Teaching Genealogists AI™
Instructor: Sylvia Vargas

AI Tools for Research and Organization

Module 1 – AI as a Research Assistant: What It Can and Cannot Do

Course Instructor: Sylvia Vargas
Organization: Puerto Rican Genealogy Group
Website: https://puertoricangenealogy.org
Date Created: 2026-03-05
Estimated Time: 20–30 minutes

Learning Objectives

By the end of this module, you will:

- Understand what AI tools can and cannot do in genealogical research

- Know how to create a Genealogist Profile prompt that improves every AI session
- Understand how the five elements of the GPS relate to AI-assisted research
- Know the privacy rules for sharing genealogical data with AI tools

---

What AI Is and Is Not

AI tools like Google Gemini and Claude are language models. They predict text. When you give them a prompt, they generate a response based on patterns learned from large amounts of text. This has real implications for genealogical research.

What AI can do:

- Compare two records you paste and identify consistencies and contradictions
- Extract structured data from transcriptions you provide
- Suggest record types that might exist for a given time period and place
- Draft research plans based on information you supply
- Identify patterns in data you share (witness clusters, naming conventions)
- Explain historical context, record systems, and terminology
- Convert unstructured notes into organized formats

What AI cannot do:

- Search genealogical databases on your behalf
- Access FamilySearch, Ancestry, or any external repository
- Verify that a record exists or that a fact is true
- Replace your judgment about whether evidence is sufficient
- Guarantee accuracy – AI can generate plausible-sounding but incorrect information

The fundamental rule: AI works with what you give it. If you give it accurate transcriptions and clear questions, it produces useful analysis. If you give it vague questions or assume it can retrieve records, it will disappoint.

AI and the Genealogical Proof Standard

The GPS has five elements. AI can assist with some and cannot substitute for others.

GPS Element	AI Can Help	AI Cannot Substitute
1. Reasonably exhaustive research	Suggest record types you may have missed	Search repositories or confirm records exist
2. Complete citations	Format citations you draft	Verify that a source exists or is accurate
3. Analysis and correlation	Identify consistencies and contradictions	Evaluate original documents or images

GPS Element	AI Can Help	AI Cannot Substitute
	across records you share	
4. Conflict resolution	Reason through conflicting evidence you present	Determine which source is more reliable without your input
5. Written conclusion	Draft narrative summaries from facts you provide	Draw genealogical conclusions independently

The GPS requires reasonably exhaustive research – meaning you, the researcher, must search the relevant repositories. AI cannot do this for you.

GPS Checkpoint

Before using AI output in your research record, ask:

1. Did I provide accurate source transcriptions, or did I summarize from memory?
2. Has the AI suggested records it cannot verify exist?
3. Have I labeled any AI-generated text as a draft requiring verification?
4. Does the AI output contain any inferences I need to confirm against original sources?

Gemini vs. Claude: A Brief Comparison

Both tools are useful for genealogical research. Neither is definitively better – they have different strengths.

Feature	Google Gemini	Claude (Anthropic)
Access	Free tier available; Gemini Advanced paid	Free tier available; Claude Pro paid
Context window	Large; handles long documents	Large; handles long documents
Strengths	Google integration; web search option	Instruction-following; long document analysis
Privacy	Review Google's data policies	Review Anthropic's data policies
Best for	Quick lookups, drafting with web context	Long transcription analysis, structured output

Use whichever tool you are comfortable with. The prompting principles in this course work with both.

Create Your Genealogist Profile

A Genealogist Profile is a reusable prompt you paste at the start of each AI session. It tells the AI your research context so you do not have to repeat it in every prompt.

Template:

> I am a genealogist researching Puerto Rican families. I follow the Genealogical Proof Standard. I am researching [family surnames] in [municipalities], approximately [date range]. My current research question is: [specific question]. I will share records as transcriptions. Please do not search the internet or suggest records you cannot verify exist. Flag any response that contains speculation.

Adjust the bracketed fields to your research. Save this as a text file and paste it at the start of each new AI session.

Before and After: The Difference a Good Prompt Makes

Vague prompt:

> Tell me about José Rivera from Puerto Rico.

The AI will generate a generic, unverifiable response about common Puerto Rican surnames and history. This is not useful for research.

Structured prompt (with Genealogist Profile active):

> I have two records that may refer to the same person. Record A: José Rivera, age 25, Utuado, 1910 census. Record B: José Rivera Torres, age 61, death certificate, Utuado, 1945. Please compare these records and identify evidence that supports or contradicts a match. List unresolved questions.

The AI now has specific data to analyze and a clear task. The output is something you can evaluate and document.

Privacy Rules for Genealogical AI Use

Before pasting any genealogical data into an AI tool, apply these four rules:

1. **Living persons:** Never paste identifying information (name, date of birth, location, Social Security number) about any living person or anyone whose death you cannot confirm.
2. **Sensitive records:** Psychiatric, criminal, and adoption records carry additional privacy risks. Avoid pasting full record text; use summaries with names removed.

3. **Data retention:** Review each tool's privacy policy. Some tools use your inputs to train future models. Use the privacy or "temporary chat" settings where available.
4. **Family members:** If a record reveals information about a living relative that they may not know (unknown parentage, health conditions), do not share that information with AI tools without that person's consent.

What's Next

In Module 2 – Record Matching: Solving Common-Name Problems, you will learn how to write a comparison prompt that produces a structured analysis of whether two records belong to the same person – including how AI handles Puerto Rican naming conventions and name variations across record systems.

Back to Course Overview · *Module 2 ->*