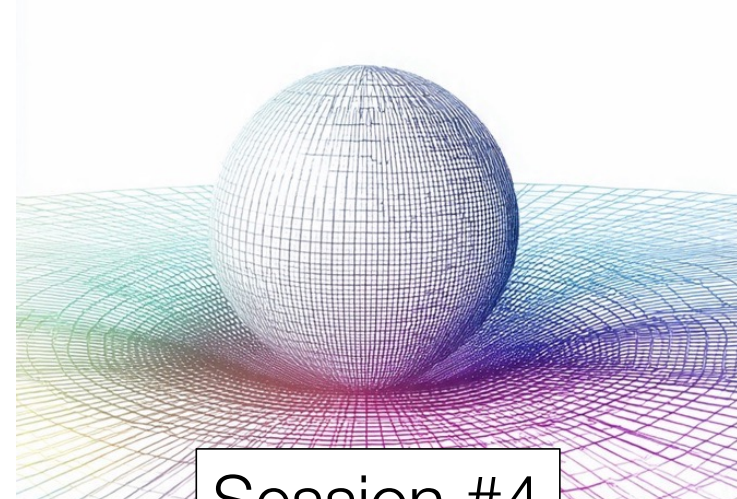


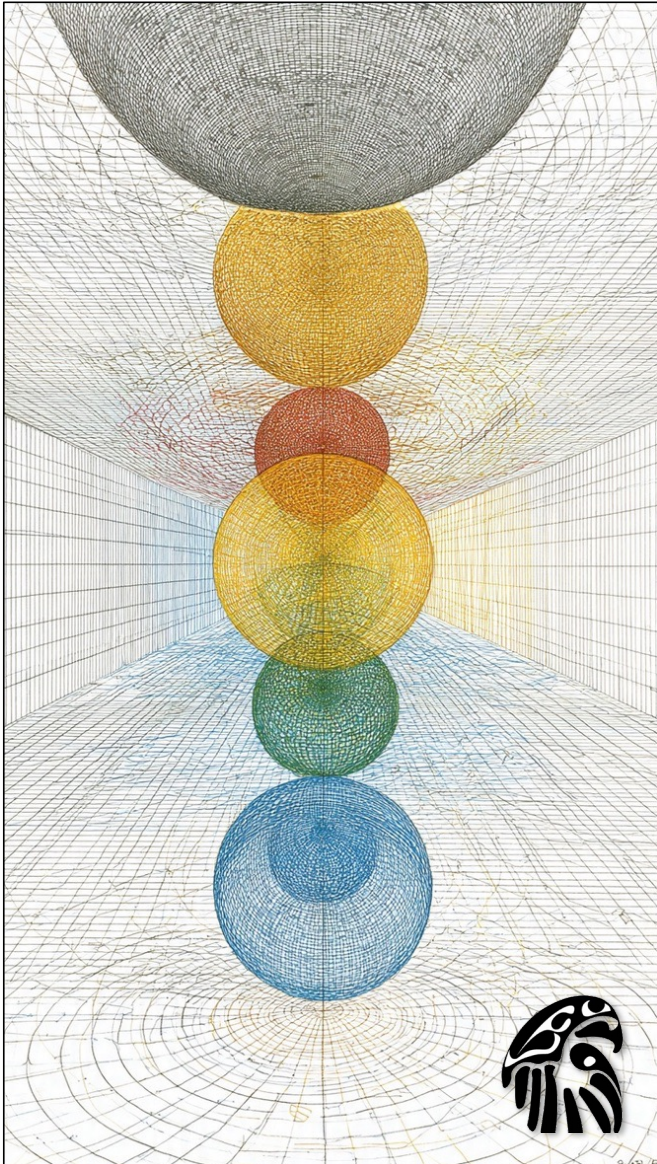
Further Experiments

In Artificial Intelligence

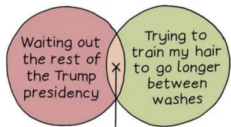


Session #4

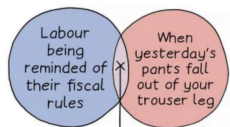
November 5, 2025



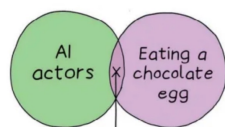
November 4, 2025



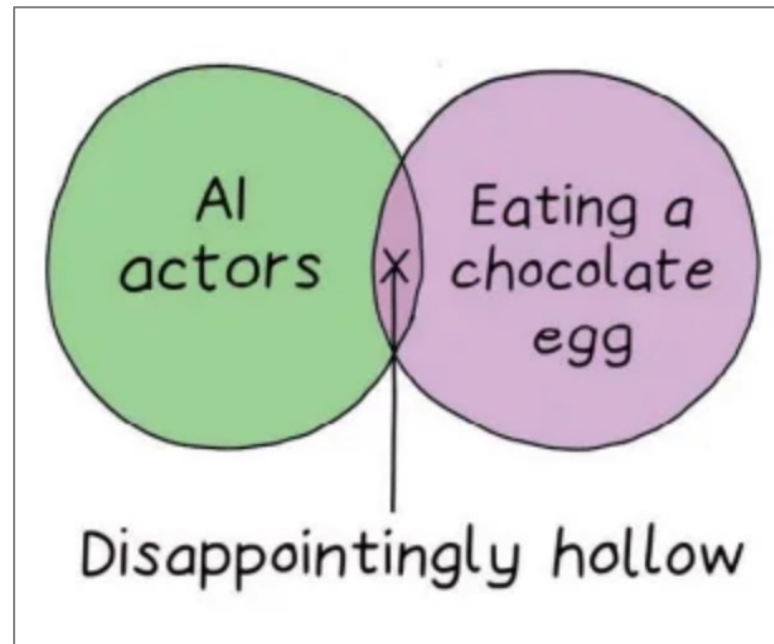
Braced for it to become increasingly putrid



Something from the past returning to embarrass you



Disappointingly hollow



Google Gemini Picture & Video Imagery Manipulation



Taxonomy of AI Hallucinations



1. Factual Errors (AI gets the facts wrong)

Intrinsic: The AI contradicts itself or the input.

- Says someone's birth year is 1980, then later 1975.
- Summarizes a story as both "happy" and "tragic."

Extrinsic: The AI invents things that don't exist.

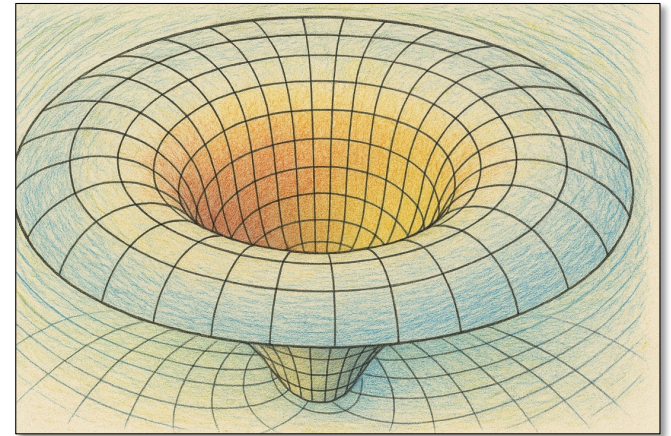
- Mentions "The Parisian Tiger" (never existed).
- Cites a Shakespeare play that doesn't exist.

Factuality: The AI makes claims that are false.

- Says "Charles Lindbergh walked on the moon."
- Claims Mount Everest is in Europe.

Factual Errors: The AI gives misleading or mistaken details.

- Says JWST took first exoplanet images (not true).
- Claims Amazon was founded in 1998 (actually 1994).



1. Factual Errors (continued)

Temporal: The AI confuses timelines or dates.

- Says "Murakami won the Nobel Prize in 2016" (he hasn't won).
- Claims Abraham Lincoln met Winston Churchill.

Overgeneralization: The AI applies patterns too broadly.

- Says "all birds can fly."
- Claims "every U.S. president was male" (ignores possibility of future change).



2. Misuse of Context or Instructions

Faithfulness: The AI twists what was actually said in the input.

- Article: "FDA approved the vaccine." → AI: "FDA rejected the vaccine."
- Recipe calls for sugar, but AI writes salt.

Contextual: The AI adds details or misreads the context.

- Input: "Nile in Central Africa." → Output: "Nile in Central African mountains."
- User means Paris, Texas, but AI assumes Paris, France.

Instruction: The AI fails to follow directions.

- Asked to "answer in Spanish," but replies in English.
- Told to list 3 points, but writes 5.

3. Logical & Reasoning Errors

Logical: The AI makes mistakes in reasoning or contradictions.

- Solves $2+2$ as 5.
- Says "if all dogs are mammals and cats are mammals, then cats are dogs."

Amalgamated: The AI blends multiple facts into one wrong answer.

- Combines two movies into one plot.
- Says "Einstein discovered DNA" (mixing Einstein with Watson/Crick).

Nonsensical: The AI gives responses that don't make sense.

- NBA discussion: switches from "Adam Silver" to "David Stern" randomly.
- Gives weather forecast for Mars when asked about New York.

4. Specialized Hallucinations (Part 1)

Ethical: The AI says harmful or defamatory things.

- Falsely accuses a professor of fraud.
- Gives unsafe advice like "drink bleach to cure illness."

Code Generation: The AI writes code that looks right but doesn't work.

- Python code missing imports.
- HTML with broken tags that won't display.

Multimodal: The AI misidentifies or invents things in images/audio + text.

- Says a cat photo shows a dog.
- Describes a "giraffe in the picture" when none exists.

Attribution/Citation: The AI invents sources, papers, URLs, or misattributes quotes.

- Cites "Johnson et al. (2023)" when no such paper exists.
- Invents a URL like "www.fake-study.org/research."
- Attributes a Mark Twain quote to Einstein.

4. Specialized Hallucinations (Part 2)

Quantitative/Statistical: The AI fabricates numbers, percentages, or statistics.

- Says "73% of doctors recommend this" with no source.
- Claims "average global temperature rose 2.4°C" (incorrect figure).
- Invents company revenue: "Tesla made \$450B in 2023."

Entity Confusion: The AI confuses or conflates people, places, or organizations.

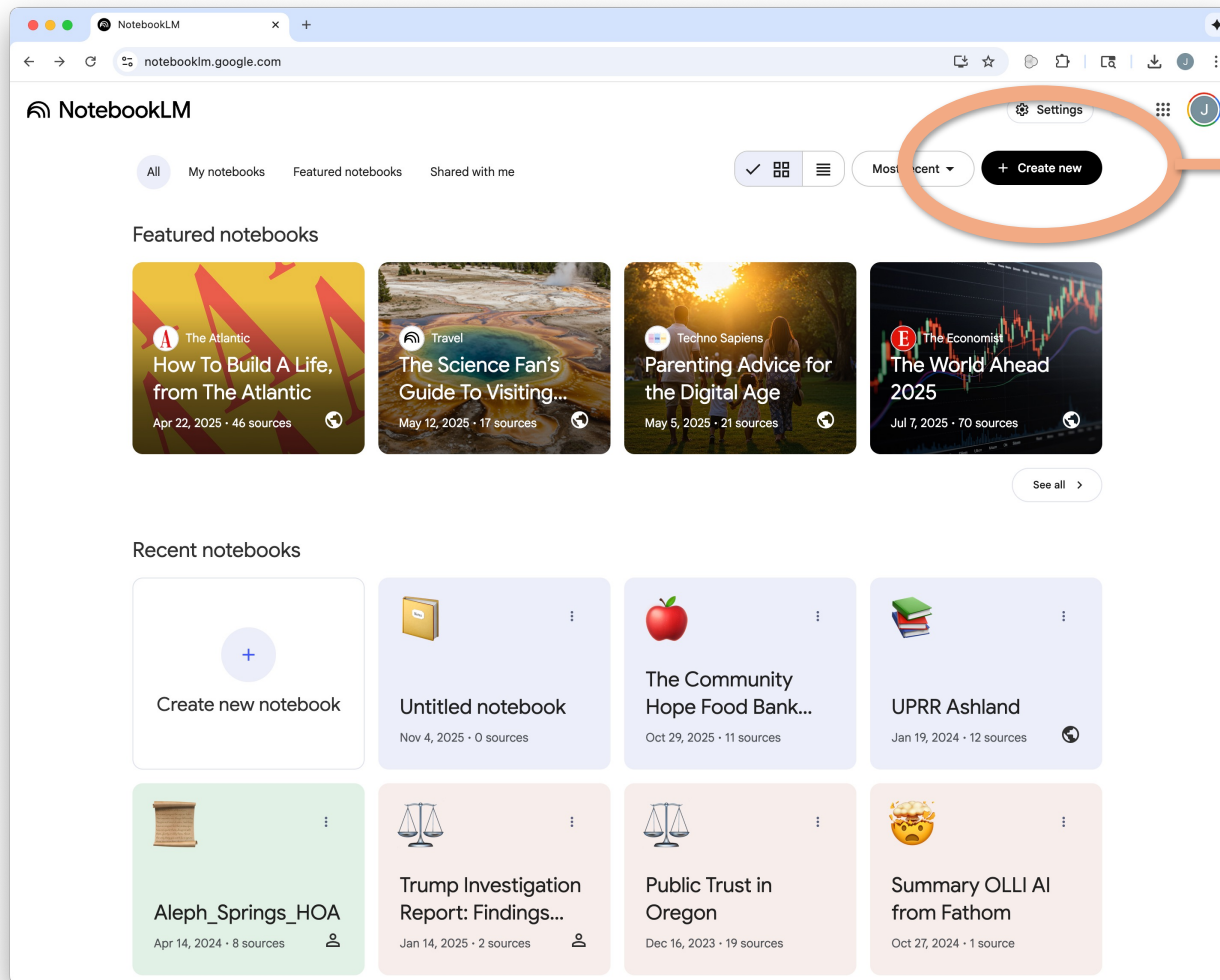
- Confuses SpaceX with Blue Origin.
- Says Jeff Bezos founded Tesla.
- Mixes up Cambridge, MA with Cambridge, UK in context.

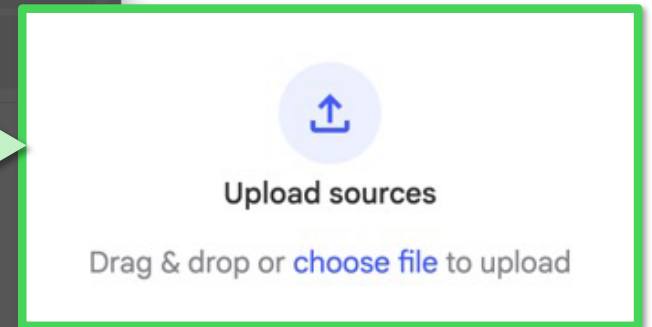
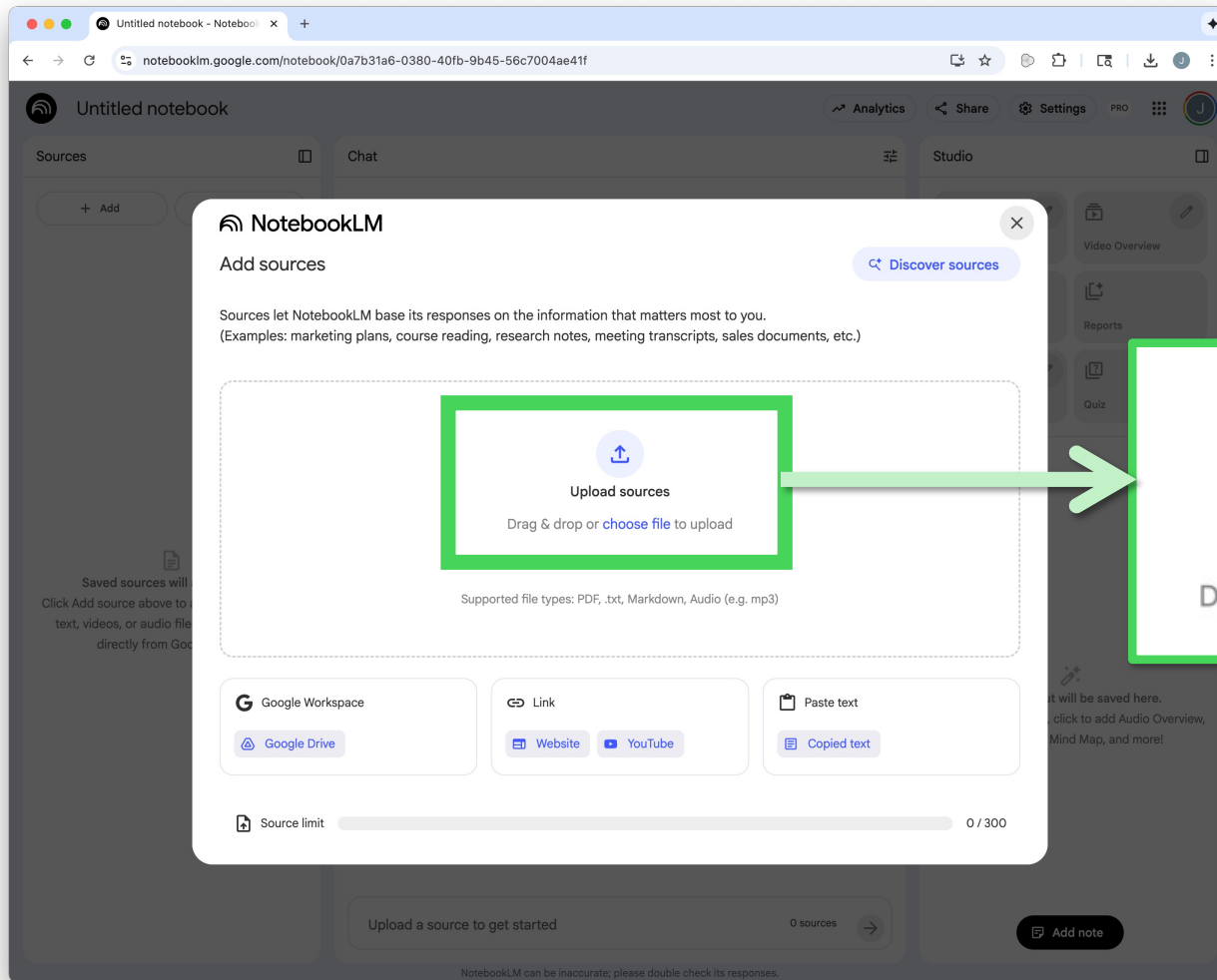
Format/Capability: The AI claims to perform actions it cannot or misrepresents capabilities.

- Says "I've attached the file" when it cannot attach files.
- Claims "I sent the email" when it has no email access.
- States "I updated your calendar" without that capability.

NotebookLM







Untitled notebook - NotebookLM

notebooklm.google.com/notebook/0a7b31a6-0380-40fb-9b45-56c7004ae41f

Untitled notebook

Sources

+ Add

Discover

Select all sources

GPO-J6-REPORT.pdf

Chat

Untitled notebook

1 source

Video Overview

Audio Overview

Mind Map

Studio

Audio Overview

Video Overview

Mind Map

Reports

Flashcards

Quiz

Studio output will be saved here.
After adding sources, click to add Audio Overview, Study Guide, Mind Map, and more!

Add note

One or more files were not uploaded. Only the following files types are supported: pdf, txt, md, 3g2, 3gp, aac, aif, aifc, aiff, amr, au, avi, cda, m4a, mid, mp3, mp4, mpeg, ogg, opus, ra, ram, snd, wav, wma

Studio

Audio Overview

Video Overview

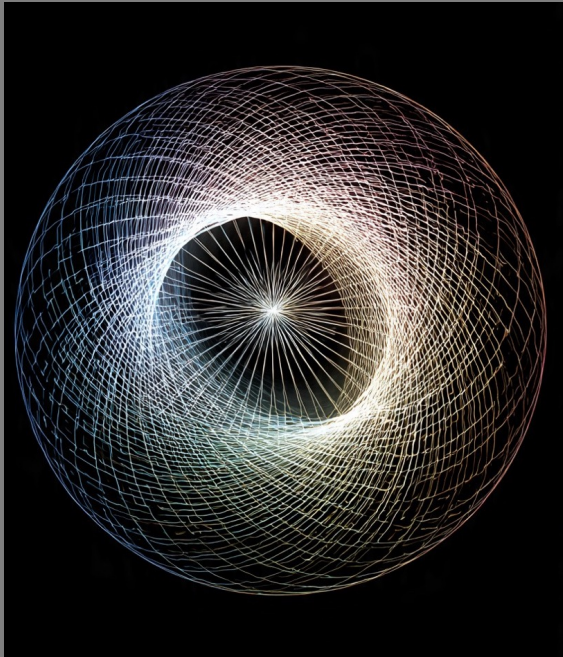
Mind Map

Reports

Flashcards

Quiz





AI Assistant Capabilities

What to Look for When Comparing AI Editors



Everyday Features You'll Use Most

Text & Conversation

- Having back-and-forth conversations with follow-up questions
- Text corrections (spelling, grammar)
- Remembering what you discussed in the current conversation
- Saving and searching past conversations
- Setting personal preferences that apply to all conversations

Files & Documents

- Uploading documents, spreadsheets, PDFs, and photos
- Creating new Word docs, PowerPoint presentations, Excel files
- Reading text from images (OCR)
- Converting between file formats

Current Information

- Searching the web for up-to-date information
- Providing sources and citations you can verify
- Getting real-time data (weather, news, prices)



Creative & Analysis Tools

Images & Visuals

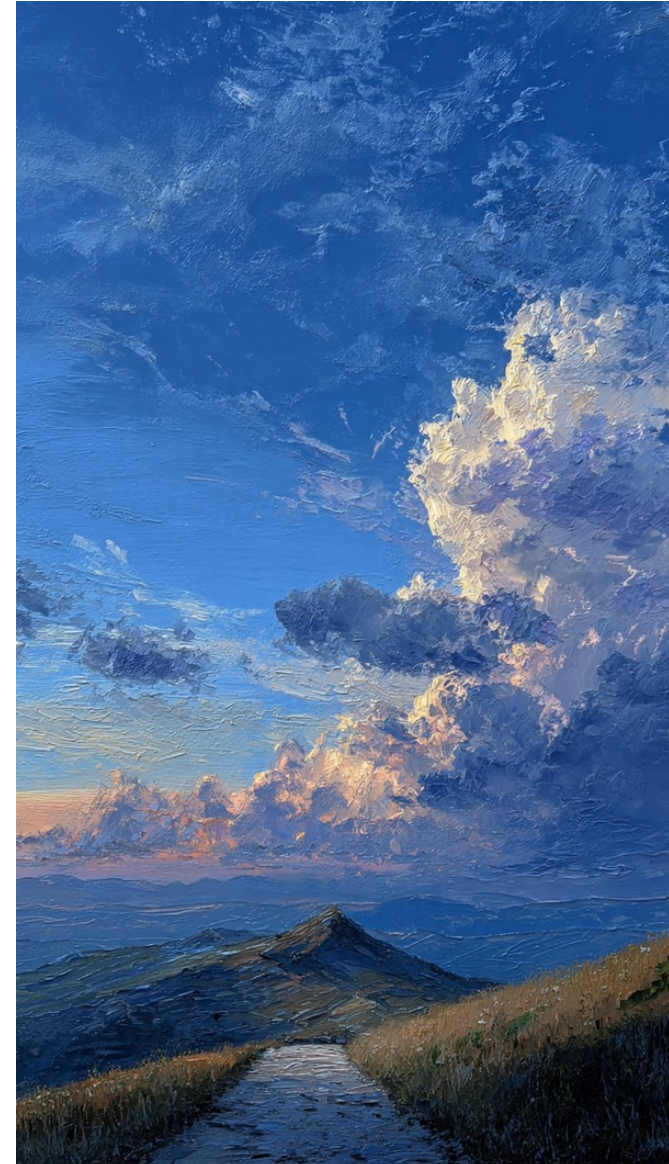
- Understanding and describing uploaded photos
- Creating new images from text descriptions
- Editing images (background removal, enhancements)
- Creating charts, graphs, and diagrams

Language & Writing

- Translating between languages
- Improving writing (grammar, clarity, style)
- Adjusting tone (formal, casual, enthusiastic)
- Summarizing long documents

Data & Analysis

- Performing calculations and solving math problems
- Analyzing spreadsheet data
- Creating comparison tables
- Making recommendations based on multiple factors



Advanced Features (Optional)

Technical Capabilities

- Writing and debugging computer code
- Working with complex technical documents
- API access for developers

Voice & Audio

- Voice input (speak instead of type)
- Audio transcription (converting recordings to text)
- Text-to-speech output

Integration & Automation

- Working inside other programs (email, Office apps, browser)
- Connecting to your calendar or email
- Automating multi-step tasks
- Batch processing multiple files

Collaboration

- Sharing conversations with others
- Team workspaces
- Exporting content in various formats



Safety Features & Comparison Checklist

Privacy & Security

- Options for private/temporary conversations
- Data deletion controls
- Content filtering for appropriate responses
- Transparency about sources and limitations

When Comparing Different AI Assistants, Ask:

1. Which everyday features does it have? (Most important)
2. How well does each feature actually work? (Quality matters)
3. Can it access current information from the web?
4. Does it show sources so you can verify information?
5. What file types can it work with?
6. Is it easy to use? (Try it yourself!)
7. What are the privacy and security options?
8. What's the cost? (Free vs. paid tiers)





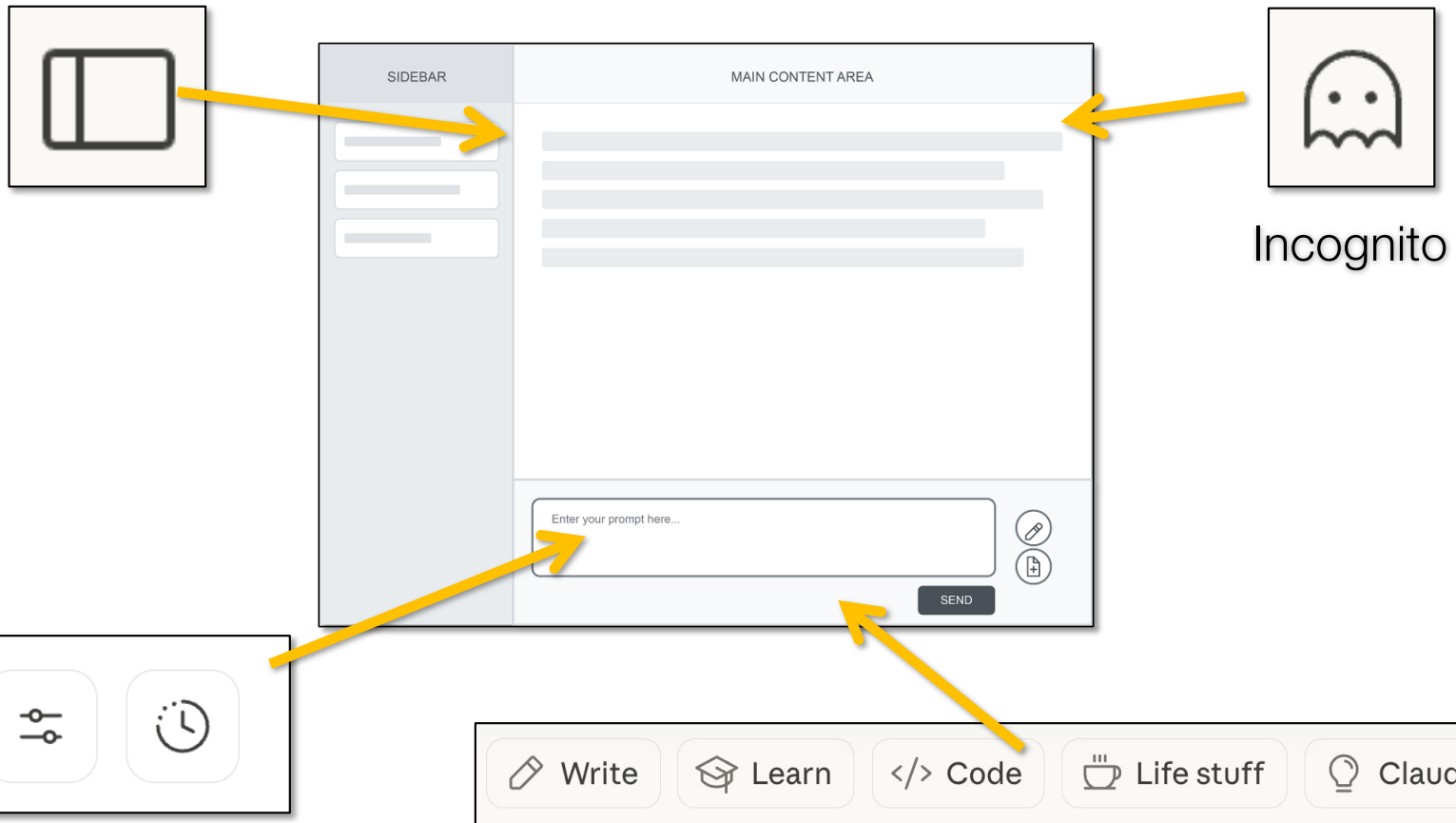
SIDEBAR

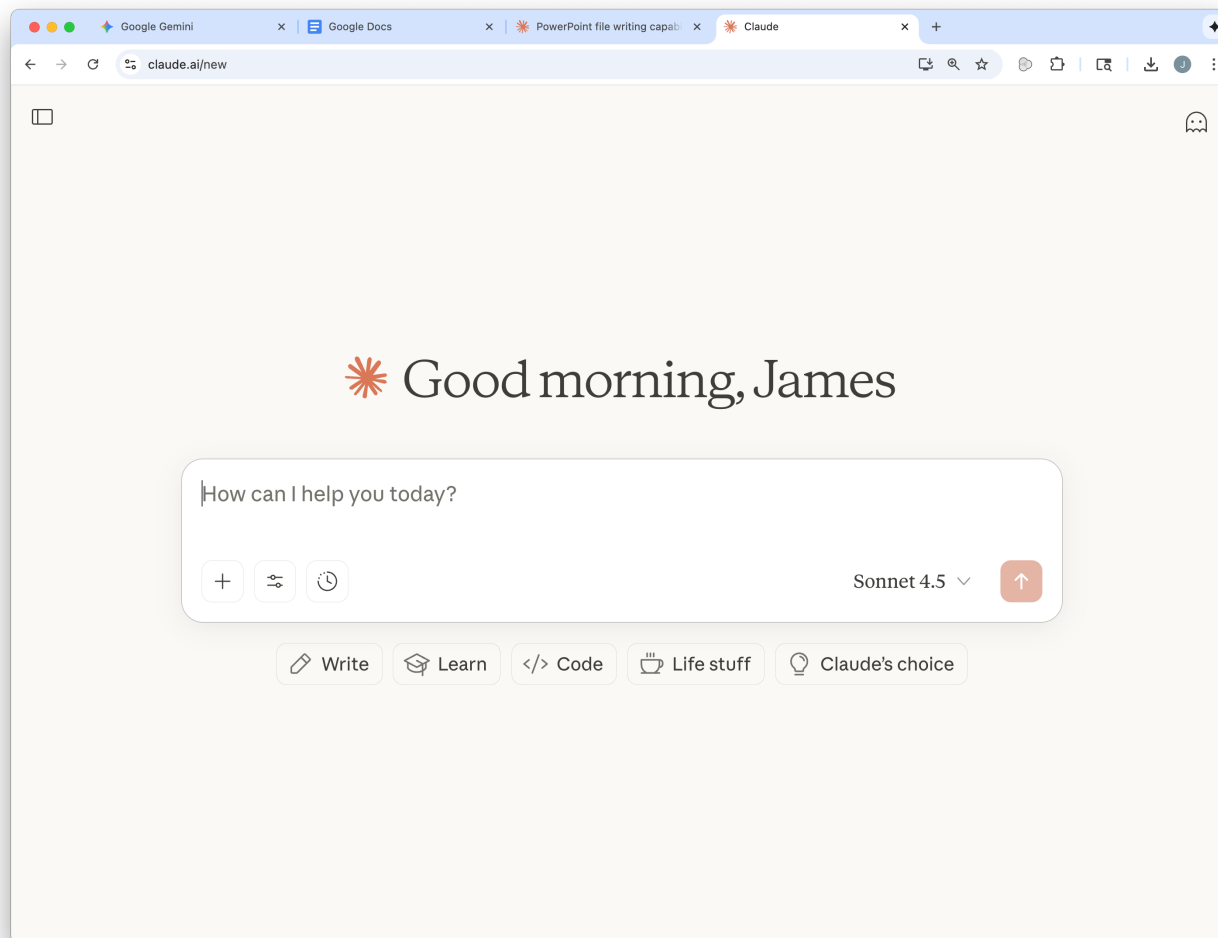
MAIN CONTENT AREA

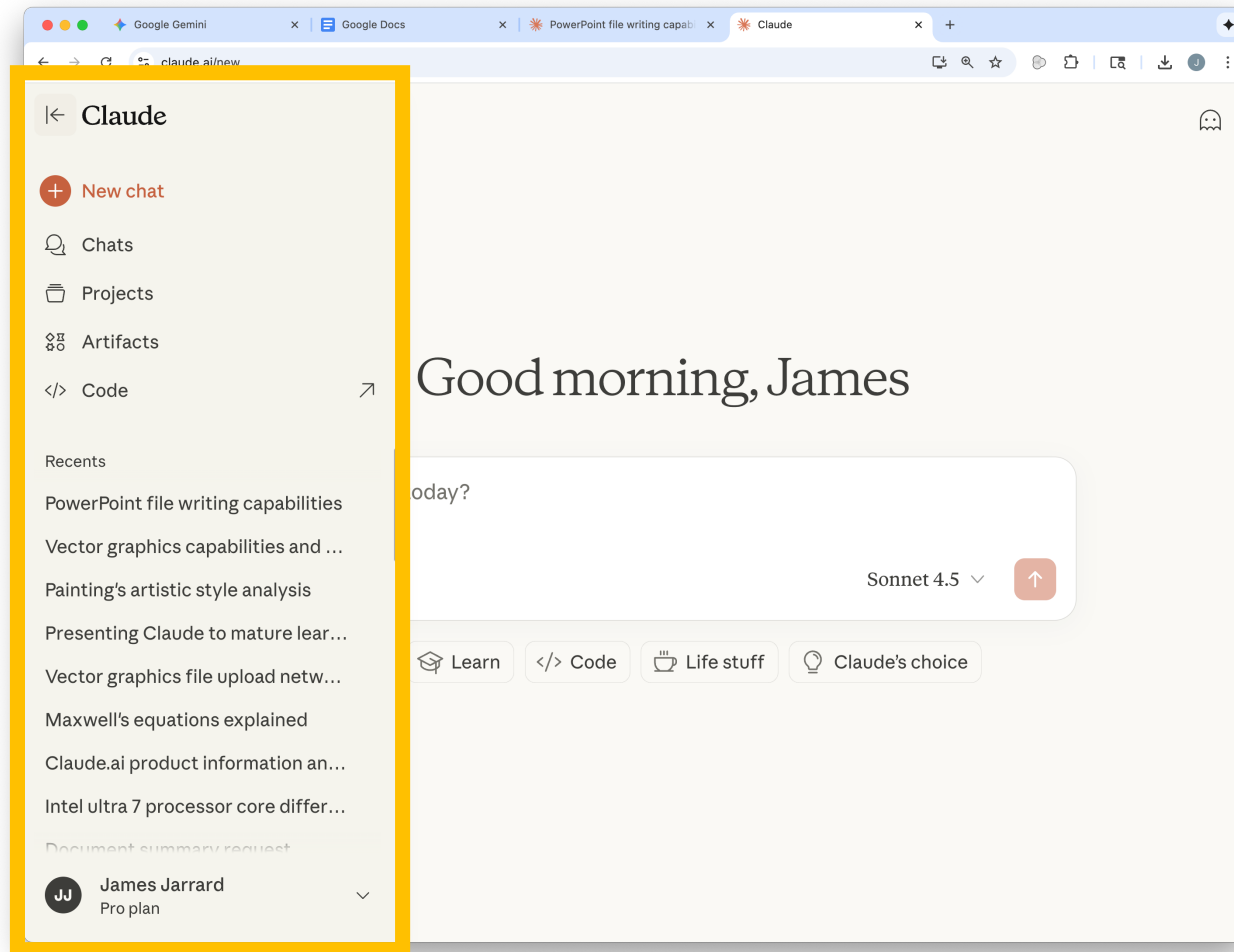
Enter your prompt here...

SEND









 Good morning, James

How can I help you today?



 Write

 Learn

</

Sonnet 4.5 ▾



Opus 4.1

Deep brainstorming model. Consumes usage faster.

Sonnet 4.5

Smartest for everyday tasks



Haiku 4.5

Fastest for quick answers

More models



 Good morning, James

How can I help you today?



Sonnet 4.5 ▾



 Search menu

 Use style >

 Extended thinking ☐

 Research ☐

 Web search ☒

 Drive search [Connect](#) 

 Gmail search [Connect](#) 

Life stuff

 Claude's choice

 Good morning, James

How can I help you today?



Sonnet 4.5 ▾



Search menu



Upload a file



Take a screenshot



Add from GitHub



Add from Google Drive

Connect 



Use a project



Life stuff



Claude's choice



©Talks with Crows

demonstration

November 4, 2025

○ Meta AI
29