AI for Genealogists: The Best of What's Available Today

© 2025-2026 Robert Cameron Weir
rob@yourigins.com | www.yourigins.com

ChatGPT (OpenAI)

— Broad knowledge base trained on a large portion of the internet
— Great for content generation, problem-solving, and reasoning
— Can interpret images
— Can translate languages
— Can teach you to read new languages and interpret new record types
— Caveats: prone to hallucinations, lacks source citations

Gemini Notebook (Google)

— You upload your own sources to create a personal research assistant
— Useful for analyzing books, family records, court documents, and reports
— Has a huge context window (~ 1 million words)
— Returns citations and supports structured research tasks
— Work with audio or video files by automatically creating a transcription
— Handle modern handwriting, but not early handwriting
— Correlates information from multiple source documents or manuscripts

Transkribus (READ-COOP)

— Designed for transcribing handwritten and printed historical documents
— Supports many languages and scripts with prebuilt models
— Can train new models based on handwriting samples that you provide
— Character Error Rate (CER) is around 5%, so it still requires review and editing.

Example Prompts and Use Cases

— ChatGPT: "Can you explain how inheritance worked in 19th century Quebec?"

— Gemini Notebook: Upload PDFs of local and state histories and ask, "List all physicians or surgeons working in New Hampshire before 1776."

— Transkribus: Scan and transcribe colonial deeds, church registers, or letters.

Sources for Sources

The most powerful uses of AI for genealogists often involve bringing specific, relevant records and documents into your AI session.

Here are some good places to find sources, beyond what you already have on your hard drive:

— The Internet Archive (https://archive.org/)

— HathiTrust (https://www.hathitrust.org/)

— Google Books (https://books.google.com/)

— FamilySearch Digital Library (https://www.familysearch.org/library/books/)

— Local, regional, and state historical societies and state universities

You Are What You Eat

Computer programmers have an old saying: Garbage in, Garbage out (GIGO), meaning that your algorithms can be brilliant, but if they are fed bad data, you can expect bad outputs. When you select the sources you are in control. But you need to choose wisely.

Weakness—If a particular book is so weak, due to bad research, uncited claims, known errors, etc., that you would not think of citing it in a report, then don't expect the AI is going to magically transform it into something useful. Feed your AI high-quality sources.

Bias—Watch out for bias introduced by your selection of sources. For example, if you are researching early colonial encounters with Native Americans and use only 19th-century narratives written by the descendants of English settlers, then AI may reflect that bias back to you.

Quality—Not all PDFs are created equal. Some are "native-digital" created by exporting from programs like Microsoft Word. Some were created decades ago with "print to PDF" drivers. Others were created from scanned images, undergoing optical character recognition (OCR). Generally speaking, modern, born-digital PDFs work very well with AI. Older ones, and ones created from scans, may or may not work well.

A quick test of file quality is to first load the document into your PDF reader and see if you can copy a page or two of text and paste it into a word processor. If the resulting text looks reasonable, then expect the AI will do a good job with it. However, if the text is all scrambled, with random characters all over the place, then the AI may struggle with it. You may be able to fix these errors by rescanning the original or re-OCRing the scanned images.

Caveats and Best Practices

— AI tools are not infallible. Always verify information.
— Be aware of knowledge cut-off dates, especially with ChatGPT.
— Results can vary across sessions and tools.
— Be careful of bias introduced by reliance on dated books.
— Use AI as an assistant, not an authority.
— Trust your genealogical instincts—AI helps but does not replace expertise.

Resources

— Book: Robert Cameron Weir, *ChatGPT for Genealogists* (Antic Press, 2026)
— Facebook Group: Ancestors and Algorithms: AI for Genealogy
— Facebook Group: Genealogy and Artificial Intelligence (AI)