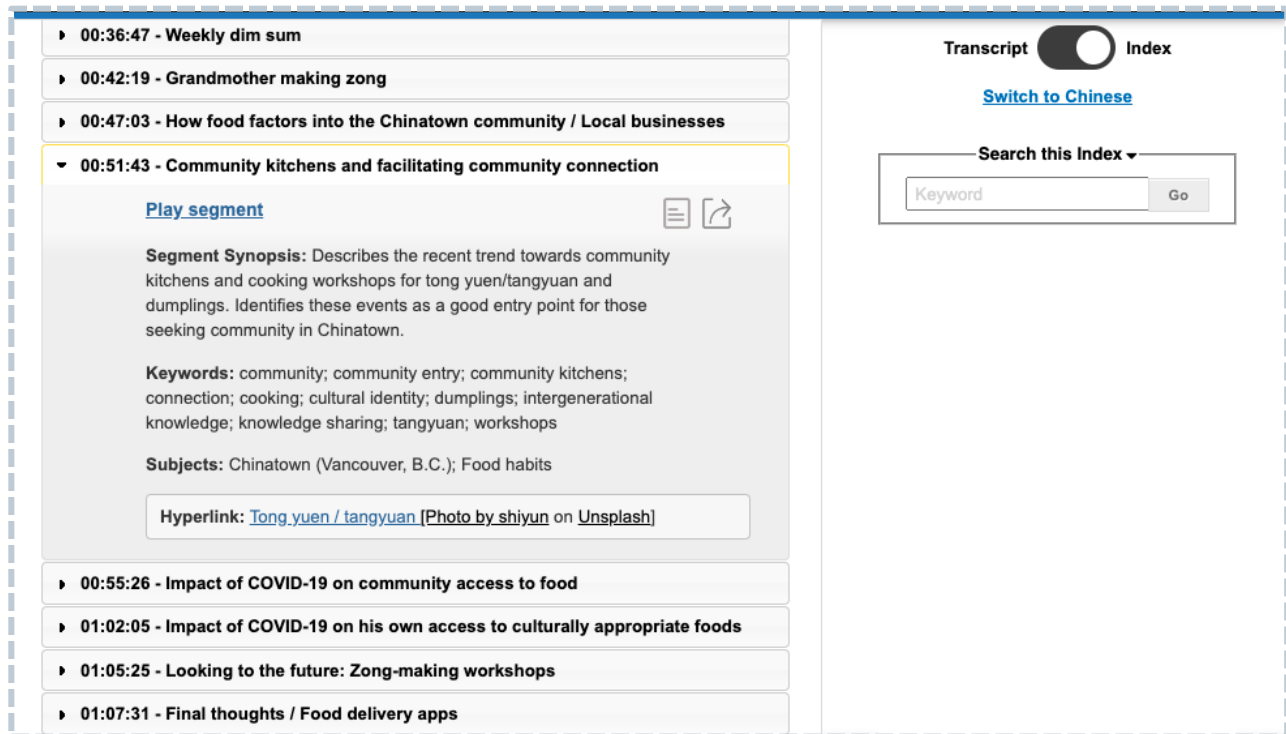


Figure 2: Expanded view of an index segment in the web viewer.

interviews alone would have been “prohibitively expensive and time-consuming.”⁵ Opting instead to index via OHMS, and through employment of library school students, the UW Law Library was able to prepare nearly 25 interviews for final review within just six months.

Doug Boyd, professor at the University of Kentucky and leader of the team that developed OHMS, notes that in 2014, the university's Nunn Center for Oral History indexed over 900 hours of interview footage on OHMS for only one-tenth of the cost of transcription.⁶

However, the potential benefits OHMS can bring to a project go far beyond the financial. As with any archive, discoverability and access are typically at the forefront of an oral history

⁴ Kristopher Turner, “Creating history: a case study in making oral histories more accessible in the digital age,” *Digital Library Perspectives* 33, no. 1 (2017): 50. <https://doi.org/10.1108/DLP-06-2016-0016>.

⁵ Ibid., 58.

⁶ Douglas A. Boyd, Janice W. Fernheimer and Rachel Dixon, “Indexing as Engaging Oral History Research: Using OHMS to ‘Compose History’ in the Writing Classroom,” *The Oral History Review* 42, no. 2 (2015): 354. <https://doi.org/10.1093/ohr/ohv053>.

project. The emergence of born-digital assets and subsequent development of tools to support them means that new ways of facilitating access are constantly being created. As one such tool, the OHMS system invites us to critically examine whose voices we are magnifying with our oral history collections, who we want these voices to reach, and how we will enable that process.

1.2 Beyond Oral History?

Despite its name, it is not difficult to imagine the scope of possibilities for applications of OHMS in an institutional repository, or digital collections in general. Conference presentations, recordings of special events, and documentary films are just a few examples of moving image materials that can be found among online holdings.

Webinars present a particularly salient example, having become an increasingly prevalent part of our reality in the wake of the COVID-19 pandemic. Many of us have encountered those eight-hour-long recordings on YouTube of the second day of a conference, and struggled to scrub to a specific presentation with a sensitive mouse on the progress bar. A light sweep of indexing on OHMS would make individual presentations much more accessible and could even display entire abstracts from conference proceedings through segments on the web viewer.

Breaden outlines a unique use case at Duke University, wherein OHMS was used to index silent films shot by filmmaker Herbert Lee Waters.⁷ As there are no audio tracks associated with the reels, and therefore nothing to transcribe, OHMS offered an opportunity to more comprehensively describe the many different segments within one film.

While this particular guide uses an oral history interview to demonstrate the many features of OHMS, it should not dismiss the possibilities that the application may bring to the other types of audiovisual collections housed in a repository.

2. On Metadata

2.1 Oral History Metadata: An Overview

You will notice much of this guide focuses on metadata—what metadata looks like in the context of oral history, recent efforts to articulate best practices, and what tools exist to assist metadata collection. This is because so much of what OHMS can offer is dependent on the degree to which resources are catalogued and indexed. In this next section, we will take a

⁷ Craig Breaden, "Words Painting Pictures: Indexing the H. Lee Waters Project using OHMS," *Provenance, Journal of the Society of Georgia Archivists* 34, no. 1 (2016). <https://digitalcommons.kennesaw.edu/provenance/vol34/iss1/15/>.

break from discussing OHMS to explore just why metadata is so critical when it comes to oral history dissemination, and how one can develop a comprehensive metadata plan that best suits the goals of a project or collection.

2.1.1 The Oral History Association Metadata Task Force

The Oral History Association Metadata Task Force [OHAMTF, not to be confused with OHMS], first established in 2014, was created with the goal of evaluating current oral history metadata practices in various types of institutions and to determine how practitioners were facilitating discoverability of and access to their collections.⁸ Due to the variability of goals and concerns across institutions, it became apparent that any recommendations for best practices following the OHAMTF multi-phase survey would have to be flexible. Task force members note,

The biggest takeaway of our survey is no surprise to those working with oral history: there is no “one size fits all” solution for metadata capture and creation. Communities of practice need an approach that balances the special nature of oral history with their local missions, goals, and resources, and one that acknowledges oral history interviews may be only a fraction of an overall repository’s or organization’s collection.⁹

In light of this discovery, the OHAMTF worked with the Metropolitan New York Library Council [METRO] to develop a metadata toolkit on their Archipelago software that would enable oral history practitioners—independent interviewers, museum curators, librarians, archivists, and so on—to evaluate metadata choices for their projects.¹⁰ Divided into sections for the typical four stages of an oral history project (Plan, Interview, Process, Disseminate), users can check off prospective elements from a comprehensive list and download their results as a PDF or CSV file.

The OHAMTF describes their element list as “schema-agnostic,”¹¹ and note that it is meant to help practitioners critically evaluate their metadata decisions through every stage of collection, rather than represent an actual input standard. The exported CSV file does, however, provide a column for mapping the checked elements to an existing schema or standard if you are determining how to integrate them into the structure in place at your institution. For those who might be unsure of where to start, the Metadata Task Force members have created “Practitioner Profiles” that introduce a few oral history practitioners, an overview of their institution and

⁸ Kata et al., 2020, “Oral History Metadata and Description: A Survey of Practices,” Metadata Task Force, Oral History Association. https://www.oralhistory.org/wp-content/uploads/2021/01/OHA-MTF-White-Paper_2020.pdf.

⁹ Ibid., 3.

¹⁰ “Metadata for Oral Histories: An Interactive Toolkit,” Oral History Association, accessed March 29, 2021, <https://oha.archipelago.nyc/>.

¹¹ Kata et al., 2020, 15.

collections, their goals with metadata, and their subsequent ideal element sets. Visitors who have tested the toolkit are encouraged to submit their own profile and element set for publication to

The screenshot shows a web interface titled "Element List for Oral Histories". At the top, there is a progress bar with eight numbered steps: 1 Welcome, 2 Plan, 3 Interview (highlighted in blue), 4 Process, 5 Disseminate, 6 Your info, 7 Preview, and 8 Complete. Below the progress bar, it says "3 of 8". The main heading is "Metadata concerning the actual interview". Below this heading is a list of eight items, each with a checkbox:

- ☐ Additional Information about interview recording room
- ☐ Age of interviewee at time of interview
- ☐ Age of interviewer at time of interview
- ☐ Community represented in the interview
- ☐ Compensation details for anyone involved in interview
For example: interviewees receiving a .wav file of their interview recording or an intern earning class credit for her role as interviewer
- ☐ Copyright holder contact info
Could be subfields: Phone Number, E-mail Address, Institutional Affiliation, Mailing Address & Social Media Handle
- ☐ Copyright holder name
Could be subfields: Full Name, Alternate Name
- ☐ Duration (recording)

Figure 3: Interface of the Element Set Builder Tool developed by OHAMTF and METRO.

guide future users.

Let's use my experience with The 2020 (Re)Collection project to consider the evaluative nature of this toolkit. It has been widely documented that since the start of the pandemic, COVID-19 has disproportionately impacted BIPOC (Black, Indigenous, People of Colour) communities.¹² A narrator speaking to their experience of COVID-19 for this project would certainly be affected by these systemic factors, but may not be aware or able to name it outright in the interview. If the narrator's race does not otherwise come up in discussion, it would not appear in the resulting transcription. Depending on how thoroughly the interviewer consulted the narrator prior to collection, the narrator's racial identity may be unspecified or generalized, which would in turn make it difficult (and unethical) for a cataloguer to assign race-related keywords or subject headings. A researcher later looking for firsthand accounts of the pandemic from people of a particular racial background may then struggle to even discover this interview data.

This may seem an obvious example (and perhaps an indication of negligence on the part of the interviewer), but if we apply this thinking to some of the more granular elements in the list, we can see how the OHAMTF toolkit might inform project planning. Religion of interviewee, native language of interviewer, and gender identity of interviewee are just a few examples of attributes listed in the "Plan" section of the element list. The sensitive nature of this data would

¹² See <https://www.cdc.gov/coronavirus/2019-ncov/covid-data/investigations-discovery/hospitalization-death-by-race-ethnicity.html> for updated data visualizations.

require thorough discussion with and permissions from the interviewee. Without consideration of these elements prior to interview collection, an opportunity to include more descriptive and potentially valuable metadata is lost.

2.1.2 Further reading

The difficulty in creating metadata records for oral histories is often in the fact that one interview “package” can contain many parts—a media file, a transcript, consent forms, controlled vocabularies, supplementary materials, and so on. This is typically compounded by the issue of legacy metadata—metadata gaps or inconsistencies in already processed interviews. Metadata systems should be able to uniquely identify single assets, draw relationships between assets associated with one interview and, at an even higher level, draw relationships between interviews within a collection.

Metadata standards for oral history interviews are often dependent on the current practices of their home institutions. Elinor Mazé gives an extensive overview of the forms this can take,¹³ while Steven Sielaff provides an example for operating within a specific content management system.¹⁴ For many institutions, Dublin Core may prove to be the established option for digital resources. Nancy Mackay offers a concise explanation of how to map parts of an oral history interview summary to Dublin Core elements.¹⁵

However, the metadata that we choose to document cannot and should not be determined solely by the structure of our institution’s content management system. Often called the “People’s History”, oral history continues to provide opportunities for us to examine how we are engaging with data, with stories, and with the people around us.¹⁶ They occupy a unique cultural space in our digital collections that encourages us to think critically about not just users, but contributors. Jaycie Vos provides an example of what this assessment can look like by covering the work done for the New Roots digital archive, which is dedicated to the oral histories of Latin American migrants.¹⁷ Vos notes that the New Roots team chose to hide self-

¹³ Elinor A. Mazé, “Metadata: Best Practices for Oral History Access and Preservation,” *Oral History in the Digital Age*, Institute of Museum and Library Services, accessed March 30, 2021, <http://ohda.matrix.msu.edu/2012/06/metadata/>.

¹⁴ Steven Sielaff, “Metadata at BUIOH: A Case Study,” *Oral History in the Digital Age*, Institute of Museum and Library Services, accessed March 30, 2021, <http://ohda.matrix.msu.edu/2015/10/metadata-at-buioh-a-case-study/>.

¹⁵ Nancy MacKay, “Appendix A,” in *Curating Oral Histories: From Interview to Archive* (Walnut Creek, CA: Routledge, 2015), 161-164.

¹⁶ MacKay, “Curating Oral Histories in the 21st Century,” in *Curating Oral Histories: From Interview to Archive* (Walnut Creek, CA: Routledge, 2015), 28.

¹⁷ Jaycie Vos, “New Roots: An Oral History Metadata Case Study at the University of North Carolina at Chapel Hill,” *Oral History in the Digital Age*, Institute of Museum and Library Services, accessed March 30, 2021, <http://ohda.matrix.msu.edu/2015/10/new-roots-an-oral-history-metadata-case-study-at-the-university-of-north-carolina-at-chapel-hill/>.

explanatory metadata elements from public view to declutter the interface for users, while also introducing “new geographic metadata, topic/theme metadata, and bilingual metadata” to “reach broad audiences outside of the academy, particularly Latino communities throughout North and South America.”¹⁸

2.2 Interview-Level Element Set

The element set in [Appendix A](#) is a sample that includes all the fields offered in the OHMS Metadata Editor, as well as a few elements that would be generally recommended for describing an oral history resource, whether that be the interview file itself or derivative materials (transcript, index, etc.). Though this sample covers the fields most likely to be associated with a digital resource and the baseline for dissemination, consider using the OHAMTF toolkit mentioned in [section 2.1.1](#) to determine additions that would best serve your user communities and facilitate long-term preservation.

¹⁸ Ibid.

3. The OHMS Application: A Practical Guide

This section will provide a walkthrough of the process of indexing and editing a metadata record for a media file in OHMS. Please note that there is a detailed user guide linked on the OHMS website that outlines workflow and technical considerations, and provides a thorough overview of import options for existing metadata records (i.e., if you are indexing legacy resources).¹⁹

The walkthrough offered here is meant to provide a plain language demonstration of the OHMS application, from start to finish, to empower potential users at any technical comfort level—students, student workers, instructors, heritage project organizers, library staff, etc.—to describe and process audiovisual materials. Anyone with further questions is encouraged to consult the latest version of the OHMS User Guide.

NOTE: In order to gain access to the application outlined in the following pages, you must apply for a free OHMS account and become the OHMS repository administrator for your institution. If there is already an OHMS repository administrator at your institution, they are able to grant access directly and should be contacted first before trying to apply through the Nunn Center at the University of Kentucky.

Administrators can assign different access levels to new users, which may be preferable for instructors looking to incorporate OHMS into course work and who want to invite students to contribute.²⁰

Upon login to the OHMS application, the first screen you will be greeted with is the *Interviews* tab, or homepage (see Figure 4). This page contains a selection of sample entries provided by the Nunn Center, or any entries for resources that have already been processed through OHMS at your institution. From this page, users can navigate to the Metadata Editor, Index module, transcript upload window, and Sync module by using the corresponding links. This is also where you can export completed entries as CSV or XML files, and any processing

¹⁹ See *Important Links* in the sidebar at <https://www.oralhistoryonline.org/>.

²⁰ For some examples of OHMS as a pedagogical tool, particularly for the digital humanities, see: Dan Royles, "The Oral History Review Teaching Digital Humanities with Oral History: The Staring Out to Sea Oral History Project and OHMS in the DH Classroom," *The Oral History Review* 43, no. 2 (2016): 408–420. Janice W. Fernheimer et al., "The Oral History Review Sustainable Stewardship: A Collaborative Model for Engaged Oral History Pedagogy, Community Partnership, and Archival Growth," *The Oral History Review* 45, no. 2 (2018): 321–341.

notes for an entry can also be controlled from this page. At any point, if you would like to see how your resource will look in the final web viewer interface, click the [\[Preview\]](#) link under an entry's title to open its preview in a new tab.

As mentioned at the beginning of this guide, **OHMS is not a repository** for media. Neither video nor audio files can be stored in the OHMS application. To get started using the Metadata Editor or any other parts of the OHMS system, a media file must first be uploaded to an external content host. At time of writing, OHMS supports files on YouTube, Kaltura, Avalon, Brightcove, SoundCloud, Aviary, and Vimeo. Direct links to media files are also supported (i.e., downloading, as opposed to streaming content), though the OHMS User Guide offers a cautionary warning to optimize files for this option, as playback performance may be affected. The interview that I will reference throughout this walkthrough is a video file (MP4) that has been uploaded to Vimeo. It has a runtime of 1:09:45.

OHMS | ORAL HISTORY METADATA SYNCHRONIZER

Change Password

INTERVIEWS PROJECTS THESAURUS IMPORT USERS REPOSITORY MANAGER

Interviews

SEARCH: [] SUBMIT CLEAR NEW

Filter by Project -- Select Batch Action --

<input type="checkbox"/>	Title Accession Number	Collection ID	Series ID	Metadata	Index	Transcript	Sync	Notes	Status	Export
<input type="checkbox"/>	Interview with Mark Lee, March 3, 2021 [Preview]			Metadata	Index	Re-Upload	Sync	Notes	In Process	XML CSV
<input type="checkbox"/>	SAMPLE 005: Interview with Fred Noe, November 14, 2013 2013oh464 bik 025 [Preview]	ohbik	bik004	Metadata	Index	Re-Upload	Sync	Notes		XML CSV
<input type="checkbox"/>	SAMPLE 003: Interview with Beulah Collins, August 1, 1983 2014OH158 GN 010 [Preview]	OHGN	GN001	Metadata	Index	Re-Upload	Sync	Notes		XML CSV
<input type="checkbox"/>	SAMPLE 006: Interview with Raul Lagos Espinosa, November 20, 2009 2011OH262 HHK 008 [Preview]	OHCHK	HHK001	Metadata	Index	Re-Upload	Sync	Notes		XML CSV

Figure 4: The *Interviews* tab serves as the OHMS application homepage.

3.1 Metadata Editor

To begin creating an entry for a media resource, click the blue *New* button at the top right of the *Interviews* homepage. This will open a new form in the Metadata Editor, where you can fill out any interview-level metadata associated with your resource. To successfully create

and work with a record in OHMS, the minimum fields that must be populated are *Title*, *Media Format*, *Media Host*, and “media connection”²¹.

In Figure 5, we can see some of the administrative, descriptive and technical metadata fields that I have entered for my interview with narrator Mark Lee. Most of the metadata in this form will not be populated automatically—every field except the two in the **Transcript** section will need to be entered manually. Project managers should ensure that a schema is in place for cataloguers to follow. As covered in [section 2.2](#), this schema should indicate which fields are required to maintain consistency across projects and collections, and should also state formatting specifics (e.g., first/last order for names, preferred date format, criteria for plain text summaries, etc.). It is important to note that not all fields in the Editor will be made visible to users via the web viewer, and are instead included to easily transport metadata for a resource in a single file.

The screenshot shows the 'Metadata Editor' interface. At the top, there's a status dropdown set to 'In Process', and two buttons: 'SAVE' and 'RETURN TO INTERVIEWS'. A note indicates that an asterisk (*) denotes a required field.

Admin and Description

- Repository: UBC iSchool
- Title: Interview with Mark Lee, March 3, 2021
- Accession Number: (empty)
- Interviewee: Mark Lee
- Interviewer: Sasha Gaylie
- Interview Date: Year 2021, Month 03, Day 03
- Date (non-preferred format): (empty)
- Collection ID: (empty)
- Collection (Name): The 2020 (Re)Collection
- Collection Link: https://blogs.ubc.ca/2020recollection/
- Series ID: (empty)
- Series (Name): Food as Resistance
- Series Link: https://blogs.ubc.ca/2020recollection/food-as-resistance/

Media and Technical

- Media Format: Video (selected)
- Media Host: Vimeo
- Embed Code: <iframe src="https://player.vimeo.com/video/522074551" width="640" height="360" frameborder="0" allow="autoplay; fullscreen; picture-in-picture" allowfullscreen></iframe>
- Duration: 1:09:45
- Media Filename: gaylie_lee_dis_20210303.mp4
- Format: .mp4
- Type: Oral History

Rights and Usage

- Rights Statement: Rights holder has given consent to the University of British Columbia to use any part of the interview (sound recordings, images and supplementary materials) for The 2020 (Re)Collection project and indexing in OHMS.

Figure 5: Top half of the form in the Metadata Editor.

²¹ This field, or series of fields, will appear under the *Media Host* dropdown list. It will look different depending on your chosen host. It may require a direct link, a streaming URL, an embed code, an account ID, a player ID, etc. For my interview, I simply needed to copy and paste the iFrame embed code from Vimeo’s Video Manager into the Metadata Editor.

Admin and Description. As my interview is part of an exhibit hosted through WordPress (UBC Blogs) and has not been accessioned by a library, it does not have an accession number. To connect my interview with its parent collection and other interviews for the project, I have instead included URLs to the exhibit site (*Collection Link*) and interview series (*Series Link*). These are displayed as hyperlinks directly above the media player in the web viewer (see Appendix B, 1). A larger text box is provided for the *Summary* element, which can include a plain text description of the interview’s contents or a curatorial statement about the overall project. Under this field are the *Keywords* and *Subject* elements.

Along with *Title*, the *Keywords* and *Subject* elements can have thesauri or controlled vocabularies assigned to them. This may be useful if you have project-specific terminology with variable spellings and want to be consistent across items. The OHMS system automatically connects the linked data service for the Library of Congress Subject Headings (LCSH) and AFS Ethnographic Thesaurus, but users may upload their own vocabularies by clicking the *Thesaurus* tab in the upper navigation bar and following the instructions on that page. The thesauri can then be assigned to the *Title*, *Keywords*, or *Subject* fields from the **Thesaurus** section of the Metadata Editor. Any thesaurus assignments will also apply to the segment-level metadata in the index, which I will cover in the next section. I chose to assign LCSH to my subject terms for this interview, and then addressed any potential gaps with uncontrolled keywords.

Figure 6: Keywords and subject headings delimited by semi-colons are shown in separate text boxes after saving and refreshing the Metadata Editor.

Clicking the red plus sign to the right of a field will add an additional text box to allow for multiple value entries to an element, though the OHMS Editor will automatically split values separated with semi-colons if entered into a single text box (see Figure 6). This is particularly useful for the *Keywords* or *Subject* elements, which tend to have more values for input than

others. After saving, a refresh of the Metadata Editor page will show the values in separate text boxes.

Note that this will not work with commas, colons, pipes (vertical bars), or any other delimiters. To ensure the values are read and separated properly, do not put a space in front of or after the semi-colon. Upon export to a CSV file, multiple values for an element will be shown in a single cell with the semi-colon delimiters.

Media and Technical. This section is how the OHMS application connects to your media file. Follow the instructions for your particular media host. For Vimeo, it is as simple as copying and pasting the embed code for your uploaded video into the *Embed Code* field of the Metadata Editor.

Rights and Usage. This section contains fields for rights and usage statements, or any acknowledgements. In the case of oral history resources, it is important to clarify whether rights have been transferred from the copyright holder for preservation and/or dissemination by the institution, and to indicate whether they have placed restrictions on usage. These statements, if filled out in the Editor, are visible in the footer of the web viewer interface (see Appendix B, 8).

Miscellaneous. The value entered for *XML Filename* will be the filename for any exported XML files after metadata creation, indexing and/or transcript synchronization has been completed. It is critical that this field be filled out with a unique identifier, especially if you are batch exporting interviews, as this filename will be the way that the web viewer later locates the resource, and will also determine the latter half of the resource's URL. You may use the *User Notes* field to communicate any messages to users directly within the web viewer—it will display as text under the media player.

You may have noticed some omissions in this coverage of the Metadata Editor. The **Transcript** and **Language** sections of the form are not required fields but do carry special functions, and will be covered in sections [3.2](#) and [3.4](#), respectively.

3.2 Transcript Synchronization

NOTE: To minimize issues with uploading transcripts to OHMS, formatting considerations should be very clear prior to even starting transcription. Special attention should be paid to timestamp requirements.

A major feature of the OHMS system is transcript synchronization. In the web viewer, a synchronized transcript enables users to click on a linked timecode that would then force the media player to skip to that section of the interview. Because the web viewer interface also includes a search bar to query text in either the interview's transcript or index, users interested in a particular subject can search for a term, find its location in the transcript, and click the

corresponding timecode to jump to that part of the audio or video. Transcripts are uploaded by clicking on the *Upload* link in the “Transcript” column on the *Interviews* homepage.

3.2.1 Importing transcripts: Formatting considerations

When I first transcribed my interview for The 2020 (Re)Collection project, before knowing that I would be utilizing the transcript synchronization feature in OHMS, I followed the same guidelines given to all project contributors. The guidelines were to include approximate timestamps for around two-minute increments, and I put each of these codes on their own line in the transcript document.

After I decided to synchronize my transcript and the time to upload the document to OHMS came, I saw that I could indicate which timecode interval best corresponded with my transcript and I selected “2 minutes” from the dropdown list. I was surprised to find that OHMS was unable to read my transcript and the upload subsequently failed. I soon realized that OHMS required timecodes that were not only exactly at the interval mark—for 2 minutes, this means at precisely 00:02:00, 00:04:00, and so on—but that were also placed **in-line** with the script text. I went through my transcript a second time and placed the timecodes in square brackets wherever they occurred in the dialogue, even if this meant plopping them in the middle of a sentence. The final rendering in the OHMS viewer will take these timecodes out of the text and display them as links to the left of the transcript.

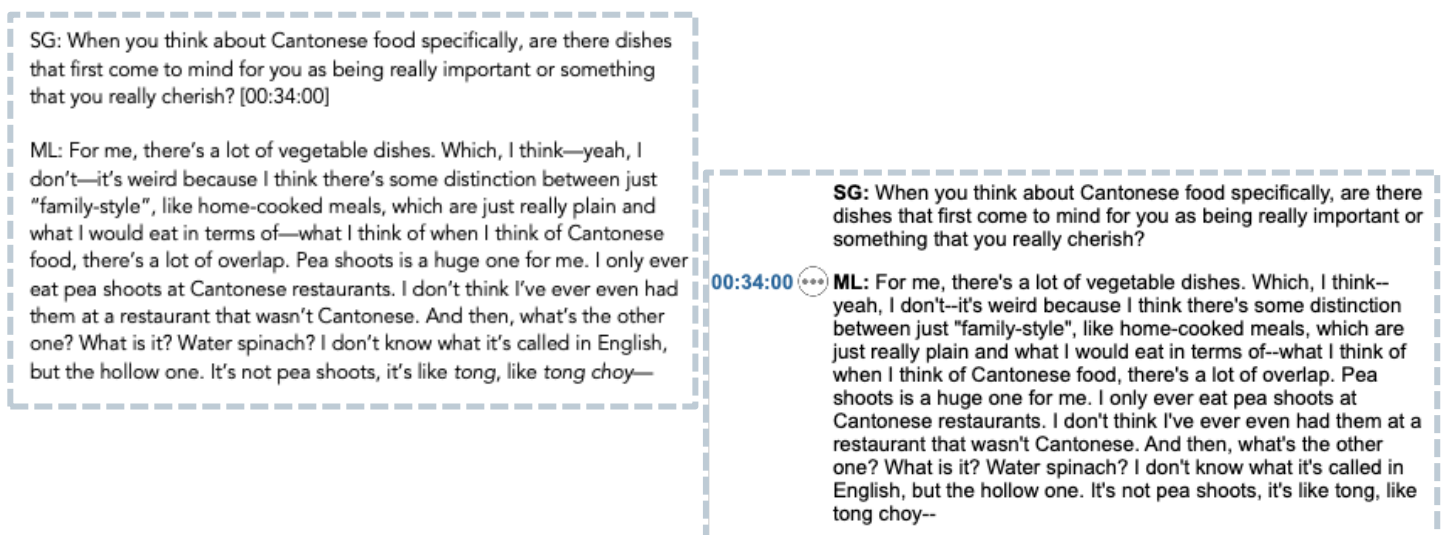


Figure 7: (L) Excerpt of transcript in MS Word. (R) Same excerpt displayed in OHMS web viewer.

Other important formatting considerations that I encountered:

- ⇒ Double dashes that had been Autoformatted to em dashes in Microsoft Word reverted to double dashes in OHMS.
- ⇒ My narrator used some Cantonese words during the interview and, as per the style guide for our project, I italicized any transliterations. These italics did not carry over to the OHMS rendering.
- ⇒ I did not put special formatting on names or initials for speakers; the OHMS viewer will automatically identify and bold these roles after upload. OHMS recommends a colon after speaker identification and a single line break between speakers.²²
- ⇒ Any page numbers, headers or footers should be removed prior to upload. My original transcript had a descriptive header containing interview metadata, but this both confused the OHMS system and resulted in some unattractive formatting in the viewer. My OHMS-compatible version of the transcript simply starts with the first speaker and the first line of dialogue.

Once you have successfully uploaded your transcript, it is recommended that you go through and confirm the timecodes accurately match the text. In the Metadata Editor, you should see that the *Transcript Sync Data* field has been automatically filled by OHMS and requires no further action.

3.2.2 The Sync module

As an alternative to timecoding your transcript manually, there is a Sync module offered in the OHMS application that will allow indexers to listen for the word that occurs at the chosen interval and simply click on it to insert a timecode. Full transcripts, without existing timecodes, must still be uploaded as a document prior to using the Sync module. A series of chimes acts as a warning that an interval is approaching—the first chime occurs at a specified lead time (ten seconds by default) so that the indexer can quickly search for upcoming text in the transcript, and then a second chime indicates the exact time interval where the indexer should click on the corresponding word. The “forward” button will then skip the media ahead to the next interval.

It is worth noting that because this form of synchronization is done directly in the OHMS application, it is liable to any browser or internet disruptions. If the browser’s refresh button or hotkeys are accidentally pressed, any unsaved timecodes are lost. If there is an internet connectivity issue and indexers are unable to save progress, these timecodes will also be lost.

²² OHMS User Guide, last updated October 1, 2020, 25. https://www.oralhistoryonline.org/wp-content/uploads/2020/11/@OHMS_user_guide_master_v3-8-3.pdf

The ease of click-inserted timecodes should be weighed against the hazards of operating within an internet browser. If you decide to use the Sync module, ensure that you save your progress frequently.

3.3 Indexing

It should be clear throughout this guide that the Index module, which offers the ability to describe media at the segment level, is where the OHMS system really shines. The module can be accessed via the *Index* link on the *Interviews* homepage. For projects that do not have time or resources for full transcription, segment indexing provides alternative means for subject access. Clips of a media file can be assigned titles, synopses, subject headings, keywords, and even links to external, supplementary material. Like a table of contents, any segments created for a resource will be displayed on the Index page of the web viewer as a list of folded headers. Each header contains the timecode for the start of the segment and the segment title. Only one segment can be “unfolded” at a time to display its associated metadata and/or links (Figures 2 and 8 show examples of unfolded segments).

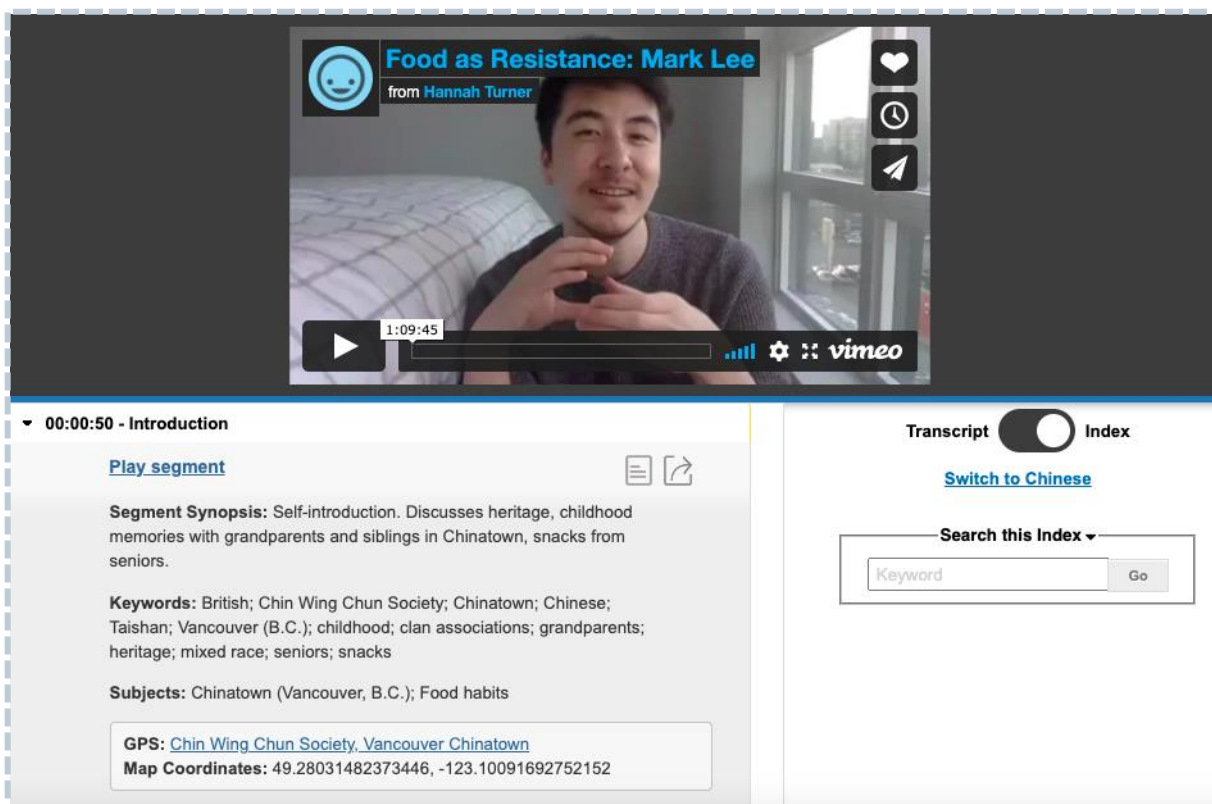


Figure 8: Full view of "Introduction" segment in the web viewer.

3.3.1 Segment description

Part of what makes OHMS such a powerful tool for access is that the Index module is so simple to use. Once you have opened the module, your media player will be displayed on the righthand side of the screen, with a blue button under it saying, “Tag Now”. Simply play your media file to the desired start time for a segment, then click the tag button. You have just created a segment! A form with several fields for text input will open up. If the timing of the segment is slightly off, there are some media player navigation buttons under the *Timestamp* field that can be used to adjust to a particular moment in the clip, and then an “Update Time” button to confirm the new timecode.

Title. This is the only field that requires a value; everything else can be indexed to a project’s desired level of completion. I opted for longer, descriptive phrases for my own segment titles to ensure they painted a clear picture of what a given segment might be about. Examples of my segment titles include “Evolving relationship with Chinatown” and “Impact of COVID-19 on community access to food.” There is a 500-character limit for this field, though it should be noted that there is a *Segment Synopsis* field available for more robust description of segments that will display under the segment’s folded header.

Partial Transcript. A segment’s first sentence(s) of dialogue can be inserted in the *Partial Transcript* field. This may be a desirable option for projects that do not provide full transcripts so that users can be sure they have navigated to the correct section of the media.

The screenshot shows a form with two main sections: "Keywords:" and "Subjects:". The "Keywords:" section contains a list of tags: Chinatown, Taishan, mixed race, Chinese, British, Chin Wing Chun Society, Vancouver (B.C.), childhood, snacks, heritage, and grandparents. Below this list, it says "Keywords should be separated by a semi-colon." and "Current Thesaurus: None". The "Subjects:" section contains two tags: Chinatown (Vancouver, B.C.) and Food habits. Below this list, it says "Subjects should be separated by a semi-colon." and "Current Thesaurus: Library of Congress Subject Headings (Linked)".

Figure 9: Interface for applying keywords or subject terms to a segment in the OHMS Index module.

Keywords/Subjects. The *Keywords* and *Subjects* elements operate similarly to how they do in the Metadata Editor, with the exception that after inserting a semi-colon delimiter in these fields, the terms will display as individual boxes (see Figure 9). Remember that any thesaurus or controlled vocabulary applied to a field (Title, Keywords, or Subjects) at the interview level will carry over to the segment level. OHMS will kindly remind you of any thesaurus that you have

applied to an element and if you begin to type in a field that has been assigned a thesaurus, a list of suggested words will appear. With access at top of mind, I used keywords to fill in any subject or geographical data that I felt LCSH did not adequately cover. Because my interview is currently the only one from The 2020 (Re)Collection to be indexed in OHMS, I did not apply a controlled vocabulary to these keywords. However, should interviews from the same project eventually be processed through OHMS, a controlled vocabulary would ensure that all

instances of, for example, "COVID-19" (e.g. "coronavirus", "covid-19", etc.) are spelled the same.

Segment Synopsis. This element is the final field for segment description and, while not mandatory, it offers the opportunity to summarize themes or topics that come up in the clip. With a full transcript accompanying my interview, I chose sentence fragments over structured paragraphs for my synopses. This is so that potential users can get a quick snapshot of what the segment is about and then navigate to the full transcript if they wish to learn more. While there does not appear to be a character limit for this text box, evaluate how this field can best improve access to the data in your resource.

Once you have pressed the "Save" button on the top right, under the media player, you should be able to see your new segment as a folded header. Clicking the header will show a preview of the segment metadata that you have just created, and also display "Edit" or "Delete" buttons. In the web viewer, each unfolded segment on the Index page will show two buttons in the righthand corner. The first, which looks like a document, will bring you to the exact part of the synchronized transcript that corresponds with that segment. The second button, a box with an arrow, will open up a text box containing a URL that can be copied and pasted for sharing that specific segment of media.

3.3.2 Expanding functionality: GPS coordinates, hyperlinks, and images

This is where it gets really fun. For those wishing to take the index one step further, there are some options for linking external data or supplementary interview material to segments. In the Index module, under the *Segment Synopsis* field, there are two fieldsets: GPS and Hyperlinks.

GPS. The GPS fieldset integrates Google Maps with OHMS. To make use of this function, latitudinal and longitudinal coordinates must be entered into the *GPS* field, and a title or description into *GPS Description* that will serve as the link label in the web viewer. The *GPS Zoom* field determines how "zoomed in" to the indicated point the Google Maps display will be if opened. The final result in the web viewer is a link that will call up a Google Maps lightbox (a foregrounded frame in the same window that also dims the background) that centers on the chosen coordinates. See Figure 10 for an example.

Hyperlinks. The hyperlinks fieldset allows for segments to be linked to any other external resources. This may include links to relevant articles, public documents, images, or websites. In the context of oral histories, it is not uncommon for narrators to be asked to provide photographs or images pertaining to the interview topics, in case it may aid in memory

recollection or serve as a supplementary resource during exhibition.²³ Like the GPS coordinates, linked images will appear as a lightbox, while any other resources will open in a new tab. Given how many dishes my narrator mentions in his interview on food and identity, I chose to use the hyperlinks fieldset to link a public domain image of each dish to the segment in which they were discussed.



Figure 10: (L) A segment containing the GPS coordinates of a historical building mentioned in the interview. (R) The interactive Google Maps lightbox that the GPS link opens.

²³ Valerie J. Janesick, "Oral history interviewing with purpose and critical awareness," in *The Oxford Handbook of Qualitative Research* (2nd ed.), ed. Patricia Leavy (New York: Oxford University Press, 2020), 462.
<https://doi.org/10.1093/oxfordhb/9780190847388.013.23>

3.4 Bilingual Indexes and Transcripts

A major update to the OHMS system was the addition of bilingual capabilities in 2016.²⁴ Synchronized transcripts and indexes can now be created and displayed in two languages for a single resource. As you might imagine, this has incredible implications for access—resources can reach much wider audiences, users can query terms from the transcript or index in either language from within the web viewer, and subject terms that are specific to a particular linguistic context can be added. Interviews conducted in various languages that belong to the same collection can be indexed in their primary language, contain translations to any target language, and still indicate that they have relationships with the same parent collection.²⁵

To enable the bilingual option, we must first return to the Metadata Editor. Under the **Language** section of the form, you will see two text fields and a checkbox. For the first field, simply called *Language*, fill in the name of Language 1. This will likely be the language that the interview or resource has been processed in. Tick the checkbox to enable transcript and index translations. Finally, put the name of Language 2 in the *Language for Translation* field. Note that the exact values entered for these two text boxes will be displayed in the web viewer for the resource. Because language codes or abbreviations often used in metadata records may not be familiar to your users, you may wish to use the full language name.

After enabling translations in the Editor, return to the *Interviews* homepage. The same dialogue window where we uploaded our transcripts for synchronization will now allow upload of a translated transcript. Formatting of the translated transcript may require some trial and error, particularly when dealing with non-Roman script.²⁶ If there are no timecodes in your translated transcript, the Sync module can be used in the same manner as with the original transcript.

For segment-level translations, head back to the Index module. Choose a header, unfold it, and click the blue “Edit” button. You should now notice a second column to the right of your original text boxes that replicates the input fields, with the name of your second indicated

²⁴ Noriko Sugimori, “War Memories: Intergenerational, Intercultural Oral History Project – About the Project,” Oral History in the Liberal Arts, August 19, 2016, <https://ohla.info/war-memories-intergenerational-intercultural-oral-history-project/>.

²⁵ For a detailed account of the challenges (and joys) of multilingual indexing in OHMS, staff from the Yiddish Book Center have shared their experience with the Wexler Oral History Project, which includes interviews in English, Yiddish, Spanish, Polish, Russian, Hebrew and more: Christa Whitney and Carole Renard, “Preserving Yiddish Language & Culture through Bilingual Oral History Access,” Oral History Review, Oral History Association, December 11, 2020, <https://oralhistoryreview.org/technology/preserving-yiddish-language-culture-through-bilingual-oral-history-access/>.

²⁶ Kiyoto Tanemura of the War Memories project has created a useful guide for handling some common formatting errors with translations: Kiyoto Tanemura, “Formatting and Uploading transcripts to OHMS (in English or Japanese),” Oral History in the Liberal Arts, July 25, 2016, <https://ohla.info/uploading-transcripts-to-ohms-english-or-japanese-on-a-mac/>.

language at the top of the column. Figure 11 shows a segment in the Index module that I have begun to translate into Traditional Chinese. All values in fields applicable to Language 1, including GPS coordinate labels and hyperlinks, can now be translated and entered into the fields in the column for Language 2. Once the resource has been made available to the public, users can switch between Language 1 and Language 2 by clicking the link underneath the Transcript/Index toggle (see Appendix B, 6).

Figure 11: Bilingual segment indexing.

Some final considerations...

If you are importing metadata created outside of OHMS rather than entering it through the application's Metadata Editor, there are templates available to download in either XML or CSV format to aid you in preparing your data for smooth import.²⁷

To download a resource's transcript and index as a single PDF, with a cover page that includes links to the collection or series, click the printer icon on the top right of *Preview mode* or in the final web viewer (see Appendix B, 3). All segment-level indexing and transcript timecodes will be automatically exported.

²⁷ See the links under *Resources* at <https://www.oralhistoryonline.org/documentation/>.

3.5 Web Viewer

Once you have finished indexing your resource, export it as an XML file through the link in the last column on the *Interviews* homepage. The XML is required for display through the web viewer. There is an option to download the metadata as a CSV file that can be used to transfer metadata to a content management system, but the CSV file is not necessary for the context of the OHMS viewer.

To make your indexed resource available to the public, the web viewer portion of the OHMS system must be installed on your server. There are detailed installation instructions for either an enterprise environment or a third-party host service on the OHMS website at <https://www.oralhistoryonline.org/documentation/>. Once the web viewer has been installed, your resource's XML file should be placed into the designated cache files directory. Because the web viewer is installed directly onto your server, you or your host institution are responsible for keeping it updated. You are strongly encouraged to consult the OHMS User Guide for directions on installing the web viewer and building hyperlinks to your final resource.

Conclusion

There were two goals in constructing this guide. The first was to highlight the advocacy nature of oral history and demonstrate how developments in metadata and digital resource management can aid practitioners in assessing access to their collections. The second was to provide a walkthrough of the metadata creation and indexing possibilities of the OHMS system by drawing on my own experience using the toolkit.

For readers interested in following up on some of the topics raised in this guide, there are links to several other blogs and articles that cover working with OHMS linked in the [Resources](#) section at the end of this document.

Appendix A

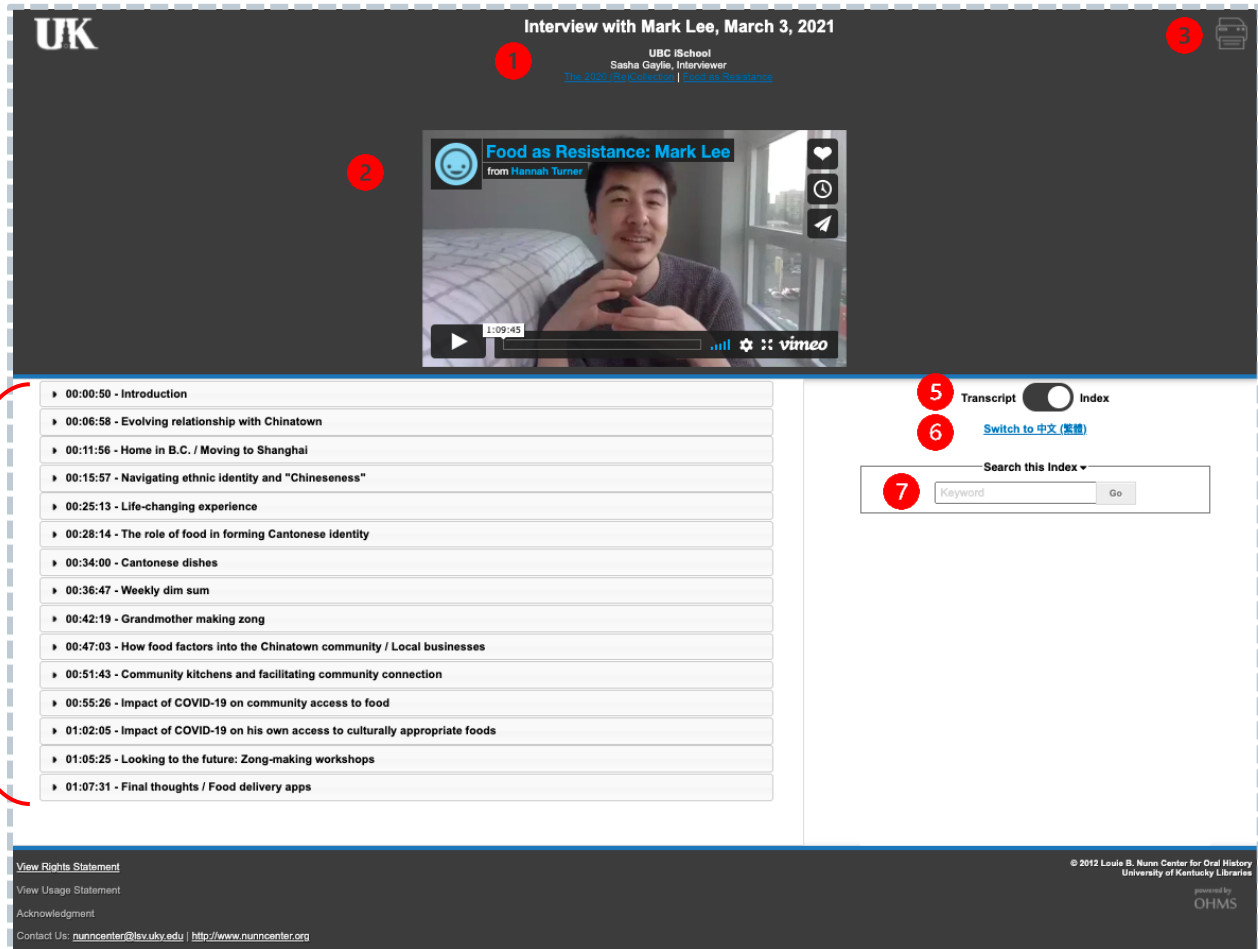
Sample Metadata Element Set

Element Name	Element Description	Values	Required?	Repeatable?
File name	Name of media or text file associated with an interview	Must include interviewer surname, interviewee surname, part number (if multiple sessions), collection date and file format. interviewer_interviewee_[part#]_[date].format	Y	Y (If multiple parts)
Repository	Name of holding repository		Y	N
Title*	Title of interview	Thesaurus assignment optional. Interviews within a collection or project should follow the same naming conventions.	Y	Y
Accession Number	Identifier for asset		Y	N
Interviewee	Name of person being interviewed	First Last	Y	Y
Interviewer	Name of person interviewing	First Last	Y	Y
Cataloguer	Name of person processing or indexing interview	First Last	Y	Y
Interview Date	Date of interview collection	ISO 8601 YYYY-MM-DD	Y	Y (if applicable)
Creation Date	Date digital resource was created	ISO 8601 YYYY-MM-DD	Y	N
Collection ID	Identifier for collection		Y (if available)	N
Collection Name	Name of collection		Y (if available)	N
Collection Link	URL to collection		Y (if available)	N
Series ID	Identifier for series		Y (if available)	N
Series Name	Name of series		Y (if available)	N
Series Link	URL to series		Y (if available)	N
Summary	Brief description of resource	Plain text description; can be interview abstract or curatorial statement.	Y	N
Keywords	Interview-level keywords	Thesaurus assignment optional. For interviews part of a collection or larger project, local controlled vocabulary is recommended.	N	Y

Subject Headings	Resource subject headings	Thesaurus assignment recommended. LCSH, AFS Ethnographic Thesaurus, or local controlled vocabulary.	Y	Y
File format	Original file format of resource	Ex. .mp4, .wav, .docx	Y	N
Item type	Type of content	Dropdown menu (media; transcript; index; supplementary material) Supplementary material can include images, documents provided by the interviewee, etc.	Y	Y
Duration	Length of media file	Timestamp format (HH:MM:SS)	N	N
Media format*	Type of interview media	(Audio; Video)	N (media only)	N
Media host*	Name of media host	Dropdown menu (Avalon; Aviary; Brightcove; Kaltura; SoundCloud; Vimeo; YouTube)	N (media only)	N
Media connection*	Link between media and media host	Dependent on media host (Ex. URL, host ID information, iFrame embed code).	N (media only)	N
Rights	Statement on rights	Plain text description.	Y	N
Usage	Statement on use conditions	Plain text description.	Y	N
Restrictions	Sharing restrictions indicated by interviewee	Plain text description.	N	N
Language	Language(s) of resource	ISO-639 Ex. eng, yue, hin	Y	Y (if translations available)
*Minimum values required to create and work with an asset in OHMS				

Appendix B

OHMS Web Viewer Interface



1. Interview title, repository name, interviewer, collection link, series link
2. Media player
3. Button to export interview index and transcript as PDF
4. Index: folded headers
5. Transcript/Index toggle
6. Link to switch languages (if translation available)
7. Keyword search bar
8. Links to rights statement, usage statement and acknowledgements

Bibliography

The 2020 (Re)Collection. University of British Columbia. April 15, 2021.

<https://blogs.ubc.ca/2020recollection/>

Boyd, Douglas A. "The Oral History Review OHMS: Enhancing Access to Oral History for Free." *The Oral History Review* 40, no. 1 (2013): 95–106.

<https://doi.org/10.1093/ohr/oht031>

Boyd, Douglas A., Janice W. Fernheimer and Rachel Dixon. "Indexing as Engaging Oral History Research: Using OHMS to 'Compose History' in the Writing Classroom." *The Oral History Review* 42, no. 2 (2015): 352–367. <https://doi.org/10.1093/ohr/ohv053>

Breaden, Craig. "Words Painting Pictures: Indexing the H. Lee Waters Project using OHMS." *Provenance, Journal of the Society of Georgia Archivists* 34, no. 1 (2016).

<https://digitalcommons.kennesaw.edu/provenance/vol34/iss1/15>

Fernheimer, Janice W., Douglas A. Boyd, Beth L. Goldstein and Sarah Dorpinghaus. "The Oral History Review Sustainable Stewardship: A Collaborative Model for Engaged Oral History Pedagogy, Community Partnership, and Archival Growth." *The Oral History Review* 45, no. 2 (2018): 321–41. <https://doi.org/10.1093/ohr/ohy052>.

Janesick, Valerie J. "Oral history interviewing with purpose and critical awareness." In *The Oxford Handbook of Qualitative Research* (2nd ed.), edited by Patricia Leavy, 456–479. New York: Oxford University Press, 2020.

<https://doi.org/10.1093/oxfordhb/9780190847388.013.23>

Kata, Lauren, Natalie Milbrodt, Steven Sielaff and Jaycie Vos. 2020. "Oral History Metadata and Description: A Survey of Practices." Metadata Task Force, Oral History Association.

https://www.oralhistory.org/wp-content/uploads/2021/01/OHA-MTF-White-Paper_2020.pdf

Mackay, Nancy. *Curating Oral Histories: From Interview to Archive*. Walnut Creek, CA: Routledge, 2015.

Mazé, Elinor A. "Metadata: Best Practices for Oral History Access and Preservation." Oral

- History in the Digital Age. Institute of Museum and Library Services. Accessed March 30, 2021. <http://ohda.matrix.msu.edu/2012/06/metadata/>
- "Metadata for Oral Histories: An Interactive Toolkit." Oral History Association. Accessed March 29, 2021. <https://oha.archipelago.nyc/>.
- OHMS (Oral History Metadata Synchronizer) User Guide. Oral History Metadata Synchronizer. Last updated October 1, 2020. https://www.oralhistoryonline.org/wp-content/uploads/2020/11/@OHMS_user_guide_master_v3-8-3.pdf
- Royles, Dan. "The Oral History Review Teaching Digital Humanities with Oral History: The Staring Out to Sea Oral History Project and OHMS in the DH Classroom." *The Oral History Review* 43, no. 2 (2016): 408–20. <https://doi.org/10.1093/ohr/ohw051>.
- Sielaff, Steven. "Metadata at BUIOH: A Case Study." Oral History in the Digital Age. Institute of Museum and Library Services. Accessed March 30, 2021. <http://ohda.matrix.msu.edu/2015/10/metadata-at-buioh-a-case-study/>.
- Sugimori, Noriko. "War Memories: Intergenerational, Intercultural Oral History Project – About the Project." Oral History in the Liberal Arts. August 19, 2016. <https://ohla.info/war-memories-intergenerational-intercultural-oral-history-project/>.
- Tanemura, Kiyoto. "Formatting and Uploading transcripts to OHMS (in English or Japanese)." Oral History in the Liberal Arts. July 25, 2016. <https://ohla.info/uploading-transcripts-to-ohms-english-or-japanese-on-a-mac/>.
- Turner, Kristopher. "Creating history: a case study in making oral histories more accessible in the digital age." *Digital Library Perspectives* 33, no. 1 (2017): 48–62. <https://doi.org/10.1108/DLP-06-2016-0016>
- Vos, Jaycie. "New Roots: An Oral History Metadata Case Study at the University of North Carolina at Chapel Hill." Oral History in the Digital Age. Institute of Museum and Library Services. Accessed March 30, 2021. <http://ohda.matrix.msu.edu/2015/10/new-roots-an-oral-history-metadata-case-study-at-the-university-of-north-carolina-at-chapel-hill/>
- Whitney, Christa and Carole Renard. "Preserving Yiddish Language & Culture through

Bilingual Oral History Access.” Oral History Review. Oral History Association. December 11, 2020. <https://oralhistoryreview.org/technology/preserving-yiddish-language-culture-through-bilingual-oral-history-access/>.

Resources

1. Playlist of OHMS tutorials uploaded to YouTube by the Nunn Center:
<https://www.youtube.com/playlist?list=PLj1ajYzHk6aSkmF74MCluzC3tcT34lSYb>
2. Instructional blog posts on different facets of the OHMS system by the Oral History in the Liberal Arts initiative: <https://ohla.info/themes/ohms/>
3. Manual of best practices for archiving oral history by the Oral History Association:
<https://www.oralhistory.org/archives-principles-and-best-practices-complete-manual/>
4. “Micro-essays” and other resources on best practices for digital oral history projects, collected by the Oral History in the Digital Age (OHDA) project:
<http://ohda.matrix.msu.edu/best-practices/>