
SpokenWeb Metadata Schema and Cataloguing Process

Version 4

SpokenWeb Metadata Task Force

Cite as:

Camlot, J., Dowson, R., Ferrier, I., Fong, D., Kail, R., Lu, E., Luyk, S., MacDonald, C., Meza, A., Neugebauer, T., Wiercinski, J., Ajeeb, Y., Barillaro, A., Chandler, N., Hannigan, L., Harris, M., Hooper, E., Kolosov, L., Knudsen, J., MacGregor, H., Mash, C., McLeod, K., Pickering, H., Roberge, A. (2025). *SpokenWeb Metadata Schema and Cataloguing Process*. Version 4.
<https://doi.org/10.11573/spectrum.library.concordia.ca.00996470>

Dec 05, 2025

CONTENTS

List of Authors/ Contributors	1
1. Introduction	3
1.1. How to Engage with a Collection	3
2. Troubleshooting (what to do when you don't know what to do)	5
3. How to Photograph Archival Assets	6
3.1. Technical Specifications (Equipment, Process)	6
3.2. How to name and format image files	6
4. Metadata Fields	7
4.1. Cataloguer	7
4.2. Classification Fields	7
4.3. Material Description (for Physical & Digital Items Fields)	10
4.4. Preliminary Material Description Fields	10
4.5. Physical Items Fields	11
4.6. Digital Items Fields	15
4.7. Title	16
4.8. Title Source	17
4.9. Title Note	18
4.10. Source Item ID / Identifiers	18
4.11. Persistent URL	18
4.12. Language	18
4.13. Production Context	18
4.14. Genre	19
4.15. Statement of Responsibility	22
4.16. Date	24
4.17. Location	25
4.18. Rights Fields	28
4.19. Related Works	30
4.20. Contents	31
4.21. Contents Notes	46
4.22. Notes	46

LIST OF AUTHORS/ CONTRIBUTORS

Version 1 authors (2019):

Jason Camlot (Concordia U)
Rebecca Dowson (SFU)
Ian Ferrier (Wired on Words)
Deanna Fong (SFU)
Roma Kail (Victoria University in the UofT)
Sean Luyk (U Alberta)
Aurelio Meza (Concordia U)
Tomasz Neugebauer (Concordia U)
Jared Wiercinski (Concordia U)

Version 2 authors (2019):

Jason Camlot (Concordia U)
Rebecca Dowson (SFU)
Ian Ferrier (Wired on Words)
Deanna Fong (SFU)
Roma Kail (Victoria University in the UofT) (also Metadata Controller)
Eva Lu (Victoria University in the UofT)
Sean Luyk (U Alberta)
Aurelio Meza (Concordia U)
Tomasz Neugebauer (Concordia U) (also Metadata Controller)
Jared Wiercinski (Concordia U)

Version 2 contributors (2019):

Ali Barillaro (Concordia U)
Lee Hannigan (U Alberta)
Cole Mash (SFU)
Holly Pickering (U Alberta)
Nathan Chandler (U Calgary)

Metadata Controllers (2019):

Roma Kail (Victoria University in the UofT)
Tomasz Neugebauer (Concordia U)

Version 3 authors (2020):

Jason Camlot (Concordia U)
Rebecca Dowson (SFU)
Ian Ferrier (Wired on Words)
Deanna Fong (SFU)
Roma Kail (Victoria University in the UofT) (also Metadata Controller)
Sean Luyk (U Alberta)
Aurelio Meza (Concordia U)
Tomasz Neugebauer (Concordia U) (also Metadata Controller)
Jared Wiercinski (Concordia U)

Version 3 contributors (2020):

Ali Barillaro (Concordia U)
Jakob Knudsen (SFU)
Linara Kolosov (SFU)
Cole Mash (SFU)
Andrew Roberge (Concordia U)

Metadata Controllers (2020):

Roma Kail (Victoria University in the UofT)
Tomasz Neugebauer (Concordia U)

Version 4 authors (2025):

Jason Camlot (Concordia U)
Corina MacDonald (Concordia U)
Tomasz Neugebauer (Concordia U)

Version 4 contributors (2025):

Yara Ajeeb (Concordia U)
Maia Harris (Concordia U)
Ella Hooper (Concordia U)
Hannah MacGregor (SFU)

1. INTRODUCTION

Welcome to the SpokenWeb metadata cataloguing project! Broadly speaking, a metadata schema is a rationalized way of describing a collection with a set of rules according to which a controlled “language” operates.¹ In this context, language means an artificial construct defined for the purposes of describing our collections in a systematic and consistent way.

1.1. How to Engage with a Collection

Collections can differ from each other significantly. Some have been born via analog media technologies (tape recorders, for example), some have been “born digital” (recorded as digital files directly onto flash or hard drives), and some may be a mixture of the two. Some have been processed to some extent already using a different kind of cataloguing or metadata schema, and others may never have been organized before. The nature and state of a collection, will have an impact on the way and ease with which you are able to catalogue the artifacts and objects in it. In most cases, in order to properly capture the information about the collection you will be describing with the SpokenWeb Metadata Schema, you will engage in the following activities:

1. You will access and handle the collection and the artifacts through your local library or research center. You will also require a computer while processing the collection so that you can input your metadata (data about the audio data) into our SpokenWeb cataloguing template.
2. Where possible, you will photograph the analogue assets (tapes, reels) and their containers. It will be important to identify and retain images about the collection as a secondary source of information, and these images may also be integrated into digital presentations of the collection down the road. More information about how to do this appears below in the “Photographing Archival Assets” section.
3. You will engage in describing each artifact in the collection using the fields that comprise the SpokenWeb Metadata Schema. This will entail following a specific style of entering information into the relevant field categories, and may also entail seeking the information that you’ll need to fill in, either by examining the artifact, listening to the audible content of the sound recording, or by consulting external sources (i.e. doing research!).
4. A note on listening: Given that some of the information that we will be including in our description will come from listening to the contents of the recordings themselves, a few initial points about how to listen, and what to listen for, are in order. In listening to an archival recording for the purpose of cataloguing, the goal, in the first instance, is to capture a sense of the nature and overall contents of the recording in question. If the digital file you are listening to was originally an analogue tape, for instance, that tape may have been used to capture a single event (say, a poetry reading that lasted an hour), or it may have been used to capture multiple events recorded at different times. During the first listening, use a transcription software of your choice and produce a time-stamped account of the contents of the audio artifact, noting with time stamps what we call discrete signature moments during the recorded audio event. By signature moments we mean moments that signify important discrete sections in the audio. For example, in a recording of a poetry reading event one may hear, first, a speaker delivers an introduction to the event, followed by the poet stepping up to the podium to read a series of poems from a book. In this example, the introduction may provide you with information that will be useful to your description of the file. The introducer may state his/her name, the date, the venue, the name of the invited poet who will soon read, and other points of information that may be required information for our metadata fields. The introduction as a whole would comprise one discrete section of the audio that should be timestamped and named at the start as a signature moment with a timestamp and a tag that identifies the nature of the speech (i.e. Introduction) and, if you have it, the name of the speaker. When the poet begins to read you would then identify that beginning with a time stamp, identifying the name of the poet, and the nature of the next section of speech (i.e. Introductory remarks). When the poet begins to read you should then identify each poem read as a discrete

¹ Dublincore.org/groups/tools/dctools2006/2MetadataGlossary.pdf

section of the audio, thus describing the contents of the audio as it progresses with timestamps and poem titles. If the recording consists of multiple events that took place at different times, these events would also be reflected in the timestamping description. The result of this approach to listening will be the production of a “Table of Contents” (TOC) of sorts, for an audio recording. Beyond the production of this audio TOC, you may also find important clues that will help you fill in your other metadata fields. While such information often appears at the start of a recording (as I mentioned in my discussion of the Introduction), sometimes clues about who is speaking, where the event took place, etc., are found randomly throughout a recording. Pay attention for such moments when names, dates, place names and poem and book titles are spoken as they can be useful to you as you work on describing the nature and contents of your audio object with as much accuracy as possible.

2. TROUBLESHOOTING (WHAT TO DO WHEN YOU DON'T KNOW WHAT TO DO)

At times you may find that the information requested by this cataloguing process document is not available, or that the information you have does not easily fit into the schema and rules of the SpokenWeb guidelines. Sometimes you will discover questions that this document does not answer. Therefore, you should document any questions, decisions or variables that are not addressed within it, or that are contrary to it. These documented questions and issues and other relevant points of information pertaining to the cataloguing process should be provided to Jason Camlot (jason.camlot@concordia.ca). Your documented questions and feedback will be integrated into our analysis of the process throughout the cataloguing activities so that we can continue to improve upon our schema and clarify our description of the cataloguing process.

3. HOW TO PHOTOGRAPH ARCHIVAL ASSETS

3.1. Technical Specifications (Equipment, Process)

If you are able to photograph the archival recordings, you should try to do so at the beginning of your cataloguing process. Photograph each side of the artifact and its case, as well as any accompanying material containing textual information. Photographs should be in .jpeg or .png file format and at least 500 pixels in width if possible. There is a 2 MB file size limit for images that will be uploaded to the SpokenWeb dataset.

3.2. How to name and format image files

In taking photos of the archival audio artifacts, you will be generating new data (digital image files) and those files will also need to be named properly for easy use. Each filename of an image should have the following syntax:

[Collection name]_[Fonds number]_[Box number]_ID_Side number.jpeg ^[*]

Examples:

- Kiyooka_Msc321_Box8_128_Side1.jpeg
- Kiyooka_Msc321_Box10_insert.jpeg
- Gilbert_21_Side2.png
- Ultimatum_5-3_Side1.png

NOTE: Not all collections will be coming from archives, so this syntax will not apply to all.

^[*] Collection name may be derived from the name of the fonds. If the fonds is author fonds, then use the author's last name ("Kiyooka", "Gilbert", etc.). If the fonds is an event series, use the name of the series ("Ultimatum", "Véhicule", etc.). Supply fonds and box numbers when available, and use any extant ID numbers if the collection has already been catalogued. The object is always the primary source of information for determining a unique ID number.

4. METADATA FIELDS

The SpokenWeb schema describes both literary activities or events and the AV assets or digital files that document them. The following guidelines provide details about the information to be entered in each field. The schema is organized into overarching descriptive categories, within each there are multiple sub-fields to be filled in.

4.1. Cataloguer

That's you. The cataloguer refers to the individual who is cataloguing a particular artifact. Having your name associated with the entry will help us find you if we have any questions about how you've filled in a field, and also will identify your contributions to this project over time. Enter the cataloguer name as **Cataloguer Last Name, First Name**. In a separate field, enter the Email address of the cataloguer.

4.2. Classification Fields

Note: Classification fields will in most cases be added by the administrators of the SpokenWeb metadata management system.

Partner institutions, collections, series, and sub-series are managed separately from individual items in the SpokenWeb metadata schema. They are organized in hierarchical relationships, in which a partner institution has one or more collections, each collection has one or more series, and each series has one or more sub-series, and so on indefinitely. When an item is catalogued as belonging to a specific level of this classification, for example when an item is assigned to a series, the relevant collection and institution will also be automatically assigned to the item.

Examples of the classification structure:

Concordia University (partner institution)
 SGW Poetry Reading Series (collection)
 Poetry 3 (series)

Concordia University (partner institution)
 Ian Ferrier fonds (collection)
 Words and Music (series)

Below are the metadata guidelines for each classification level.

4.2.1 Partner Institution

Label (mandatory): Enter the name of the institution that holds the collection. For collections that are not affiliated with any institutions, enter "Community Collection". This label will be used in the front-end applications.

URI: Enter a URL or other permanent identifier if possible (e.g. Wikidata, institutional website).

Description: You may enter a description of the institution if desired.

4.2.2 Collection

Label (mandatory): Enter title proper for a fonds or collection to provide sufficient information to identify and find the original resource, for example: Daphne Marlatt fonds.² This label will be used in SpokenWeb front end applications.

If the collection is not deposited in an institutional archive but has been made available by a community organization via Internet Archive or another open access repository, use the name of the collection provided by that organization. If there is no clear title provided to a collection, a title must be created at least two of the following:

- Name of individual(s) who created the collection
- Name of venue and place (town or city) of events documented in the collection
- Primary genre or descriptor identifying nature of events documented in collection
- Time span of the events documented in the collection
- Platform where they are available

Example: Christopher Hand, Zeke's Gallery Montreal Reading Series, 2004-2007, via Internet Archive

URI: Enter a URL or other permanent identifier if possible (e.g. Wikidata, collection website).

Contributing Unit: Enter the Contributing Unit, for example, Collections and Rare Books (if available).

Source Collection: This will often be the same as the Label entry: enter a title proper for a fonds or collection to provide sufficient information to identify and find the original resource, for example: Daphne Marlatt fonds.

Source Collection Description

Copy/paste source collection description from online library catalogue into the corresponding field or develop one with your supervisor (it should include some interesting things about a collection, such as names of poets featured on artifacts, year of broadcast, etc.).

If you are creating a collection description from scratch (for example, for a community collection), consider including the following:

- Who made this collection?
- Where was it made?
- How did it come into your possession?
- If this is a series or a donation, what time span does it cover?
- What genres are presented?
- What format(s) comprise the collection?
- Who were some of the readers/presenters?
- Does the collection represent people from a specific region?
- How was it recorded? If from analogue sources, are there digital copies?
- Who houses the collection now and where?
- What are some highlights of the collection?
- Are the recordings indexed? How?
- Are the recordings copyrighted?
- If so, are there any special permission notices or rights restrictions attached to the collection?
- What makes this collection of interest to you and/or other students of literature?

² "Rules for Archival Description," Canadian Archival Council, Bureau of Canadian Archivists, July 2008, https://archivescanada.ca/wp-content/uploads/2022/08/RADComplete_July2008.pdf.

Source Collection ID Number: A reference code or id number of the related resource from which the described resource is derived, for example: MsC 142.

Persistent URL: Provide a link to the collection that you are describing, if available.

Collection Image URL: Provide a link to an image file that will represent the collection.

4.2.3 Series

A series describes documents arranged in accordance with a filing system or maintained as a unit because they result from the same accumulation or filing process, or the same activity; have a particular form; or because of some other relationship.³

SpokenWeb sometimes uses series to capture the way in which reading series have been named and branded.

Example of using the series field

General Example:

Collection Name: SpokenWeb Archive of the Present

Series Title: Listening Practices

Sub-series: Virtual Listening Practice 2020

General Example:

Collection Name: Véhicule Art (Montréal) Inc. fonds (P027)

Series Title: A/V Material

Specific Example:

Collection Name: SGW POETRY READING SERIES

Series Title: Poetry 1

Label (mandatory): Enter the Series title. This label will be used in SpokenWeb frontend applications.

URI: Enter a URL or other permanent identifier if possible (e.g. Wikidata, series website).

Description: Enter a description of the series if available.

Wikidata URL: Enter relevant Wikidata link(s). The basic Wikidata search bar can be found at this [link](#).

4.2.4 Sub-series

Label (mandatory): Enter the Sub-series title. This label will be used in SpokenWeb frontend applications.

URI: Enter a URL or other permanent identifier if possible (e.g. Wikidata, sub-series website).

Description: Enter a description of the series if available.

Wikidata URL: Enter relevant Wikidata link(s). The basic Wikidata search bar can be found at this [link](#).

³ Committee on Descriptive Standards. "ISAD(G): General International Standard Archival Description." INTERNATIONAL COUNCIL ON ARCHIVES, 2000. https://www.ica.org/app/uploads/2024/01/CBPS_2000_Guidelines_ISADG_Second-edition_EN.pdf .

4.3. Material Description (for Physical & Digital Items Fields)

Definition: The Material Description for Physical and Digital Items provides information about the material specifications of the artifact or object being described. Some of the questions this field will answer include: What is it made of? What are the technical specifications according to which it functions? What are its affordances or capacities? What's its condition? Etc. The term Item refers to the original source item of the digital audio signal, or the original digital audio file (if born digital). Related items (such as migration copies) may be accounted for in the Notes field.

General Guidelines:

- Provide information about any important physical or material characteristics, or technical requirements that affect use of the item of description or understanding of the item.⁴
- Describe the item. Material Description also allows for multiple entries associated with one performance, for example, in cases when one event had been recorded on multiple assets. If the latter, you have an option to provide individual metadata for each of the physical items that are a part of the event.
- As noted by most schemas consulted, “technology for digital audio files and mass storage systems is [and was] still evolving.”⁵ Therefore, include “at least a minimum of information and description of the physical characteristics for Local Access Files that will allow the configuration of equipment to play back the files correctly.”⁶ System requirements and modes of access should be described in the Note field.

Chief Source(s) of Information:

- Most of the information you will seek can be found on the physical units of description (the material things you're describing), including the container, labels and other materials that might accompany the Item. For example, a reel to reel tape (physical unit of description) may come in a tape box (container) and have a card inside the box (other material) with additional technical info about the recording typed on it.
- If information is taken from sources other than the Chief Source of Information (the Item and its container/components), please identify the sources of this externally-found information in the Notes.

4.4. Preliminary Material Description Fields

Recording Type: Determine if the original unit of description is analogue or digital. (This will probably be self-evident. Basically determine if it's a thing or a digital file). Select Analogue or Digital. If it is Analogue, select the appropriate AV Type and then go to the **Physical Items Fields** to continue entering the data. If it is digital, select the appropriate AV Type and then go to the **Digital Items Fields** to continue entering the data.

AV Type: Determine and indicate if the item contains audio or video (with audio). Again, this will likely be self-evident, but should be indicated in the AV TYPE field.

Identify the AV Type with either Audio or Video.

⁴ “Physical characteristics and technical requirements,” ISAD(G): General International Standard Archival Description, INTERNATIONAL COUNCIL ON ARCHIVES, 19-22 September 1999, 3.4.4, p.29, https://www.ica.org/sites/default/files/CBPS_2000_Guidelines_ISAD%28G%29_Second-edition_EN.pdf.

⁵ “Area 5: Introduction,” Physical Description, IASA Cataloguing Rules, The International Association of Sound and Audiovisual Archives <https://www.iasa-web.org/cataloguing-rules/50-introduction>.

⁶ IASA, “Area 5: Introduction”

4.5. Physical Items Fields

- Material designation
- Physical composition
- Storage capacity
- Extent
- Side
- Playing speed
- Track configuration
- Playback mode
- Equalization
- Generations
- Tape brand
- Sound quality
- Physical condition
- Conservation
- Other physical description
- Accompanying material
- Other

Material Designation: Assign a specific Material Designation to the Item. The specific material designation identifies the particular class of material to which the item belongs. For our schema, we have identified a set number of material designations that we believe will be the most common within our collections. For material designations not found on our list, please contact Jason Camlot (jason.camlot@concordia.ca).⁷

- Reel to Reel
- Cassette
- Microcassette
- Wire
- DAT
- Disc
- Floppy Disk
- CD
- Minidisc
- Cylinder
- U-Matic
- Betamax
- VHS

⁷ “Appendix D Glossary,” IASA Cataloguing Rules, The International Association of Sound and Audiovisual Archives, <https://www.iasa-web.org/cataloguing-rules/appendix-d-glossary>.

- Hi8
- MiniDV

Physical Composition: This refers to the material medium used to capture the audio. For example, reel to reel will usually refer to magnetic tape, but may also refer to a medium of magnetic wire. A flat disc record may be made of vinyl, but it may also be made of shellac or aluminum. The most common materials for this category are as follows⁸:

- Aluminum
- Magnetic Disk
- Magnetic Tape (usually Reel to Reel, 8-Track Cartridge or Cassette)
- Vinyl (usually discs)
- Lacquer (usually discs)
- Shellac (usually discs)
- Wax (usually cylinders)
- Laser (this includes CDs and videodiscs)
- Magneto-Optical (this includes minidiscs)

Storage Capacity of Artifact: In the case of some media formats, like audio cassette tapes, there will sometimes be a playing time duration indicated on the artifact itself. This refers to the total time storage capacity of the artifact when the media format is played at the standard speed associated with the medium.⁹ For example, was the recording made on a 60 minute tape or a 90 minute tape? The time storage capacity should be written in the following manner:

T

indicates that a time value follows. (Any value with a time must begin with T).

hh

specifies a two-digit hour

mm

as part of time, specifies a two-digit minute

ss

specifies a two-digit second

Use this format:

T hh:mm:ss

For example:

90 minutes = T01:30:00

45 minutes = T00:45:00

3 minutes and 21 seconds = T00 :03 :21

Extent: Extent refers to the size of the media storage material.¹⁰ For tapes this refers to the width of the tape. For discs, the size of the record's diameter dimension. For example, common sizes are:

Tape:

- 1/8 inch (audio cassette tapes)

⁸ IASA, "Appendix D Glossary"

⁹ IASA, "Appendix D Glossary"

¹⁰ ISAD(G), 3.1.5, p.16

- 1/4 inch (most portable reel to reel tapes)
- 1/2 inch
- 3/4 inch
- 1 inch
- 2 inch

Discs:

- 12 inch (the diameter of traditional “LPs”)
- 10 inch (often 78s were in this diameter,
- 7 inch (the diameter of traditional “45s”

Side: For certain media such as audio cassette tapes that divide the length of the magnetic tape into half-tracks, and for two sided discs (vinyl, etc.) You should indicate the Side (A or B) that the content of the digital file refers to. So if a digital file has captured audio from Side A of an audio cassette tape, you should mark A in the Side field. If a digital file combines audio from both sides of an audiocassette or record, then you can mark A and B in this field.

Playing Speed: Give the playback or playing speed of the physical unit as follows¹¹:

For analogue discs: give the playback speed in revolutions per minute (rpm). Examples of common playback speeds for analogue discs follow:

- 16 2/3 rpm
- 33 1/3 rpm
- 45 rpm
- 78 rpm

For sound tape reels: give the playback speed in inches per second (ips). Round the playback speed to the nearest full figure, if appropriate. The most common speeds are as follows:

- 1 7/8 ips
- 3 3/4 ips
- 7 1/2 ips
- 15 ips
- 30 ips
- 15/16 ips

4.5.1. Notes for Physical Items

Where possible, include the following additional information in the Notes field:

Track Configuration: For audio tapes, if possible, give the number of tracks, unless the number of tracks is standard for the unit being described. If necessary, give the track configuration. For example: Half-track. 2 track. 4 track, 24 track¹².

¹¹ IASA 5.C.2 who based it on AACR26 .5C3, 7.5 C5 from RAD8.5 C3

¹² “RAD,” 8.5C6, p. 8-17

Playback Mode: Give the playback mode [or number of sound channels] if the information is readily available, using one or more of the following terms as appropriate¹³. Some types of playback mode follow¹⁴:

- Mono (1 channel)
- Stereo (2 channels, or “hi-fi”)
- binaural stereo (also known as dummyhead)
- quadrophonic (4 channels)
- surround sound

Equalization (EQ): Indicate if EQ (equalization) adjustments have been applied to the audio. If the parameters of the adjustments applied to the frequency response of the signal are known, that information can be indicated, as well.

Generations: Distinguish between different generations of material.

Example:

- Copy from an original loaned by UTARMS

Tape Brand: Where indicated on the artifact, or reliably on the container, record the tape brand and the specific type number, for example, Ampex 456 or Scotch 250. This makes all the difference in being able to track degradation issues (e.g. sticky shed syndrome) from the item metadata. Knowing if it’s Scotch 200, or Scotch 250, or Scotch 300 is relevant for the preservation purposes. When the info is available, it’s usually written clearly on the box. Older tapes will often have named lines of tape, sometimes on a shiny sticker or something that might say, for example, “Radio Mastering Extraordinaire”. This can be helpful to know, and would be great to record, if it’s readily available. If it’s not obvious, then write down “Unknown acetate”.

Sound Quality: Based on broad categories of clarity and audibility, grade the audio quality of the recording as either Excellent, Good, or Poor.

Physical condition: Make notes on the physical condition of the unit being described if that condition affects use or understanding of the unit.

Indicate any important physical conditions, such as preservation requirements, that affect the use of the unit of description (ISAD G 3.4.4) or understanding of it.¹⁵

Examples:

- Sticky shed syndrome
- Fungus growth impairs playback
- Rejected cuts scratched through by operator
- Several tracks scratched through by operator
- Recorded with a constant audible hum
- Speed varies due to weak batteries at time of original recording

Conservation: If the unit being described has received any specific conservation treatment, briefly indicate the nature of the work.

Other physical description:

Examples:

- Impressed on rectangular surface 20 x 20 cm Reproduced from inner to outer grooves
- Recorded with clip-on microphone

¹³ “RAD,” 8.5C7, p. 8-17

¹⁴ Based on IASA 5.C.6 expanded AACR2 6.5 C7

¹⁵ “ISAD(G)” 3.4.4, p.29

- Recorded on 1 side of 1 audio disc

Accompanying Material:

Example:

- Issued with illustrated sleeve and liner notes, so liner notes could be entered in this field.

Other: Add any other descriptive information about the material that you deem relevant.

4.6. Digital Items Fields

- Digitized file name
- Digitized file path
- Digitized file URL
- Public access URL
- Contents
- Notes
- Channels
- Sample rate
- Precision
- Duration HH:MM:SS
- Size
- Bit rate
- Encoding
- Title
- Caption
- Credit
- Content type
- Featured

Digitized file name: Enter the name of the digital file

Digitized file path: Enter the path where the file is currently stored or will eventually be exported. If there is an existing folder structure for the digitized files, we need to be capturing where in the folder structure the Digitized File currently is. Alternatively, we need to be capturing where in the folder structure the Digitized File will be placed if it is to be exported out.

Digitized file URL: Enter the URL where the file is stored.

Public Access URL: Enter the URL where the file can be accessed by the public (if applicable).

Contents and Notes: These descriptions apply to the individual part(s) of a multiple part item. For example, the Contents field 4.18 applies to the whole item and does not allow for detailed description of individual parts.

Channels: The number of audio data streams, for example, a mono recording of a microphone has 1 channel.

Sample rate: The sample rate is measured in hertz (Hz) and represents the number of digital samples captured per second in order to represent the audio waveform.¹⁶

Precision (bit depth): The number of bits of information in each sample, e.g. 16 bit, 24 bit.

Duration: Use the format HH:MM:SS (Hours:Minutes:Seconds) to enter the duration if applicable.

Size: The size of the digital file in kilobytes, megabytes, or gigabytes. E.g. 1.66 GB

Bit rate: The number of computer bits conveyed or processed per unit of time in an audio or video file. Normally expressed in kilobits per second (Kbps), or megabits per second (Mbps).

Encoding: The algorithmic process of transforming analog information into digital form. Digital audio files are often encoded in formats WAV or MPEG-3, image files in JPEG or PNG, and video files in MPEG-4.

Title: The title of the digital image file.

Caption: A short description and/or explanation of the content of a digital image file.

Credit: List the people or institutions responsible for the production of the image, including their roles (i.e. photographer, designer, painter, etc.).

A section for **Content Type** is used to indicate the item as either:

Sound Recording: process of electrical or mechanical inscription of sound waves¹⁷

Video Recording: single work, or take, made using the medium of video¹⁸

Poster: any piece of printed paper designed to be attached to a wall or vertical surface¹⁹

Photograph: image created by light falling on a light-sensitive surface²⁰

Document: preserved information²¹

Artefact: artifact created by humans which gives information about the culture of its creator and users²²

Featured: If you have more than one image you can identify the file to be used as the main visual link to the record on a frontend web application.

4.7. Title

A word or phrase, usually appearing on an artifact (either **digital file** or **analogue artifact** or **container**), naming the item or the work (often as a group of individual works or recorded sounds) contained in it²³.

Procedure:

The Title field has two objectives: 1) to identify the artifact and 2) to describe it. The primary role for the Title is to identify the artifact. If the information on the artifact is useful for this purpose, it should be used as a Title. If the information on the artifact does not allow to identify the item (for example, if all of the artifacts in the collection have the same information written on them), then a descriptive Title should be generated to identify each artifact in a collection. If sufficiently descriptive, format it like this: [Name of Speaker] at [Venue] and [Year].

¹⁶ Federal Agencies Digital Guidelines Initiative. <https://www.digitizationguidelines.gov/glossary.php>.

¹⁷ "Sound Recording." Wikidata. Accessed January 19, 2021. <https://www.wikidata.org/wiki/Q5057302>.

¹⁸ "Video Recording." Wikidata. Accessed January 19, 2021. <https://www.wikidata.org/wiki/Q30070675>.

¹⁹ "Poster." Wikidata. Accessed January 19, 2021. <https://www.wikidata.org/wiki/Q429785>.

²⁰ "Photograph." Wikidata. Accessed January 19, 2021. <https://www.wikidata.org/wiki/Q125191>.

²¹ "Document." Wikidata. Accessed January 19, 2021. <https://www.wikidata.org/wiki/Q49848>.

²² "Cultural Artifact." Wikidata. Accessed January 19, 2021. <https://www.wikidata.org/wiki/Q1791627>.

²³ "International Standard Bibliographic Description for Non-Book Materials ISBD(NBM)," International Federation of Library Associations and Institutions (IFLA), 1987, p.7, https://web.archive.org/web/20240418162441/https://archive.ifla.org/VII/s13/pubs/ISBDNBM_sept28_04.pdf

Capitalize the first letter of the first word and of the first letter of proper nouns and additional words according to the appropriate usage in the language in which the material is catalogued.²⁴ Put square brackets around the descriptive title.

Example: [Phyllis Webb at Sir George Williams University, 1966]

If [Name of Speaker], [Venue], or [Year] are missing then only include information that is available.

Use the Title Source field to cite the sources of information.

The following is the order of preference for the source of title information²⁵:

1. the **item** itself (including any permanently affixed labels, or title frames, or the audio itself);
2. accompanying textual material (e.g. **cassette** insert, **CD** slick, inlay or booklet, recording/project accompanying documentation such as correspondence, donor agreements, recordist's worksheets, script, transcript, cue sheet);
3. a container that is an original part of the **item** (e.g. sound **cartridge**, video **cassette**, sleeve, container for video); or from
4. a secondary source such as reference or research works, a publisher's or distributor's brochure, **broadcast** program schedule, abstract, index or other available finding aid, container which is not an original part of the item (e.g. a **film** can be used to store a **reel** of **film**, tape box for storing audio tape), or the audiovisual content of the item itself. If the information is taken from a secondary source, cite the source in a **Title Source**.
5. For the unidentified information, listen to the recording.

For listing titles of individual works that are read within a given recording, see procedures for timestamping in the **Contents Field** (below).

4.8. Title Source

Indicate Title Source using one of the two options described below, in order of preference:

Option 1. Please specify a URI or other permanent identifier if possible, for example, if the title was retrieved from an online archival catalogue. Example: <https://concordia.accesstomemory.org/artist-as-worker-ideas-on-work>

Option 2. If no URI is possible, please cite the secondary source in free text

Decision Making Grid

Example How to Code Free Text in the Title Source Field	Appropriate when
Speaker is introduced at the beginning of this recording (include the time-code)	Such material is available
Speaker identification is based on cataloguer's expertise after having listened to multiple recordings. Publisher's brochure	For example, you as an expert have positively identified the voice on a tape and attributed a speech sample to a person on the basis of its perceptual properties (spoken phrase, emotions, additional ambient noise) Such material is available

²⁴ IFLA, 0.8, p.16

²⁵ Anglo American Cataloguing Rules (2005)

4.9. Title Note

Transcribe any handwritten additional information written on the container.

4.10. Source Item ID / Identifiers

This is an unambiguous reference to the resource within a given context:: 123 (if available).

4.11. Persistent URL

Provide a link to the Item that you are describing, if available.

4.12. Language

Enter the language(s) of a recording. More languages can be added as needed.

Examples:

English

French

English, French

English, French, Spanish

English, Cree

4.13. Production Context

This refers to the production scenario of the recording and is determined by the circumstances under which the recording was produced, as well as its intended audience and purpose.²⁶

Enter the appropriate Production Context, see definitions below (note that only one Production Context should be applicable to a single artifact):

- Audiobook: A recording of a oral reading of a book designed for commercial distribution and consumption.
- Broadcast: A recording of a radio broadcast.
- Classroom Recording: A recording of a lecture in a classroom setting.

²⁶ IASA. Sound Archives - A Guide to their Establishment and Development. <https://www.iasa-web.org/sound-archives/cataloguing>

- Documentary recording: A recording of a sound made outside of a controlled studio environment or professional performance venue that is generally unedited and typically made with portable equipment.
- Home recording: A sound recording produced in a private home.
- Internet recording: A recording produced on an online platform.
- Lab recording: A sound recording produced in a speech or language lab.
- Studio recording: A sound recording produced in a professional recording studio.
- Podcast: A program (such as music, news or arts programs) that are like a radio or television show but that are downloaded over the Internet.

Note that most frequently used Production Context would be: Documentary recording, Home Recording or Studio Recording. This could be determined after you listened to it.

4.14. Genre

In our usage, genre is distinguished from recording type or kind, which we refer to and define in terms of the production context.²⁷ The recording type refers to the production scenario of the recording, whereas genre refers to the audiotextual forms audible (discernible) within the recording.²⁸ In this way, we are establishing a metadata field that is descriptive of content, from a generic perspective.

Definition: Genre is a term or terms that designate a category characterizing a particular style, form, or content, such as artistic, musical, literary composition, etc.²⁹ In the generic terms we have chosen it is assumed that the genre refers to an audible source produced through speech or by other means. You will need to listen to the recording to determine genre.

NOTE: A single recording can contain multiple genres. If the audiotext you are listening to consists of more than one genre, list them separated by comma. However, in listing the genres of a recording, you should concentrate on the most *prominent* or *dominant* generic features and content of the audiotext.³⁰ For example, if an hour long recording of a poetry reading has a moment or two of conversation about the microphone at the start of the recording, the genre for this recording should be poetry, and not poetry, conversation. You will have to use your judgment in determining the audiotextual genres most appropriate for your each recording you listen to. We have provided a series of terms that will assist you in this work.

Genre should be chosen from the following controlled vocabulary of terms.³¹ Note that more than one genre might be applicable to a single artifact, and multiple terms are allowed. Again: how do you make a decision on when to assign a specific genre? It has to be among the most salient audible features of an artifact.

Here is the list of terms you must draw from:

- Ambient sounds
- Autobiographical sound recordings
- Conversation
- Interview
- Letter

²⁷ "Introduction to Library of Congress Genre/Form Terms for Library and Archival Materials," Library of Congress (LOC), April 2019, <https://www.loc.gov/aba/publications/FreeLCGFT/2019%20LCGFT%20intro.pdf>

²⁸ LOC, "Genre/Form Terms"

²⁹ LOC, "Genre/Form Terms"

³⁰ LOC, "Genre/Form Terms"

³¹ LOC, "Genre/Form Terms"

- Meeting
- Oral history
- Oral storytelling
- Music
- Performance
 - Improvised speech
 - Sound poetry
 - Spoken word poetry
 - Sound poetry, Music
- Podcast
 - Audio Essay
 - Narrative
- Reading
 - Drama
 - Fiction
 - Poetry
 - Non-Fiction
- Soundscapes
- Speeches
 - Commemorative works
 - Panels
 - Question-and-answer period
 - Talks
- Text-Sound Compositions
- Workshops (seminars)

Definitions of genre terms³²

Ambient sounds: Recordings of sounds of the surrounding environment external to an audio system that are often recorded separately and mixed into other recordings to enhance live effect.

Autobiographical sound recordings: Based on the narrower terms of the LOC subject heading “Autobiographies,” this term includes sound recordings of memoirs, confessions, personal memoirs and egodocuments.³³

Conversation: The informal exchange of ideas and information between two or more people by spoken words.

Interview: Recordings of formal meetings at which information is obtained (as by a reporter, radio commentator, or researcher) from a person.

Letter: Recordings of written text or extemporaneous speech directed or sent to a person or group of people.

³² www.merriam-webster.com/; other dictionaries and common sense

³³ This genre term is an original variation of the standard LOC narrower terms for “Autobiographies” (<http://id.loc.gov/authorities/genreForms/gf2014026047.html>).

Meeting: Recordings of research and development-oriented task force and team “business” meetings including planning discussions, project progress reports, and work sessions. Meetings of organizers of literary events or projects discussing their organizational work would also apply.

Music: Sonic works produced with musical instruments and/or the human voice that order tones or sounds in succession, in combination, and in temporal relationships.

Oral history: Recorded accounts of historical information about individuals, families, important events, or everyday life, derived through planned interviews.

Oral storytelling: Oral narrative stories delivered by one or more speaker(s) that may draw on or adapt traditional folk story forms. Storytelling differs from oral histories in that the content is generally told to an audience or community with the purpose of engaging and/or entertaining and/or sharing a lesson or knowledge with them in the delivery of a narrative. Further, they are not the result of an interviewer-interviewee dynamic, but are delivered by a storyteller who is self-consciously inhabiting that role on his or her own.

Performance: Recordings of creative works designed specifically for oral performance.

Improvised speech: Recordings of extemporaneous speech produced in the context of a performance.

Sound Poetry: Poetry meant to be performed that emphasizes sounds instead of the semantic meaning of the words themselves.

Spoken word poetry: Poetry that is meant to be performed and that is heavily stressed, metrically regular, and characterized by improvisation, free association, and word play.

Readings: Recordings of the recitation of a literary work.

Drama: A composition in verse or prose intended to portray life or character or to tell a story usually involving conflicts and emotions through action and dialogue and typically designed for theatrical performance.

Fiction: Readings of literature in the form of prose, especially short stories and novels, that describes imaginary events and people.

Poetry: Readings of literature in the form of verse or other literary forms identified with this genre of literature.

Non-Fiction: A wide-range of read materials including criticism, biography, history, etc.

Soundscapes: Compositions that consist of natural or synthetic sounds from specific locations that are sometimes manipulated electronically.

Speeches: A formal address or discourse delivered to an audience.

Commemorative works: Speeches (as in toasts, roasts, eulogies, and encomiums) that act as a memorial or mark of an event or a person.

Panels: Recordings that feature discussions of topics by panels of speakers or experts.

Question-and-answer periods: Recordings that feature speakers or experts responding to questions posed by a live or remote audience.

Talks: Recordings that feature talks or lectures by individual speakers or experts.

Text-sound compositions: Musical compositions consisting primarily of electronically and/or computer-altered or computer synthesized spoken words.

Workshops (seminars): Writing workshops, especially creative writing, or other methods and techniques based gatherings, but also “relatively small instructional sessions or classes emphasizing demonstration and practical application of skills and principles in a specialized field or occupation” that can include a seminar, as in academic (graduate) seminar.³⁴

³⁴ “Art & Architecture Thesaurus Full Record Display (Getty Research).” n.d. Www.getty.edu. Accessed October 3, 2021. <http://www.getty.edu/vow/AATFullDisplay?find=workshop&logic=AND-e=&subjectid=300069765>.

4.15. Statement of Responsibility

Statements of responsibility describe the persons or corporate bodies responsible for the intellectual or artistic content of a work. This definition should be interpreted as broadly as possible to include any and all entities that contributed to the creation, performance or realization of a work. This is similar to the concept of “authorship” but is intentionally much broader.

Categories include:

- Creators of the intellectual or artistic content of the work
- Performers of whose participation is confined to performance, execution, or interpretation
- Performers, narrators, and/or presenters
- Persons who have contributed to the artistic and/or technical production of a resource
- Persons, families, or corporate bodies responsible for the production, publication, distribution, or manufacture of a resource

Special attention should be paid to include the different kinds of contribution relevant to audiotextual works:

- Recordists
- Series organizers
- Collectors
- Archivists

Responsibility can be extended to include not just voices/speakers on a given recording, but other creators/contributors not present. For example, a performance of a poem by another author would constitute a kind of responsibility.

This schema includes two categories of responsibility³⁵ :

Creators: those persons or bodies responsible for the creation, organization, or programming of the recorded event such as a Series Organizer or Producer, and/or the creation of the recorded asset, such as a Recordist;

Contributors: Normally, contributors would include all those who have participated in any capacity, such as a Reader, Performer, Narrator, Interviewer, Interviewee, Speaker, etc.

Normally there will be a limited number of Creators and an indefinite number of Contributors. Both Creators and Contributors may have one or more assigned Roles.

Creator and Contributor Roles

Assign roles to both creators and contributors where known. Role terms should be drawn from the following list:

- Author
- Performer
- Narrator
- Presenter
- Interviewer

³⁵ Our schema draws from the definitions of Creator and Contributor found in the “Dublin Core Metadata Element Set Version 1.1: Reference Description,” Dublin Core Metadata Initiative, Accessed 12 October 2018, v.1, 1, <https://www.dublincore.org/specifications/dublin-core/dces/2012-06-14/>. & “Outline of Elements and Attributes in MODS Version 3.7,” Metadata Object Description Schema (MODS), Library of Congress (LOC), July 31, 2018, Version 3, <https://loc.gov/standards/mods/mods-outline-3-7.html#name,%20v.3>.

- Interviewee
- Producer
- Publisher
- Distributor
- Manufacturer
- Recordist
- Series organizer
- Collector
- Archivist
- Reader
- Speaker
- Storyteller
- Elder
- Donor
- Host
- Sound designer
- Transcriber

A creator or contributor may only have 1 role listed/entry. For repeated roles (e.g. author and series organizer), create separate creator or contributor fields with a role as required.

Creator and Contributor URI Fields Authorized names of creators and contributors should be drawn from established authority lists where possible.

- [Library of Congress Name Authority File](#)
- [Wikidata](#)
- [Dictionary of Canadian Biography](#)
- [Virtual Internet Authority File \(VIAF\)](#)

Enter the URL to the applicable authority record in the corresponding URI fields. For example, if using VIAF, for Irving Layton, choose “Personal Names” for fields to search in VIAF, and then take the permalink from the Irving Layton record <http://viaf.org/viaf/66482092>. For WikiData, provide the QID, for example, search for Leonard Cohen and then copy the correct URL: <https://www.wikidata.org/wiki/Q1276>.

Creator / Contributor Nation: Use this look-up field in Swallow to select Creator/Contributor Nation. The field is used to identify the specific Indigenous Nation(s) with which a creator is affiliated, drawing terminology directly from the First Nations, Métis, and Inuit Indigenous Ontology (FNMIIO)³⁶. Multiple entries are supported, ensuring creators/contributors can be described in relation to all of their community affiliations.

Data Entry Syntax

- In both the creator and contributor fields the following format should be used: Last, First
- For each creator and contributor fields, enter YYYY (birth)-YYYY (death/ - for living creators/contributors). Where exact dates are not known, add a question mark, e.g. 194?-2007

³⁶ First Nations, Metis and Inuit – Indigenous Ontologies (FNMIIO). National Indigenous Knowledge and Language Alliance/Alliance nationale des connaissances et des langues autochtones NIKLA-ANCLA. June 21, 2019. <https://nationalindigenousknowledgeandlanguagealliance.home.blog/2019/06/21/first-nations-metis-and-inuit-indigenous-ontologies-fnmiiio/>.

- Where a creator or contributor is unknown, record as Unknown [role], e.g. Unknown Speaker

Sources of Information

- Creator and contributor fields should be transcribed from the item (the recording) and any accompanying materials (e.g. programs) first, if possible and if the information is deemed reliable/accurate.
- Secondary sources may be used as well (e.g. research works).

Levels of Description Statement of responsibility can apply to different levels of a given resource:

- An entire recording (e.g. Series Organizer Jason Camlot)
- A section of a recording (e.g. a reading by Robert Creeley)

Sample Records (based on various entries from [Robert Creeley Penn Sound author page](#))

EXAMPLE 1: from Exact Change Yearbook c.1963, broadcast by Paul Blackburn on “Contemporary Poetry”						
Creator		Date	Role	Contributor	Date	Role
Creeley, Robert White	Robert	1926-2005	Performer	Blackburn, Paul	1926-1971	Broadcaster

EXAMPLE 2: Ballade pour Robert Creeley, c. 1993						
Creator		Date	Role	Contributor	Date	Role
Creeley, Robert White		1926-2005	Author	Darras, Jacques	1939-	Recordist

Creator/Contributor Notes

- It may be necessary to include creator and/or contributor information in other fields such as a title, general note or table of contents where additional information is required, or the use of a role term is not desirable/complete. For example:
 - From recording Creeley sent to Paul Blackburn, 1963
 - Creeley discusses his life and work and reads poems, with Pierre Joris, to the improvised jazz of Steve Lacy
- It is acceptable to duplicate information in a creator/contributor field with the more detailed explanatory information found elsewhere

4.16. Date

We want our items to have dates associated with them so that we can understand their significance within historical timelines, both in relation to other literary or historical events, and in relation to each other. Determining a date may seem simple, but that is not always the case.

Finding the Information: First, think about where your information is coming from. Chief Source(s) of Information is the source from which the Date is taken or determined. In the case of our project, the chief source of information is, ideally, the sound recording being described, or the unit of description. This includes the object itself as well as any labels, notes or accompanying material. The Chief Sources of Information are one or more of the following resources.

- a. the item itself, including any labels, etc., that are permanently affixed to the item or a container that is an integral part of the item
- b. the container itself (e.g., a box)

- c. accompanying material (e.g., technical recording slips)

If the information is taken from a source other than one of these sources, this *must* be stated in the **Date Notes** field. No square brackets should be used in the Date field to indicate a supplied date. Both the source and an explanation of the supplied information must therefore be provided in the Notes. Even if one or more of the Chief Sources of Information are used, it is still recommended that the source be provided in the Notes field.

Procedure: Perform the following steps as closely as possible in order to catalogue the item:

1. Decide and select which one of the following two Date Types best describes the work:
 - Production Date– when the recording was produced
 - Publication Date- when the recording was broadcast, distributed or first made public
 - Performance Date- when the reading/event was performed
 2. The Date Field is required: this means that the elements of this field cannot be left empty – some value must be entered. Leave all elements blank if the date cannot be determined.
 3. Enter as outlined below according to the prescribed syntax and punctuation. It is very important that all date entries use this specific syntax:
Year: YYYY Example: 1997
Year and month: YYYY-MM Example: 1997-07
Complete date: YYYY-MM-DD
Example: 1997-07-16
 4. Enter into the **Date Notes** field any explanations or additional information that pertains to the date of the item that is not reflected in the date field.
-

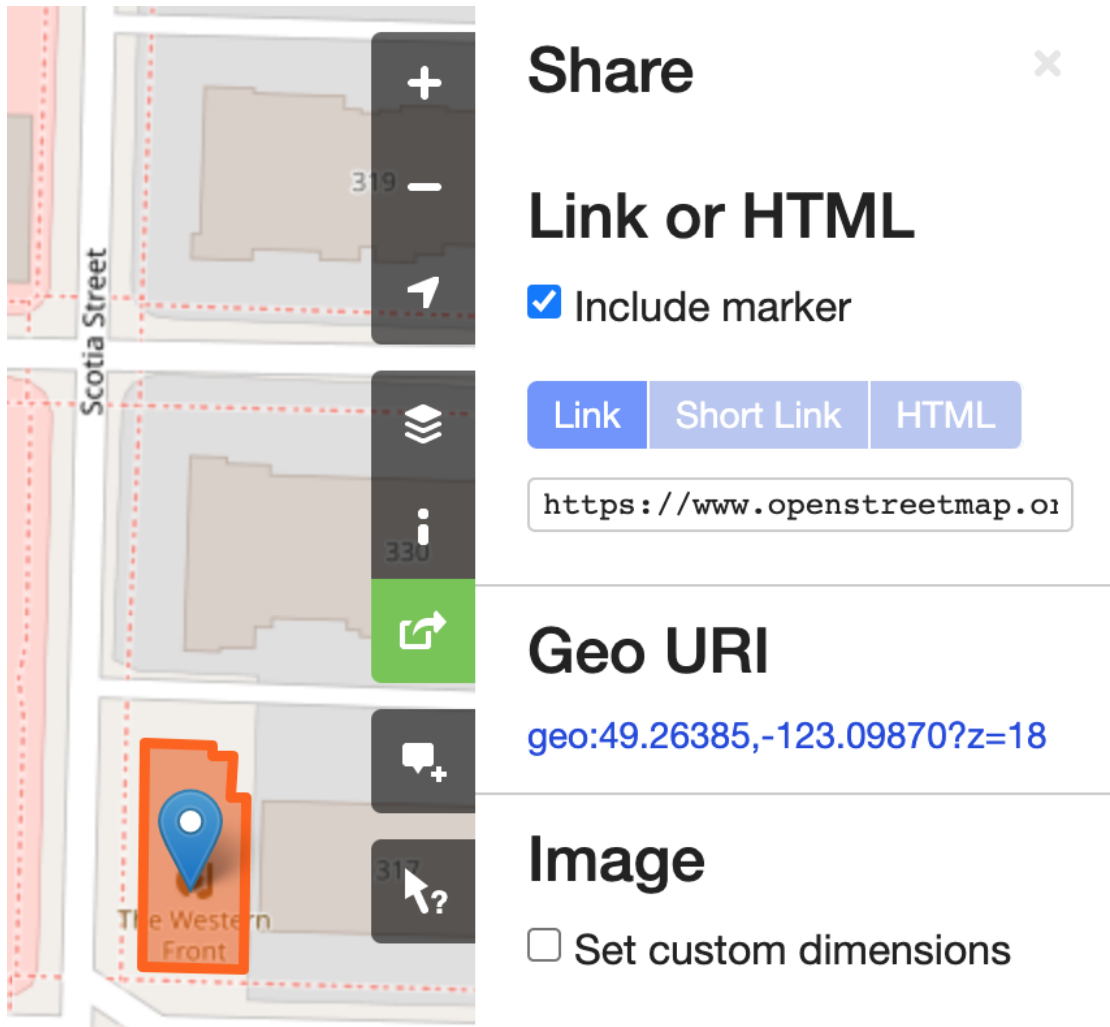
4.17. Location

Where was a recording made? Answering this question may provide us with interesting information about where literary events and activities occurred across the country. We will be using a few methods of capturing location information, and will be entering this data according to a set syntax. The three primary fields related to location are the **Address** (which refers to the official street address of a location), the **Venue** (that is, the name of the venue where something was recorded), and then, as a recommended field, the **Latitude** and **Longitude** of the location as well as a **URI** for that location.

OpenStreetMap includes specific “node” links for entities in OSM that have a unique latitude and longitude, as well as more complex entities such as a **street**, region, city area, country, etc. The more important objective of Location cataloguing is to enter the correct latitude, longitude, address and venue name. The link to OSM is optional, and to be added only if you can find a stable “node” link, or a specific coordinate link.

A “**node**” link contains the word “node” in the link, followed by an ID, for example: <https://www.openstreetmap.org/node/1296620055>

For the link by coordinates, you can use the interface of OSM, as described here: <https://wiki.openstreetmap.org/wiki/Browsing>, so clicking on “share” on the map, and then adding a marker and copying the “link” from there.



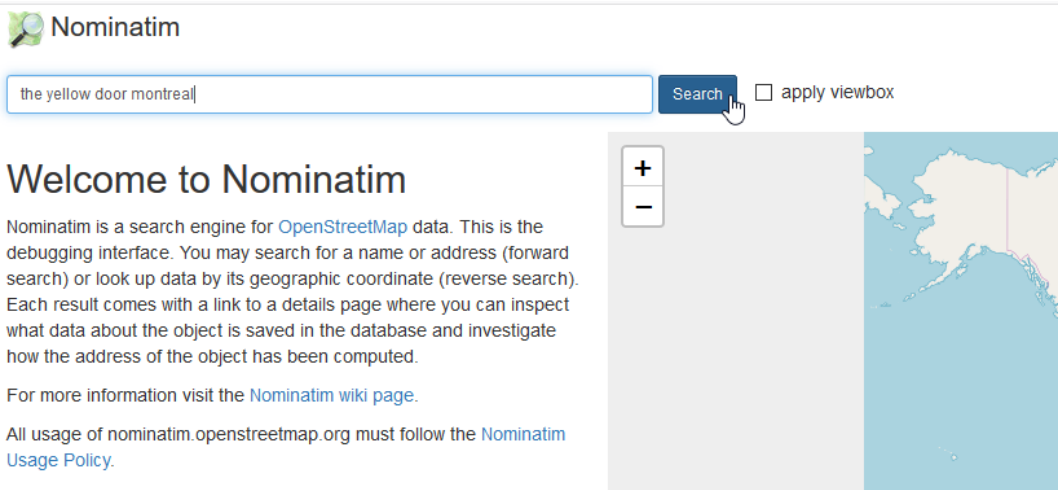
Where to find your information:

Address, Latitude and Longitude and URI: For the Address, LL and URI, use the **OpenStreetMap**

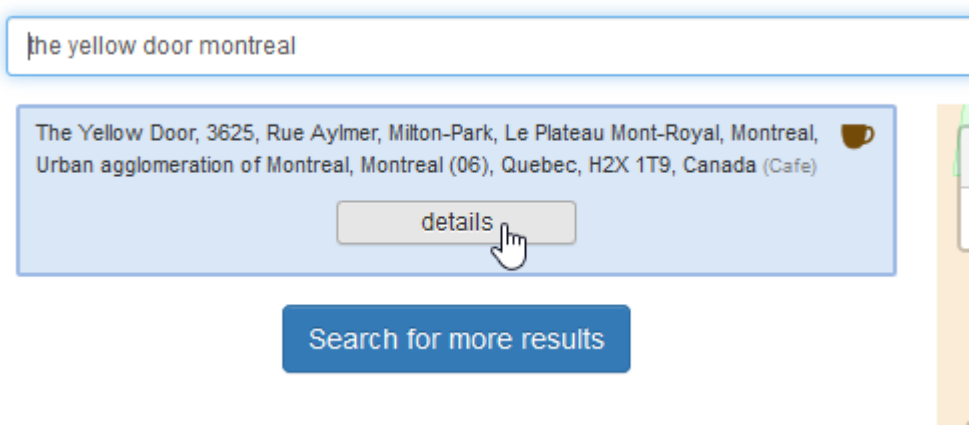
For example, to find the Address of The Yellow Door in Montreal:

- Example: Go to <https://nominatim.openstreetmap.org/>

1. Search -> The Yellow Door Montreal ->



2. Click on “details”:



3. Copy/paste information from the entry for the location in OSM:

The Yellow Door [link to this page](#)

Name	The Yellow Door (name)
Type	amenity:cafe
Last Updated	2020-12-05T21:19:59+00:00
Search Rank	30
Address Rank	30 (house / building)
Coverage	Point
Centre Point (lat,lon)	45.5088401,-73.5781434
OSM	node 5919155489 link to OSM
Place Id (on this server)	66560086
Computed Postcode	H2X 1T9
Address Tags	3625 (housesnumber) Rue Aylmer (street)
Extra Tags	https://www.openstreetmap.org/node/5919155489

Venue: For Venue, you may find a specific location name from the source material.

Cataloguing Procedures and Syntax:

Address	Venue	Latitude	Longitude	URI
required	recommended	additional	additional	additional
3625, Rue Aylmer, Montreal, Quebec, Canada	The Yellow Door (Montreal)	45.508841	-73.578142	https://www.openstreetmap.org/node/5919155489

For **Address**, we prefer the use of House Number, street name, city name, State Province, Country as it appears in OSM (<https://nominatim.openstreetmap.org/>). If it is not found in the OSM database, please supply as much information as is known, for example, the Country name where the recording was made.

The order in which the pieces of the address are given is important, please use the following syntax:

- House Number, Street Name, City, State/Province, Country
 - Example: 1455, Cypress Street, Vancouver, British Columbia, Canada
 - Example: Canada
 - Example: Toronto, Ontario, Canada

For **Venue**, transcribe what is on the source, followed by the name of the city in round parenthesis, for example:

- Example 1: R2B2 Bookstore (Vancouver)
- Example 2: Roy [Kiyooka]'s House (Vancouver)

For **Latitude and Longitude**: Copy and paste the LL numbers you find from Open Street Map.

URI: Copy and paste the URI from Open Street Map.

4.18. Rights Fields

The Rights section includes the Rights field to express the copyright status of a digital object, as well as a License field for information about how you can access and re-use the objects. The Rights statements are designed to be used by cultural heritage institutions to communicate the copyright and re-use status of digital objects to their users. Works that are known to be free from copyright or licensed under a known Creative Commons license should be labelled as such using the License field.³⁷

Required Field: **Rights**

Additional Field: **License**

Additional field: **Access** Options include Closed, Streaming, Streaming and download.

Additional Field: **Rights Note**

Controlled Vocabulary: Use the following two controlled vocabularies for specifying the Rights and License fields:

³⁷ <https://rightsstatements.org/en/documentation/#use-by-cultural-heritage-institutions>

- Rights statements³⁸
- Creative Commons licenses³⁹

One of the following 13 statements should be specified, using a URL:

1. The Public Domain Mark (PDM) - Digital Objects which are no longer protected by copyright. Objects that are labelled as being in the public domain can be used by anyone without any restrictions.

Specify the URL: <https://creativecommons.org/publicdomain/mark/1.0/>.

2. No Copyright - non commercial re-use only (NoC-NC) - public domain Digital Objects which have been digitised as an outcome of a public-private partnership, where the terms of the contractual agreement limit commercial use for a certain period of time.

Specify the URL: <http://rightsstatements.org/page/NoC-NC/1.0/?language=en>.

In addition, in the Rights Note, where possible, publish the first calendar year in which the Digital Object can be used by third parties without restrictions on commercial use, as noted in the contractual agreement.

3. No Copyright - Other Known Legal Restriction (NoC-OKLR) - public domain Digital Objects that are subject to known legal restrictions other than copyright which prevent their free re-use.

Specify the URL: <http://rightsstatements.org/page/NoC-OKLR/1.0/?language=en>.

In addition, in the Rights Note, a link to a page detailing the legal restrictions that limit re-use of the object, or a free text description of the restriction.

4. In Copyright (InC). use with in copyright Digital Objects which are freely available online and where re-use requires additional permission from the rights holder(s).

Specify the URL: <http://rightsstatements.org/vocab/InC/1.0/>

5. In Copyright - Educational Use Permitted (InC-EDU). in copyright Digital Objects which are freely available online and where the rights holder(s) have allowed re-use for educational purposes only.

Specify the URL: <http://rightsstatements.org/vocab/InC-EDU/1.0/>

6. Copyright Not Evaluated (CNE) - use with Digital Objects where the copyright status has not been evaluated.

Specify the URL: <http://rightsstatements.org/vocab/CNE/1.0/>

7. The Creative Commons CC0 1.0 Universal Public Domain Dedication (CC0) - used to waive all the rights in a Digital Object. By applying this waiver, all possible existing rights in the content are waived, and the objects can be used by anyone without any restrictions.

Specify URL: <http://creativecommons.org/publicdomain/zero/1.0/>

8. Creative Commons - Attribution (BY). lets others distribute, remix, tweak, and build upon the licensed work, even commercially, as long as they attribute the rights holder as described in the licence. CC BY is recommended to enable access, discovery and use of licensed works.

Specify the URL: <http://creativecommons.org/licenses/by/4.0/>

9. Creative Commons - Attribution, ShareAlike (BY-SA). lets others remix, tweak and build upon the licensed work, even for commercial purposes, as long as they attribute the rights holder as described in the licence, and license their adaptations of the work under the same terms. All new works based on the original licensed work will carry the same licence, so any derivatives will also allow commercial use.

Specify the URL: <http://creativecommons.org/licenses/by-sa/4.0/>

³⁸ <http://rightsstatements.org/page/1.0/?language=en>

³⁹ <https://creativecommons.org/licenses/>

10. Creative Commons - Attribution, No Derivatives (BY-ND). licence allows for redistribution, including commercial and non-commercial use of the work as long as no alteration is made to the work and the rights holder is attributed according to the specifications of the licence.

Specify the URL: <http://creativecommons.org/licenses/by-nd/4.0/>

11. Creative Commons - Attribution, Non-Commercial (BY-NC). Lets others remix, tweak, and build upon the licensed work for non-commercial use. Any new works created and based on your work must be attributed to the rights holder as specified in the licence, and may be available for non-commercial use only.

Specify the URL: <http://creativecommons.org/licenses/by-nc/4.0/>

12. Creative Commons - Attribution, Non-Commercial, ShareAlike (BY-NC-SA). Licence lets others remix, tweak, and build upon the licensed work for non-commercial use as long as they attribute the rights holder of the work under the terms specified in the licence, and license new creations under identical terms.

Specify the URL: <http://creativecommons.org/licenses/by-nc-sa/4.0/>

13. Creative Commons - Attribution, Non-Commercial, No Derivatives (BY-NC-ND). The most restrictive of the six Creative Commons licences, only allowing others to download the licensed works and share them with others as long as they attribute the rights holder as specified in the licence, but users cannot change the work in any way or use them commercially.

Specify the URL: <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Cataloguing Procedures: The rights statements fall in four categories:

1. Statements for works that are in copyright (Choose #4 or #5)
2. Statements for works that are not in copyright (Choose #1, #2 or #3)
3. Statements for works where the copyright status is unclear or unknown (Choose #6)
4. Creative Commons. All Creative Commons licenses and legal tools can only be applied by, or with the permission, from the rights holder. If the rights holder granted a Creative Commons license (Choose #7, #8, #9, #10, #11, #12, #13) In addition, select the known Creative Commons license in the License field.

Example (Unknown/unclear):

<https://montreal.spokenweb.ca/sgw-poetry-readings/phyllis-webb-at-sgw-1966-roy-kiyooka/>

Rights: <http://rightsstatements.org/vocab/CNE/1.0/>

4.19. Related Works

Related works refers to books or other publications that are either referred to or read from in a literary recording. Citations of works should be provided in the MLA 9th Edition style guide.

The basic MLA format for a book title is:

Author's Last Name, Author's First Name. Title (in Italics). Publisher, Year.

You may refer to the [Purdue OWL](#) (Online Writing Lab) for citation instructions for different kinds of publications.

URL: If possible, provide the link to the correct edition of the item as it appears in the [WorldCat library catalogue](#).

If a link to the item is not available in WorldCat, provide any other link that offers information about a work, or a direct link to the publication or item in question.

Examples:

Foster, Edward Halsey. "Gay Literature: Poetry and Prose". The Oxford Encyclopedia of American Literature. Jay Parini (ed). Oxford University Press, 2004.

URL: https://www.worldcat.org/title/oxford-encyclopedia-of-american-literature/oclc/769478515&referer=brief_results

Ward, Geoff. "John Joseph Wieners, poet, Jan. 6th 1934 - March 1st 2002". The Independent, 15 March 2000.

URL: <https://www.independent.co.uk/news/obituaries/john-wieners-9143191.html>

4.20. Contents

The contents field will be developed through a process that involves listening, timestamping important moments in the recording, and, at times, research in order to determine correct names and titles relevant to the recording. This combination of timestamped titles and names will function as the equivalent of a table of contents for a sound recording and may eventually be used to facilitate the online navigation of a recording.

The basic procedure for generating a timestamped contents list entails using a transcription software that allows you to control the listening speed of a recording, enter notes through an automated timestamping mechanism provided by the software, and then export that information so that it can be pasted into the content notes field.

As you are producing your timestamped contents description of the recording you should timestamp and thus signal the start and end of each event of discrete significance. For our purposes this will usually mean changes in speaker and discrete literary works read. For example, If you are listening to a reading that has someone introducing the reader, you would first timestamp the beginning of the introduction and title this with the term Introduction and then the name of the Introducer. For example, Warren Tallman introduces Robert Creeley. Then, when the reader steps up to the microphone you would indicate that with a new timestamp followed by the name of the reader. If the reader is introducing a poem, you would follow his/her name by Introduces and then the title of the poem, for example, Dorothy Livesay introduces Outrider. Then when the reading of that actual poem begins you would timestamp that moment with the name of the reader and the title of the poem, so, Dorothy Livesay reads Outrider.

4.20.1. Overview and Purpose

The Contents field serves to describe the audible or audiovisual (AV) content (speech and other sounds, video image) of the audio asset. By simple analogy, it can be understood as equivalent to the table of contents (TOC) of a book, using time-stamps instead of page numbers, but it has the potential to be much more detailed than a typical TOC, and to be linked to other data sources, if the cataloguer has the time and resources to make it so. **For the SpokenWeb schema, this field may exist on a continuum from containing no data at all to full transcription and detailed description of the sonic or AV signal, with added Wikidata links, where possible.** In principal, it is always more useful to have some information about what a recording contains than none, and the more information provided the more useful this field will become. That said, there are particular formatting requirements that we have established for providing information in the Contents field. Some contents information that a cataloguer holds may be more suitable to the Contents Note field than the Contents field, proper, if the information does not conform to the grammar of the field, as outlined in this guide.

Beyond providing basic information about the content of an audio or AV recording, the primary purpose of this descriptive work is to make a recording easier for a user to navigate according to access points of information about the identities of speakers, what has been said, and actions are audibly and/or visibly registered in the recording. By timestamping such points of information along a timeline of the audio- or AV-documented event, the Contents field may be used to help users move across segments of a recording according to the points of greatest interest to them. It creates unique access points. This approach to documenting the contents of a recording, according to a set syntax, grammar

and punctuation, also facilitates searching for, and locating, data points at particular moments within a sound recording and across a wide range of recordings.

4.20.2. General Process of Creating Contents Data

The Contents field is developed through a process that involves listening, looking, timestamping important moments in the recording, and, at times, research in order to determine correct names, titles, and other kinds of information relevant to the recording. The basic procedure for generating a timestamped contents list entails using a transcription software that allows you to control the listening and viewing speed of a recording, enter notes through an automated timestamping mechanism provided by the software, and then export that information as a text file so that it can be pasted into the Contents field in Swallow.

There are many tools for Mac and PC that may be used, some for audio-only, and some that can handle both audio and AV assets⁴⁰. Tools may facilitate manual timestamped transcriptions or generate automated transcripts. In either case, the goal is to produce an accurate and informative transcript of the actions and speech of a recorded event with time stamps that will help users navigate the contents of the recording. The transcription process will be indicated as a note in the Contents Note field (see 4.21 below). Details about the transcription tool and process used may also be added as a Preservation Note (see 4.22 below).

Our schema accommodates either Linear or Nested approaches to timestamping. A Linear approach timestamps sonic or AV events as sequential points on a line. A Nested approach may introduce hierarchies between sonic events, so that a series of smaller events can be framed within a larger set of labelled categories.

Linear Timestamping

As you produce your timestamped contents description of the recording, you should timestamp and thus signal the start of each discrete event of significance. For our purposes this will usually mean changes in speaker (as in a conversation or interview); changes in reader and discrete literary works read, and possibly sounds of audience response such as applause and laughter (as in a poetry reading event); changes in scene or significant physical actions (in the case of video documentation); and, other sound and AV events that seem to warrant their own timestamped segment. In the linear timestamping mode, timestamped segments function on a continuous line of discrete descriptions, and subsegments are not nested within larger ones. In other words, if an action occurs in the middle of an ongoing action (say, spontaneous audience laughter in the middle of the reading of a poem, the dropping of a wine glass during a recorded conversation), the interrupting action would be labeled as its own described event (the laughter, the glass breaking), and then the description of the previous event (the reading of that single poem, the thread of the conversation) would resume at the next time stamp. No explicit hierarchy is established between the time-stamped events in this approach.

Nested Timestamping

We have mostly used linear timestamping for SpokenWeb collections. However, the SpokenWeb team at the University of Alberta worked with a nested timestamping format through ERA A+V. For a detailed discussion of their process

⁴⁰ Transcriba for Mac was the first transcription software used for timestamping SpokenWeb audio collections contents. As of 2020 it is no longer supported, but there are other options available including:

[ELAN \(MPI\)](#) – An open-source annotation tool widely used in linguistics and oral history research. It allows users to create detailed, multi-layered, time-aligned transcripts for audio/video files. Requires downloading and installation.

[oTranscribe](#) – A free, open-source web app for manual transcription. It integrates playback controls with a text editor, making it easy to pause, rewind, and insert timestamps without leaving the editor.

[Parlatype](#) – A lightweight open-source transcription player (Linux) that syncs audio playback with text entry and allows quick insertion of timestamps. Requires installation.

[Transkriptor](#) is a commercial tool that is user-friendly and efficient for generating time-stamped transcripts quickly.

see^{41, 42}.

Degrees of Granularity

As already mentioned, the time-stamped contents field may range in granularity of description from

- minimal timestamping of speakers or sound events,
- to more robust time-stamped identification of speakers, titles, non-speech sound events, and content-originating keywords (see *Principle of Keywords*),
- to complete time-stamped speaker identification, title identification, non-speech sound events, content-originating keywords, and full transcription of all speech content.

The three basic degrees of granularity just described correspond roughly to the “three levels of indexing” outlined by the [guidelines for oral history indexing](#) of the IMLS Oral History Metadata Synchronizer tool.

Normally, all entries of a specific collection will be described at the same level of granularity. The cataloguing team will thus need to make some decisions about how detailed their Content entries for a collection will be prior to cataloguing it..

⁴¹ ERA A+V allows users to add hierarchical navigational structures to any file. These structures can be created in XML, or through a graphical user interface that generates XML when changes are saved. When working with the graphical interface, users can see the waveform of the file in question and easily start and stop the recording as they work through the timestamping process.

The highest-order label is the “Title,” which is the name of the event. After that, subsequent labels called “Headings” or Divs are given to each speaker if there are multiple speakers in an event. Within each Heading are “Timespans” or Spans, which are the specific utterances made by a speaker over time—a poem, for example. Each of these Timespans has its own label, such as the title of a poem. Each Span has a beginning (when a poem begins) and an end. In ERA A+V, distinct sections of a file (or, in the case of audio interviews or literary readings, different speakers within an event) can be given their own distinct labels (divs). The higher-level labels or Divs do not have their own total timestamp encompassing the contents within it (the Spans). However, the spans themselves are clearly marked individually.

In the example of a poetry reading below, the title is “Margaret Atwood and Dorothy Livesay, 1969-02-20,” which appears at the top of the graphical interface, and is one of the first lines in XML. All subsequent Headings (the names of the speakers) are subordinate to the title. Atwood and Livesay are the only two performers here; each gets a Heading or div. Each performed poem is represented by a subordinate Timespan. Below, the first Timespan is given the label “The Shrunk Forest,” which is the name of the first poem that Atwood reads. The Timespan begins at 00:00:00.00 (the beginning of the event) and ends at 00:01:07.00, with the poem’s conclusion. After a brief pause, which has no Timespan, Atwood introduces the next poem, “Two Versions of Sweaters,” and the process repeats. Later in the event, when Livesay takes over, she gets her own Heading or Div, her poems are represented as Timespans and given labels for their titles, and the process continues.

⁴² Example XML representing nested timestamping in ERA A+V:

```
<item label="Margaret Atwood and Dorothy Livesay, 1969-02-20">
  <div label="Margaret Atwood">
    <span label="Atwood performs &quot;The Shrunk Forest&quot;" begin="0:00:00" end="0:01:07"/><span label="Atwood introduces and
reads &quot;Two Versions of Sweaters&quot;" begin="0:01:09" end="0:02:24"/><span label="Atwood performs &quot;Woman Skating&quot;"
begin="0:02:28" end="0:03:53"/><span label="Atwood performs &quot;Even Here in the Cupboard&quot;" begin="0:03:56" end="0:04:20"/><span la-
bel="Atwood introduces &quot;Christmas Tree Farm, Oro Township&quot;" begin="0:04:21" end="0:04:37"/><span label="Atwood performs &quot;Christ-
mas Tree Farm, Oro Township" (in five sections)" begin="0:04:38" end="0:04:48"/><span label="Atwood performs &quot;II&quot;" be-
gin="0:04:49" end="0:05:13"/><span label="Atwood performs &quot;III&quot;" begin="0:05:14" end="0:05:31"/><span label="Atwood performs
&quot;IV&quot;" begin="0:05:32" end="0:05:49"/><span label="Atwood performs &quot;V&quot;" begin="0:05:50" end="0:06:21"/>
  </div>
  <div label="Dorothy Livesay">
    <span label="Livesay introduces Nisei" begin="0:06:50" end="0:08:09"/><span label="Livesay performs selections from Nisei" be-
gin="0:08:10" end="0:13:51"/><span label="Livesay introduces another section of Nisei" begin="0:13:52" end="0:14:33"/><span label="Livesay
performs another section of Nisei" begin="0:14:34" end="0:18:12"/><span label="Unknown Host's concluding remarks" begin="0:18:23"
end="0:19:33"/>
  </div>
</item>
```

4.20.3 Preparatory Decisions Prior to Content Description

In addition to choosing between a Linear or Nested approach to timestamping, there are two primary decisions that should be made about the approach that will be taken to describing the audio/AV contents of a collection. As mentioned, the first pertains to the level of granularity of the description. The second pertains to how the digitized or digital files to be described will be handled in relation to the events they might document and the primary entity of description. Let's take a moment to think about some factors and considerations surrounding these important preparatory decisions.

1. Granularity of Description Depending on human and other resources available, and the cataloguing team's sense of the relative usefulness of the kind of description that would be most useful for research and teaching with a collection, the cataloguers may decide that only a very basic description of the contents of a recording is necessary or possible. Ideally, this would provide at least some basic information about what the recording contains, such as the name(s) of reader(s) or speaker(s), the titles of works read, and/or a brief list of Keywords capturing subject matter content spoken in the audio itself (see *Principle of Keywords*). In its most basic form, such information would be provided without timestamps but simply as a general description of a recording's content without intent to identify "where," or "when," in the recording particular audible or visible events occur. This approach might be taken for a collection that consists of a very large number of recordings, or if the complexity of the content is such that a more detailed, timestamped description is deemed out of scope with the resources available for cataloguing the collection.

If resources are available, and a more detailed, timestamped description seems justified, then the cataloguer(s) should decide how granular the description will be. In making this decision it will be useful to ask yourself which of the following kinds of information are a priority, and feasible:

- a. Speaker identification
- b. The inclusion of titles of works read
- c. The inclusion of titles of books from which works read have come
- d. The inclusion of Wikidata Item Identifiers (Q-codes) for titles of books
- e. The inclusion of Keywords from content heard and seen
- f. The inclusion of Wikidata Q-codes for selected Keywords
- g. The identification of non-speech sound events such as applause, laughter, etc.
- h. The inclusion of full transcription of all extra-poetic speech (speech other than reading)
- i. The inclusion of full transcription of all speech heard on the recording (whether read, performed, spoken, etc.)

As mentioned above, decisions concerning the granularity of Contents description made for a particular collection may be determined by a variety of factors including the degree of complexity of the audiovisual content, the relative value of certain layers, kinds and categories of description in relation to needs identified for research and teaching, the size of the collection, and, the resources (human labour, tools, infrastructure) available for the work. We recommend that that same level of granularity, or detail, be maintained in the description of all assets from the same collection

The Inclusion of Wikidata Item Identifiers (Q-Codes) for Names, Titles and Keywords If resources permit, we recommend that cataloguers include Wikidata Q-Codes in square brackets next to names of people, places, book titles, and other Keywords that may have Wikidata entries. Including such linked data in timestamped descriptions will make our descriptions all the more useful and discoverable to researchers and students.

The basic Wikidata search bar can be found at this [link](#).

2. Relationship of Digital Files to the Primary Entity of Description In addition to planning how detailed your Contents description will be, it is also important to make some basic decisions about the relationship between the digital files that contain the content, and the primary, organizing entity that defines what, exactly, is being described in a Swallow entry (and its Contents field). This is especially important in cases where the primary entity of description (say, an event that took place on a particular date, over a particular period of time) exists on two or more digital files.

To some degree we are applying the definition of entity used by AtoM: "An entity is an object about which an information system collects data." We are, however, extending this definition to include the use of an entity as a primary organizing

principle. In other words, our schema allows an entity to function as a means for organizing the data related to it. In this schema, a primary, organizing entity could be a material asset(s) or digital file(s). It could also be an activity, a segment of an activity, or a particular group of activities, also known as an event.

Using an event as the umbrella or top level category to organize related files or assets aligns with traditional archival description—multi-level and hierarchical. Once you choose the top-level, or organizing entity, you provide detailed descriptions of the subordinate records, which might be analog and/or digital. Further, if choosing a digital or analog file(s) or asset(s), as the primary, organizing entity, there may be no need to organize the records further, hierarchically.

Based on these definitions, let's start with a more straightforward example, first. Say that a collection of recordings captures readings from a poetry series. Each event in the series lasted one hour and each was captured as a unique mono (single track) reel to reel tape recording. There were ten events, and so, ten reels of tape. Each of the ten reels was then digitized by producing a single digital WAV file that captured the contents of each of the ten reels. So: ten WAV files, capturing the contents of the ten mono tape reels, that document ten discrete poetry reading events (in a series). In this example, there is a one-to-one correlation between analogue asset, digital file, and documented event. The Contents field will describe the contents of the digital file, and in effect will also be describing the contents of the original analogue asset (the single reel of tape) and the single event from a series of ten that was captured on tape. You may decide that the primary, organizing entity of description in a Swallow entry is the original analogue asset, or this historical event, or the digital file. In each case, the primary entity of description is identified with a single digital file. While both analogue and digital assets will be described in Swallow, the selection of the primary, organizing entity will determine the object that determines a hierarchy of description, and which entity comes to organize subsequent decisions about description, what gets described where, and to what degree of detail. This principle of an organizing entity of description is based on an interest in access. It responds to the question: what entity will make these audiovisual materials most useful in a digital presentation of the content to a user?

When multiple digital and analog assets are associated with a single event, such as a reading series, more complicated scenarios may arise. There may be, for example, multiple digital assets associated with one or more analogue recordings associated with a particular event. In such cases, the cataloguer will need to decide whether they are creating unique Swallow entries for each digital file available, or, whether either the analogue asset (if there is one) or the original event (or some portion of each event) will function as the primary, organizing entity of description that determines a Swallow entry. Later we will illustrate scenarios in which analog or digital files are the primary, organizing entities of description. Briefly, here, however we provide three examples that illustrate an event as the primary, organizing entity:

Example A) In the case of the Sir George Williams Poetry Series collection, a reading event was documented on between one and three reels of tape. When digitized, each reel resulted in a unique digital file. In Swallow, the overarching entity that defines an entry is the reading event itself. Therefore, for each entry in Swallow between one and three digital files is described in the contents field.

Example B) The Words and Music Show (Ian Ferrier) collection was partly digitized from MiniDiscs and partly delivered as born-digital files on hard drives. For this collection there may be as many as two or more digital files that, together, document the performances of a single evening's event. For this collection, the dated event serves as the entity that is documented in a single entry in Swallow, and the contents field may contain time-stamped descriptions of two or more digital files, with the timestamps for each file beginning at 00:00:00.

Example C) Much of The Ultimatum (Alan Lord) Collection was recorded on 1/4" 8-track reel to reel tape, meaning, we have multi-track renderings of the performances. In this instance, there may be as many as eight individual tracks, each one rendered as a unique digital file, for a single artist's performance. Further, the event of a single evening comprised of multiple performers, may have been recorded over multiple reels, with some performances using only a few tracks and others more. In this case, the cataloguing team decided to use "the performance set" (the slotted performance of a single artist or act) as the primary, organizing entity of description. Each Swallow entry describes a single set, noting the analogue assets and digital files associated with that particular performance set, and providing a timestamped contents description of a single digital file that consists of a multitrack mixdown of the individual tracks that documented that set.

Deciding the primary, organizing entity of description for an entry in Swallow prior to cataloguing allows for consistency in managing the relationships of assets to events in the description of entities that comprise a collection.

The following sections will explain, with examples, the prescribed grammar (the rules about standard terms, punctuation, and other structural elements) for the Contents Field. The Contents field grammar begins with the core elements of a discrete time-stamped descriptive entry. In all cases, the time-stamped Contents description is built around three key elements: The Speaker or Descriptor, the Numerical Timestamp, and the Descriptive Label. Our explanation of the Contents field's grammar begins with definitions of these three elements.

4.20.4. Contents Field Grammar and Controlled Vocabulary

1. Core Elements of a Time-Stamped Contents Field Description A timestamped description signals and provides information about a sonic or audiovisual event. It is composed of three elements: (A) The Speaker or Descriptor, (B) The Numerical Timestamp, and (C) The Descriptive Label.

Example:

Heather Hermant (A)

00:34:43 (B)

Introduces “The Long-Distance Runner”, with music by Tom Walsh. | Keywords: sentimental; 1999; Budapest; Trafo House of Contemporary Arts; Swifty Lazarus; Grace Paley. (C)

A) The Speaker or Descriptor (short name: Descriptor)

This first element identifies the agent behind the time-stamped sound. When you know the full name of a speaker, repeat it for every discreet timestamp attributed to them. Stage names and aliases function like full names. Ideally, the name of a speaker should correspond to a contributor listed in Swallow. To help with linkability, you can note aliases in the speaker's contributor field.

When several individuals perform under one name, that group name should be the recurring Speaker. If individual group members' names are known, they can be listed in square brackets in the first timestamp (but do not need to be listed thereafter). If a single, identified group member speaks around the performance, the timestamp should be attributed to the individual.

Examples:

Swifty Lazarus [Todd Swift, Tom Walsh] 00:00:28 Performs “Love” from The Envelope Please.

Todd Swift 00:02:00 Thank you!

There will also be situations where the linkable name of a speaker is unknown. First names, nicknames or other identifiers can be used where they are available. When you have exhausted these options, you can list an unidentified speaker as “Unknown Speaker”. You should assign a sequential number to every discernable unknown speaker within the content of an asset or file, in order of their appearance (ex. “Unknown Speaker 2”).

Descriptors are used for sonic events that are not attributed to a speaker. One prominent instance is “Audience”, to which you can attribute applause. Likewise, “Unknown” indicates that the source of the sonic event is unknown, but implies that the sonic event is not speech. (eg. Audience, unknown, end)

B) The Numerical Timestamp (short name: Timestamp)

The timestamp marks the beginning of the descriptive entry. It follows the format HH:MM:SS (Hours:Minutes:Seconds). The end-time of a sonic event is not required in the Linear approach of the SpokenWeb schema. (When text is converted to XML the end timestamp will be assumed to be the beginning of the next timestamp on the line.) In the Nested approach, end timestamps are intentionally marked. Some time-stamping software will also include milliseconds following the HH:MM:SS numbers, so the timestamp would read HH:MM:SS:mm. Including milliseconds is not required, but is acceptable.

The cataloguer should do their best to time-stamp a described sonic or AV event as accurately as possible within the pre-determined parameters of granularity. There is no set rule about the required minimum or maximum length of a time-stamped segment; decisions about what counts as a sonic or audiovisual-event, apart from the separation of one literary work from another, and speech that is expository or explanatory (extra-poetic speech) from read or performed material (poetic speech), are at the discretion of the cataloguer, in coordination with any particular rules that may have been developed in relation to the specific collection that is being described.

C) The Descriptive Label (short name: Label)

The label holds the description of the sonic or audiovisual event as well as keywords. When both are used, a pipe [|] separates both portions. The descriptive portion exists on a spectrum from short description to full transcript. The keywords are a list of linkable data points (access points) present in the described sonic or AV event. There should always be something in the label. For instance, you can use markers of silence or uncertainty such as [silence] or [?] to avoid leaving the field blank. The one exception is the END timestamp which may be left blank. While the format of the label might differ between institutions, cataloguers should normally maintain the same degree of specificity or granularity across a given collection.

2. Overview of Controlled Vocabularies and Grammar There is no exhaustive list of terms to use in every possible timestamping context and many descriptive situations will be formulated at the cataloguer's discretion. However, in order to create cohesion across a wide range of collections catalogued at different sites, we have developed a select controlled vocabulary to be followed whenever possible.

Describing Key Actions: In the majority of cases when sound is attributed to one speaker (whether in sound or AV recording), the label should begin with a present tense verb, followed by one or a series of nouns providing essential information regarding that action. In performative contexts the verb "Performs" should be used primarily, while other terms that imply a more specific type of performance should be used when directly referenced in the recording.

Examples:

Introduces _____ (event, names, titles)

Performs _____ (all encompassing/avoids presumption of intention)

Reads _____ (if indicated)

Sings _____ (if indicated)

Resumes _____ (used when a previously identified sound event [say, a reading of a particular poem] resumes following interruption by another identified sound event [say, applause or laughter])

Asks _____

Addresses _____ (used when a speaker is addressing an individual or the audience as a whole directly)

Discusses _____

Announces _____ (for example, announcing intermission between sets, announcing end of event, etc.)

Promotes _____ (used when host or artist promotes a work or event, i.e. a book for sale at a book table)

After the initial verb and accompanying description in a label, subsequent descriptive language in the same annotation is not required to follow the same formula.

Example:

Robin Blaser
00:14:55
Asks question, exchange with Warren Tallman follows

Annotations of audible moments of communal responses like applause and laughter from the audience do not require the use of a present tense verb in the label and should be treated as follows:

Audience
01:44:36
Applause

Audience
02:03:33
Laughter

For sounds made by specific but unknown individuals from the audience, the chosen speaker should be specified by a number.

Example:

Audience Member 1
00:33:09
Addresses Warren Tallman

If a notable sound cannot be attributed to a particular agent, the speaker should be named “Unknown.” Most often this will apply to the label “Ambient Sound.” If the cataloguer wishes, any additional remarks about the nature or quality of the sound can be written between square brackets. This formula can also be used for notable moments absent of sound, using instead the label “Silence”.

Examples:

Unknown
00:01:22
Ambient Sound [loud bang]

Unknown
00:02:30
Ambient Sound [voices]

Unknown
00:02:30
Silence [pause, or muted, or erasure, etc.]

In the production of both full transcripts and timestamped descriptions, it may be necessary to note when human speech becomes inaudible or difficult to discern. If you are unable to work out what is being said, use the term “unintelligible” between square brackets. When you are able to make an educated guess about something difficult to hear, the word or sequence of words should be sectioned off with square brackets with the addition of a question mark in parentheses.

Examples:

Ian Ferrier 00:10:14 I am going to read [unintelligible].

OR

Ian Ferrier 00:10:14 Introduces [unintelligible]

Ian Ferrier 00:10:14 I am going to read that other [poem (?)] later.

Describing Video A group called the Audio Description Coalition (ADC) was formed in 2006 to document best practices and standards for video description, producing “Standards for Audio Description and Code of Professional Conduct for Describers”, initially published in 2007 and updated in 2009. The document is available online at: https://www.perkinselearning.org/sites/elearning.perkinsdev1.org/files/adc_standards.pdf.

The founding members of ADC were actively involved with live description of performances and museum exhibits. This document is intended to assist those practicing and learning to become professional audio describers. It outlines the basic principles of providing audio description that “helps to ensure that people who are blind or have low vision enjoy equal access to cultural events by providing the essential visual information”. Although the context of this document is that of improved access for those who are blind or have low vision, the basic principles described can also be applied, with common sense and practice, to our context of describing video for research use. The primary audience for our descriptions is comprised of humanities scholars, so cataloguers may calibrate their focus in description with this audience in mind. The basic idea is to describe what you see, as objectively as possible. That means describing physical appearances and actions, rather than motivations and intentions. The gestures and facial expressions of characters are visible and so should be described, but motives and reasoning are not visible and so are not subject to description.

Basic practical approaches and actions we recommend for describing video:

- We suggest that the cataloguer begin with a quick scan (by scrubbing across the video) to get a sense of the main transitions that exist in the video content to be described. This will help the cataloguer gauge the number of video description timestamps that will be required in describing a recording, and to decide upon the degree of granularity of the description to be performed.
- Once the content of the video as a whole is assessed, the first description may be used to provide a full account of the scene or setting (as with the opening set description in a play), allowing for subsequent timestamped descriptions to be shorter and more action-oriented. This will be so especially in videos that document an event that takes place in the same setting throughout the action. In such cases, the opening description may contain more information, and be longer, than subsequent time-stamped descriptions.
- The primary descriptive mode should consist of indicators of what is visible on the video only, and not what is presumed to be happening. The focus should be on actions, verbs (standing, jumping, swaying) and things, nouns (holding a microphone, holding a book, smashing a plate). The use of nouns and verbs that offer precision concerning what is seen are welcome, but the cataloguer should be careful not to project their own assumptions onto what is seen through the nouns and verbs they chose to use.
- The cataloger should focus on descriptions of what are determined to be the most significant actions of agents, descriptions of the most significant things observed. As a rule, timestamped video description can proceed at a high level (not overly granular). Do not attempt to describe everything. Think about what would be most relevant to a user-base of literary and cultural history scholars, the primary audience for our descriptions.
- The cataloguer should avoid using adjectives and adverbs that offer value judgements, presumptions and interpretive assessments of what is seen. For example, instead of saying that a poster, photograph, or a performer’s clothing is beautiful (this is a matter of opinion, a value judgment), do your best to describe the things observed that may have caused you to make that assessment. Describe the colors and text that appear on the poster; describe what is seen in the photograph (“a human figure standing before a house”); describe what the clothes look

like (“a black dress”).

We suggest the following two approaches within a grammar for integrating video contents descriptions into the timestamped Contents description of an AV artifact.

Approach 1: The first approach allows the cataloguer to timestamp a visual event on its own, as a distinct contents event (that is to say, distinct from audio content). The cataloguer, in this case, selects the primary “visual event” that warrants the timestamped description, and describes it in square brackets before the timestamp. An event may be an object, an agent (speaker, reader), or an action of the video camera. This short, bracketed description of the timestamped visual event is followed by a carriage return, the timestamp, and then, a longer description of the visual event, following another carriage return. Any visual content description must be signalled first with the phrase “Video Description” followed by a colon [:] .

[Short Indicator of Visual Event]

00:10:14

Video Description: The content of the video description

Examples:

[Ceiling Fan]

00:10:14

Video Description: A ceiling fan spins.

[Ian Ferrier]

00:10:14

Video Description: Ian Ferrier bends to pick up a microphone.

[Camera Pan]

00:10:14

Video Description: The camera pans from left to right.

Approach 2: The second approach may be used when adding video description to a timestamped segment of audio. In this case, one adds video description to the timestamped audio description or transcription by inserting a pipe | sign, followed by the phrase “Video Description” and a colon [:] after any audio description that has already been provided.

The generic form of this grammar is as follows:

Speaker

Timestamp

Audio description | Video Description:

Example:

Ian Ferrier

00:10:14

I am going to read a brand new poem. | Video Description: Ian Ferrier strums an electric guitar.

As noted above, the opening description of a scene or setting may be more detailed than subsequent descriptions of the setting, and subsequent description of actions within that first described setting need not repeat details of this first description. Subsequent descriptions should emphasize new information. You begin with a general description of the scene, and then refer to specific events within that scene.

Example of a video description sequence:

[background noise]

00:00:00

[Music and crowd voices]. | Video Description: Grey visual noise

[background noise]

00:00:35

[Music and crowd voices. Previous song cuts out and a new one begins.] | Video Description: Event poster held by two hands. Poster reads “Les Mardis de L’Oeil Rechargeable Ultimatum Presentent de Londres Kathy Acker.” Photographic image of Kathy Acker. Picture of a woman in black and pink. Half of her face is lit.

Kathy Acker

00:01:19

Performs “x”. Performs. | Video Description: Quick cut to portrait shot of Kathy Acker performing into a microphone. Half of her face is lit by a spotlight. Short cropped hair. Long earring dangling from right ear. Several piercings along lobe of left ear. Necklace.

[Camera zooms out on Kathy Acker]

00:13:27

Video Description: As camera zooms out, Kathy Acker, is holding pages, looking out to audience. Background art becomes visible behind her.

[Camera zooms in on Kathy Acker]

00:13:50

Video Description: Close up of Kathy Acker reading.

Kathy Acker

00:15:36

Kathy Acker finishes reading [audience applause]. | Video Description: Camera zooms out as Kathy Acker picks up a backpack and walks off the stage. Camera zooms in on background art.

Example of a sequence with distinct video segments:

[Two men in room]

00:00:27

[Electronic music] | Video Description: Fisheye lens view, black and white video of two men in office, back to back, sitting on chairs, typing on keyboards into computer terminals.

[Video Art]

00:01:15

[Electronic music] | Video Description: Pixelated digital art featuring shapes, images and words, changing rapidly. Including [summary of things that appear] Telephone, reel to reel tape machine, geometrical shapes, human figures, cartoon figures, words, Hitler [Q code], Mussolini [Q code].

END

00:15:34

Recommended terms for use in the description of camera framing and movement:

Camera Framing:

- Extreme Long Shot
- Long Shot
- Full Shot

- Medium Long Shot
- Medium Shot
- Medium Close-Up
- Over the Shoulder
- Close-Up
- Extreme Close Up
- Up Shot
- Down Shot

Camera Movements:

- Pan (left, right)
- Zoom (in, out)
- Dolly (in, out)
- Tilt (up, down)
- Boom (up, down)
- Truck (left, right)

3. Marking the End of a Digital Audio File

In the Linear mode of timestamped description it is necessary to add an extra timestamp to mark the end of an audio file. At the end of every timestamped description of an audio file, insert a final timestamp with the Speaker/Description filled in as END, leaving the label blank, unless you wish to include a descriptive annotation referring to the nature of the ending, which should appear in square brackets. You can additionally use the square brackets to note whether there is a link between the end of one audio file and the beginning of another, as in cases when a single event has been recorded over multiple analogue assets that have been digitized as unique digital audio files.

If you are using ERA AV to produce a Nested timestamped description, it is not necessary to add this closing timestamp manually as the ERA AV system will do so automatically,

Examples:

END 01:44:49

END 01:30:55 [cut off abruptly]

END 02:11:45 [recording of event continues on tape 2, file WM100499_02.WAV]

Principle of Keywords

Keywords may be used to reflect content in lieu of a full transcript. Keywords should be words or phrases inherent to the content and not interpretive additions. In other words, they should be **derived from vocabulary found in the audible content itself**. While there is no limit to how many of the words may be used (ranging from none to full transcript), keywords should be selected on the basis of their utility in signalling an important individual, object, point, theme, idea, or subject raised in the passage. Such might include the name of a person, organization, or title of a work (Atwood, Black Mountain Group or Night Poem), a generic form or place (sonnet, lyrical, Concordia University or Montreal), or an adjective, verb, or description that captures tone or scope (humorous, ironic). Proper nouns, which include names, titles, places, and particular things, will typically warrant the designation of Keyword, and a basic hierarchy for the parameters of keywording in a description may move from names of individuals, to titles of works, to place names, to other categories of designation.

Examples:

In the following transcription, the words that are in bold represent terms that might be selected as keywords if the approach to description involves keywording rather than transcription.

Allen Ginsberg 00:18:23 **George Bowering**, who I've known a long time, asked me to read a poem that I haven't read through but once before, called "**Angkor Wat**." So I'll try that. It's middle-sized, like, ten minutes, probably. What it is, is **notations** taken down in the course of one night in **Cambodia**, in **Siem Reap**, which is outside of Angkor Wat, a town outside of the ruins.

When formatting, the list of keywords should follow the content description and be separated by a pipe [|]. The pipe should be followed by the word "Keywords", which should be followed by a colon [:]. Items within the Keywords list should be separated by semicolons.

Allen Ginsberg | 00:18:23 | Introduces "Angkor Wat" | Keywords: George Bowering; notations; Siem Reap; Cambodia; Angkor Wat.

Square brackets should be used to designate additional, specific information from the cataloguer. If further specificity is required, cataloguers can insert parentheses within square brackets. Whenever possible, keywords should be accompanied by their corresponding Wikidata link (**in the first instances of their appearance**), to facilitate networked searchability:

Allen Ginsberg | 00:18:23 | Introduces "Angkor Wat" [from Angkor Wat (<https://www.wikidata.org/wiki/Q96035194>)] | Keywords: George Bowering [<https://www.wikidata.org/wiki/Q1239280>]; notations; Siem Reap; Cambodia; Angkor Wat.

Wikidata links allow for an interconnected web of information. Prioritize linking people, particularly writers and performers, and artistic works. However, Wikidata moderators do have a specific notability policy and you may not find entries for every item you wish to include.

Example:

Margaret Atwood

00:00:00

I should **apologize** to begin with for my **voice**. I don't usually sound quite this much like **Tallulah Bankhead**. I have the **Montreal** plague. The first poem is called "**This is a Photograph of Me**," and it's the first poem in **The Circle Game**. I love **Kafka's** **The Trial**, but prefer **Wordsworth's** "**I Wandered Lonely as a Cloud**".

Keywords:

Margaret Atwood

00:00:00

Introduces "This is a Photograph of Me" from The Circle Game [<https://www.wikidata.org/wiki/Q7723073>]
Keywords: voice; Tallulah Bankhead [<https://www.wikidata.org/wiki/Q255815>]; apologize; Montreal; Kafka [<https://www.wikidata.org/wiki/Q905>]; The Trial [<https://www.wikidata.org/wiki/Q36097>]; Wordsworth [<https://www.wikidata.org/wiki/Q45546>]; "I Wandered Lonely as a Cloud" [<https://www.wikidata.org/wiki/Q2295398>]; London [Ontario].

4. Summary of Typographical Rules

To facilitate research and to ensure proper conversion of the Contents field to XML when necessary, timestamping and description entered in the Contents field must follow certain typographical rules. This section outlines the main typographical markers and their functions for use in contents description.

The key typographical markers in the Contents field are:

- Return [`<_`] [NOTE: This is not a visible marker, but the insertion of a carriage return resulting in the separation of terms by pushing an item to the next line.]
- Comma [`,`]
- Semi-colon [`;`]
- Pipe [`|`]
- Question mark in parentheses [`(?)`]
- Double Quotation Marks [`“ ”`]
- Square Brackets [`[]`]
- Round Parentheses [`()`]
- Period [`.`]
- Ellipsis [`...`]

Proper use of these markers according to the established conventions will ensure the searchability and operability of the Contents field. As a general rule, the principle elements of a timestamp should be listed in the following typographical format:

Ex.

Name of Speaker
00:00:00
Performs “...”

Return

The three core components of a timestamp are divided by carriage returns, and each timestamp is also separated by a carriage return. The Return key is used only for those purposes.

Comma

No commas will succeed the verb, unless you wish to list further descriptors or activity [`,`]. In other words, commas should be used sparingly, and only where grammatically warranted. See rules regarding the use of semicolon for further clarification.

Double Quotation Marks

Double quote marks are reserved for specific citational use [`“ ”`]. Titles of works should be indicated by quotations and followed by the word ‘from’ when designating known publications. Additional descriptors should be identified in square brackets and contain linked data where possible [`[]`].

Ex.

Margaret Atwood
00:25:06
Reads “Siren Song” from You are Happy [Oxford UP, 1974].

Ex.

Allen Ginsberg
00:54:29
Sings “The Little Boy Lost” and “The Little Boy Found” from Songs of Innocence and Experience
[<https://www.wikidata.org/wiki/Q20713959>].

Separate items/agents within square brackets with a semi-colon [`;`]. All additional information, possibly emerging from research and added at the cataloguer’s discretion, should be enclosed within square brackets. To include further details, use parentheses within square brackets [`()`].

Ex.

Swiftly Lazarus [Todd Swift (vocals); Tom Walsh (saxophone, computer)]

00:05:32

Performs “West of an Idea/ Hlinka Guard” from The Envelope Please [CD].

Descriptions of audible content will range from single word descriptions, to detailed representations of sonic material through the use of Keywords, to full transcriptions. When providing Keywords, insert a pipe symbol [|] and separate items with semicolons [;]. It is strongly recommended to provide Q-codes from Wikidata in square brackets for Keywords, where possible. So, square brackets indicate information being brought to interpretation by the cataloguer, or any number of qualifiers about the nature of the sound signal, such as [poem] or [CD], as shown in the examples:

Ex.

Allen Ginsberg

00:18:23

Introduces “Angkor Wat” | Keywords: George Bowering [<https://www.wikidata.org/wiki/Q1239280>];

“Angkor Wat” [poem]; notations; Siem Reap [<https://www.wikidata.org/wiki/Q11711>]; Cambodia

[<https://www.wikidata.org/wiki/Q424>{}]; Angkor Wat [<https://www.wikidata.org/wiki/Q43473>].

Basic Rules for Transcription When transcribing speech or other audible materials, use established vocabulary to replace silences [silence] or inaudible speech [unintelligible] within square brackets. Or, label what you cannot hear clearly but choose to interpret within square brackets, and append a question mark in parentheses to demonstrate uncertainty in deciphering sound signals [(?)]. An unintelligible phrase or silence might also signal a [cut] in recording.

The use of ellipses in square brackets [...] indicates audio content that has not been described. Sections that are left out should be clearly timestamped. Ellipses may replace descriptions of audible materials.

Include periods [.] throughout and at the end of every entry. When providing full transcription, use appropriate punctuation such as commas, question marks and other symbols where necessary. When using brackets, place punctuation after the closing bracket. Similarly, periods follow quotation marks and any other punctuation marker.

Symbols to Avoid Avoid using <> or &.

Insertion of END time stamp To signal the end of recording (in the linear timestamping method), the timestamp must conform to the Contents field grammar and maintain the established typographical formula. If you wish to describe something about the ending, or signal the connection of this recording to another recording, use square brackets.

Ex.

END

00:37:52

[Cut out]

END

00:27:31

[File 2 of 3]

5. Instructions for identifying assets.

Given that a single Item can comprise multiple recordings, the cataloguer should clearly identify each asset in question at the start of every Contents field entry (in addition to documenting it in the Digital File Description Contents entry). This practice will minimize confusion and mark transitions between separately timestamped or transcribed assets which is especially useful if compiled one after another in the same box of the Contents field. Indicate the filename or chosen title, followed by a description of the type of recording and its position in relation to the subsequent recordings, followed by a comma and any further specifications **when necessary** (ex: Tracks #s or Sides) between square brackets.

Examples:

STE-001.wav [File 1 of 2]

Unknown 00:01:22 Ambient Sound [loud bang]. ... END 00:20:30

STE-002.wav [File 2 of 2]

Audience 00:00:06 Applause.

For timestamping or transcription purposes, when several recordings have been edited together (either one after another or as a multitrack mix) to form one file, just the filename can appear at the start of the entry. However, it is important to indicate in the Content Note, that the digital file described is a combination or mix of several files, and to list all of the original components/assets.

Example:

Mario Campo at Ultimatum 1985 Night 1 [Tape 1, Tracks 1-4] Formula: Ultimatum_Mixdown_Name_Tape_TTracks_Tape_TTracks.mp3 Ex: Ultimatum_Mixdown_Daniel Guimond_U-1_T1_T2_T3_T4_T5_T6_T7_T8.mp3

4.21. Contents Notes

A short description or abstract of the contents of the recording should be provided here. Any other information concerning the AV content that is not included in the Contents field may be added here, as well. Include information about the transcription process, including the tools used. If AI transcription software was used, we require that the outputs be human-verified. Please document your process in this field.

4.22. Notes

Notes allow cataloguers to input additional information regarding the item. There are several different note types that could be applicable:

General: General notes about the item

Related Version: Different published versions or manifestations of the item that are related or can be linked together

Cataloguer: A note on the different people involved in cataloguing the item or new cataloguers who have taken over the cataloguing of the item

Preservation: Specific notes about preservation details typically found from third parties such as, who did it, what happened to the object, or extra XML information.