

The AI Detective: Citation Audit & Fact-Check Blueprint

Catch Fake AI Sources, Phantom Quotes, and Invented Theorems in 4 Steps

Student Name:

Date:

Period / Grade:

Subject / Course:

Research Paper Topic:

Study Partner:

1. The 4-Corner Citation Audit Protocol

Corner 1: The AI Claim & Suggested Citation	Corner 2: Database Cross-Verification
Record the exact claim and citation generated by AI. <i>Example: 'Smith (2021) showed 40% memory boost in Journal of Brain Science.'</i>	Search Google Scholar, JSTOR, or Perplexity for the exact title and DOI. <i>Action: Check if the journal and author actually exist in reality.</i>
Corner 3: Primary Quote Match (Ctrl+F)	Corner 4: Final Integrity Verdict
Open the real paper PDF and search for the claimed statistic. <i>Action: Verify if the paper truly supports the claim or was misquoted.</i>	Mark as: [VERIFIED], [HALLUCINATED], or [MISATTRIBUTED]. <i>Action: Replace hallucinated references with authentic library sources.</i>

2. Citation Fact-Check Audit Log (Includes Worked Example)

#	Citation / Claim (AI Output)	Database Checked	Status & Verification Notes
EX	'Marie Antoinette said "Let them eat cake" during 1789 flour shortage'	Library of Congress • Rousseau Confessions (1782)	[HALLUCINATION] Apocryphal myth. Quote appeared 7 years prior in Rousseau Book VI. Corrected in draft.
1			[] Verified [] Hallucinated [] Misquoted Notes:
2			[] Verified [] Hallucinated [] Misquoted Notes:
3			[] Verified [] Hallucinated [] Misquoted Notes:

3. Socratic Adversarial Red-Teaming Workshop

Paste this prompt to interrogate your AI research partner before turning in your paper:

My Red-Team Prompt: "Act as a skeptical academic reviewer. Look at my draft claims: [Paste claim]. Identify the single weakest assumption, any potential historical or scientific misattributions, and provide 2 primary sources that contradict it."

What contradiction or correction did the AI reveal?

Authentic Primary Source located to replace it:

Author: _____ **Title:** _____ **Year:** _____ **DOI/Link:** _____

4. Metacognitive Reflection & Peer Evaluation Rubric

Evaluation Criteria	Needs Practice (1 Pt)	Good Progress (2 Pts)	Mastery Champion (3 Pts)
Citation Authenticity	Copied AI citations without clicking links or checking ISBNs.	Verified author names, but did not confirm page numbers.	100% verified via DOI/library index with active source links!
Primary Quote Precision	Assumed AI quote was accurate without text matching.	Found original document, but quote was slightly altered.	Direct Ctrl+F quote match verified inside official primary text!
Attribution Transparency	No disclosure of AI tool usage in bibliography.	Basic footnote mentioning AI tool without prompt details.	Complete MLA/APA AI methodology disclosure included!

5. Academic Integrity & Scholarly Honor Pledge

I proudly uphold scholarly integrity with this personal commitment:

[X] **Verification Rule:** I will never submit an AI-generated source without verifying its physical existence in a real library database.

[X] **Zero Plagiarism:** I use AI to brainstorm and fact-check, never to write essays or pass off machine text as my own voice.

[X] **Privacy Protection:** I will never input my peers' personal data, grades, or private information into external AI tools.

Student Signature: _____ **Date:** _____