How to Use AI as a Lost Innovator

A practical method for doing credible work — and earning a hearing — without a credential.

William Stafford, ADN, LI-AIast3 Independent researcher

Version 1.2 — July 2026 *Preprint. Deposited on Zenodo for a citable DOI: 10.5281/zenodo.20722113 (this DOI represents all versions).*

This guide is the practical companion to the standard described in *Earned Trust: Verifiable Disclosure as a Credential-Substitute for AI-Assisted Work* (https://doi.org/10.5281/zenodo.20719927). The standard explains *what* to disclose; this guide explains *how* to do the work so that the disclosure means something.

Who this is for

A **Lost Innovator** is someone doing serious intellectual work who lacks the conventional credential the field treats as a license to be taken seriously — no relevant degree, a degree below the bar, a degree in another field, or no institutional affiliation. If that's you, you have probably felt the gate: ideas that go unread not because they are weak but because no one has a reason to spend attention on someone without the badge.

This guide is a way through that does not require the badge. The trade is simple and demanding: **where a credential buys trust in advance, you earn it afterward — by showing exactly how the work was made and letting anyone check.** You will not be asking to be believed. You will be making your work *checkable*, which is a different and more honest thing.

One promise up front, and one warning. The promise: this method can give genuinely good work a real chance of being evaluated on its merits, without waiting for a gatekeeper's permission. The warning: it earns you a *hearing*, not a *verdict*. It does not make weak work strong, and it is not a substitute for actually knowing your subject. What it does is get the work into the room. Whether it survives is up to the work.

Though written for those outside the gate, the method serves anyone who would rather have their AI use be legible and checkable than assumed and suspected — credentialed researchers included.

The one rule that everything else serves

You are the author and the last check. The AI is an instrument, not a co-author. Every method below exists to keep that true. The moment you let a model do your thinking unchecked, or steer it until it agrees with you, you have produced something you cannot stand behind — and the disclosure that should protect you will instead expose you. Most of what follows is therefore aimed inward as much as outward: the steps that let a reader check your work are the same ones that keep you from fooling *yourself* — from mistaking a fluent draft for a true one, or a model's agreement for confirmation. Hold this rule and the rest is mechanics.

The method, step by step

Step 1 — Set a base, add specialists

Pick one model you genuinely think *with* — the one whose back-and-forth helps you find words and structure. That is your **base**. Do your thinking and drafting there. Then bring other models in for specific jobs they're suited to, treating each as a distinct instrument:

- *Ideation / drafting* — your base model, where the thinking happens.
- *Adversarial review* — a different model (ideally two), used to attack the work.
- *Verification / search* — a model or tool for finding sources, which you then check yourself.

Using more than one provider matters: a second, independent model catches what the first is blind to, and independence is what makes a later disclosure meaningful.

Two properties matter more than brand loyalty when you choose these systems. **Architectural diversity:** your critics should not share your drafter's blood — models from different companies fail differently, and a weakness one family shares can sail through critique by its cousins. **Reasoning tier:** for critique and verification, use the strongest reasoning tier you can access. In the author's experience the stronger tiers catch what the lighter ones miss, and the difference shows up exactly where it matters — in the attacks on your argument and the checking of your sources. Where you can afford one, a subscription is the price of a serious second opinion — and it is honest to note that access is tiered: the door AI opens is not free of tolls. Record which service, tier, and version you used as you go; models change under the same name, and your disclosure (Step 7) is only as precise as your notes.

Step 2 — Draft with your base model

Feed it your half-formed thoughts and let it help you turn them into clear prose. Ask it what your idea resembles in the existing literature — it will name thinkers and books you may never have encountered, which is part of how the gate is lowered. Draft in sections. Build as you go.

Two cautions even here. Keep the work pointed at *your* thesis; if the model drifts into writing a different, blander paper, pull it back — you steer. And treat everything it tells you about the outside world (a date, a finding, a source) as **unverified** until you have checked it yourself. More on that in Step 4, because it is the step that matters most.

Step 3 — Attack your own work

This is the step most people skip, and it is the one that earns trust. Take your finished draft to a *different* model and instruct it to try to break it. A prompt that works:

"Act as a hostile peer reviewer who wants to reject this. Find the strongest objections, the weakest links in the argument, and anything that is factually wrong or overstated. Do not praise it. Tell me where a smart skeptic would attack."

A weak hostile review is generic — "add more sources," "tighten the intro." A good one names a specific claim and says exactly why a critic would reject it. If the model stays shallow, push it: ask for the single strongest objection, ask it to steelman the opposing view, then ask it to attack its own critique. Iterate until the objections are concrete enough to act on.

Do this with at least two independent models. Then **fix what they break, or defend it on purpose.** A review that only flatters you is worthless — it tells you nothing and, if you later count it as scrutiny, it is a lie. The reviews that count are the ones that changed the work or forced you to defend a point for a stated reason.

Step 4 — Verify every source yourself

This is non-negotiable, and it is where credibility is won or lost. AI tools fail in quiet ways that look like real research until you look closely:

They name a real source — a journal, an institution — but get its *finding* wrong, sometimes reversing it. They hand you "example" citations, formatted and ready to paste, that are fabricated. They slide from one job to another inside a single conversation — you think you're being audited, and the tool has quietly started drafting.

A real example from building this very project: a model reported that a 2021 study had found the internet *increased* research productivity; the actual paper found close to the opposite. Nothing flagged the claim as wrong — it read like ordinary research until the source was opened and read. That is the danger: the failures are quiet, not loud.

The defense is the same every time: **go and read the primary source.** Concretely, for each factual claim: (1) get the *specific* source — exact title, author, year, ideally the page — not "a Stanford study"; (2) find it yourself by searching the title, not by trusting the model's link, which may be invented; (3) read the relevant passage and confirm the claim matches in *direction and magnitude*, not just topic; (4) if you can't locate or confirm it, cut the claim or mark it explicitly as unverified. Treat any citation you have not personally opened as not yet real. One unverified claim that misrepresents another author's work can sink an entire piece — so this is not optional diligence, it is the core of the method. You are the verification layer no model reliably provides.

Step 5 — Integrate honestly

You decide what goes in. As you integrate, hold to one discipline that separates real inquiry from motivated reasoning: **you control the direction of the work, but you follow the evidence within it.** Steering your inquiry is legitimate — the thesis is yours. *Redirecting a model until it confirms what you already believed, and keeping only the agreeable answers,* is not; it is the failure the whole method exists to prevent. Keep the disconfirming passages. A record in which the AI only ever agreed with you is a warning sign, not a credential.

Step 6 — Keep the receipts

Save your transcripts — the drafting sessions, the hostile reviews, the verification checks. Keep them organized enough that you could produce the relevant ones if someone asked. You are not posting your entire chat history; you are making it *available on request*. The standing possibility that you could show your work is what disciplines it, and what lets a skeptic check rather than just doubt.

Step 7 — Disclose how it was made

State, plainly and in a checkable form, which AI systems you used and what each did. Use the **AIast** notation from the companion standard:

Byline: *Your Name, [LI-]AIast(n)* — where *n* counts the user-facing AI services used in substantive roles — a count of services, not a score. Method note: *"Drafted with [model]; independently critiqued by [models]; sources retrieved and checked via [model] and against primary sources. The author directed the work, integrated all contributions, and is solely responsible for every claim. Working transcripts available on request."*

Function tags (in the method note) say what each did: *i* ideation, *d* drafting, *c* critique, *v* verification, *s* source retrieval. The LI- prefix is optional — use it if the definition fits you. Disclose honestly: a single deeply-used system is AIast1, not dressed up as several. If no AI meeting the standard's disclosure threshold touched the work, you can say so affirmatively with AIast0. Name model versions and dates where the service shows them; where it does not, record the service and the dates — your saved transcripts carry the rest.

Step 8 — Publish in the open

Put the work where it has a permanent, dated, citable home — a free open repository such as **Zenodo** or **OSF**. This does two things a gatekeeper would otherwise control: it makes your work *citable* (it gets a DOI, a permanent identifier), and it *timestamps* you as the originator. Be clear about the limit, though: a DOI gives permanence, citability, and a dated claim to priority — *not* an audience. Publishing makes the work findable and referenceable; getting it actually *read* is a separate effort — sharing it, posting where your readers are, inviting response. The repository opens the door; it does not fill the room. Briefly: create an account, click *New upload*, add your file, fill in title / author / abstract / keywords, choose a license (CC BY is a good default), and publish. The DOI is minted instantly. If you revise, publish a new version — the repository keeps the history, which is itself a kind of receipt.

The mindset

Three ideas hold the method together.

Earn the hearing; let the work face the verdict. You are not claiming to be right. You are making the work legible and testable, and asking only to be read and engaged. That is all transparency can buy — and it is enough to get in. **Be specific enough to be wrong.** A claim sharp enough that others can test and even refute it has contributed something, because it enters the record where the field can take it up. Vague, unfalsifiable, hedged-to-death work contributes nothing, however safe. Aim to be usefully, checkably wrong rather than uselessly vague.

Transparency is not expertise. Showing your work does not make you an authority, and it does not make a flawed argument sound. It removes the *excuse* for ignoring you, nothing more. The rest is on you to actually understand your subject. Keep learning the field; the disclosure only gets you to the table.

Quick checklist

1. Choose a base model; line up independent reviewers from different providers, at the strongest tier you can access.
2. Draft with the base; treat every external fact as unverified.
3. Have two different models try to break it; fix or defend.
4. Read every primary source yourself; cut what you can't verify.
5. Steer the work, but follow disconfirming evidence; keep it.
6. Save and organize transcripts — available on request.
7. Disclose with the AIast notation and a method note.
8. Publish to an open repository for a dated DOI.

Disclosure (per the AIast standard)

William Stafford, ADN, LI-AIast3 — *AIast: Claude (i/d/c/v/s) · GPT (c) · Grok (c).* This guide was produced with Claude (Anthropic) as interlocutor, drafter, critic, and editor, and was independently critiqued by GPT (OpenAI) and Grok (xAI). For version 1.2, Claude (Anthropic, July 2026) served as editor for the conformance revision and the model-selection guidance, with every change directed and approved by the author. The author directed the work, made all decisions, and is solely responsible for its content. It documents a process the author carried out in practice. Working transcripts available on request. The count of three reflects the user-facing AI services used in disclosure-worthy roles — a count of services, not a measure of rigor.

License

Companion works: *Earned Trust: Verifiable Disclosure as a Credential-Substitute for AI-Assisted Work* — the AIast disclosure standard (https://doi.org/10.5281/zenodo.20719927); *The Lost Innovators Hypothesis: A Single-Case Documentation of AI-Mediated Theoretical Inquiry by a Non-Credentialed Author* — the case behind it (https://doi.org/10.5281/zenodo.20721730).