

AILLA Self-Deposit Tool Training

January 2018

Susan Kung & Ryan Sullivant
Archive of the Indigenous Languages
of Latin America
aila@aila.utexas.org

Archive of the Indigenous Languages of Latin America

ailla.utexas.org

ALLA is a digital language archive specializing in multimedia materials that are in and about the indigenous languages of Latin America.

ALLA has parallel user interfaces in both English and Spanish, as well as metadata fields (for descriptions and titles) reserved specifically for indigenous languages.



What constitutes a collection?

Collections are made of resources (i.e., folders) which contain media files

Media files in a resource are related, for example, a set of files that are

- the results of a single speech event
- collected for a single purpose
 - surveys, experiments
- curated along a theme, e.g.
 - semantic domains
 - grammatical categories
 - toponyms

Example collection:

[Sochiápam Chinantec Whistled Speech Collection of Mark Sicoli](#)

Example resource:

[Whistled speech in fields 2](#)

Example media file:

[CAM20303-01.mpg](#)

Archival Collections

Archival (mostly static) data collections ≠ dynamic working data collections.

No deleting files or resources.

No updating files or resources by replacement.

Not all media formats can be archived.

Collections must be organized and documented for people unfamiliar with your practices, workflows, and purposes.

Progressive Archiving

We recommend **progressive archiving** for working collectors.

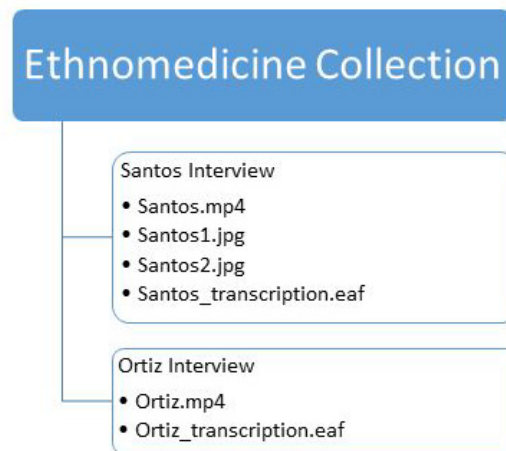
Immutable portions of archives (i.e. “raw data”) are deposited first.

Additional files are added as they are finalized.

Archiving tasks are...

- spread out over time;
 - done closer to each data collection trip or production step.
-

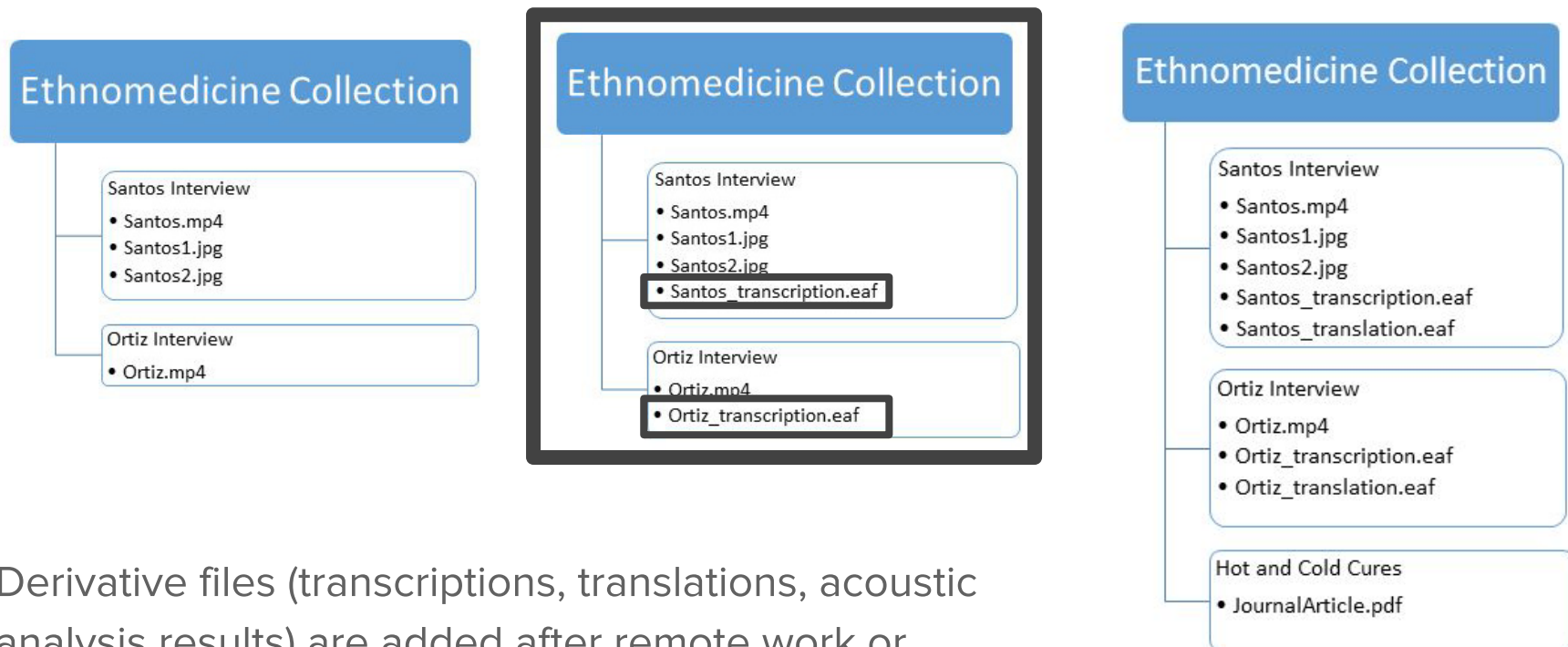
Progressive Archiving



Primary data (videos, audios, photos) are archived after each field trip.



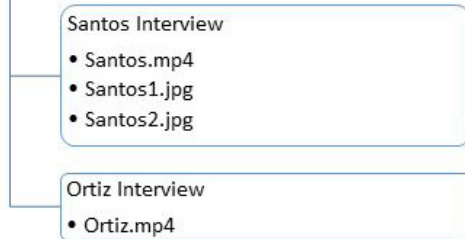
Progressive Archiving



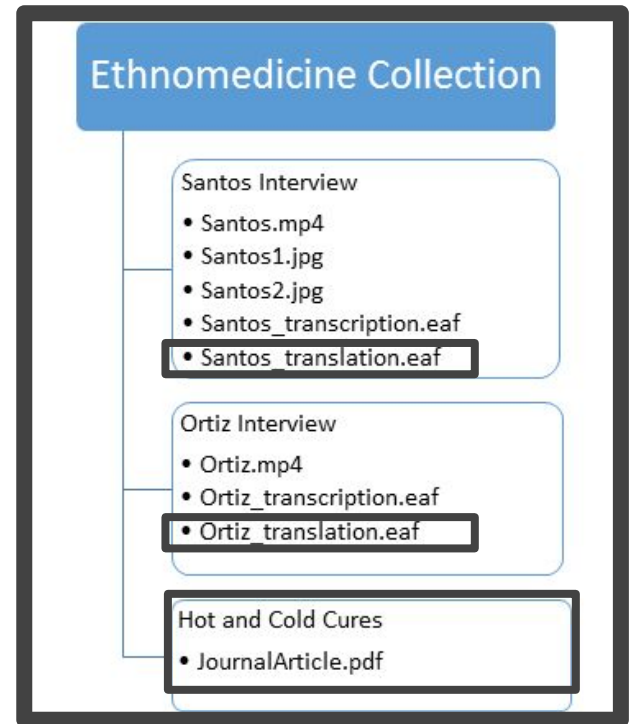
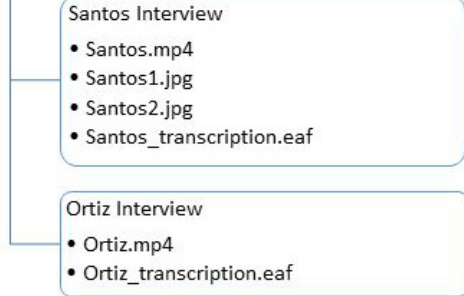
Derivative files (transcriptions, translations, acoustic analysis results) are added after remote work or later trips

Progressive Archiving

Ethnomedicine Collection



Ethnomedicine Collection



New materials (journal articles, lexicons, etc.) can be added to new or existing resources.

How to register as an AILLA depositor

- [Create an Account](#).
 - read and understand [Conditions for Use of Archive Resources](#).
 - Sign and Complete a [Depositor License Agreement](#) form.
 - Sign a [Self-deposit agreement](#) form.
 - Submit [Metadata Spreadsheet](#) with information about collection, languages, and contributors.
 - AILLA staff will create your AILLA collection and add any new or unique contributors and/or languages to AILLA;
 - AILLA staff will upgrade your account to Depositor status.
-

Other types of ALLA records

- Languages: Every language that is listed in an ALLA resource appears in the Languages Collection.
 - Countries: Every country that is listed in an ALLA resource appears in the Countries Collection.
 - Persons: Every named person that participated in the creation of any of the materials in ALLA appears in the Persons Collection.
 - Organizations: Any organization that was directly responsible for the creation of these materials appears in the Organizations Collection.
-

What you can NOT do with self-deposit tools

- Create a new collection.
 - Create or the record for
 - A person
 - An organization
 - A language
 - A country
 - Delete objects or records.
 - Move media between resources.
-

What you CAN do with self-deposit tools

- Edit your collection metadata.
 - Create resources and edit resource metadata.
 - Ingest media files
 - < 1GB
 - Certain formats
 - Create and edit metadata for media files.
 - Edit access levels.
-

3 Types of Form Fields

1. Free text
 - a. You can write anything in this field.
 2. Autosuggestions
 - a. Indicated by a small circle at the beginning of the text box/field.
 - b. You must write slowly (to give the index time to match the text and make suggestions from the saved records), then you must choose from the suggested options. NOTHING will be saved if you simply type you text here.
 3. Drop-down menus
 - a. A list of controlled vocabulary in English and Spanish.
 - b. You must choose one of the offered options.
-

Tips for adding media files

Not just any file can be added to ALLLA.

Be sure to

- Check your formats (see next slide)
- Check your file sizes (about 1GB limit; compress and/or split larger files into parts if necessary)
- Run virus scan software on all media files
- Avoid special characters in filenames (esp. space, colon, period*, diacritics, etc.)

* A period should only be used to separate the filename from the format type, filename.wav

Allowed media file types

Content Model	Allowed Formats
Audio	WAV, MP3
Video	MPG, MPEG, MP4, MOV, AVI (M4V, MKV)*
Binary	text and XML-based formats (XML, EAF, TRS, TXT, CSV, DB, etc.)
Basic image	JPG, PNG, GIF
PDF	PDF/A (an archival specification)
Large Image	single-page TIF, TIFF

AILLA's access levels

Access Level	Restriction	Description	Backend Level
Level 1	Public	Authenticated Access Default	0
Level 2	Password protected	Password only Curation in Progress	1
Level 3	Password protected	Time limit Temporary embargo	2
Level 4	Password protected	Contact depositor Controlled access	3

Choosing an Access Level

Commercial use (except with explicit permission from the copyright holder) is prohibited for all access levels.

Users must agree to ALLLA's Conditions of use prior to gaining access to media files.

Public (Level 1): All materials that can be shared publicly.

Choosing an Access Level

Curation in progress (Level 2)

- Administrative documents
 - Depositor License Agreement
 - Self-deposit agreement
 - Metadata sheets
 - Files undergoing AILLA staff curation
 - Materials that were accepted before our Public Access Policy was necessary.
-

Choosing an Access Level

Temporary Embargo (Level 3)

- Embarrassing or incriminating material
 - today's gossip → yesterday's history
 - Some copyrighted material
 - Check your author agreement
 - Dissertation material in progress
 - 2-3 yr embargo for degree completion
 - Consider making a few resources public for the community
 - Other embargo (e.g., publisher requirement on preprints)
-

Choosing an Access Level

Controlled Access (Level 4)

- Material that can only be accessed according to special protocols.
- Collections where community and/or depositor wishes to maintain control of access.
- For both scenarios:
 - ALLLA cannot vet individuals!
 - You, a community representative or someone else must do that.

You/community representative **must** maintain your contact information and be prepared to respond to requests from people you do not know. ALLLA reserves the right to alter your access level if you do not respond to requests in a timely fashion.

Notes about Access

Metadata is always open access and visible to browser search engines, even for restricted resources.

If a resource can never be shared with anyone, it should not be archived at ALLA.

All restrictions are subject to periodic review by ALLA staff.

Navigate AILLA by following the BREADCRUMBS



The Archive of the Indigenous Languages of Latin America

[Collections](#) [Languages](#) [Countries](#) [Announcements](#) [Organizations](#) [Persons](#)

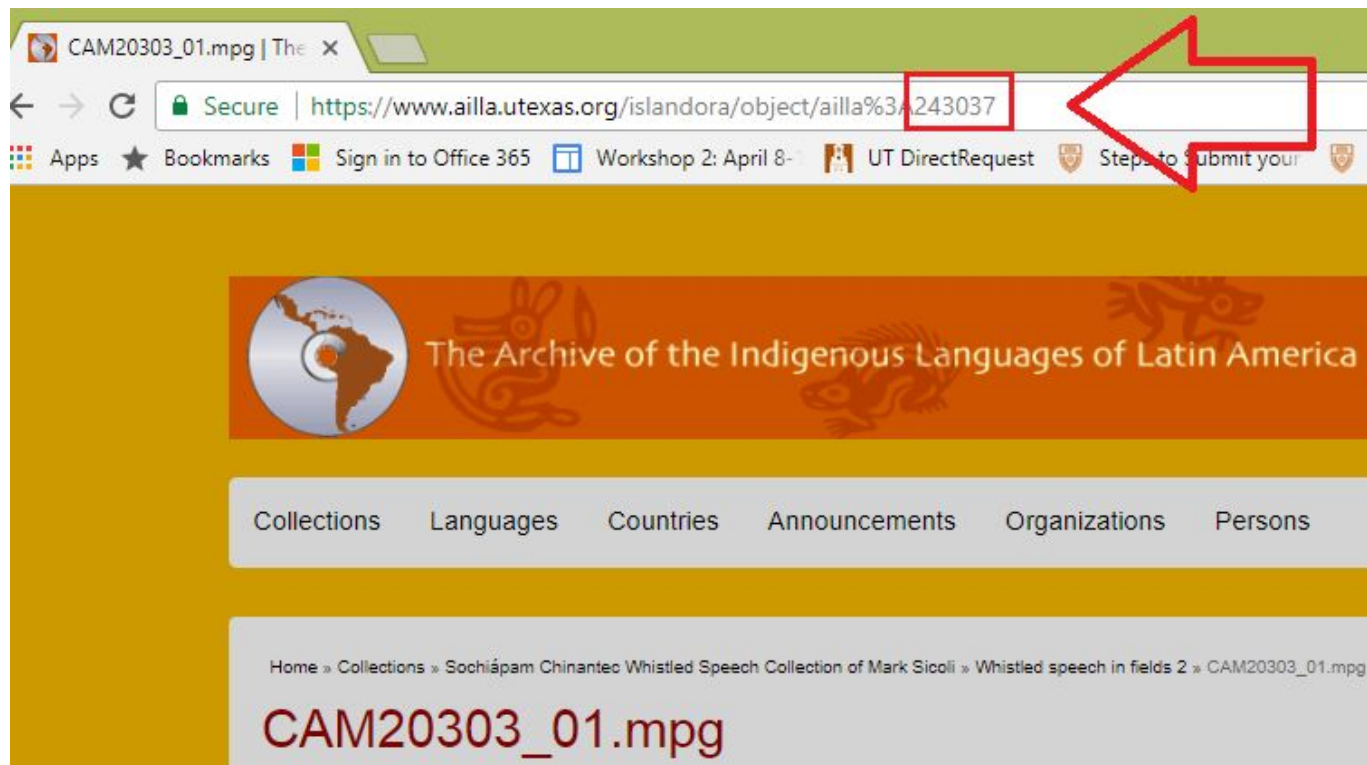
Collection Title	Resource title	Filename
Home » Collections » Sochiápam Chinantec Whistled Speech Collection of Mark Sicoli	» Whistled speech in fields 2	» CAM20303_01.mpg

CAM20303_01.mpg

Object Details

Language(s)	Chinantec, Sochiápam Spanish English
-------------	--

PID (Persistent Identifier):
6-digit number at the end of the AILLA URL



The screenshot shows a web browser window with the following elements:

- Browser tab: CAM20303_01.mpg | The x
- Address bar: <https://www.ailla.utexas.org/islandora/object/ailla%3A243037>. The number 243037 is highlighted with a red box, and a red arrow points to it from the right.
- Navigation bar: Collections Languages Countries Announcements Organizations Persons
- Breadcrumbs: Home » Collections » Sochiápam Chinantec Whistled Speech Collection of Mark Sicoli » Whistled speech in fields 2 » CAM20303_01.mpg
- Page title: CAM20303_01.mpg

Using the self-deposit tools

The following steps can be done only after

- You have submitted all of the required paperwork to ALLLA;
 - ALLLA staff have created your collection and upgraded your account to Depositor status.
-

Step 1: Edit your collection metadata

1. Navigate to your collection.
2. Edit the Collection metadata:
Manage > Datastreams > ID Column, MODS row > Edit
 - a. Make edits on the form.
 - b. Click Update to save changes.



Step 2: Create a Resource

From your collection page:

1. Select Manage > Add an object to this collection
 2. Complete metadata form with information from the first row/resource listed in the "Resources - Recursos" tab on your spreadsheet.
 3. Click Ingest to create resource and save the metadata.
 4. Note (write down) the PID in the appropriate column.
(We suggest adding a column for the PID to the Resources and Media Files tabs of your metadata spreadsheet if there is not one already. This will help you keep track of what you have archived vs what you have not.)
-

Tips for adding media files

Before adding files, be sure to

- Check your formats (see above)
- Check your file sizes (about 1GB limit, compress and/or split larger files into parts if necessary)
- Run virus scan software on all media files
- Avoid special characters in filenames (esp. space, colon, period*, diacritics, etc.)

* A period should only be used to separate the filename from the format type, filename.wav

Step 3: Add a media file to the resource

Navigate to the resource page that you just created:

1. Select Manage > Add an object to this Resource.
 2. Select the media file's Content Model (i.e., media type).
 3. Complete the metadata form.
 4. Choose the file to upload.
 5. Click Ingest.
 6. Note (write down) the PID.
-

Setting access levels

Access levels can be set from the

- File level - one file at a time
- Resource level - entire resource or selected files
- Collection level - entire collection at once

File/resource/collection page > Manage > Set access levels



Step 4a: Setting the controlled access

From the file:

1. Select Manage > Access Levels.
 2. Select Level 3: Contact depositor.
 3. Provide a password and any other information
 - a. Optional password hint
 - b. Justification for restriction
 - c. Depositor (or contact) name
 - d. Depositor (or contact) email
 4. Click Commit Changes.
 5. Click View.
-

Step 4b: Setting the time embargo

From the file:

1. Select Manage > Access Levels
 2. Select Level 2: Time limit
 3. Provide a password and any other information
 - a. Optional password hint
 - b. Justification for restriction
 - c. Release date
 4. Click Commit Changes
 5. Click View
-

Step 4c: Changing access levels (Resources)

1. Navigate to the resource (using breadcrumbs)
 2. Select Manage > Access Levels
 3. Select the files to unrestrict
 4. Select the desired level
 5. Click Commit Changes
 6. Click View
-

A well-curated collection

As you have seen, self-deposited collections are built piece by piece, with limits on the sizes of files that can be added.

Think about what parts of your materials are best for your collection, and prioritize those parts that are of greatest value for future users.

