

Decoding the Past

AI-Powered Photo Analysis for Genealogists

© 2026 Robert Cameron Weir
rob@yourigins.com | www.yourigins.com

How Photographs Become Genealogical Evidence

The *source-information-evidence* framework is at the core of genealogical analysis. Photographs are sources. They can be interpreted to extract information, and this information can be applied as evidence for arguing an answer to a genealogical question.

QUESTIONS

- Who is this?
- What occasion?
- Where was it taken?
- When was it taken?
- How old was she?

INFORMATION

- Physical: Size, medium, and format
- Mounting: Studio or lab markings, frame style, edge style
- Annotations: Content of any notes, as well as the style of handwriting
- Hair or facial hair style
- Clothing
- Accessories: Hats, bows, neckties, jewelry, fans, book, diploma, sword, etc.
- Setting: Backdrop, props, etc.
- Pose: Individual or clustering within a group

EVIDENCE

- Approximate date the photograph was taken
- Approximate age of the sitter
- Approximate date of birth
- Marital status
- Socio-economic status
- Where the photograph was taken
- Occasion of the photograph

The Strange World of AI Vision

Today's familiar chatbots, like ChatGPT (OpenAI), Claude (Anthropic), and Gemini (Google) are all *multimodal*, meaning they can process other forms of data, like images, audio, and video, in addition to text. Via a process of *embedding*, they encode details, impressions, and overviews of the image, simultaneously, and relate those to other information the model has encoded about these same topics.

The details of how this comes about are quite complicated. Fortunately, we don't need to understand how it works "under the hood." The important thing is to realize that it is quite strange. We should avoid assuming that the AI will see things as you and I do. It will often overlook extremely obvious details, while at the same time noticing things that we miss. It is not necessarily better than we are, but it is very different than we are, and that brings value as a source of second opinions.

Digitizing Photographs

Ideally, old photographs will be scanned once and then safely stored. This avoids repeated handling and exposure to the bright light of scanners. A typical process creates a high-resolution archival copy of the image that can then be further processed for different uses: a restored copy for printing, a lower-resolution copy for sharing online, etc. AI analysis fits within this process and benefits from specific additional secondary copies made from the archival image.

A typical archival process for digitizing photographic prints looks something like this:

1. If safe to do so, lightly brush the original to remove dust
2. Scan on a flatbed scanner
 - ▶ In 8-bit color (even for monochrome prints)
 - ▶ At least 600 pixels per inch (ppi)
 - ▶ Higher resolution (1200 ppi+) might be justified for very small images (passport photos) or where more intensive restoration is needed
 - ▶ Save in a format like TIFF, never a "lossy" format like JPEG.
 - ▶ Include the entire photograph, front and back, including the original mounting, frame, or case
3. Avoid further processing of the archival digital image. Treat it as a read-only master, and do any further processing on a copy.
4. Keep a detailed log, perhaps in a spreadsheet, associating the digital file name with the source of the photograph, e.g. "Sally (Miller) Henry's blue scrapbook, page 23, upper left."
5. Back up the archival images and logs, ideally to more than one place, one of which should be in a different place than where the original photographs are stored. (You want to avoid a situation where a single event, like a fire or flood, destroys both the original and the digital archive at once.)

Note: The above describes digitizing photographic prints. Steinhoff's book, listed in the Resources section, discusses how to digitize negatives and slides.

Preparing Digital Images for AI

Ideally, you will start with a high-resolution archival digital image. It will represent the best quality reproduction you could obtain. Anything far less than this may result in subpar results. In particular, images downloaded from most websites are best avoided. They are typically low-resolution, cropped to cut off the frame and studio markings, and are often digitally edited. Just as a genealogist knows to seek out the most original form of a record possible, seek out the best quality digital reproduction you can.

If you have followed the digitization steps described in the previous section, you would take the archival TIFF file and convert it to one or more PNG or JPEG files to upload to the AI.

Why convert? The LLM chatbots cannot read TIFF files. And even if they could, the chatbots also have limitations on pixel size and file size for how large an image they can process:

LLM	Size before downscaling	File size limit	Number of attachments
ChatGPT	~ 2000 pixels	20 MB	10
Gemini	~ 2000 pixels	100 MB	10
Claude	~ 8,000 pixels	500 MB	20

(These limits are current as of today, but expect them to change from release to release, often in an upward direction.)

The practical solution is to produce a single overview image, in JPEG or PNG format, around 2,000 pixels on the longest side, as well as up to nine additional cropped images of relevant details in the photograph. Here you need to exercise your human judgment as to which details in the photo matter most. For example, for a military uniform, the details are often in the cap, the shoulders, the collar, the medals, the buttons, and the belt. For a woman's portrait the relevant details are hair style, bows, collars, sleeves, and jewelry. For men, the style of facial hair and the necktie are often important details.

Prompting Techniques

1. Start off by making sure you are using a “thinking” model in your chatbot set to a “high” level of effort. Do not use the default setting, which is usually an less-capable “instant” model.
2. Upload your overview image as well as any detail crops you have.
3. Include an image of the reverse as well, if anything is written there
4. Include any relevant background information that you know about the photograph, based on provenance, etc.
5. Enter your prompt.

I'm not an advocate of lengthy prompts. I'd rather lead the AI through the analysis, step-by-step and interactively, responding and reacting to what it says at each step.

Some specific techniques I've found useful include:

- Say “take a detailed look” or “examine carefully” or similar language to encourage it to focus on details rather than an instant impression.
- First ask for a detailed analysis of the photograph, describing in detail everything seen, before then following up with higher level questions about dates and places.
- If you are suspicious about the veracity of the information written on the back of the photo (perhaps it was written many years later), don't include that info in your prompt. Instead, see if the AI, analyzing the photo alone, comes up with a different interpretation.

Note: There are two basic ways to date a photograph: 1) Based on internal evidence such as fashion, 2) Based on knowing birth of at least one person in the photo and then estimating their age in the photo.

Known young children work particularly well for that second kind of dating, since the difference between a 2 year old and a 7 year old is more evident than the difference between a 62 and a 67 year old.

Note also that dressier (formal) photos are easier to date than snapshots in work clothes, due to more fashion clues.

Verification

AI responses must be verified. You should never just accept them as true. When analyzing a photograph, AI is making observations, “seeing” details of the image, as well as making inferences about those details. Both observations and inferences should be verified:

- If the AI says a woman is wearing a wedding ring, then zoom in and see if you see it as well. Is it on the left hand? Ring finger? Apply your human judgment here.
- If the AI says a man is wearing a “bowler” hat, and you don’t know what that is, then look it up on Google and look at some examples. Then judge whether the AI classified the hat correctly.
- If the AI makes an inference that such-and-such a fashion was mainly popular in the 1880s, then you may need to dig a little more to verify. Some of the books listed in the Resources can help identify what fashions were popular at what times.

Boundaries and Risks

- AI does not have superhuman vision. If you cannot see it, the AI probably cannot either. Details that are unclear to you will be unclear to it.
- Limitations of the original or the scan are common culprits. Snapshots are often not in sharp focus. Studio photographs are often better.
- Limitations of training data lead to limitations in the model, including cultural and class bias.
- Specialized forensic software can compare faces, do age progression, estimate whether people are related, etc. But the ability for general purpose AIs like ChatGPT to do this accurately is unproven.

Resources

- Ctein, *Digital Restoration from Start to Finish: How to Repair Old and Damaged Photographs*. Focal Press, 2006.
- Fitzpatrick, Colleen. *The Dead Horse Investigation: Forensic Photo Analysis for Everyone*. Rice Books Press, 2008.
- Mannon, Melissa. *The Unofficial Family Archivist: A Guide to Creating and Maintaining Family Papers, Photographs, and Memorabilia*. ArchivesInfo Press, 2011.
- Moorshead, Halvor (ed.), *Dating Old Photographs 1840-1929*. Moorshead Magazine Ltd., 2000.
- Moorshead, Halvor (ed.), *More Dating Old Photographs*. Moorshead Magazine Ltd., 2004.
- Reilly, James M. *Care and Identification of 19th-Century Photographic Prints*. Eastman Kodak Company, 1986.
- Severa, Joan. *Dressed for the Photographer: Ordinary Americans and Fashion, 1840-1900*. Kent State University Press, 1995.
- Steinhoff, Sascha. *Scanning Negatives and Slides: Digitizing Your Photographic Archives*. Rocky Nook, 2007.
- Taylor, Maureen A. *Photo Organizing Practices: Daguerreotypes to Digital*. Picture Perfect Press, 2014.
- Taylor, Maureen A. *Searching for Family History Photos: How to Get Them Now!* Picture Perfect Press, 2014.