

## DESCRIPTION OF FIELDS AND CONTENT FOR METADATA RECORDS OF MULTIMEDIA FILES

COLUMN	COLUMN HEADER	DESCRIPTION OF CONTENT
1	MULTIMEDIA FILE	Full filename of multimedia file (audio [.wav] or video [.mp4]) <b>NOTE: For filename conventions see the relevant pdf file in the \doc folder.</b>
2	TRANSCRIPTION AND TRANSLATIONS	List of files that contain the transcription and, if completed, translation of the audio or video. If there is both a Transcriber (.trs) and ELAN (.eaf) file associated with the same multimedia file then they are both listed here, separated by a semi-colon. That that for audio often the first transcription was done in Transcriber, the file was then imported into ELAN, proofread and translated into Spanish.
3	ASSOCIATED FILES	<b>Video only:</b> In the case of mp4 video an audio .wav file is extracted from the video to facilitate transcription in ELAN. In these cases the "derived" .wav file is listed.
4	UNIQUE IDENTIFIER	The unique identifier is unique and never-changing to each multimedia file. It is often the date (yyyy-mm-dd) followed by a hyphen and a letter (rarely two). Note that the UID should not be taken to indicate anything but a somewhat arbitrary code for unique identification. The "date" format was chosen as it offers no chance of repetition. Occasionally typos or other changes in the .wav/.mp4/.trs/.eaf files might be found. These may be corrected but the UID never changes. The UID is a very easy way to find files via the filename.
5	DURATION	This is the duration (mm:ss) of the multimedia file.
6	FILE SIZE (multimedia)	This is the size (in kb) of the multimedia file.
7	LANGUAGE (GENERAL)	This is the general language name (i.e., 'Nahuatl') which in classifications such as Glottolog or ISO 639-3 often represent a higher level language group.
8	LANGUAGE (SPECIFIC)	This is the specific language and corresponds to a Glottolog or ISO 639-3 code, given in the following field.
9	GLOTTOCODE AND ISO 639-3	See codes at <a href="https://glottolog.org/">https://glottolog.org/</a>
10	PRIMARY COMMUNITY (SPEAKER)	This is the primary community of the recording. Occasionally two speakers from different communities (or even languages) will be recorded. This is what is considered the primary one.
11	COMMUNITY COORDINATES	These are the decimal coordinates (N and W, the latter with a minus sign following the conventional representation for areas west of the Greenwich Prime meridian).
12	MUNICIPALITY	This is the municipality of the Primary Community.
13	STATE	This is the state of the Primary Community.
14	COUNTRY	This is the country of the Primary Community.
15	RECORDED BY	This is the individual (last name, first name) who recorded the audio or video.
16	DATE OF RECORDING	This is the date of the recording.
17	RECORDING FORMAT	This is the recording format (e.g., sampling rate and bitrate)
18	MICROPHONE	This is the name of the microphone that was used.
19	POWER SOURCE	This is the power source (when the data was recorded) usually AC (outlet), DC (battery), or AC (battery pack with AC outlet).
20	CONSULTANT #1	This is the name (last name, first name) of the consultant. Demographic data (village of origin, sex, birthyear) is available in a different database.
21	CONSULTANT #1 ROLE	This is the role of Consultant #1.
22	CONSULTANT #1 SEX	This is the sex of Consultant #1.

23	CONSULTANT #1 ORIGIN	This is the village or residence or origin of Consultant #1.
24	CONSULTANT #1 TRACK	This is the track on which Consultant #1 is recorded. For two channel recordings it will be Left or Right. If one channel it will be 'Mono'.
25	CONSULTANT #2	This is the name (last name, first name) of the consultant. Demographic data (village of origin, sex, birthyear) is available in a different database.
26	CONSULTANT #2 ROLE	This is the role of Consultant #2.
27	CONSULTANT #2 SEX	This is the sex of Consultant #2.
28	CONSULTANT #2 ORIGIN	This is the village or residence or origin of Consultant #2.
29	CONSULTANT #2 TRACK	This is the track on which Consultant #2 is recorded. For two channel recordings it will be Left or Right. If one channel it will be 'Mono'.
30	GENRE	This is a controlled vocabulary for genre, i.e., the content/theme of the recording. By far the most common is Botany.
31	SUBGENRE	This is a further division of the thematic content of the multimedia files.
32	INDIGENOUS TITLE	This is an Indigenous title to the recording.
33	ENGLISH TITLE	This is an English title to the recording. It is not necessary a direct translation of either the Indigenous or Spanish titles.
34	SPANISH TITLE	This is an English title to the recording. It is not necessary a direct translation of either the Indigenous or Spanish titles.
35	DESCRIPTION	This is a description, often very detailed, of the content of the recording. In many cases, such a Botanical Field recordings, the description is meant to provide easy access to the substantive content of the recording thus facilitating an understanding of the ethnographic content and not relying on evaluation of the transcription or translation, at least for an initial understanding of the content.
36	TRANSCRIBED BY	This is the name of the person who transcribed (and corrected or proofed) the transcription.
37	TRANSLATED BY	This is the name of the person who translated (and corrected or proofed) the transcription.