

# Workshop 2: Prepping content

---

COMMUNITY-DRIVEN METADATA, FILE ORGANIZATION, AND  
PREPARING YOUR ITEMS

# Goal: Prepping community contributions for the digital archive

---

- **Steps**

- Think about metadata and its community value
- Transcribe oral history interviews
- Crop and edit photos

- **Why not just go straight to Github?**

- **Shared community representation**

- To give communities time and space after a community event to talk about how the digital exhibit reflects community memories and values, *in aggregate*

- **Technical efficiency**

- Prep work outside of Github makes it easier to transfer info about each community contribution to the Github digital exhibit web page.

# Metadata

---

# Crafting Metadata in Community Settings

In technical terms, metadata is

- **Info about something** that we can see before we open that thing
- **Flexible but structured.** Just a file name, or lots of info in lots of different categories
- **Easily searchable** to make that thing easier to find.

In community archive processes, metadata lets us

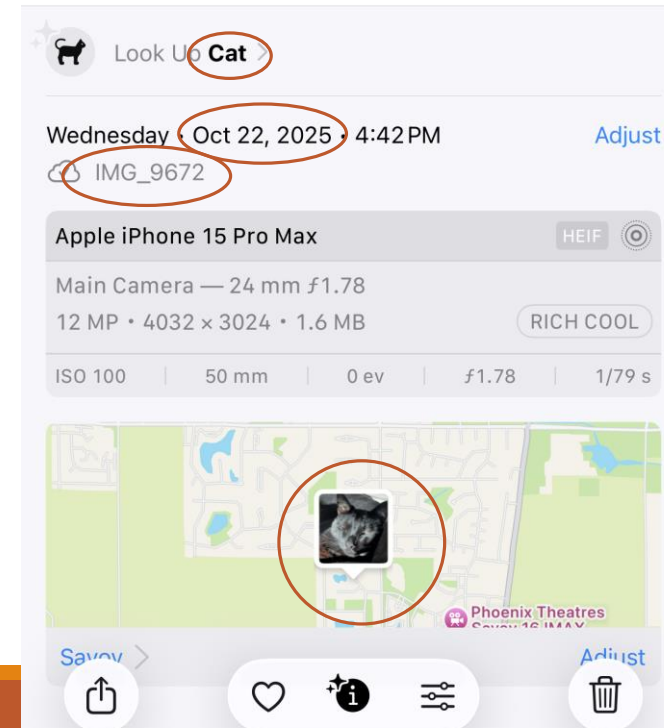
- Confirm who we are and what matters to us
- Have a say in how we are represented online, both to our community members and to other people

Three points in time for conversations about metadata

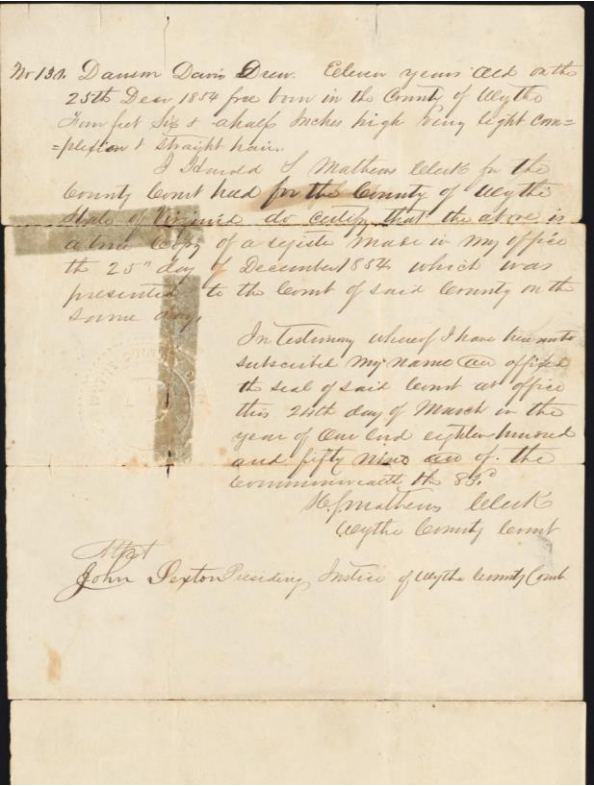
- **Before oral history interviews:** craft an interview process that lets community members contribute their own metadata
- **After interviews are done:** remind us of how metadata contributes to community representation
- **If no interviews were done:** thoughtfully represent metadata from community members about their contributions to the archive



Add a Caption



# Lived Experience in Metadata: The Jennie Elder Suel Collection at Miami University



## Miscellanea Collection

Collection Identifier: Mss. Coll. Miscellanea

Walter Havighurst Special Collections Finding Aids | Miscellanea Collection

Collection Overview Collection Organization Container Inventory

### Dates

Creation: 1776-2002

### Extent

6 cubic feet

### Language of Materials

English

Collapse All

#### Finding Aid & Administrative Information

| Title                        | Guide to the Miscellanea Collection              |
|------------------------------|--|
| Author                       | Finding aid prepared by Betsy Butler (2007-2009) |
| Date                         | 2016   |
| Description rules            | Finding Aid Prepared Using Dacs                  |
| Language of description      | English  |
| Script of description        | Code for undetermined script                     |
| Language of description note | Finding aid written in English                   |

#### Repository Details

edom certificate for Dawson Davis Drew Stewart, 24 March 1859

ton, John

rk Office of Wythe County Court; Mathews, Harold L.

sl, Jennie Elder

i4-12-25

rk Office of Wythe County Court

jinia--Wytheville

edom certificate of 11-year-old Dawson Davis Drew (Stewart). Dated December 25, 1854. The document describ and a half inches high. Very light complexioned and straight hair. Dawson relocated to Butler County in prior to 1 ages

x 21 cm

e Black people  
can American girls  
very

wart, Dawson D. (Drew)  
thews, Harold L.  
ton, John

edom  
ntier

i0-1859

ebellum Period  
very and Freedom  
ly Migration

idom certificates  
timonies  
l court records

t

jlsh

lter Havighurst Special Collections & University Archives. Oxford, Ohio.

inie Elder Suel Collection. Miscellanea Collection

lwer 3

# Institutional vs. Community-Centered Metadata Creation

---

## **Institutional Metadata**

Planned early in project design

Built into data collection tools (e.g., oral history prompts, History Harvest forms)

Designed to serve institutional organization, access, and control

Emphasizes standards, compatibility, and consistency

## **Community-Centered Metadata**

- Often emerges later, after materials have been collected
- Motivated by meaning-making, storytelling, and sharing within the community
- Driven by relationships, genealogical connections, and cultural context
- Serves the community's own priorities—not institutional systems

## **Key Differences**

- Timing and purpose
  - Institutions use metadata to organize knowledge for preservation.
  - Communities use metadata to interpret and reclaim knowledge for self-definition.

# Community-Centered Collaborative Metadata

---

## Title

- Prompt: "What would you call this item?"
- Intention: Community defined naming

## Description

- Prompt: "Tell us the story behind this."
- Intention: Centering personal meaning

## Date

- Prompt: "When was this from? Approximate is okay."
- Intention: Valuing memory over precision

## People

- Prompt: "Who is in this photo/story?"
- Intention: Naming and honoring individuals

## Place

- Prompt: "Where did this take place?"
- Intention: Grounding materials in place

## Emotions/Significance

- Prompt: "How does this item make you feel?"
- Intention: Centering affect and cultural value

## Permissions

- Prompt: "How should this item be used/shared?"
- Intention: Honoring agency and consent

## Tags or keywords

- Prompt "What words would help people find this?"
- Intention: Enabling community-generated vocabularies

# A quick metadata-prep walkthrough

---

You all have access to this at

<https://digitalarcplatform.github.io/documentation/docs/postEvent/metadata/>

It's long, so we'll do a quick walkthrough today to cover the big points.



# Questions to ask as you start prep

---

What metadata is easy to get?

What metadata will need to be gathered in community?

Where will you store your prep work?

- o Google Sheets
- o In "Plain text" in Google Docs?

# Transcription

---

# Initial decisions

---

Do you have access to software that can transcribe for you?

- Microsoft Word and Google Docs have the capability of auto-transcribing audio, but with limitations.

What to include in the transcript?

- Do you want to include all sounds spoken, like “umm” or “like”?
- Will you include duplicated words?
- Will you represent non-verbal sounds like laughter?

How will you label the speakers in the text file?

- Keep in mind the contributor's consent at the time of the recording.

Will you include timestamps in the text file?

# Steps

---

1. Download your audio file to your computer
2. Create a document for the transcription
  1. In Microsoft Word or Google Docs
3. Begin the transcription process
4. Clean up/ edit your transcript
5. Upload file to website
  1. This may include the audio file as well, depending on the contributor's consent

# Recommendations when transcribing

---

Create a style guide that your community will follow. This will ensure consistency from one transcript to another.

- Format of the document
- Punctuation
- How to note unintelligible language

Accurately represent what was said in the recording, avoid excluding anything, unless the contributor requests that a section be omitted.

Establish a workflow for checking the transcripts

- Who can assist when something is in another language or unintelligible?
- Who gives a final look over of the transcript?

# Google Docs vs. Microsoft Word

---

## Google Docs

- Pros
  - Free to anyone with a Gmail account
  - Can be stored in your community's Gmail account
- Cons
  - Speech-to-text is not very accurate. It requires very clear audio.
  - The speech recognition sometimes stops working all of a sudden

## Microsoft Word

- Pros
  - Transcripts are more complete. Includes time stamps and can distinguish between different speakers
- Cons
  - Requires a paid subscription to Microsoft Office products

# Break!

## Take 5

Stretch, get water, do a little dance in the kitchen

# Organizing Your Files

---

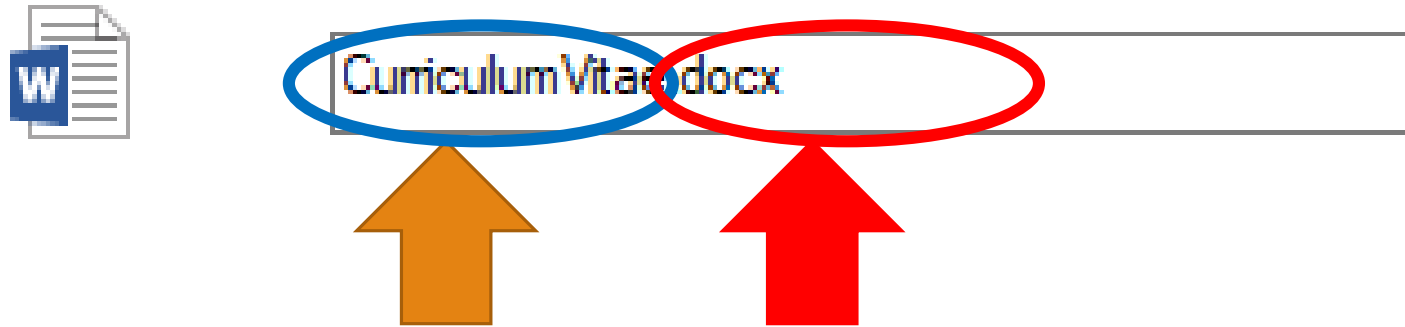
A.K.A “HOW TO KEEP TRACK OF THE NUMEROUS FILES THAT MAKE YOUR ARCHIVE POSSIBLE”



# Naming Your Files

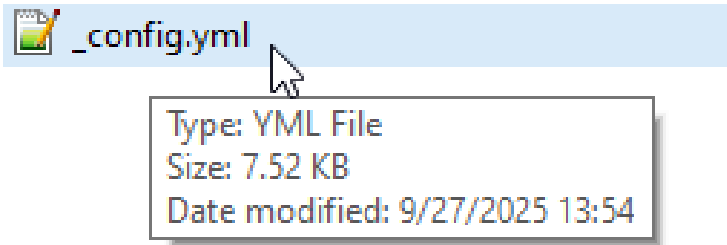
---

## Anatomy of a File Name

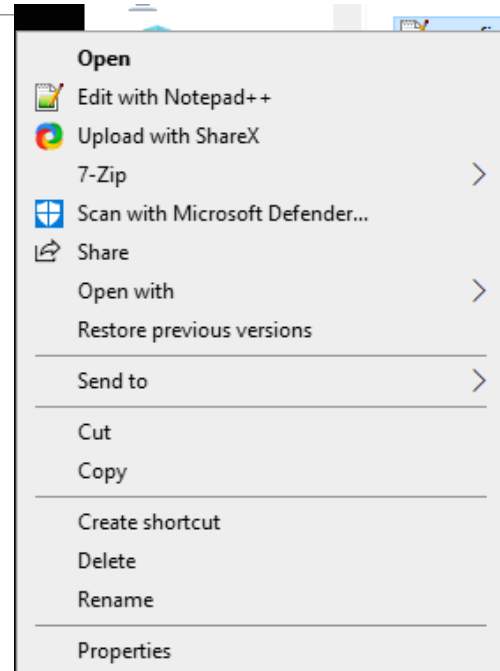


- The name you (the user) gives to the file
- Can and should change
- The file extension which sets the type of the file
- Generally do not want to change

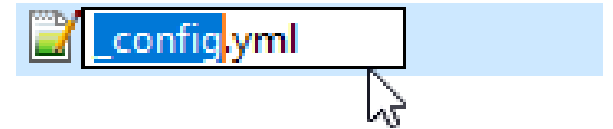
# Naming Your Files



Step 1: Hover over your file in Explorer/Finder
























Step 2: Right-click and select “Rename”



Step 3: Assign the file a new name **but** do not change the file type

Note: This holds across almost every major file platform (e.g. Google Drive, OneDrive, etc.)

# Naming Your Files

| PC > Downloads  |                  |          |        |  | Search Downloads |  |
|---|------------------|----------|--------|--|------------------|--|
| Name  | Date modified    | Type     | Size   |  |                  |  |
|  464027715_10230186726004899_87616096...   | 10/19/2024 11:39 | JPG File | 75 KB  |  |                  |  |
|  464029623_8331673440241805_161277309...   | 10/20/2024 18:25 | JPG File | 60 KB  |  |                  |  |
|  464072893_1263391544690201_817656477...   | 10/25/2024 16:45 | JPG File | 745 KB |  |                  |  |
|  464085247_8873112252719857_820322770...   | 10/21/2024 14:02 | JPG File | 115 KB |  |                  |  |
|  464101964_8629725643751069_628486027...   | 10/19/2024 23:43 | JPG File | 115 KB |  |                  |  |
|  464143473_10101007301314017_90909909...   | 10/20/2024 12:15 | JPG File | 96 KB  |  |                  |  |
|  464214508_957902693036649_7424446612...   | 10/26/2024 12:12 | JPG File | 476 KB |  |                  |  |
|  464227741_965806362246500_8414588766...   | 11/2/2024 21:15  | JPG File | 131 KB |  |                  |  |
|  464275083_2190335551351959_158878888...   | 10/24/2024 08:37 | JPG File | 97 KB  |  |                  |  |
|  464276646_10229917522776892_50435875...   | 10/25/2024 16:38 | JPG File | 203 KB |  |                  |  |
|  464327650_10225477057009855_79188892...   | 10/31/2024 16:15 | JPG File | 360 KB |  |                  |  |
|  464341970_1998831683890728_750178260...   | 10/22/2024 20:01 | JPG File | 400 KB |  |                  |  |
|  464373014_9083431785014690_156474688...   | 10/25/2024 13:18 | JPG File | 466 KB |  |                  |  |
|  464383545_540887542239140_8132480683...  | 10/24/2024 20:16 | JPG File | 64 KB  |  |                  |  |
|  464384902_1090760999333758_841400520... | 10/25/2024 16:40 | JPG File | 218 KB |  |                  |  |
|  464391570_10235041442720185_84676509... | 10/24/2024 20:41 | JPG File | 130 KB |  |                  |  |
|  464399768_965806268913176_2862537696... | 11/2/2024 21:15  | JPG File | 161 KB |  |                  |  |
|  464568359_1023963862863322_839813345... | 10/26/2024 23:13 | JPG File | 45 KB  |  |                  |  |
|  464578464_10162005836136940_77243289... | 10/31/2024 17:21 | JPG File | 58 KB  |  |                  |  |
|  464580647_27357620983881511_35785129... | 10/26/2024 15:15 | JPG File | 174 KB |  |                  |  |
|  464614597_9004328459619041_489237469... | 10/26/2024 23:19 | JPG File | 469 KB |  |                  |  |

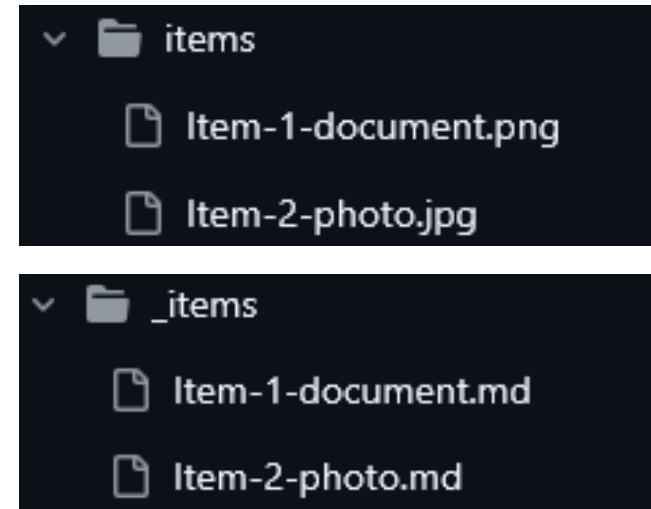
- Unclear Names
  - Also too long
- Not Organized
- Consistent(ish?)
- How do we improve?

# Naming Your Files

---

## Principles

- 1. Be specific
  - e.g. date\_itemid.ext : 2025-10-25\_01.jpg
- 2. Be consistent
  - DigitalArc *depends* on this
- 3. Be organized
  - Have a system in place for where to keep images, audio, etc. and metadata files



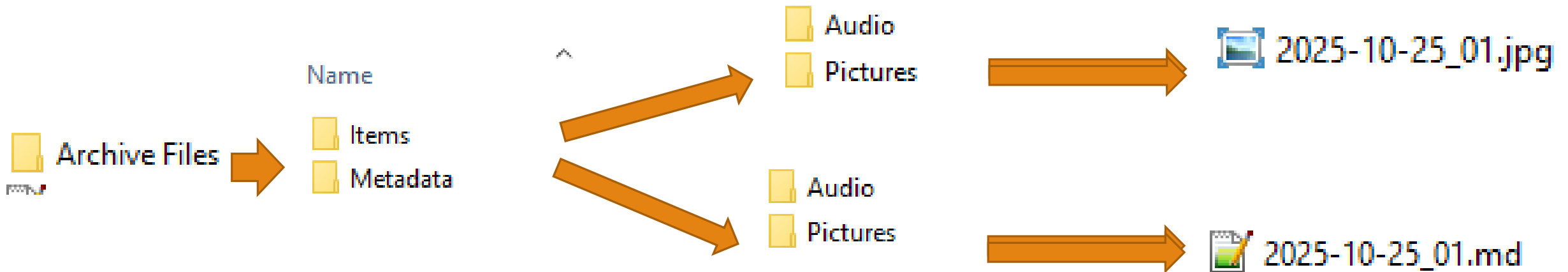
Note: You can change the names of text files in GitHub, **but not images or other media**

# Staying Organized

---

How do you keep track of all the files?

Consider:

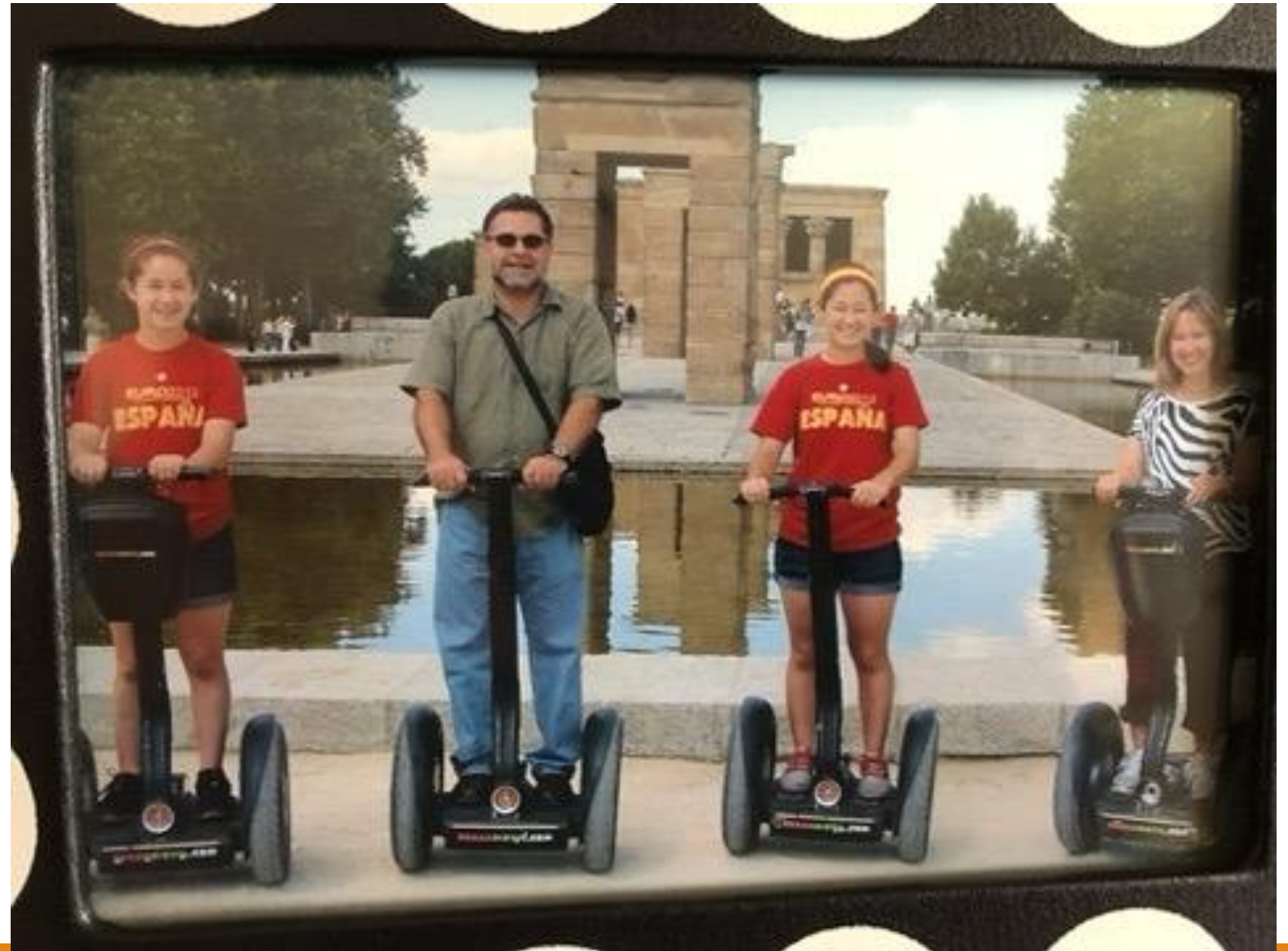


***Remember: The more work you do before uploading to GitHub, the less you have to do afterwards!***

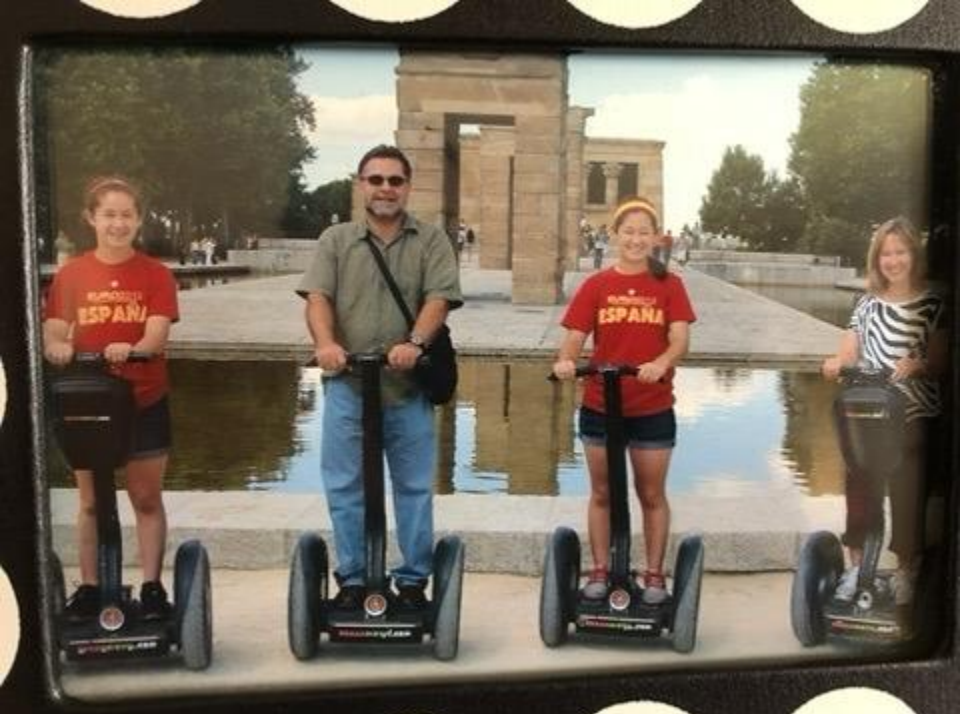
# Photo Editing: Fundamentals, Do's, and Don't's

---

layout: item  
format: photo  
title: Segway Tour  
creator: Luis González  
contributor: Luis González  
creationdate: July 15, 2012  
type: printed photograph in  
frame  
short desc: Luis, Arlene,  
Maria, and Claudia take  
segway tour of Madrid.







**Don't: add filters!**

Filters can be distracting, and more importantly, erase key details about the photograph of the item.

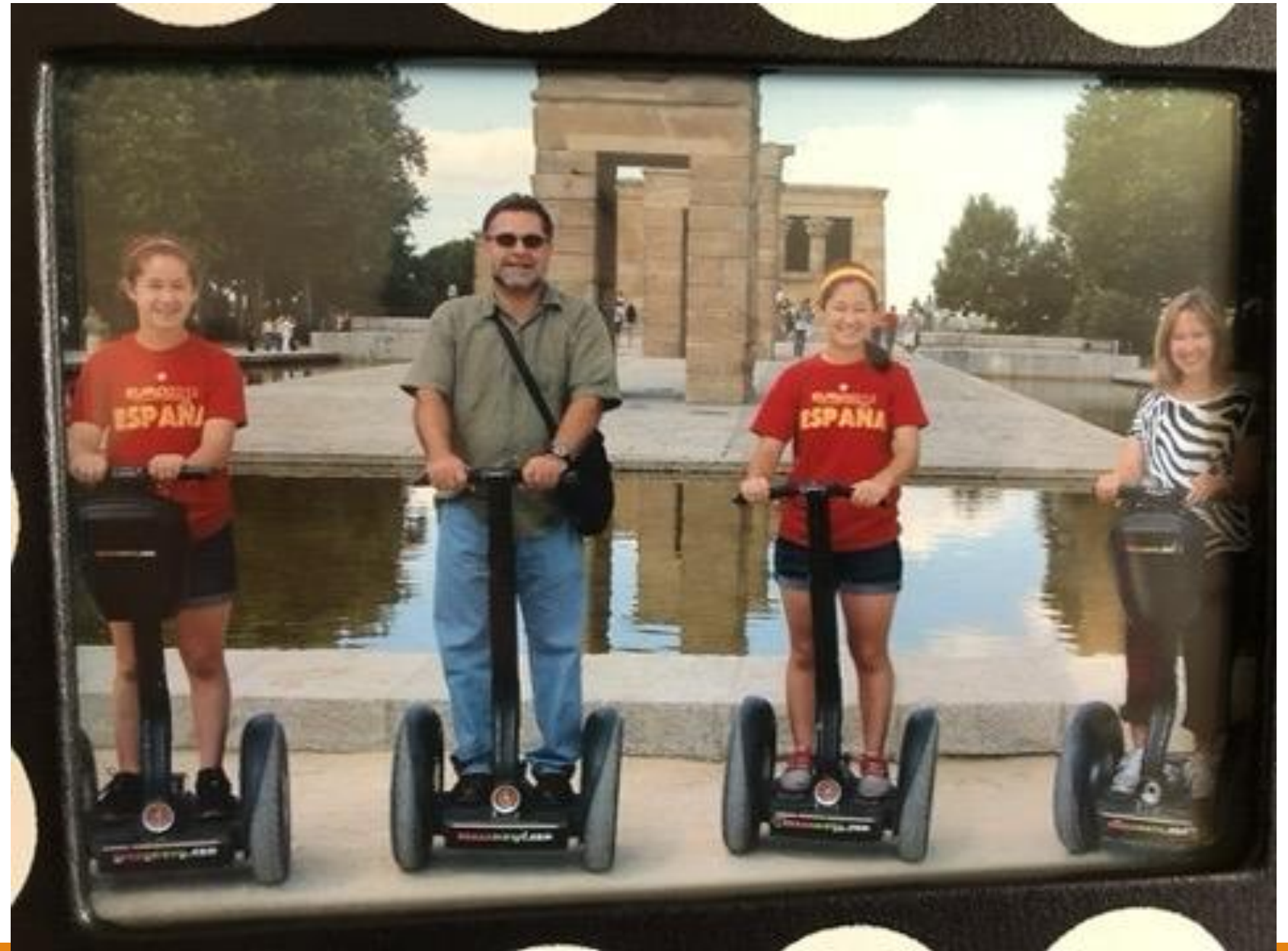




**Don't:** save your photograph as HEIC or PDF

**Do:** save your photo as .jpg or .png

Consider making a copy of your photograph before editing it.



**Don't:** edit, crop, or blur out the image so much that it becomes misleading

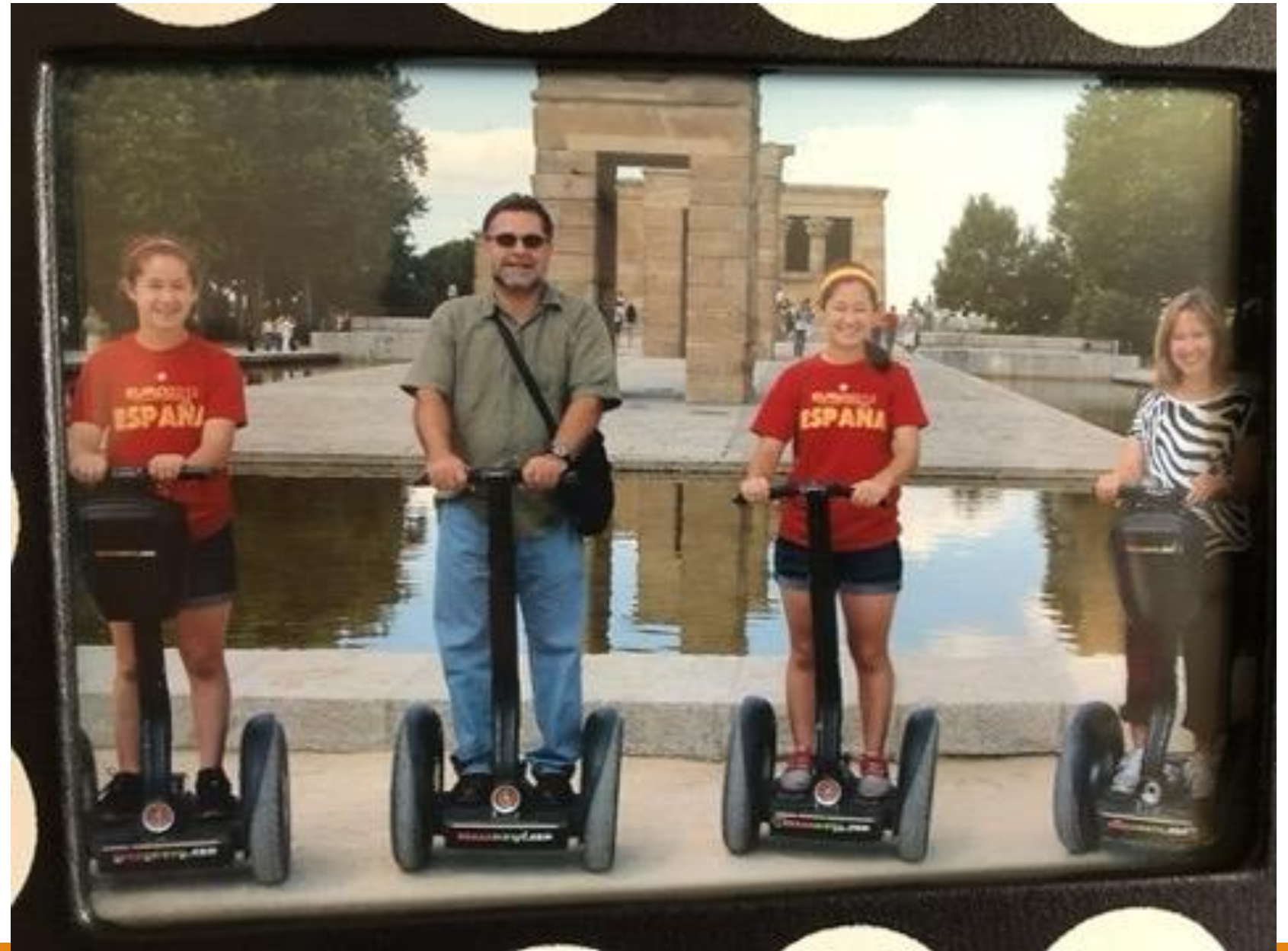
**Do:** blur out sensitive material such as mailing or email addresses





**Don't:** save your photograph as HEIC or PDF

**Do:** save your photo as .jpg or .png



**Don't:** Use photos from the internet or other sources that do not come from you, your collaborators, or the participants in the history harvests







2024-12-30-000.jpg



2024-12-30-001.jpg

**Do:** use a naming convention

For example:  
Yyyy-mm-dd-###

This includes saving the images together in a folder.



2024-12-30-002.jpg



2024-12-30-004.jpg



Original image: 2024-12-30-005.jpg

**Do:** crop the photo (but not too much)







2024-12-30-001.jpg

**Tip:** think of yourself as a photojournalist who is communicating information through photography, not an artist who is expressing herself.

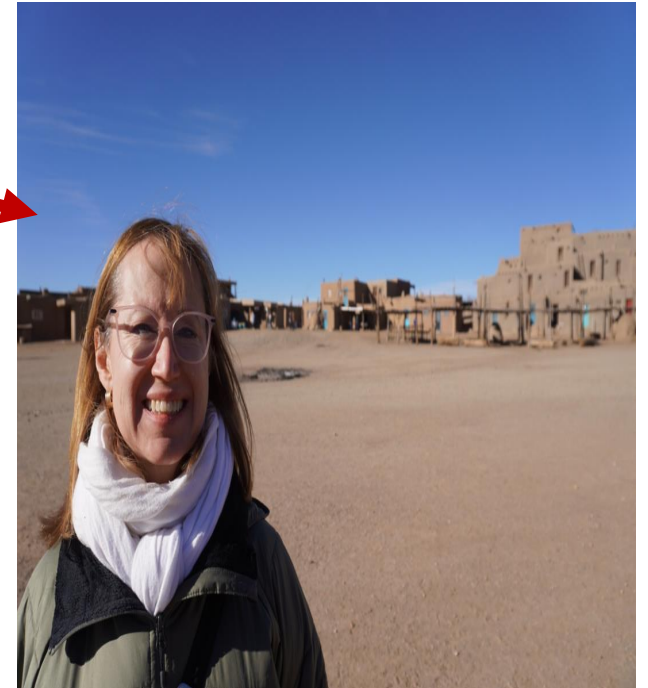
**Do:** subtly adjust brightness, contrast, and saturation to bring out details of the photograph.





**Don't:** disrupt the aspect ratio.

2024-12-30-000.jpg





before



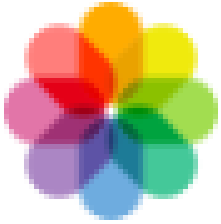
**Do:** rotate photos so that the horizontal and vertical lines in the photo are parallel to the edges.

after



Which Mac software lets you crop and save photos without downloading anything?

1. Photos app.



2. Preview.



Tip: these options are available to users with IOS 16 or later.

# Official documentation on how to edit:

---

Apple: <https://support.apple.com/guide/iphone/edit-photos-and-videos-iphb08064d57/ios>

Android: <https://www.samsung.com/us/support/answer/ANS10003229/>

# Documentation to support you

---

The DigitalArc Team has draft written instructions for everything we covered in today's workshop. You all have access to this information through the following links:

## Post-Event Clean-up Overview

- <https://digitalarcplatform.github.io/documentation/docs/postEvent/>

## File Naming

- <https://digitalarcplatform.github.io/documentation/docs/postEvent/naming/>

## Transcription Editing

- <https://digitalarcplatform.github.io/documentation/docs/postEvent/transcription/>

## Photo Editing

- <https://digitalarcplatform.github.io/documentation/docs/postEvent/photo/>

## Creating Metadata

- <https://digitalarcplatform.github.io/documentation/docs/postEvent/metadata/>

# Technical Support (Office Hours)

---

The DigitalArc Team is here to support you all in your technical needs. In the coming weeks, we will offer virtual technical support in case you have questions or need assistance.

Our availability is:

- **Tuesdays** 5-6 pm ET/ 4-5 pm CT & 8-9 pm/ 7-8 pm CT
- **Wednesdays** 11 am-12 pm ET/ 10-11 am CT
- **Thursdays** 9 – 10 pm ET/ 8-9 pm CT
- **Saturday, November 8**, 12-1 pm ET/ 11 am – 12 pm CT

Use this link to schedule a meeting with us:

- <https://calendly.com/digitalarcplatform/30min>

If our availability does not work with your schedule, please email us at [digitalarcplatform@gmail.com](mailto:digitalarcplatform@gmail.com) to find an alternate meeting time.

Tell us we need to add documentation

<https://github.com/DigitalArcPlatform/documentation/issues/new>

# Next Steps

---

In the coming weeks, we hope that you will be able to continue to apply what you learned today to some of the material that your community has been able to collect. Consider

The goal would be to process transcripts, images, and metadata for 2-4 objects in your collection.

Our next workshop will be sometime **in January** and will focus on taking the content that you've prepared for those 2-4 objects and **uploading to publish** on your community sites.

Exact date for workshop is not set yet; we are waiting on finalizing work schedules for the new year.