

Researching a Family History with Claude

The Complete Kit

A three-part guide to turning your curiosity into proof a stranger can trust
The Guide · The Toolkit · An Annotated Sample

DeWayne Lee Searcy
2026

Companion to Writing a Family History with Claude — the narrative kit.

Contents

Part One — The Guide.....	3
Before you start: what this is, and who it's for.....	3
The one promise that makes it trustworthy.....	3
What you need in hand before you begin.....	3
Setting up your workspace.....	4
The evidence discipline: five tiers.....	4
Numbering the family: the Ahnentafel.....	5
Working the documents.....	5
The neighbors are evidence: the FAN method.....	5
DNA as its own witness.....	6
Proving a line: the society-application path.....	6
Keeping it findable.....	7
When the machine can reach further.....	7
What only you can do.....	7
Part Two — The Toolkit.....	7
1. Project setup checklist.....	8
2. The standing-instructions block.....	8
3. The five-tier evidence cheat sheet.....	8
4. The Ahnentafel quick-reference.....	8
5. The source-citation template.....	9
6. The DNA intake checklist.....	9
7. The neighbor-cluster ledger.....	9
8. The application-audit checklist (society-standard proof).....	9
9. The file-naming convention.....	10
10. The session-handoff note.....	10
11. The tie-out (verification) checklist.....	10
Part Three — A Sample, Annotated.....	10
The records it stands on.....	10
The problem.....	11
The neighbor-cluster walk.....	11
What only you could do.....	11
Why this is the whole method in one case.....	11
Appendix: the documentation underneath.....	12

Part One — The Guide

A practical guide to rigorous family-history research with Claude

Before you start: what this is, and who it's for

This is a practical guide to doing serious family-history research with an AI collaborator — the kind of work that used to take a professional genealogist a career, done at the pace of an evening, without giving up any of the rigor that makes it true.

It is written for two readers. The first is the newcomer: someone with a family name, a few names on a chart, and the wish to know where they came from — who has never heard of an Ahnentafel and does not need to, yet, to begin. The second is the family that inherits the work: the daughter or grandson who will one day open the archive and want to carry it forward. Everything here is written so that either reader can pick it up cold.

A word about what “with Claude” changes and what it does not. The AI does not replace the researcher. It removes the drudgery — the reading of a hundred census pages, the drafting of a citation, the cross-checking of a date against three records — and it hands the judgment back to you. Which conflict to resolve, which lead to trust, which record actually settles a question: that stays human, and this guide will keep saying so, because the moment you forget it is the moment the work stops being true.

The one promise that makes it trustworthy

Every research method needs a promise it will not break. This one has a single promise, and the whole method hangs from it:

Every claim ties to a source you can pull yourself, and the AI proposes while you decide.

That is the entire trust model. You do not believe a conclusion because the machine is confident; you believe it because it names a record — a census page, a will, a marriage bond, an image number — that you can go open and read for yourself. Secondary sources (family trees online, a cousin's newsletter, a published lineage) are treated as *leads*, never as proof, and they are labeled that way so no one later mistakes a rumor for a record. And nothing that matters — no change to your master workbook, no promoted conclusion — happens without your sign-off. The AI drafts, tiers, and proposes; you post.

If you hold that one promise, everything else in this kit is technique. If you break it, no amount of technique will save the result.

What you need in hand before you begin

You can start with almost nothing — a surname and a rough place — but the work goes faster if you gather these first:

A **starting chart** of what you already believe, however thin: names, approximate dates, and places for as many ancestors as you can. A handful of **home sources**: photographs of documents, a family Bible, old letters, a DNA test if anyone has

taken one. And a **place to keep everything** — a single folder, backed up, that will become your archive. That is enough to begin. The method grows the rest.

Setting up your workspace

The workspace has three parts, and setting them up once saves you a hundred small frustrations later.

The project. Create a dedicated Claude project for your family. A project remembers across sessions — the findings you save into it are there next time, so the work compounds instead of starting over. When a project fills up (it will, eventually, if the family is large), you start a successor project and carry a short “state of the project” summary into it. Think of each project as a volume; the family history is the shelf.

The workbook. Your system of record is a single spreadsheet — the master workbook — with a tab for the direct ancestors, tabs for the DNA, tabs for open questions, and a change log. Everything else in the archive is evidence that feeds it. The workbook is the one file that must never move without your say-so; it is the ledger, and the ledger is sacred.

The standing instructions. At the top of your project, paste a short block of standing instructions the AI reads every session (the Toolkit gives you the exact text). It says, in effect: *primary sources first; grade every claim; label leads as leads; never invent to fill a silence; propose changes to the workbook but never make them without asking*. This is the prime directive, and it is the difference between a research assistant and a rumor mill.

The evidence discipline: five tiers

The heart of rigorous genealogy is not finding facts. It is being honest about how well you know them. Every claim in your work carries a tier, and the tier travels with the claim wherever it goes:

- **DOCUMENTED** — proven by a primary record (or the convergence of several) that you can re-pull. The birth certificate, the will naming the child, the census that puts the family in the county.
- **PROBABLE** — the weight of evidence points one way, but a single clinching record is missing. Most working genealogy lives here honestly for a long time.
- **POSSIBLE** — a plausible lead worth pursuing, not yet weight-bearing.
- **SPECULATION** — an idea, a hunch, a tree someone else built. Useful to record so you don’t chase it twice; never to be leaned on.
- **DISPROVEN** — actively ruled out. Worth keeping, because a disproven line saved is a wrong turn no descendant has to take again.

The discipline is simple and unforgiving: a conclusion may never quietly rise a tier. A PROBABLE does not become DOCUMENTED because it has sat there a year and started to feel true. It rises only when a new record earns the promotion, and the record is named. The AI is good at proposing tiers; you are the one who ratifies them, because the tier is a judgment and judgment is yours.

Numbering the family: the Ahnentafel

You cannot keep a large family straight in prose. The tool that makes it manageable is a numbering system called the Ahnentafel (German for “ancestor table”), and it is cleverer than it looks.

You are number 1. Your father is 2, your mother is 3. Every person’s father is double their number; every person’s mother is double-plus-one. So your father’s father is 4, his mother 5; your mother’s father is 6, her mother 7 — and so on, back through the generations. From the number alone you can read two things instantly: which generation a person sits in, and which side of the family they are on. A single integer carries the whole coordinate.

You do not have to compute these by hand — the AI maintains them in the workbook — but understanding the logic lets you read the master like a native. When you see #142, you know without looking that it is a specific great-(times four)-grandparent on a specific line. (The one refinement worth knowing: where a line reaches into an external register with its own identifiers — a Swiss family book, say — that outside number rides *alongside* the Ahnentafel number as a cross-walk, and never replaces it. The Ahnentafel number is always the primary key.)

Working the documents

Most of genealogy is documents: censuses, wills, deeds, marriage records, pension files, death certificates. The method for handling them has three moves.

Triage. When a batch of new material arrives — say forty scanned pages — you do not read them one at a time and lose the thread. You cluster them: which documents speak to the same person, the same event, the same question. The AI reads across the batch and proposes the clusters; you get an ordered queue of analytical threads instead of a heap.

Citation. Every fact that lands in the workbook carries a citation precise enough to re-pull — not “a census,” but the year, the county, the enumeration district, the dwelling number, and where the image lives. A citation you cannot follow back is not a citation. The Toolkit gives you a template; the rule behind it is that a stranger with your citation should be able to open the same record you did.

The workbook. Confirmed facts flow into the master workbook with their tier and their source. This is the annotation-only discipline in practice: the AI proposes the row, you approve it, and only then does the system of record move. The workbook is where the family becomes a structure instead of a story.

The neighbors are evidence: the FAN method

Here is the technique that breaks the hardest walls. When a person goes dark in the record — a maternal line that vanishes between two censuses, a common name you cannot pin down — you stop looking at the person and start looking at the people *around* them. Genealogists call it FAN: Friends, Associates, and Neighbors. Families did not move alone. They moved in clusters — in-laws, former neighbors, fellow church members — and the cluster leaves a trail even when the individual does not.

The mechanic is concrete. For a census where your family appears, you read not just their household but the nine households before and the nine after, across every decade you can find them, and you build a small ledger of who was near them and when. Sort that ledger one way and you see *proximity* (who lived next door); sort it another and you see *persistence* (who kept turning up beside them across the years and the miles). A neighbor who moves with your family from one county to the next across three censuses is not a coincidence — they are evidence, and often they are the thread that identifies your ancestor when the ancestor's own record is silent.

This is patient work, and it is exactly the kind of work the AI transforms: it can read the neighbor households and build the ledger in an evening that would have taken weeks by hand. But note the division of labor, because it is the whole lesson of this kit in miniature: the AI reads and tabulates the pages; *you* decide which persistent neighbor matters, and what their persistence proves. The reading compresses; the judgment does not.

DNA as its own witness

DNA is a witness that documents cannot cross-examine and cannot lie to, and it comes in three kinds, each answering a different question.

Autosomal DNA (the common test) covers all your lines but fades within a handful of generations; it is best for confirming recent connections and for clustering matches into the family groups they descend from. **Y-DNA** follows the strict paternal line — father to father to father — and is read through a terminal “haplogroup” marker and a match list that place your patriline on a branching tree; a matching surname on a close branch can independently corroborate a paper line. **Mitochondrial (mt)** DNA follows the strict maternal line the same way, mother to mother to mother.

The workflow is end-to-end: test, upload the raw data to the platforms that allow it, read the matches into clusters, and — crucially — *integrate* the genetic evidence with the paper evidence rather than treating it as a separate hobby. A genetic cluster can confirm what a documentary cluster predicts, and a documentary cluster can explain what a genetic match means. But DNA has the same tier discipline as everything else: a genetic inference is corroborative, and absent a documentary anchor it caps a suggested link at PROBABLE, not DOCUMENTED. The double helix is a powerful witness; it is still a witness, not a verdict.

Proving a line: the society-application path

Sometimes you want to prove a line to a standard higher than your own satisfaction — a lineage society such as the DAR, where a registrar who has never met you must be convinced, generation by generation, with a primary record for every link. This is the highest-rigor path in genealogy, and it is worth understanding even if you never file, because it is the standard everything else is measured against.

The path is a chain of proof: each parent-child link in the line from you to the qualifying ancestor must be documented by a record a registrar will accept, with no gap papered over by a family tree. The method here is the same as everywhere else, turned up to its strictest setting — DOCUMENTED at every rung, citations a stranger can re-pull, and a pre-filing audit that hunts for the weakest link *before* the registrar does. If you can build a line to society standard, you can trust it for anything.

Keeping it findable

An archive no one can navigate is a drawer of loose photographs. A consistent file-naming convention — subject, type, date, version — turns thousands of files into something you and your heirs can search. Name files so that sorting them alphabetically groups them sensibly; version them so the latest is obvious; and keep a folder structure that separates working files from the finished record. The Toolkit gives you a convention; the principle is that the person who inherits this — who is not you and does not know where you put things — should be able to find any record from its name alone.

When the machine can reach further

The newest capability is worth naming plainly, because it changes what the AI can do and therefore what you must watch. An AI collaborator can now, with your permission, reach into your own archive of files directly and drive your own logged-in browser through the subscription archives you already pay for — pulling and reading records in your session, as you would see them, at a pace no person could match.

This is real leverage, and it comes with a real obligation. When the AI gathers evidence you did not personally watch it gather, your trust can no longer come from having looked over its shoulder. It has to come the way an auditor's trust comes: from re-performance and a documentary trail. So the discipline tightens by exactly one notch. Every finding still names a re-pullable primary. You re-pull a sample yourself — two or three cited records per finding — and confirm they say what the finding claims. You keep a running register of the honest gaps the AI discloses. And the workbook stays annotation-only, so the system of record never moves without you. The machine reaching further is a gift; the audit posture is the price, and it is a price worth paying.

What only you can do

Everything the AI does, it does under your judgment, and there are things it cannot do at all — not as a limitation to be fixed, but as the permanent shape of the work:

You decide what question is worth asking. You decide which record settles it and which merely flatters it. You keep the prime directive when it would be faster to break it — the coherence of a beautiful inference is not a substitute for a page, and where the record is silent, the honest thing is to say so and leave the silence open rather than fill it with invention. You keep faith with the people in the record, especially those the record barely names, and refuse to make anyone a prop for a tidy story. And you carry the responsibility that the work is true — because your descendants will trust it precisely as far as it deserves, and not one inch further.

That is the part that does not compress. Guard it, and the rest of this kit will serve you.

Part Two — The Toolkit

Copy-paste checklists, templates, and prompts

Copy-paste templates and checklists. Adapt the bracketed parts to your family. Everything here assumes the one promise: sources you can re-pull, and the AI proposes while you decide.

1. Project setup checklist

- ☐ Create a dedicated project for the family; name it for the line and a version (e.g., “Smith Ancestry 1.0”).
- ☐ Create the master workbook: tabs for Direct Ancestors (Ahnentafel), DNA, Open Research, Sources, and a Change Log.
- ☐ Paste the standing-instructions block (below) at the top of the project.
- ☐ Create the archive folder with the file-naming convention (§9).
- ☐ Load your starting chart and any home sources.
- ☐ Write a one-paragraph “current state” note you can carry into a successor project later.

2. The standing-instructions block

Standing instructions for this project. Primary sources first: every material claim must cite a primary record I can re-pull (year, place, identifier, image location). Grade every claim on the five-tier scale (DOCUMENTED / PROBABLE / POSSIBLE / SPECULATION / DISPROVEN) and never promote a tier without a new record that earns it. Treat online trees, newsletters, and published lineages as leads, never proof, and label them as leads. Never invent to fill a silence — where the record stops, say so. Propose changes to the master workbook but never make them without my explicit approval (annotation-only). In every findings memo, include an honest “gaps” section listing what you could not verify.

3. The five-tier evidence cheat sheet

Tier	Means	Test
DOCUMENTED	Proven by a re-pullable primary (or converging primaries)	Can a stranger open the record and see it?
PROBABLE	Weight of evidence points one way; clinching record missing	What single record would promote it?
POSSIBLE	Plausible lead, not weight-bearing	Worth an hour to chase?
SPECULATION	Idea, hunch, someone’s tree	Recorded so it isn’t chased twice
DISPROVEN	Actively ruled out	Keep it, so no descendant repeats the wrong turn

4. The Ahnentafel quick-reference

- You = 1. Father = $2 \times$ a person’s number; mother = $2 \times + 1$.
- Even numbers are men; odd numbers (except 1) are women.
- Generation = $\lfloor \log_2(\text{number}) \rfloor + 1$ (you are Generation 1).
- An external register’s identifier rides *alongside* the Ahnentafel number, never replaces it.

5. The source-citation template

[Record type], [year]. [Jurisdiction: county, state/country]. [Specific locator: enumeration district + dwelling/family no. | volume + page | file/certificate no.]. [Repository / database]; [image identifier or URL]. [Archive file name in your system]. **Tier:** [DOCUMENTED/...]. **Supports:** [claim + Ahnentafel number].

6. The DNA intake checklist

- ☐ Record the kit: person, Ahnentafel number, platform, kit number, test type (autosomal / Y / mt).
- ☐ Upload raw data to the platforms that accept transfers; note where each lives.
- ☐ Y-DNA: record terminal haplogroup, the match/block-tree set, and any surname-matching close branch.
- ☐ mtDNA: record the matriline haplogroup and the maternal wall it reaches.
- ☐ Cluster autosomal matches into descending family groups.
- ☐ Integrate with paper: note where genetic and documentary clusters confirm each other.
- ☐ Tier every genetic inference (caps at PROBABLE without a documentary anchor).

7. The neighbor-cluster ledger

For a person who goes dark in the record. Build one row per neighboring household, per census.

Census year	County/ District	Dwelling #	Head of household	Surname	Notes (relationship, later persistence)
- Read 9 households before and 9 after your family, every decade you can find them. - Sort by dwelling # to see proximity; sort by surname + year to see persistence. - A neighbor who moves <i>with</i> the family across counties/decades is evidence — flag them.					

8. The application-audit checklist (society-standard proof)

- ☐ List every parent-child link from the applicant to the qualifying ancestor.
- ☐ For each link, name the primary record proving it (no link rests on a tree).
- ☐ Identify the weakest link and attack it *before* a registrar does.
- ☐ Confirm every citation re-pulls.
- ☐ Note any gap and the record that would close it.

9. The file-naming convention

[Subject]_[Type]_[Descriptor]_[Date]_[vN].[ext] e.g.,
Smith_John_Will_Granville-NC_1801_v1.pdf

- Sort-friendly: alphabetical grouping should be sensible.
- Version the latest so it's obvious.
- Separate working files from the finished record in the folder tree.
- Name so an heir who doesn't know your habits can find a record from its name alone.

10. The session-handoff note

Write this at the end of a working session so the next one (or a successor project) starts clean.

Where we are: [current workbook version; last confirmed finding]. **Open questions:** [the live CHASE targets, tiered]. **Next move:** [the highest-value next record to pull]. **Watch items:** [any continuity issue to hold in view]. **Gaps:** [what remains unverified].

11. The tie-out (verification) checklist

Run this on any finding the AI gathered that you did not personally watch it gather.

- ☐ Re-pull 2-3 cited primaries yourself; confirm each says what the finding claims.
 - ☐ Challenge anything tiered DOCUMENTED that is really a transcription or a tree.
 - ☐ Add the finding's disclosed gaps to your running gap register.
 - ☐ Confirm the workbook has not moved without your approval.
 - ☐ Retain the gathered source copies with the dated finding — they are your workpapers.
-

Part Three — A Sample, Annotated

The whole method holding together in one worked case

One case, worked, to show the whole method holding together. The names and county are real to the method; read it for the moves, which transfer to any family.

The records it stands on

A maternal line in northeast Mississippi went dark in the middle of the nineteenth century. The family — call them by the surnames that actually cluster there, the Searcys, Rories, and Marlars — appears in the 1860 census and again, changed, in 1880; but the 1870 enumeration, the one that would bridge them, seemed to swallow the household whole. The records in hand were ordinary: federal censuses for 1850 through 1900, a scatter of county marriage and probate entries, a family Bible or two, and a handful of tombstone photographs. Nothing that, read one household at a time, closed the gap.

The problem

The specific question was a common and brutal one: was the family in the 1870 census at all, and if so, under what mangled spelling and in whose household? A name that a census-taker heard and spelled by ear can hide a family completely. Read the target household alone, and the trail ends. This is exactly the wall the neighbor method was built for.

The neighbor-cluster walk

The move was to stop hunting the family and start mapping their neighbors. For each census where the family clearly appeared — 1850, 1860, 1880 — the nine households on either side were read into a ledger: who lived near them, and when. The AI read the pages and built the ledger across the decades in an evening; by hand it would have been weeks of squinting at microfilm.

Sorted by persistence, the ledger surfaced a neighbor who kept appearing beside the family across census after census and across more than one county — the kind of stubborn co-occurrence that is never chance. *(Annotation: this is the FAN principle doing its work. The neighbor was not the goal; the neighbor was the handle. — Tier at this stage: the neighbor's persistence is DOCUMENTED from the census images; what it implies about the missing family is still PROBABLE.)*

With the persistent neighbor as an anchor, the 1870 enumeration was re-read around *that* household rather than the family's own name — and there the family was, a few dwellings off, under a spelling the eye had slid past. The neighbor set had identified the household its own surname could not. *(Annotation: the identification rests on the neighbor cluster plus the age-and-composition match of the household. That is strong, but it is the neighbors carrying it — so the promotion goes to PROBABLE-strong, and only rises to DOCUMENTED when an independent record — a marriage, a probate naming the same people — confirms the household is the family. The AI proposed the tier; the researcher set it.)*

What only you could do

Every mechanical step compressed: the reading of the neighbor households, the building and sorting of the ledger, the re-reading of 1870 around the anchor. What did not compress was the judgment at each seam — deciding *which* persistent neighbor was signal rather than noise, deciding that the mangled spelling plus the age match plus the neighbor cluster was enough to identify the household but not yet enough to call it proven, and refusing to let the satisfying coherence of the reconstruction stand in for the confirming record that had not yet been found. The machine found the household. The researcher decided what it meant, and how far it could be trusted.

Why this is the whole method in one case

Look at what this single case exercised: the **five tiers** (a claim held honestly at PROBABLE rather than rounded up to true); the **Ahnentafel** (each identified person taking their number and side); **source citation** (every census read named to the dwelling); the **FAN method** (neighbors as evidence when the person is silent); the **annotation-only discipline** (the workbook moved only when the researcher approved each promotion); and the **audit posture** (the AI gathered, the

researcher re-pulled and ratified). One wall, and the whole kit was in it. That is the point: the method is not a menu of separate techniques but a single discipline, and any real problem calls on all of it at once.

Appendix: the documentation underneath

Behind the two paragraphs of the walk sit the workpapers that make it auditable: the neighbor-cluster ledger itself (every household, every decade, sortable by proximity and by persistence); the census citations, each precise to the dwelling and the image; the tier assigned to every claim with the record that would promote it; and the honest gaps — the confirming probate not yet pulled, the maternal wall the line still hits. None of it is hidden, because the promise the whole method hangs from is that a stranger can check it. Hand this appendix to a genealogist you have never met, and they can walk the same path and reach the same wall — which is the only kind of family history worth leaving to the people who come after you.

Researching a Family History with Claude — The Complete Kit. Companion to Writing a Family History with Claude. The research kit teaches how to find the people and weigh the evidence; the writing kit teaches how to turn that evidence into a book a stranger will feel. Together they are the whole practice: true, and worth reading.