The Transition Accelerator | L'Accélérateur de transition

AI Adoption Training

Workshop#4: AI for Advanced Research Techniques

December 2nd, 2025

Today's Agenda

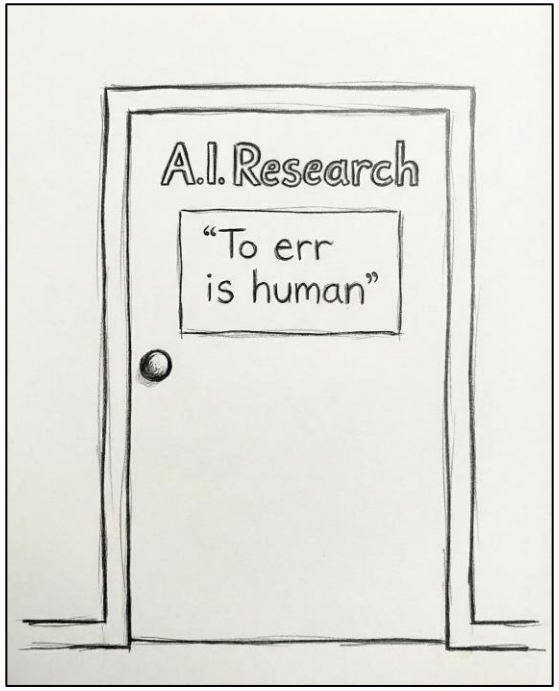
1. **Intro** (5 mins)
2. **Finding Information with LLMs** (10 mins)
3. **Deep Research Features** (15 mins)
4. **Research Tools: NotebookLM and Zotero** (15 minutes)
5. **Research Techniques** (10 mins)
6. **Wrap Up** (5 mins)

Materials are available at: https://github.com/mullinsean/ai_training

From Search to Synthesis to Analysis

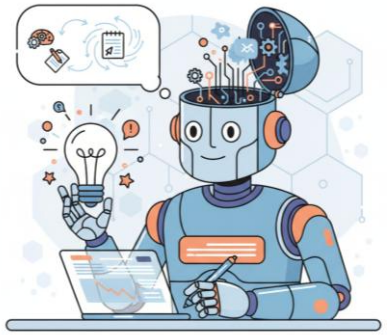
Using LLMs to find, organize and review information

- Traditional research: hours of searching, reading, note-taking
- LLM-enhanced research: accelerated discovery and synthesis
- Key question: How do we maintain research quality while increasing productivity?
- Also, LLMs can potentially increase the breadth and depth of research if used properly
    - Next step beyond Google Scholar or other web-based search tools.
- Today's focus: practical techniques that complement (not replace) critical thinking



Web Search & Document Extraction

Basic techniques for information discovery



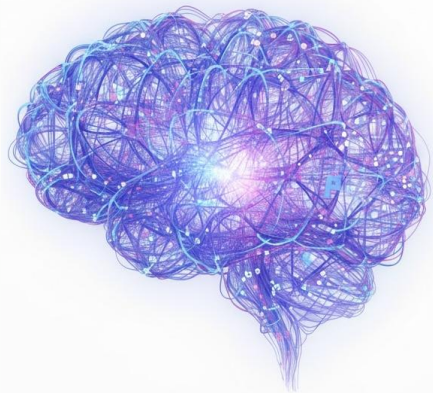
- LLMs excel at extracting and summarizing information from documents. Equipped with web search, they can be excellent tools for finding and parsing information on the web.
- Best practices include:
  - Prompting matters: Be specific about what you are looking for. Use the guidelines for good prompting
  - Specify recency when it matters (eg: “most recent” or “within last year” if those timelines matter)
  - Request multiple sources (prevents LLMs from stopping searches too early)
  - Ask for source quality indicators (eg: prioritize government sources, peer-reviewed research; avoid blog posts and opinion pieces)
- Essentially, you now have a fully conversational, intelligent interface to web search, instead of just typing in key words.

Document Extraction

In addition to asking questions about a document, LLMs can be used to surgically extract specific data or information requests across multiple documents. For example:

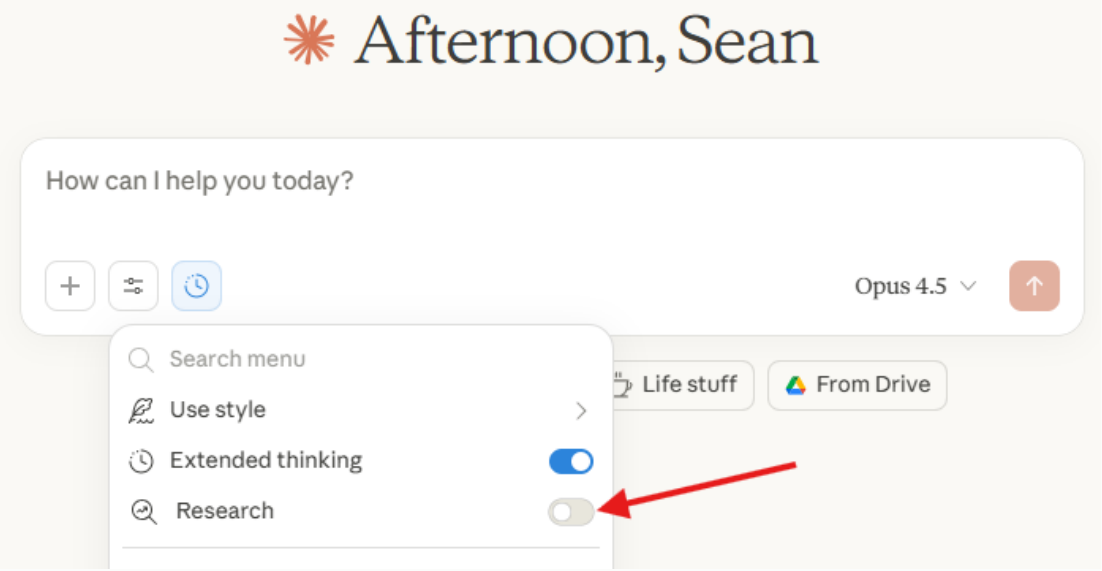
- Extracting all the quantitative findings in a report
- Asking for data in structured forms (eg: creating a table or .csv file)
- Extracting all the references in a document and matching with URLs for source documents.

Deep Research Feature



Advanced “agentic” tool creates sophisticated research reports

- All major commercial LLMs have a “deep research” feature, including ChatGPT, Claude, Gemini and Perplexity.ai
- How it works:
  - LLMs spends 10+ minutes autonomously researching a topic
  - Breaks down the research task into subcomponents, compiles information and then synthesizes into a structured report
  - Includes source refences and links back to original documents
- Extremely useful for: literature reviews, background research, policy landscape scans
- Perplexity.ai can specialize in academic papers or financial documents; Gemini can access your Google Drive files, if desired.



Deep Research – Example Applications

Wide variety of use cases

Use Case	Example	Model	Output
General overview of new topic or context	Overview of the geopolitical situation in South Korea	ChatGPT	Transcript: chat Output: pdf
Deep dive into legislative process around a bill.	History of Bill C-27 (Digital Charter Act)	Gemini	Transcript: chat Output: pdf
Summarize and categorize stakeholder reactions	Media and Stakeholder Reception to Budget 2025	Gemini	Transcript: chat Output: pdf
Financial analysis of a sector or a topic	Financial Analysis of Companies in the Nuclear Sector	Perplexity.ai	Transcript: chat Output: pdf

Deep Research – Prompting Techniques

Getting the most out of Deep Research

Prompting for Deep Research:

- Be specific about scope, audience, desired format
- If desired, direct it to particular sources (eg: Hansard transcripts, recent industry whitepapers)
- Review the LLM generated “research plan” and/or answer any clarifying questions before initiating query
- Optional: ask an LLM to create a prompt for your Deep Research query and then paste into new chat session to run it
- Break large research projects into multiple sub-problems and run Deep Research on each of them

Example Prompt:

"Act as a strategic consultant. Conduct a deep research dive into the current state of Solid State Battery manufacturing. I need to know the top 3 bottlenecks preventing mass adoption.

Constraints: *Focus only on technical and supply chain hurdles, ignore consumer marketing.*

Sources: *Prioritize recent (2024-2025) industry whitepapers and patent filings over general news.*

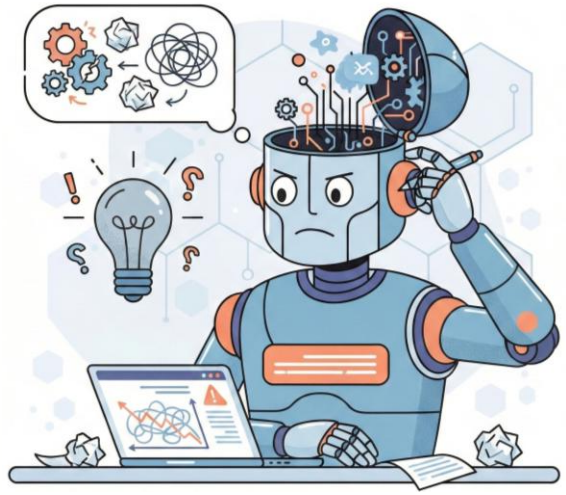
Verification: *Explicitly state if consensus is lacking on any point.*

Output: *A structured memo with a 'Technological Readiness Level' assessment for each major player."*

Deep Research features are agentic—they autonomously plan, search multiple sources, browse pages, and synthesize findings. Your prompt needs to act as **a project brief**, not a simple question.

Deep Research – Limitations

Understanding the limits of Deep Research queries



Deep Research features are not without flaws or limitations:

- Only searches what is on the web: will miss academic databases, paywalled content, internal documents
- Not comprehensive: summarizes what shows up in web searches; may miss key sources
- Often follows conventional wisdom: may not surface contrarian or cutting-edge perspectives
- Limited number of uses per month: even for paid plans, the high compute cost of Deep Research queries result in a limited number of queries per month; use sparingly

Think of Deep Research as a smart, very eager second-year undergrad who can tirelessly research any topic in 5-10 minutes.

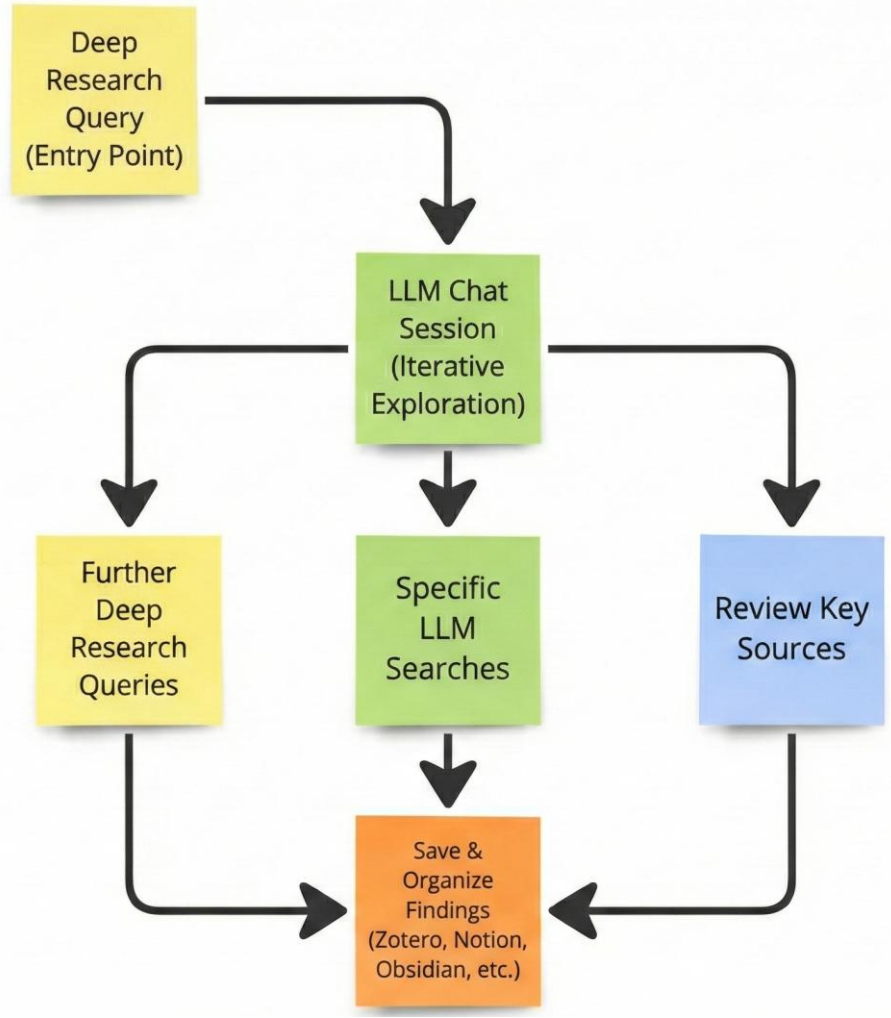
Can be very useful, but assignments should be carefully thought through and output should be reviewed with a critical eye.

Building a Body of Knowledge

Combining Deep Research and iterative LLM conversations

Ultimately, you can build up a body of knowledge by combining iterative LLM conversations with selective Deep Research queries.

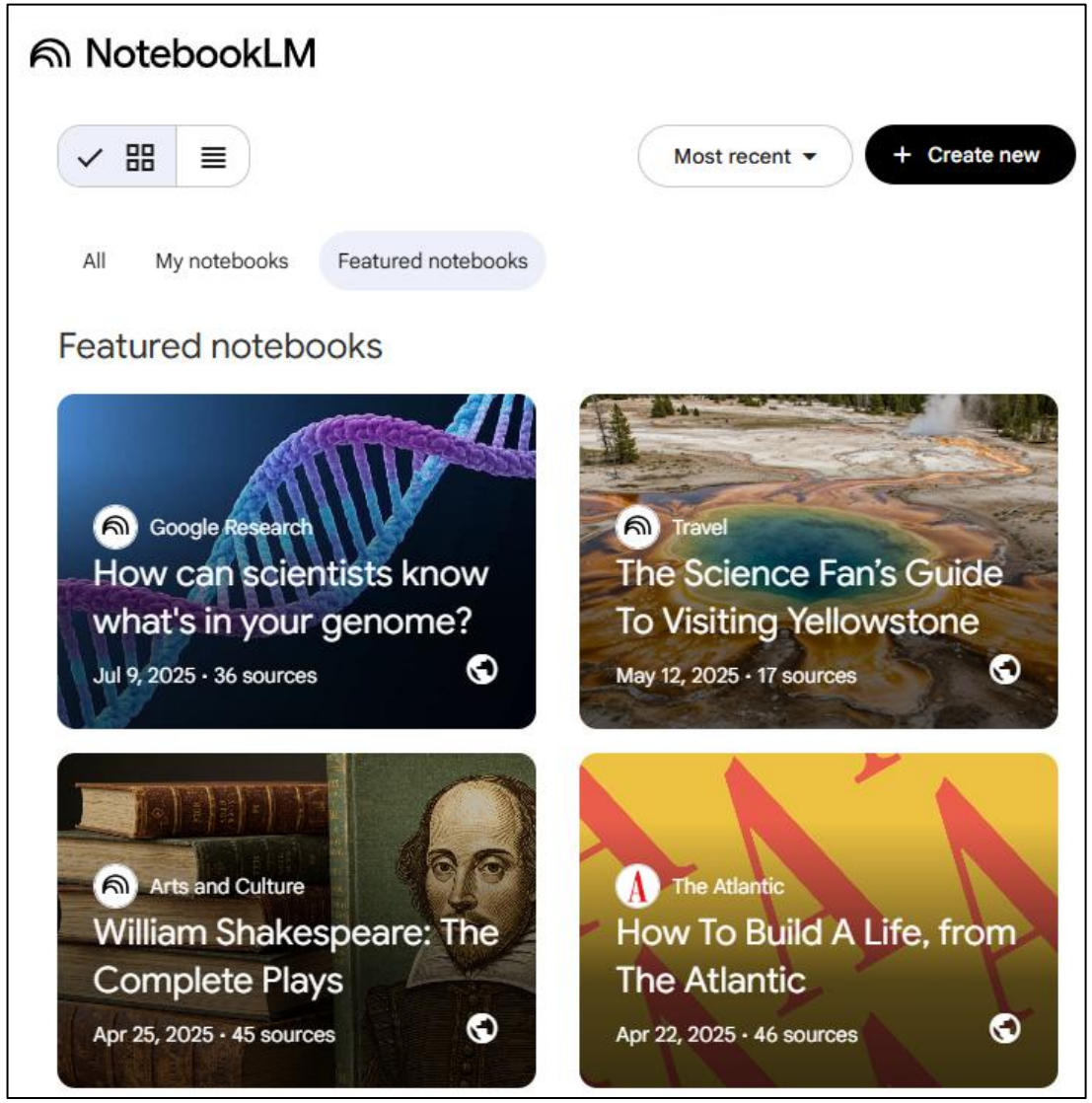
- A Deep Research query can be an entry point into a project; and then can serve as a context document for more in-depth exploration of the topic via a chat session
- These can branch into further Deep Research queries, more specific LLM searches, or reviewing the key sources directly
- Save and organize your findings as you go (notes, documents, chat links):
    - Zotero, Notion, Obsidian or your favourite note organizational tool can be very helpful



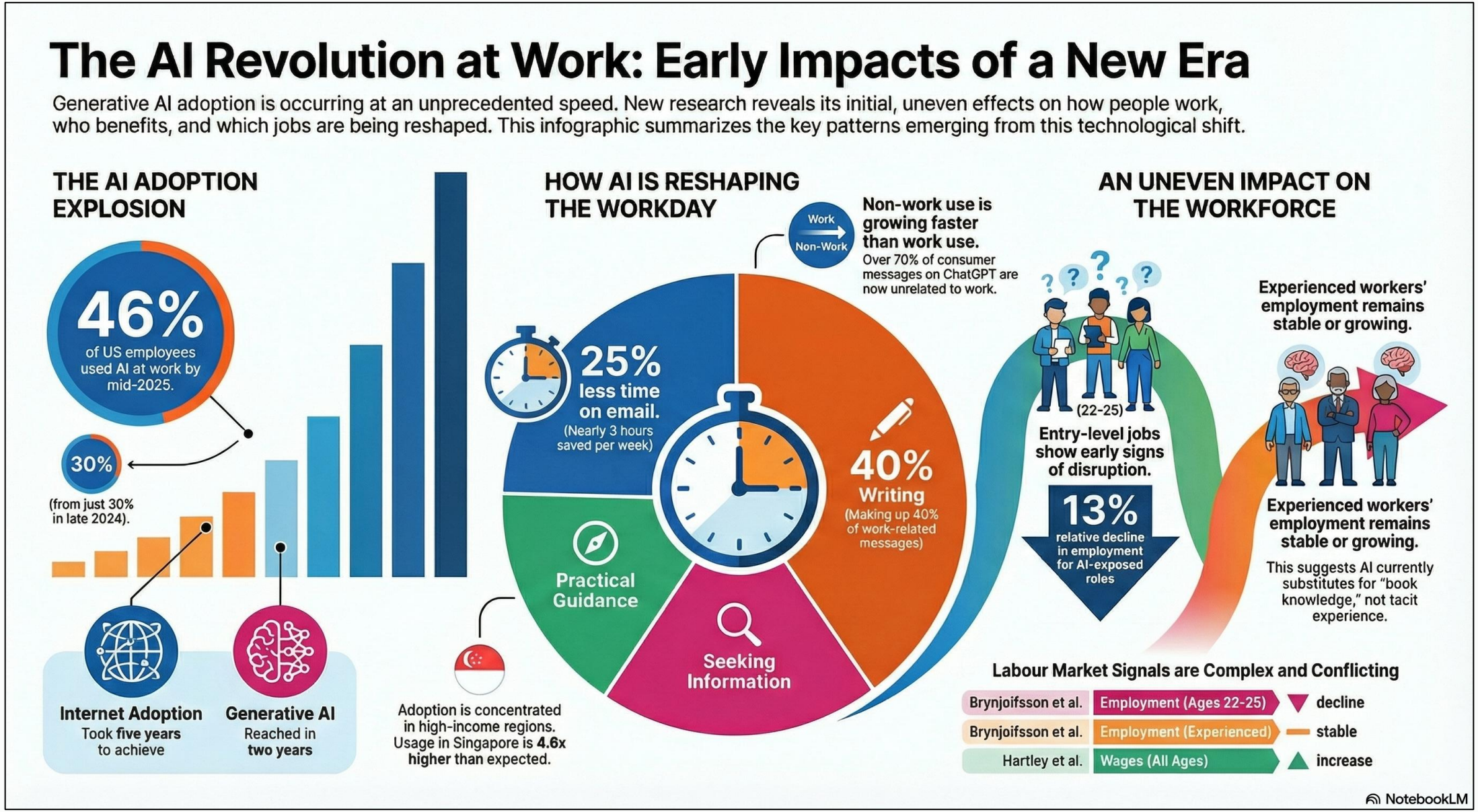
Research Tool: NotebookLM

Google’s specialized research tool for multi-document analysis

- Google’s NotebookLM free-tier allows you to explore and summarize up to 50 original sources: PDF, text, Word, even audio (podcasts or speeches) or video (Youtube)
- Useful for exploring data sources, surveying across different papers, and for self-education (eg: text books, reference manuals)
- Provides direct citations back to source documents for verification.
- Since last workshop, has now been further enhanced with Gemini Pro 3.0 image generation, allowing the creation of infographics, slide decks and other visual media.



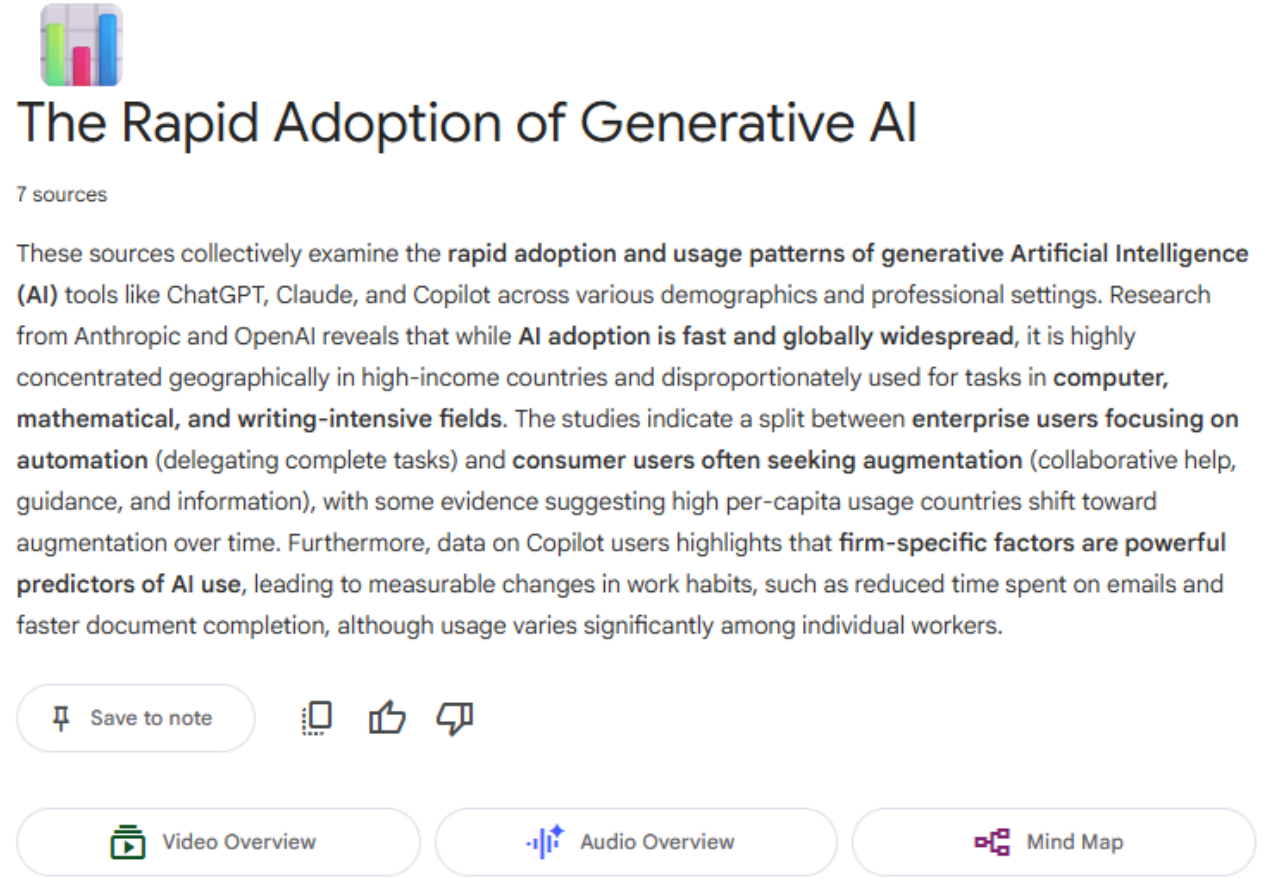
Example of NotebookLM generated infographic:



Research Tool: Notebook LM

Key Features of NotebookLM

- Summarize and Interrogate: Generate an overview, ask questions across all documents – and save answers as notes.
- Mind Map: Visual representation of the themes and connections across sources
- Audio overview: Generate a “podcast” discussion of your documents
- Infographics and Slide Decks: Still early stage, but initial results are impressive
- Can run regular LLM searches or Deep Research queries inside of Notebook to generate additional sources



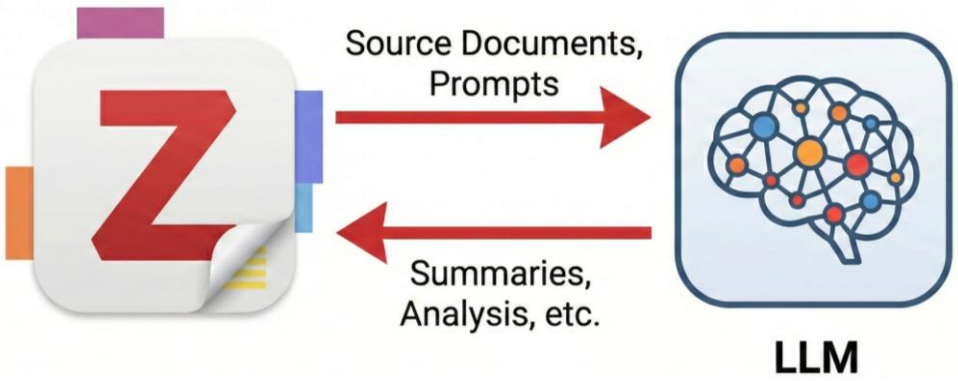
(example)

Research Tool: Zotero + LLMs

Leveraging open-source research tool for LLM use

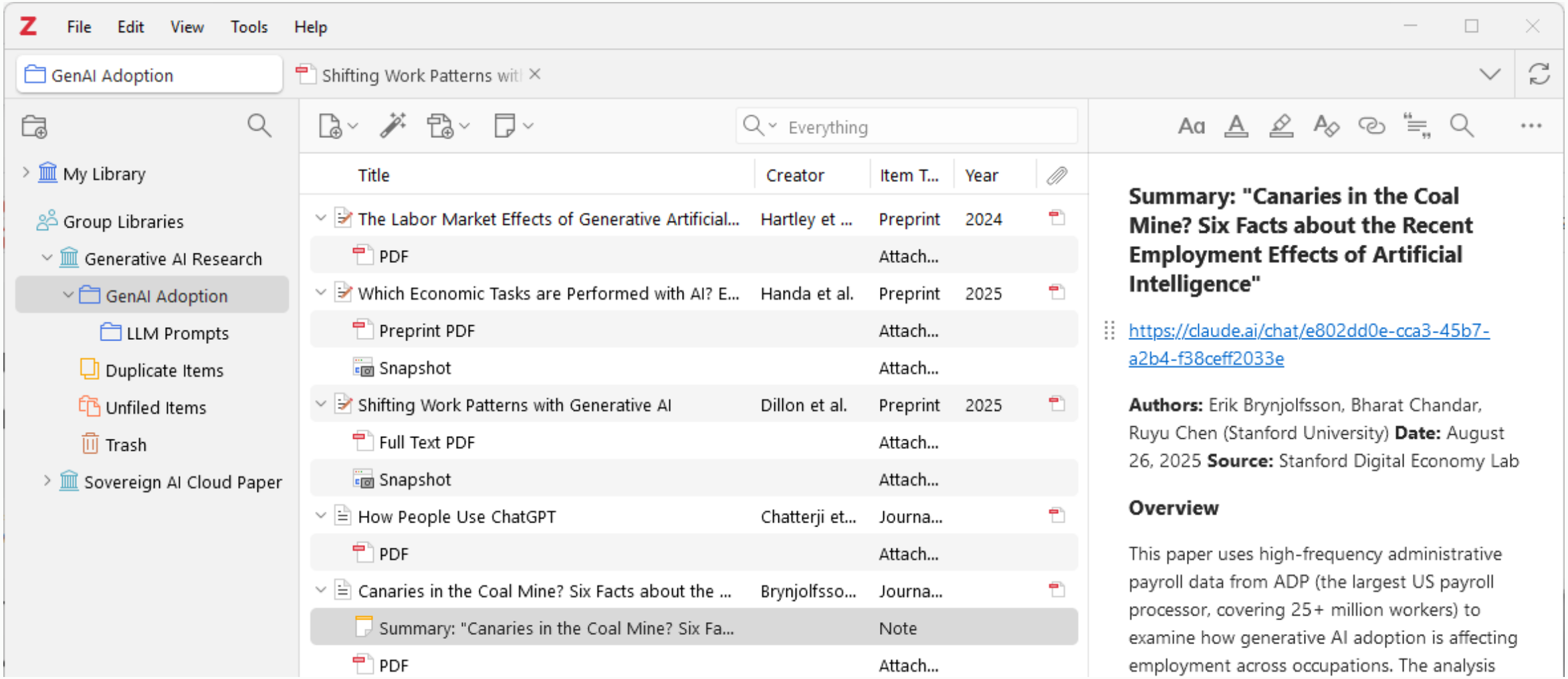
- Zotero is a free, open-source research tool for organizing primary research sources.
    - Not an AI-tool directly, but can be very useful for LLM-assisted workflows
- Build collections of sources (PDFs, webpages, videos, interview transcripts) for different projects
    - Can also include Deep Research query results
- Easy to drag PDFs from Zotero into LLMs or NotebookLM for analysis
- Save LLM-generated notes and summaries back to Zotero
- Maintains a “source of truth” for your research outside of chat session
- Zotero can also save project-specific LLM prompts for key tasks (eg: “summarize these documents with respect to X”)

zotero



Research Tool: Zotero + LLMs

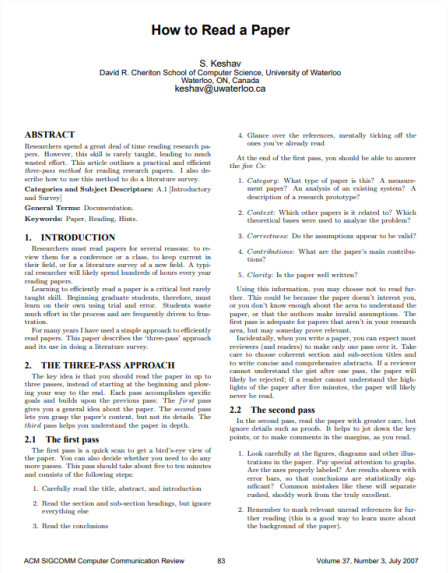
Example Zotero library integration with LLMs



Reading and Reviewing Papers with LLMs

An LLM-assisted update to the “three-pass framework” for reading papers

- Keshav (2007) articulates a helpful, “three-pass” framework for reading academic papers.



How to Read a Paper

S. Keshav
David R. Cheriton School of Computer Science, University of Waterloo
Waterloo, ON, Canada
keshav@uwaterloo.ca

ABSTRACT

Researchers spend a great deal of time reading research papers. However, this skill is rarely taught, leading to much wasted effort. This article outlines a practical and efficient *three-pass method* for reading research papers. I also describe how to use this method to do a literature survey.

Categories and Subject Descriptors: A.1 [Introductory and Survey]

General Terms: Documentation.

Keywords: Paper, Reading, Hints.

1. INTRODUCTION

Researchers must read papers for several reasons: to review them for a conference or a class, to keep current in their field, or for a literature survey of a new field. A typical researcher will likely spend hundreds of hours every year reading papers.

Learning to efficiently read a paper is a critical but rarely taught skill. Beginning graduate students, therefore, must learn on their own using trial and error. Students waste much effort in the process and are frequently driven to frustration.

For many years I have used a simple approach to efficiently read papers. This paper describes the ‘three-pass’ approach and its use in doing a literature survey.

2. THE THREE-PASS APPROACH

The key idea is that you should read the paper in up to three passes, instead of starting at the beginning and plowing your way to the end. Each pass accomplishes specific goals and builds upon the previous pass: The *first* pass gives you a general idea about the paper. The *second* pass lets you grasp the paper’s content, but not its details. The *third* pass helps you understand the paper in depth.

2.1 The first pass

The first pass is a quick scan to get a bird’s-eye view of the paper. You can also decide whether you need to do any more passes. This pass should take about five to ten minutes and consists of the following steps:

1. Carefully read the title, abstract, and introduction
2. Read the section and sub-section headings, but ignore everything else
3. Read the conclusions
4. Glance over the references, mentally ticking off the ones you’ve already read

At the end of the first pass, you should be able to answer the *five Cs*:

1. *Category*: What type of paper is this? A measurement paper? An analysis of an existing system? A description of a research prototype?
2. *Context*: Which other papers is it related to? Which theoretical bases were used to analyze the problem?
3. *Correctness*: Do the assumptions appear to be valid?
4. *Contributions*: What are the paper’s main contributions?
5. *Clarity*: Is the paper well written?

Using this information, you may choose not to read further. This could be because the paper doesn’t interest you, or you don’t know enough about the area to understand the paper, or that the authors make invalid assumptions. The first pass is adequate for papers that aren’t in your research area, but may someday prove relevant.

Incidentally, when you write a paper, you can expect most reviewers (and readers) to make only one pass over it. Take care to choose coherent section and sub-section titles and to write concise and comprehensive abstracts. If a reviewer cannot understand the gist after one pass, the paper will likely be rejected; if a reader cannot understand the highlights of the paper after five minutes, the paper will likely never be read.

2.2 The second pass

In the second pass, read the paper with greater care, but ignore details such as proofs. It helps to jot down the key points, or to make comments in the margins, as you read.

1. Look carefully at the figures, diagrams and other illustrations in the paper. Pay special attention to graphs. Are the axes properly labeled? Are results shown with error bars, so that conclusions are statistically significant? Common mistakes like these will separate rushed, shoddy work from the truly excellent.
2. Remember to mark relevant unread references for further reading (this is a good way to learn more about the background of the paper).

ACM SIGCOMM Computer Communication Review 83 Volume 37, Number 3, July 2007

“How to Read a Paper”, Keshav (2007), http://ccr.sigcomm.org/online/files/p83-keshavA.pdf

Pass	Regular Framework	LLM-aided Framework
First Pass – Quick Review	Quick scan: focus on title, abstract, introduction, type of paper, section and sub-section headings and conclusions 5-10 minutes total	Generate an audio overview or structured LLM summary focused on key areas of interest
Second Pass – Understanding of Paper	Carefully read paper to understand core arguments and key findings; go section by section, pausing to absorb data and findings	READ CAREFULLY – no shortcuts! :) But, you can interrogate the paper with an LLM as you read it.
Third Pass – Deep Understanding of Methodology and/or Peer Review	Virtually recreate the paper, including redoing proofs, rerunning data analysis or reviewing code	Use LLMs to verify proofs, inspect data, review code, check calculations

Advanced Techniques: Beyond Summaries

Sophisticated methods for extracting insights from documents

- **Comparison Matrix:** Build comparison tables across multiple documents (e.g., "Create a comparison matrix with three reports as columns. Compare: cost projections, deployment timeline, infrastructure barriers")
- **Red Teaming:** Ask the LLM to adopt a critical persona (e.g., "Skeptical Finance Minister") to critique your findings
- **Gap Analysis:** "Based on these documents, what perspectives or data points are missing?“
- **Formatting:** eg: Convert documents to LaTeX; format references in APA style
- **Data Cleaning:** clean and organize datasets (either directly or by writing code)
- **Reviewing Proofs or Code:** Paste proofs or code into LLMs for careful verification

Example Conversation:

Using Gemini to compare four papers, build a comparison chart, provide critiques and identify missing gaps.

Finally, summarize conversation and save as a note in Zotero.

(chat)

Additional Reading: AI for Research

Prof. Anton Korinek's work on LLMs for (economics) research

- Anton Korinek (UVA, Brookings): Leading researcher on AI applications in economics
- "AI Agents for Economic Research" (August 2025): Guide to using agentic AI tools for research tasks
- "Generative AI for Economic Research: Use Cases and Implications" (December 2024): Comprehensive overview of LLM applications
- Although focused on the economics profession, many of the techniques generalize to other academic and applied research fields.

Table 1
PRIMARY USEFULNESS OF LLM MODEL TYPES FOR RESEARCH TASKS

Research Category	Traditional LLMs	Reasoning Models	Agentic Chatbots
Ideation & Feedback	Good for initial brainstorming.	**Best for structured feedback and identifying logical flaws.**	Actively scans literature for novelty and grounding.
Writing	**Excellent for drafting, summarizing, and rephrasing existing text.**	Ensures logical flow in complex arguments.	Incorporates real-time web information.
Background Research	*Shortcoming: No live web access.*	Synthesizes info from provided texts.	**Best for live web searches and up-to-date literature reviews.**
Coding	Generates basic code snippets.	**Excellent for writing and debugging complex algorithms.**	**Executes code, tests hypotheses, and interacts with data files.**
Data Analysis	*Shortcoming: Cannot execute code.*	Helps interpret data and suggest approaches.	**Best for end-to-end analysis: data cleaning, coding, visualization.**
Math	*Shortcoming: Unreliable for complex math.*	**Solves multi-step problems and formal proofs at PhD level.**	Can leverage external computational tools to ensure accuracy.
Promoting Research	Drafts initial promotional content.	Tailors summaries for specific audiences.	**Automates creating summaries and posts for multiple platforms.**

Note: The most useful model type in each research category is bolded.

Korinek (2025)

Thank You!

Sean Mullin

sean.mullin@gmail.com

www.seanmullin.com