

Literature Research with AI

From Survey to Synthesis

Scientific Inquiry with AI

Student Guide

Why This Lesson

Every research project begins inside a conversation that other researchers started before you. A literature review helps you learn what the field already knows, what it still argues about, and where a new question might fit.

AI can help you identify the background information necessary to become an informed participant in extending the conversation: finding papers, comparing them, mapping citations, and drafting first-pass syntheses. But AI can also sound confident while being wrong. It can misrepresent a finding, overstate a claim, or cite something that does not support the sentence. It is your responsibility to assess the information provided by an AI just as you would when you gather information from any source.

Conducting literature reviews is an essential skill for any scientist or engineer. You will use this skill repeatedly while you are in school and throughout your career. This lesson helps you build this skill in the age of AI.

This lesson will run over two class periods.

Learning Goals

Upon completion of this investigation, you should be able to:

Phenomena:

- Read a field rather than a paper: identify what a research community treats as settled, what it treats as contested, and the signals that distinguish the two.

Experimental Practices:

- Formulate a focused, researchable question by hand, then use AI as a critic to refine it — and understand why question quality determines search quality.
- Execute an AI-assisted research workflow — claim survey, citation mapping, synthesis matrix, and grounded synthesis — using structured, reusable prompts.
- Design the dimensions of a synthesis matrix appropriate to your own research question and use it to evaluate and compare sources.
- Use the integrated deep research tools that are native to most foundation models and recognize how these tools combine the workflow you did in a step-by-step manner.

- Use a curated corpus (your verified papers, plus the deep research report as an unverified reference) in a grounded-Q&A tool such as Gemini Notebook to deepen your own understanding.

Habits of Mind:

- Verify AI-generated claims against primary sources and articulate why this is the same discipline scientists have always applied to any intermediary source.
- Identify gaps, contested points, or extensions that are not found in the current literature — the first step of original inquiry.
- Recognize when a fully automated “deep research” tool is the right way to get oriented in an unfamiliar topic, and when its output still needs further investigation or refinement to meet your objectives.

What You Will Need

- **A general-purpose AI model.** ChatGPT Plus is provided for this course and is the model this guide assumes. You are welcome to use a different model (Claude or Gemini, for example) if you prefer but recognize free accounts may not have access to all the capabilities you will need for this lesson.
- **Google Scholar** — free.
- **ResearchRabbit** — free.
- **Gemini Notebook** — free with a Google account.
- **Your institutional library proxy login** — the route to full-text papers behind paywalls.

The Core Idea

- **AI output sits on a spectrum of trust.** Both traditional and AI search tools provide you real links to papers. A language model also gives you written summaries, and those you have to check for accuracy before you rely on them to support a claim.
- **This lesson outlines how to use AI in exploring scientific literature.** As the AI does more searching, organizing, and synthesizing, your job shifts to directing and checking its work. The most important decisions still stay with you: what question to ask, what claims to trust, and what gap to pursue.
- **Building a synthesis matrix of your sources becomes a valuable discovery tool:** Each cell should contain one small claim from one source, so you can explore the full landscape in a single view. The matrix allows you to effectively identify agreement, disagreement, and knowledge gaps (empty cells) for your topic of investigation.

The driving question for the demonstration used in this lesson: “Does using AI tools for problem-solving improve undergraduate STEM students’ learning and long-term retention, or does it create a ‘cognitive offloading’ effect that reduces skill-building — and does the outcome depend on how the AI is used, such as answer-giving versus Socratic tutoring?”

Our working hypothesis is: AI may improve learning and retention when used as a Socratic tutor but may reduce skill-building when used only to get direct answers.

The five-step workflow you will be introduced to in the prelab and use in class. Steps 1-4 break down the workflow into distinct processes where you work through each step with the help of AI. In Step 5, we ask the AI to conduct all these steps as a single integrated task.

Step	What you do	What the AI does	What you must check
1. Claim survey	Frame the question and the hypothesis you want tested	Searches the literature and returns papers with their findings	Whether each summarized finding matches the paper
2. Citation mapping	Choose a seed paper and read the structure around it	Maps what the paper builds on and what has built on it	Whether the stated relationships are real
3. Synthesis matrix	Decide which columns your question requires	Fills one row per paper, leaving unknowns blank	Every cell you intend to rely on
4. Grounded synthesis	Read the matrix and form your own view first	Synthesizes agreement, disagreement, and gaps	Whether its synthesis matches what you saw
5. Deep research	Pose the question, then judge what comes back	Runs the whole workflow automatically	The claims and sources you intend to use

Part 1 — Pre-Class Preparation

~2 hours

A. Manual Literature Research (~50 minutes)

We start by conducting a manual literature search in this pre-class activity. Conducting the search steps manually will prepare you to more effectively guide an AI-assisted search that you will conduct in class.

1. Define your research question (10 minutes)

Pick a topic you genuinely care about or adopt and adapt one from the starter list below. Then write a focused, researchable question without the help of AI. A strong question names a system or population, a mechanism or variable, and a comparison.

Weak: “How does caffeine affect the body?”

Stronger: “Does caffeine improve reaction time on a simple visual stimulus task in habitual coffee drinkers compared with non-drinkers, and if so, is the effect dose-dependent?”

Starter topics (If you are stuck, you can build your question around one of this starter ideas.):

- Does spacing your studying out over time actually help you remember?
- Does listening to music while you study help you learn?
- Does using fast chargers shorten a rechargeable battery's life?
- Does the color of an object influence human response?
- Is genetic modification of food, animals, or humans beneficial or harmful?

Why carefully crafting your question matters: Question quality determines search quality. A vague question will return a vague list of publications with dubious relationships between them.

Create a hypothesis: Write down your own hypothesis for your research question that you will test against the sources you will identify. The hypothesis does not need to be correct, our goal is to find out what is understood about the question through our literature research.

2. Manual search — Use Google Scholar (20 minutes)

Keyword search, no AI. Try different search terms if you are not getting the results you desire. Review the returned sources and find five papers that seem most relevant. Read titles and abstracts.

What to look for during your manual search: The friction. Take notes as you go: did your search terms effectively capture your research question, what was hard to find, what turned out to be wasted effort, were you able to review every plausible source, and do you suspect relevant work exists under search terms you did not have time to try.

Why it matters: The friction is the point. These notes are the reference experience against which every AI-assisted result in this lesson gets judged, and class opens with them.

3. Careful manual read of one source (15 minutes)

Pick the paper that appears most relevant to your research question and read it without AI (you may only have time to read some of the sections, target doing as much as you can in 10 minutes). Write a three-sentence summary in your own words, plus one question the paper alone cannot answer.

4. Save PDFs of the 2–3 most relevant papers (5 minutes)

Get the full text of the two or three papers most central to your question. These seed the synthesis matrix you build in class; you will add to them with whatever the AI-assisted search surfaces.

Routes to a paywalled paper, in rough order of preference: your institutional library proxy; the publisher page reached through the library's link resolver; preprint servers (arXiv, bioRxiv, medRxiv, PsyArXiv, ChemRxiv); Google Scholar's "All versions" link, which often surfaces an author-hosted PDF; the author's own webpage; and, as a last resort (not as part of this exercise), a polite email to the corresponding author — which works more often than you would expect, and is itself a small lesson in scholarship as conversation.

Judgment call: If a paper resists access after a genuine attempt, document what you tried and move on. An abstract is an acceptable fallback — marked as such in your matrix. You should only do this if you determine this source is a key reference and other more complete sources are not as useful in addressing your research question.

B. AI-Assisted Literature Research (~40 minutes)

1. Watch the demonstration video (~20 minutes, including pauses)

The video (found on the class website) runs the complete five-step workflow on a shared demonstration question about AI in education. It outlines the workflow we will use in class. It contains two writing pauses:

- **Pause 1 (before Step 1 is revealed):** If you had a human research assistant surveying this literature, what exact instructions would you give them? Include what kinds of papers to find, what details to extract, and how they should handle uncertainty. Write it before you resume the video.
- **Pause 2 (before Step 4 is run):** Write one claim you would bet is settled (this can be from your hypothesis) and one claim you would bet is contested. What signal(s) would tell you the difference? Write both before you resume the video.

Bring both written notes from the pauses to class. The Socratic AI tutor in the next section will also ask about them.

2. Socratic AI tutoring session (~20 minutes)

Upload the Prelab Tutor Document (found on the course website) to your AI model and follow its lead as it reviews and builds upon the ideas in the video.

The tutor will ask questions instead of giving a lecture. It will help you think through four key ideas: 1) how much to trust primary sources identified by your searches versus AI generated summaries, 2) what each step hands over to AI, 3) why verification matters, and 4) what decisions still stay with you.

Answer honestly. There are no wrong answers here. Saying “I’m not sure, but here’s my thinking” is an acceptable approach. At the end, you will get an engagement score as a personal target, not a grade, plus a short roadmap for what to focus on next. Remember that part of this course is exploring the various uses of AI in education so evaluating the interaction you have with the AI is part of the lesson. We are learning together how to make AI tutoring more effective. (If you want to see the rubric for how the AI grades your interaction you can look at the Prelab Tutor Document.)

Share the chat log with your instructor when you finish. Only completion is recorded, the grade is not used for assessment purposes.

C. Additional Pre-Class Preparation (~15 minutes)

Written summary (~10 minutes). After the video and the tutor session, write a brief summary of the key terms and core concepts in your own words: the five steps and what each one delegates, and the difference between a search result and a generated summary. Consolidating this in written form solidifies your understanding in preparation to carry out the workflow in class.

Tool access verification (~5 minutes). Confirm all the following before class:

- Your AI platform works and its deep research mode is available to you. If you are using something other than ChatGPT Plus, confirm it can also generate files — Steps 2 and 3 ask for one.
- ResearchRabbit opens (you use it in Step 2).
- Your Google account opens Gemini Notebook (you will use for the in-class part of this lesson).
- Your library proxy login works — you may have already used it to get your PDFs.

D. Review of In-Class Investigation (~10 minutes)

Skim the steps below in preparation for the in-class session. You will move quickly in class, so arriving with the sequence already in your head will allow you to more effectively focus on the results.

Part 2 — In-Class Investigation

220 min
2 class periods

The Investigation, Step by Step

You will run the full workflow on your own research question while the instructor guides the class. This document gives you the prompts, steps, and decisions you need to make. Your instructor will set the pace and the class will proceed through each step as a group. Use a single AI conversation across both class periods and keep it for future review and reporting.

Activate — Introduction to Literature Research (~15 minutes)

Class opens with the history of innovation and the importance of transmitting knowledge from one generation to the next. We will discuss the technical advances that have moved the bottlenecks of sharing knowledge, and how AI can help us address the current challenge — an abundance of information.

Set up — Refine your question with AI (~15 minutes)

Run the refinement prompt on the question you wrote by hand in the prelab. Read the critique and the sharpened versions it proposes.

Prompt 0 — Question refinement (run before everything else)

Here is my research question: [your question]. Critique it as a literature-review question: does it name a specific system or population, a mechanism or variable, and a comparison? Is it answerable from existing literature, or is it really several questions bundled together? Propose 2–3 sharpened versions with different scopes, and for each, say what kind of literature it would point toward. Do not choose one for me — give me the trade-offs and I'll decide.

Decide: Choose a version of the question you like: your original, one AI suggested, or a mix of both. The AI gives feedback, but it does not choose for you. Your research question should still reflect what you care about and what you can realistically investigate.

Set up — Launch deep research (~10 minutes)

Open your AI's deep research mode, paste your refined question, and start the run. A full run takes ten to twenty minutes. Do not open it until Step 5 — seeing the automated answer now would contaminate the judgments you make between here and there. This deep research feature built into AI foundation models is very powerful and something you will probably use many times in the future.

Step 1 — Claim survey (~20 minutes)

Fill in the claim-survey prompt with your own details in the brackets, then run in your chatbot. Use the hypothesis you wrote as part of the prelab or use a revised version if your question changed in the last step. Skim the results, choose the two or three most relevant papers, and check which ones you can access as full-text PDFs. These are possible sources for you to analyze in building your synthesis matrix in Step 3.

Prompt 1 — Claim survey

I'm researching [your research question]. Search the literature and return 8–10 relevant papers. For each, give me: full citation with DOI or stable link, population or scope, methodology in one sentence, key finding in one sentence, and whether the finding supports, contradicts, or qualifies [your hypothesis]. Include at least two papers from fields adjacent to [primary field] that study the same phenomenon under different terminology and tell me what terms those fields use for it. If you cannot verify a detail from the paper or abstract, mark it as 'not confirmed' rather than inferring it. Do not provide a summary analysis; we are going through a deliberate process and will get to synthesizing the information later.

How AI search differs from traditional keyword searches: This step illustrates the difference between manual search and AI-assisted search. AI can provide a much more nuanced search that includes concepts and questions beyond matching keywords. It can also provide summaries and search different fields intentionally. One caution: while the AI summaries can be helpful, they still need to be checked against the original sources.

Checkpoint: If more than half the results seem off-target, your question's key terms are probably ambiguous. Revise the question and rerun — refinement is iterative, and the refinement is often key to assuring you connect to the most important information.

Step 2 — Citation mapping (~30 minutes)

Open ResearchRabbit and paste in one of your central papers from Step 1 or from your prelab search using Google Scholar. Explore the visualization: the cluster around your paper, the earlier work it builds on, and the later work that builds on it.

What to look for: You are not just collecting more papers. You are looking at how the field is connected. Older, highly connected papers may be foundational. Clusters may show different methods or schools of thought. Newer papers near the edges may show where the field is still developing.

Why it matters: This is a map of a research conversation, not a search. The goal in this step is awareness, not mastery — fully exploring a literature with a tool like this takes far longer than the time you have here. Know that this deeper mode of exploring a literature exists, and know what kind of question it answers.

Do now: Note one or two papers the map surfaces that Step 1 did not.

Using your AI directly for citation mapping: As an alternative to a customized tool like ResearchRabbit, try using a generalized prompt to conduct citation mapping.

Prompt 2

Take this paper [citation]. Identify the foundational papers it builds on (including the sources it references) and the recent papers that have built on it. For both references and citations, give me the citation, its relationship to the seed paper, and a one-sentence summary. Limit the total number of references to the twenty you believe are the most relevant to my research question [research question]. Include a graphical representation of the relationships if that helps communicate the story. If you are inferring the relationship rather than confirming it directly, say so. The result should be a well-formatted PDF report.

Step 3 — Synthesis matrix (~30 minutes)

Build your source set to five papers: two prelab PDFs you saved, two of the best papers from Step 1, and one from Step 2. Use full-text PDFs when possible, and mark any abstract-only sources clearly. Before running the AI prompt, decide which columns your matrix should include.

Decide: The columns of your synthesis matrix are an analytical decision, and they depend on your question. A core set applies almost everywhere: method or study design, system or population, independent variables, dependent variables, key finding, stated limitations. Then add one or two columns specific to your own question (e.g., dose levels for a caffeine question, type of exercise for a mental-health question). The determination of the additional columns is critical for helping the AI find the information that you identify as most important for your research topic.

Prompt 3 — Synthesis matrix

I've uploaded 5 papers on [topic]. Build a synthesis matrix with one row per paper and these columns: method/study design, [system or population], independent variables, dependent variables, key finding, and stated limitations, plus these question-specific columns: [your additions]. Add any further columns you think my question [research question] requires and tell me why. Ground every cell in the uploaded papers, and leave a cell blank rather than inferring a value the paper doesn't state. Where a source is an abstract only, mark its row as abstract-based. Only build the matrix. Do not analyze or synthesize across the papers yet. I want to study the matrix myself before you draw any conclusions. Create the matrix as both an Excel spreadsheet and as an HTML application that allows me to see the full matrix but expand individual cells to reveal the contents.

If your model cannot generate files: ask for the matrix as a plain table you can paste into a spreadsheet, and read Prompt 2's output on screen rather than as a PDF.

What to look for: Blank cells. The prompt tells the model to leave a cell empty rather than infer, so an empty cell — nobody in your reference papers measured X — marks a potential knowledge gap for future research.

Your own read first: Take five minutes and create a short assessment of the matrix — where the papers agree, where they disagree, which blank cells look like gaps — in your own words, before the AI offers any synthesis at all. You compare the two in the next step.

Step 4 — Grounded synthesis and gap analysis (~30 minutes)

Run the grounded-synthesis prompt with web search on. The prompt directs the AI to conduct the synthesis using two different approaches. Include the papers you reviewed in Steps 1–3 so the AI can compare its synthesis with your source set. Read the response carefully: look for what seems well-supported, what seems uncertain, and what you may need to check next.

Prompt 4 — Grounded synthesis (and quick AI check)

Two parts: 1) Using only the 5 papers I uploaded for the synthesis matrix: synthesize across them — where do they agree, where do they disagree, and what is the current synopsis of the field on my research question? Point to any blank cells or absences in the matrix that suggest a gap. 2) Create an independent 200-word assessment of the current state of research on [my research question] based on an independent web search and selection of primary sources. Separate what is well-established, what is contested, and what appears to be emerging. Cite primary sources and flag any places where the evidence base is thin.

The “deep research” report you started earlier does a more extensive version of this same process.

What to look for: Does the settled/contested split match the bets (i.e., your hypothesis) you wrote at the video's second pause? Score yourself honestly.

Compare: An open synthesis built from sources the AI chose is being checked against a closed corpus built from sources you chose. How did they compare? The two modes triangulate. If they agree, you have independent corroboration. If they disagree, either your curation missed something or the model's search pulled weak sources — and working out which is a key judgment skill this lesson is helping to build for your future research.

Verify — The spot-check (~10 minutes)

Pick one specific, checkable claim the AI made about one of the papers in your synthesis matrix — a number, a method, a population, a finding — and go to that paper's full text. Read the relevant section. Record three lines: the claim as the AI stated it; what the source actually says; and your verdict (confirmed / overstated / wrong / couldn't access).

Why it matters: This is not framed as defense against AI failure. It is the standard discipline of working with primary sources — the same one scientists have always applied to a review article, a textbook, or a colleague's summary. Any summary is a pointer toward the primary source, not a

substitute for it. If the AI turns out to be right every time you check, good: the habit is the point, and the checking is what makes the claim genuinely yours to build on.

Step 5 — Deep research (~20 minutes)

Quickly review the deep research report with the following questions in mind. Can you see the steps you just practiced within the report — where is it surveying, where is it synthesizing? How do the references compare to those you uncovered in your stepwise process? How does the deep research report compare/add to the analysis you have done step-by-step?

Judgment call: In the second part of Step 4 you created a quick summary. How does that short assessment compare with the deep research report? Step 4 costs seconds and is the everyday probe; deep research costs many minutes. They also take significantly different amounts of time to read and digest. In the future, you will have to make a judgment on what type of research is appropriate for different needs.

AI Tool for Exploration of Sources — Gemini Notebook (~30 minutes)

The research steps above helped you find what the field knows; the next step is to extend your learning from the sources you collected. Gemini Notebook is an example of an AI tool that helps turn information into understanding and long-term knowledge. After a short instructor demo, build a notebook using your verified papers plus the deep research report. Note the asymmetry: the papers you have checked, the report you have not — treat its claims as leads to verify rather than as established. Explore the notebook and click at least one inline citation to check the exact source passage.

Gemini Notebook only knows the sources you upload. That makes it more grounded, but not perfect. It may still miss context or combine ideas incorrectly, so always click citations to check the original source.

Gemini Notebook can be a powerful organizer for the material associated with a class. Give it a try on one of your courses uploading readings, class notes, class handouts, homework, quizzes, tests, ... You might decide to create notebooks for all your classes!

Share: If you find a feature worth showing, show it. The class shares discoveries at the end of the block.

Evaluate: Did you find limitations in the tool? What features would you add if you were designing a similar tool?

Reflect — Reading your field (~10 minutes)

Write one-line answers to the three questions below. Some will be shared aloud.

- What is the consensus claim in your area — what does the field treat as basically settled?
- What is contested or unresolved — where are researchers still arguing? And how could you tell?
- What is a question you could investigate that no paper you have found has directly answered?

Before you leave: save or export your AI conversation, your synthesis matrix, and your notebook link. The after-class work builds directly on all three.

Part 3 — Post-Class Work

~2.5 hours

A. Review In-Class Work and Generate a Summary Table (~30 minutes)

Read the deep research report closely if you did not finish reviewing it in class. Then review your AI conversation from class with a focus on your synthesis matrix and summary.

Create the table below using your favorite tool (spreadsheet or text editor) and complete it thoughtfully (your own words, don't ask AI to complete it) — this will often require multiple sentences for each element in the table. Generate a PDF as a final product that you will submit.

Step or tool	What it helped me find/do	What I verified	What I would trust it for
Manual search (prelab)			
Claim survey (Step 1)			
Citation map (Step 2)			
Synthesis matrix (Step 3)			
Grounded synthesis (Step 4)			
Spot-check			
Deep research report (Step 5)			

B. Literature Research Summary Document (~90 minutes)

Create a literature research summary document that builds on the research papers you identified and the synthesis matrix you created in class. You will move through the workflow again but have a chance to refine your research question and expand the number of sources included in your literature review. Include these 5 elements in your summary:

1. **The question and its evolution.** Include the evolution of your research question — the prelab version you wrote by hand, the AI-critiqued version from the start of class, and a final version after seeing the literature (if you decide to change it) — with one line explaining what changed at each point and why.

2. **Synthesis matrix.** Expand your matrix to 6–10 papers using citations/papers that were already uncovered in previous searches but not used in your initial matrix. Use full-text PDFs when possible. Verify the cells you rely on, leave unknown cells blank.
3. **A small map of the field.** 6–10 papers with one-sentence characterizations you have personally verified against the primary source (not against an AI summary). You can add your sources to your notebook to assist in this process.
4. **Gap statement.** Use your matrix to propose a possible research gap — this could be one you already identified, or a new gap uncovered by the expanded synthesis matrix. Look for blank cells, disagreements between papers, or missing populations/conditions. A real gap should matter and should point to evidence that would help resolve it.
5. **A reflection.** Here are some questions that you can think about when writing your reflection: Which steps were most useful for which sub-tasks? What did the notebook add that reading alone did not — including any features your classmates showed you? Did the deep research report find the gap you ended up identifying, miss it, or suggest something you had not considered? What surprised you? And how did going to primary sources change your understanding compared with relying on AI summaries?

If you prefer to apply the workflow to a different research question, you are welcome to do so.

Don't spend more than 90 minutes on this activity. It is meant to give you practice on the process and not be a perfect literature research report. Just note what you were unable to complete in the 90 minutes.

C. Retrieval Practice / Course Review (~30 minutes)

Spend about thirty minutes on integrated course review using the course's AI-enhanced retrieval tool. This is the standard routine in every lesson: it spaces your practice out over time and interleaves this lesson with earlier ones, which is what makes the material stick rather than fade.

Summary of Deliverables

Your key deliverables are listed below with the time they should be completed. The prelab items are to be completed before the in-class session because the in-class work depends on them. The **bold** items will be **turned in**, and the other items should be available to show the instructor upon request.

Due before class, with your prelab

- Your research question — the version you wrote without AI.
- Your manual Google Scholar search notes, including the friction.
- Your three-sentence deep-read summary, and the one question that paper alone could not answer.
- Full-text PDFs of your two or three central papers, or a note documenting the access routes you tried.
- Your two written pause answers from the demonstration video.
- **Your tutor session chat log, shared with your instructor. Only completion is recorded.**
- **Your written summary from Part C of the prelab.**

Due at the end of the in-class investigation

- The research question you settled on at the start of class, with one line on what you changed and why.
- Your first synthesis matrix (5 sources), with any abstract-only rows marked.
- Your own written read of the matrix from Step 3, written before you saw the AI's synthesis.
- Your spot-check record: the claim as the AI stated it, what the source actually says, and your verdict.
- Your one-line answers to the three closing questions: the consensus claim, what is contested, and a question no paper you found has answered.

Due after class

- **The step-comparison table, completed, as a PDF.**
- **Your Literature Research Summary Document — all five elements, with the expanded 6–10 paper synthesis matrix.**
- **Your Gemini Notebook of curated sources, or a link to it.**

Acknowledgements

Created by the LLNL Summer 2026 STEM Education Research Team.

Student Team: Ramina Amino, Jahanvi Chamria, Arya Ferozy, Bryanna Gonzalez, Tai Le, Zedikiah McAdams, John Navarra, Joshua Sarabia, Abdurrahman Raza

Faculty Team: Praveen Pathak, David Rakestraw, David Strubbe, Brian Utter

This lesson was developed as a partnership between Lawrence Livermore National Laboratory and the University of California, Merced.