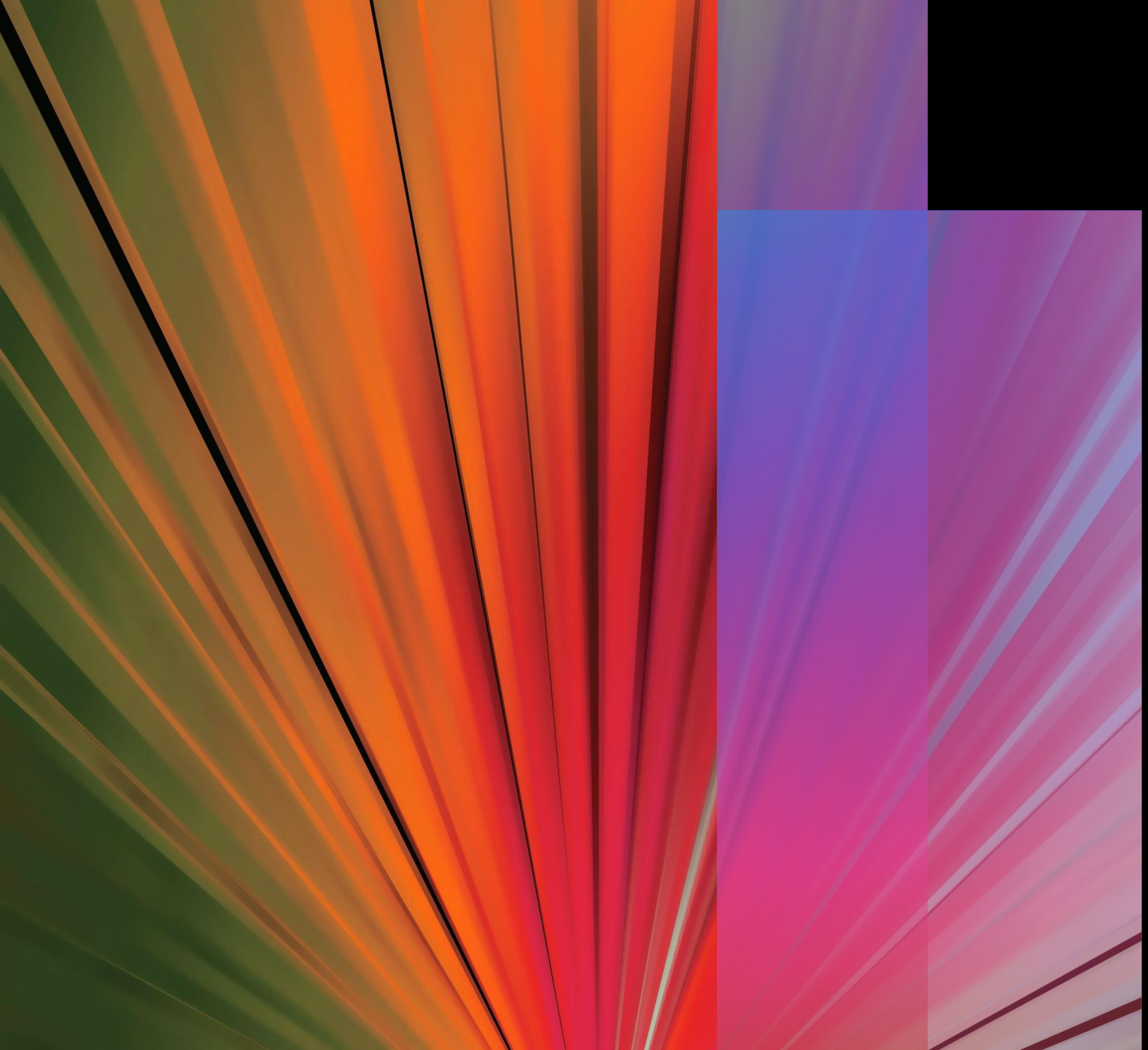


AI: From Zero to 60

CMDA 2026

Steven J. Willing MD

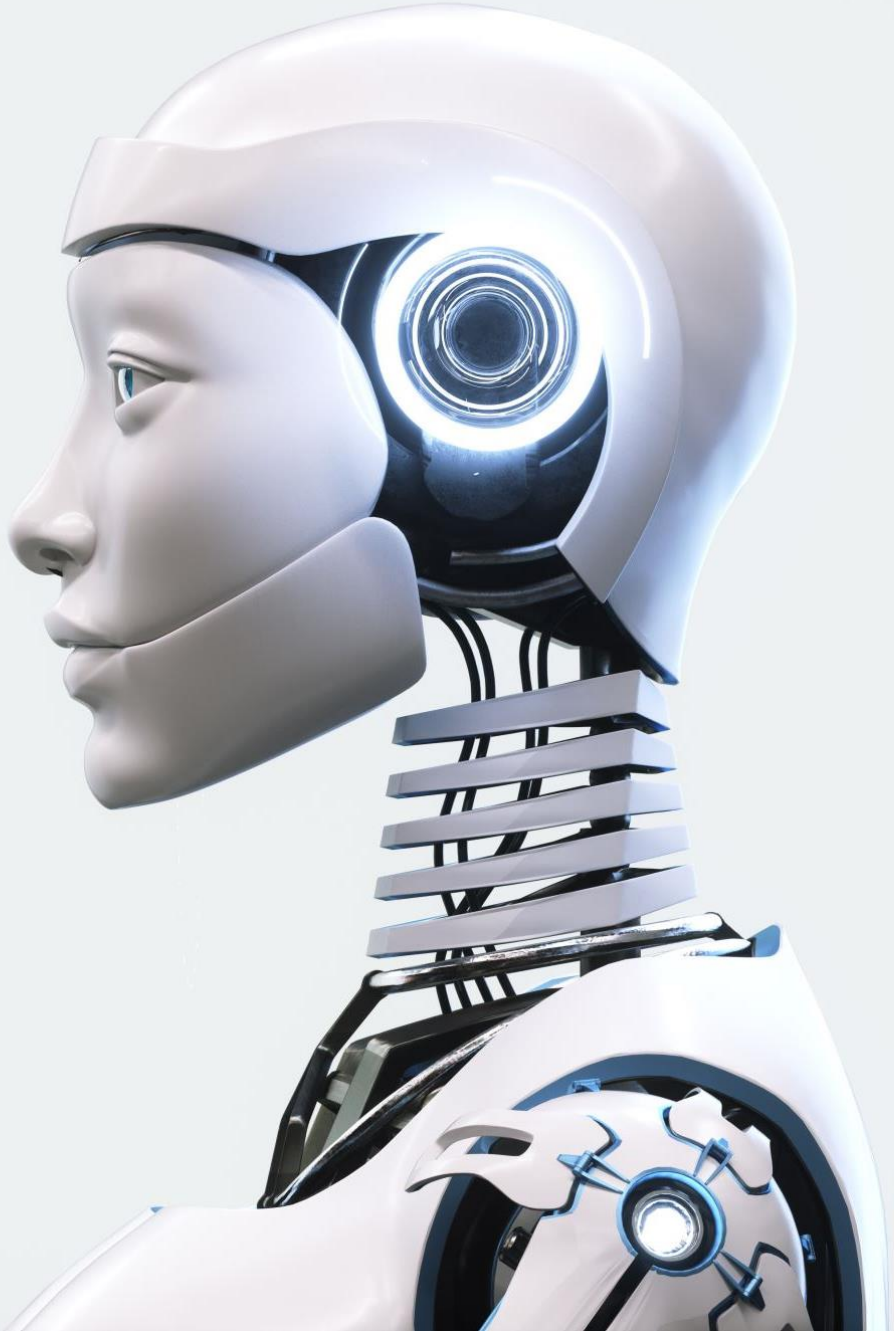
Birmingham, AL

Financial Disclosures:

Your name/logo could be here!

Flavors

- Generative AI
- Chatbot
- Predictive AI
- Companion bot
- Agentic AI
- Personal AI



You’re probably already using it

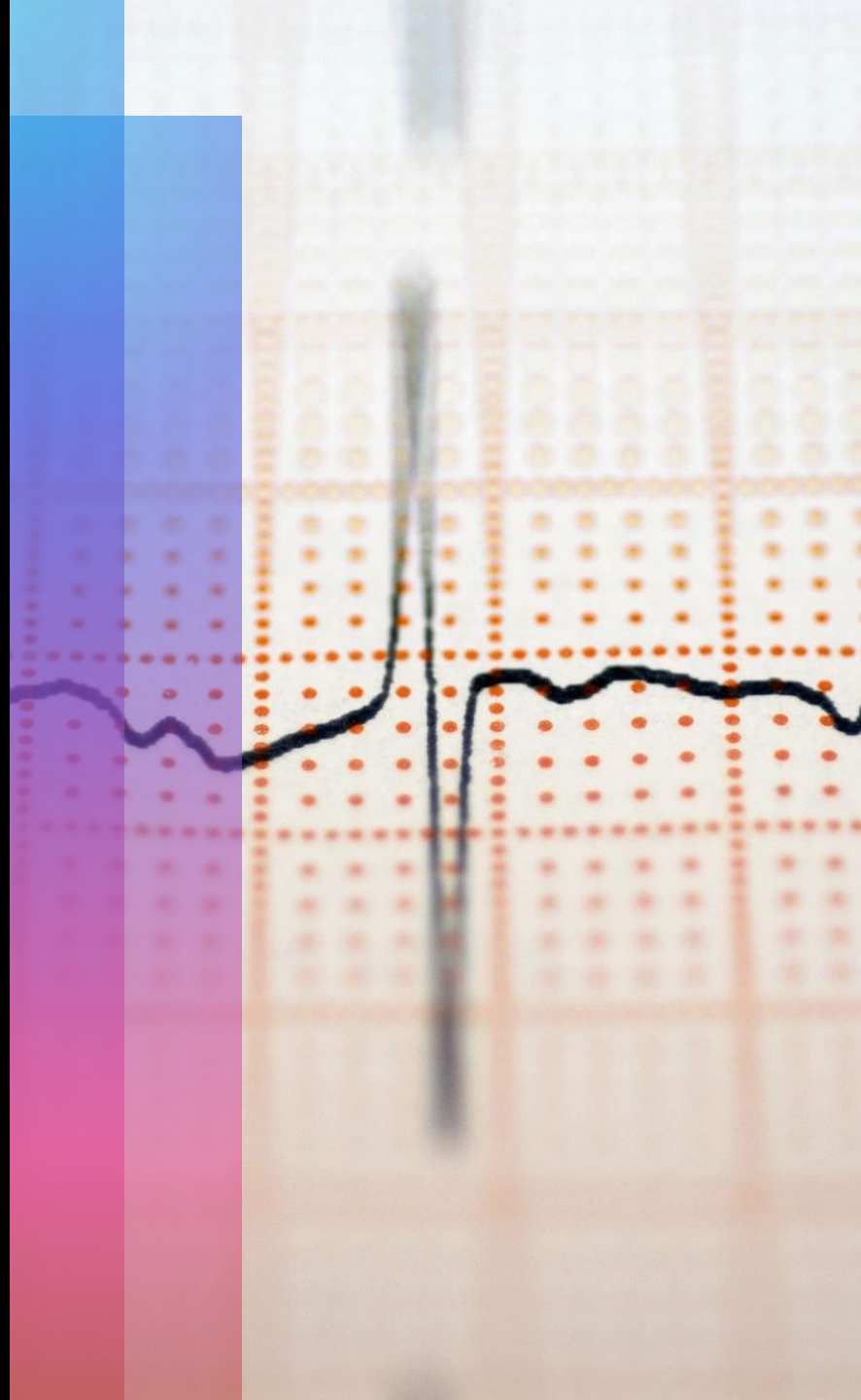
>80% of physicians using AI professionally

>340 FDA-approved tools currently in use

Major EMRs already include AI features

AI in Medicine Didn’t Start with ChatGPT

- Image analysis in radiology
- Deep learning pattern recognition
- ECG-based prediction models



Where AI Helps Physicians Most

Document generation

Clinical inquiry

Pattern detection

Risk stratification

Biomedical research

1980s: Gutenberg Age

Library Books & Journals

Index Searches

Manual Article Filing

2000s: Internet Age

Online Journals

Web Searches

Tedious Result Sorting

2020s: Dawn of AI

Instant Knowledge

Natural Language Queries

Automated Assistance

What took hours in the *Gutenberg Age*, *minutes in the Internet Age*, *takes seconds in the Age of AI.*

Common objections

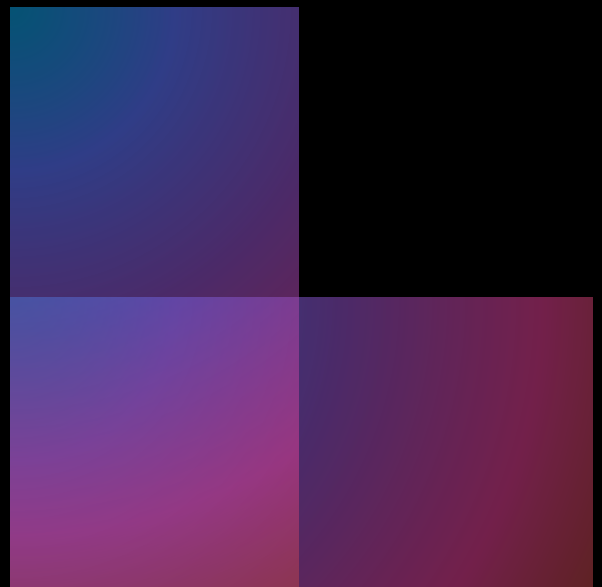
- It hallucinates
- It will replace physicians
- I don’t trust it
- I don’t have time

AI Use Maturity Curve

Stage 1 – Convenience
Stage 2 – Capability
Stage 3 – Cognitive Extension

1. Getting started: from Zero to Useful

- Know what AI does well
- Start with low-risk tasks
- Prompting 101: how to ask for what you actually want
- Use AI for focused clinical inquiry
- Know your tools



What AI is actually good at

Synthesis

Drafting

Explanation

Structured inquiry

When not to use AI

- When stakes are high but you don’t have the time to verify
- Final authoritative citations without checking source
- When you can see the data and it can’t
- Substitute for clinical judgment

More efficient, more comprehensive

Traditional workflow

- Hunt for sources
- Manually integrate information
- Then document
- Linear and time-heavy

AI-enhanced workflow

- Start with structured questioning
- Iteratively improve the output
- Stress-test the reasoning
- Then apply judgment

The right mental model

- AI is not an oracle — it's an affable colleague who is usually right but sometimes wrong with equal confidence. Verify when stakes are real.

Your first practical tasks

- Notes
  - Note-taking: Doximity Scribe, PowerScribe
- Letters
  - Writing authorization letters
- Summaries
- Differential diagnosis
- Current standards and recommendations
- Prescribing information
- Generating patient handouts

Benefits of AI-assisted documentation

- Turning rough narrative into structured note
- Generating patient instructions
- Drafting referral summaries
- Drafting prior authorization letters

Prompting 101

- Specify context (physician vs. patient)
- Request a format or structure (bullets, table, SOAP)
- Include all pertinent details
- Add constraints (length, tone). A structured prompt produces a structured answer.
- No names, dates, or identifiers unless HIPAA-compliant. The model doesn't need identity to reason well.

Use AI for focused clinical inquiry

- What is the differential diagnosis?
- What are current standards of care?
- What are points of controversy?
- Make the query as specific as you like

Know your tools

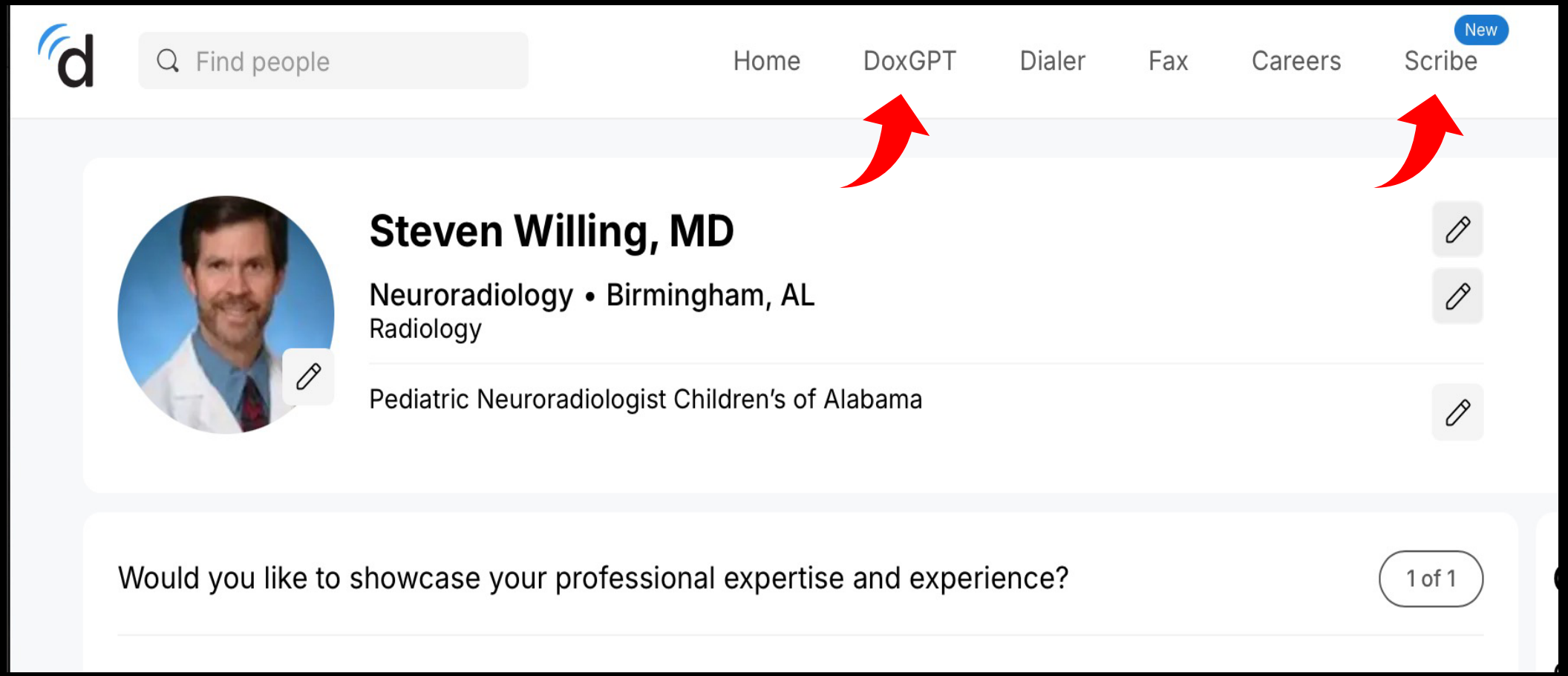
General

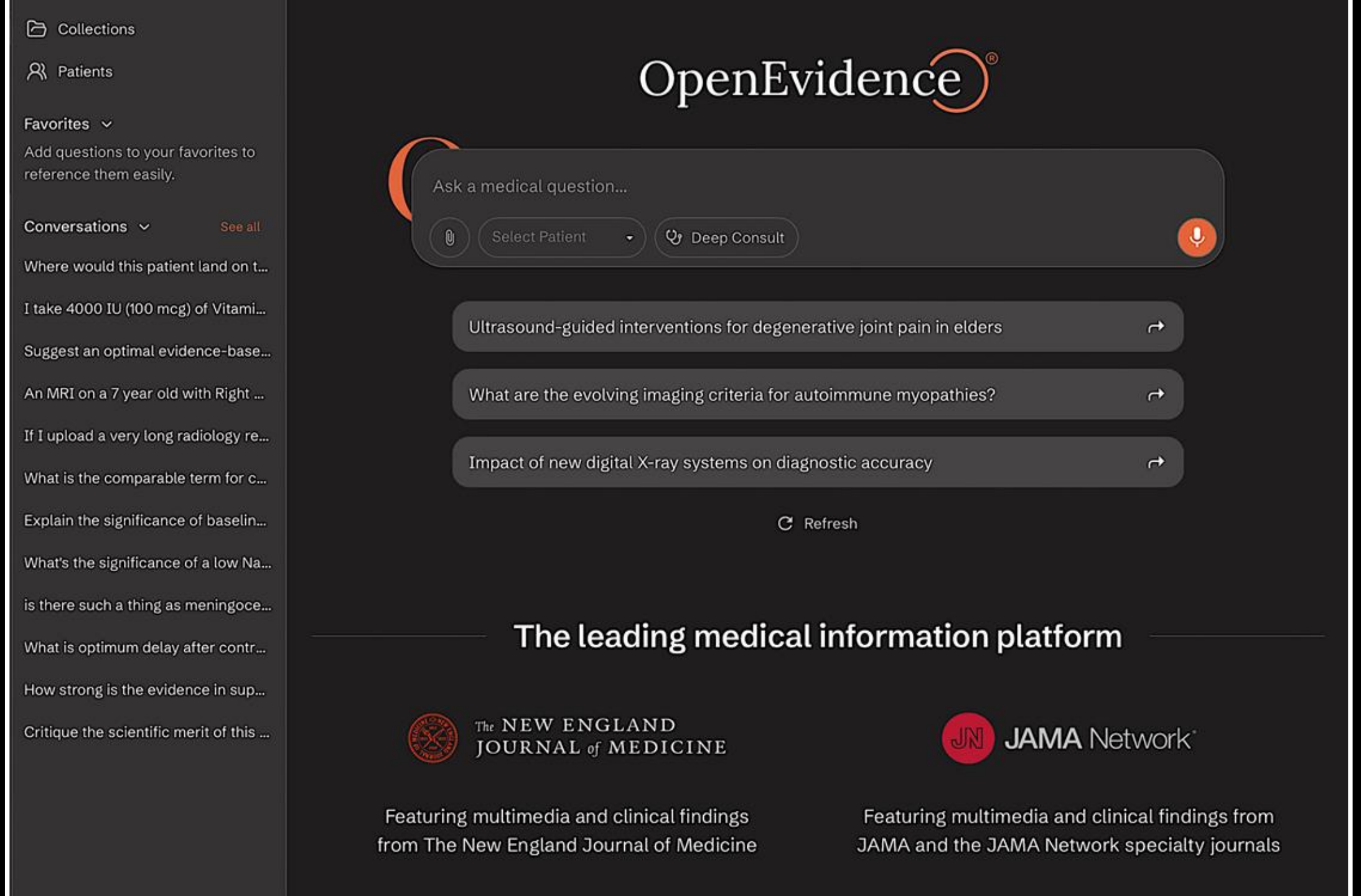
- Claude
- Gemini
- Copilot
- ChatGPT

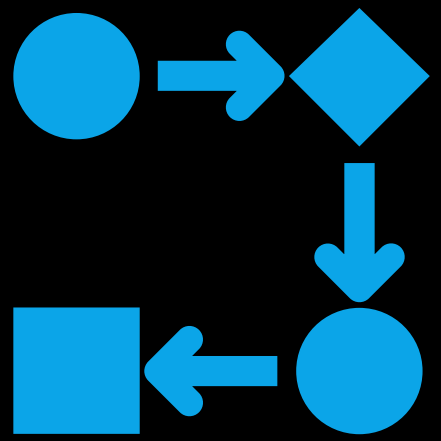
Medical

- OpenEvidence
- DoxGPT

Doximity.com




Collections
Patients
Favorites
Add questions to your favorites to reference them easily.
Conversations
See all
Where would this patient land on t...
I take 4000 IU (100 mcg) of Vitami...
Suggest an optimal evidence-base...
An MRI on a 7 year old with Right ...
If I upload a very long radiology re...
What is the comparable term for c...
Explain the significance of baselin...
What's the significance of a low Na...
is there such a thing as meningoce...
What is optimum delay after contr...
How strong is the evidence in sup...
Critique the scientific merit of this ...
OpenEvidence
Ask a medical question...
Select Patient
Deep Consult
Ultrasound-guided interventions for degenerative joint pain in elders
What are the evolving imaging criteria for autoimmune myopathies?
Impact of new digital X-ray systems on diagnostic accuracy
Refresh
The leading medical information platform
The NEW ENGLAND JOURNAL of MEDICINE
JAMA Network
Featuring multimedia and clinical findings from The New England Journal of Medicine
Featuring multimedia and clinical findings from JAMA and the JAMA Network specialty journals



2. Upping your game: getting reliable, repeatable value

- Iterative prompting: refining instead of restarting
- Custom instructions: setting defaults so you stop repeating yourself
- Spotting confident nonsense
- Use simple decision rules for platform choice

Iterative prompting: refine, don’t restart

- In general, the chatbot has no memory or access to your previous chats
- When the first answer is close but not right, continue the conversation – ask it to adjust tone, shorten, reformat, or reconsider. One good session beats five restarts.

But don’t iterate forever.

- **Context window**
- Start afresh if:
  - AI starts contradicting itself
  - It starts forgetting earlier context
  - If the answers start getting worse
  - The task is fundamentally different
  - Tip: ask AI to generate summary, then paste it into a new chat.

Custom instructions: setting defaults so you stop repeating yourself

- Tell the platform your role, preferred verbosity, and audience level so you stop repeating yourself every session. This is the single biggest time-saver for regular users.
- Things to consider
  - Never sycophantic
  - Be concise
  - Offer what I may have overlooked
  - Stick to my question, only do what I ask

Spotting confident nonsense

- Watch for suspiciously tidy answers, overly specific citations, and responses that agree with everything you say. Fluency is not accuracy.
- A finely honed BS-detector is your friend and ally.

AI Citation Hallucinations: Still a Real Problem

53

papers

NeurIPS 2025 papers contained hallucinated citations — missed by multiple peer reviewers

50+

submissions

ICLR 2026 submissions with fake citations, each reviewed by 3–5 experts who missed them

34%

more confident

Models use stronger language ("definitely," "certainly") when generating incorrect information

Fluency is not accuracy. Citations require independent verification — especially for clinical and academic use.

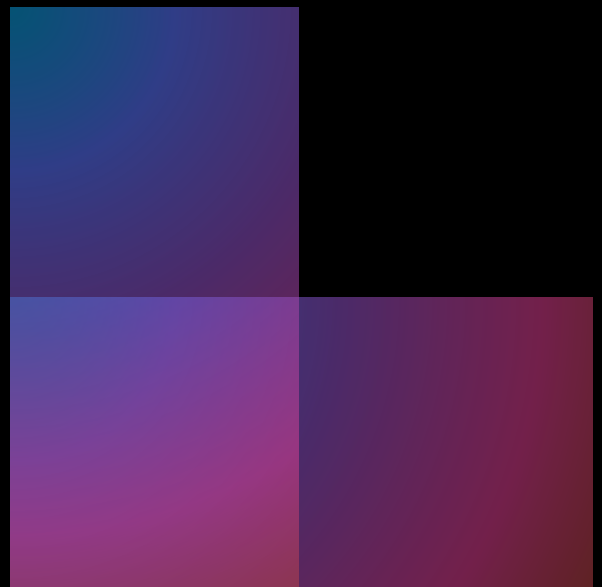
3 simple decision rules for platform choice

Need a cited clinical answer → medical platform.

Need synthesis or writing → general model.

Answer surprises you → verify elsewhere.

These three rules cover 80% of decisions.



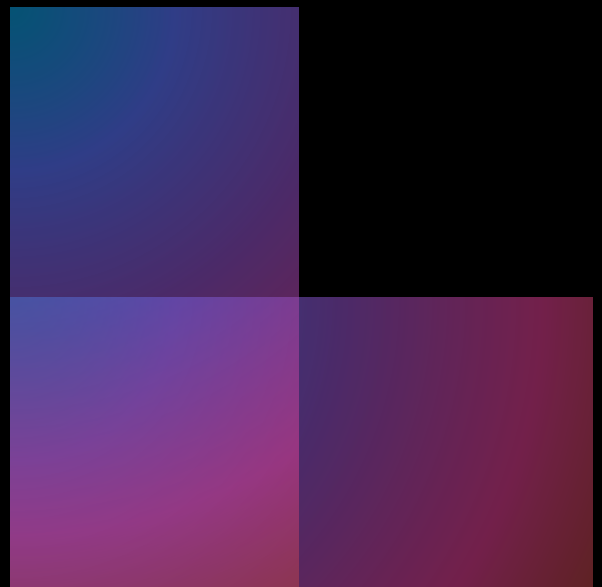
When the rules don’t cover it

If I need reliable citations → OpenEvidence

If the question is not specifically medical, or I want to think “outside the box” → use a general model (Claude, ChatGPT).

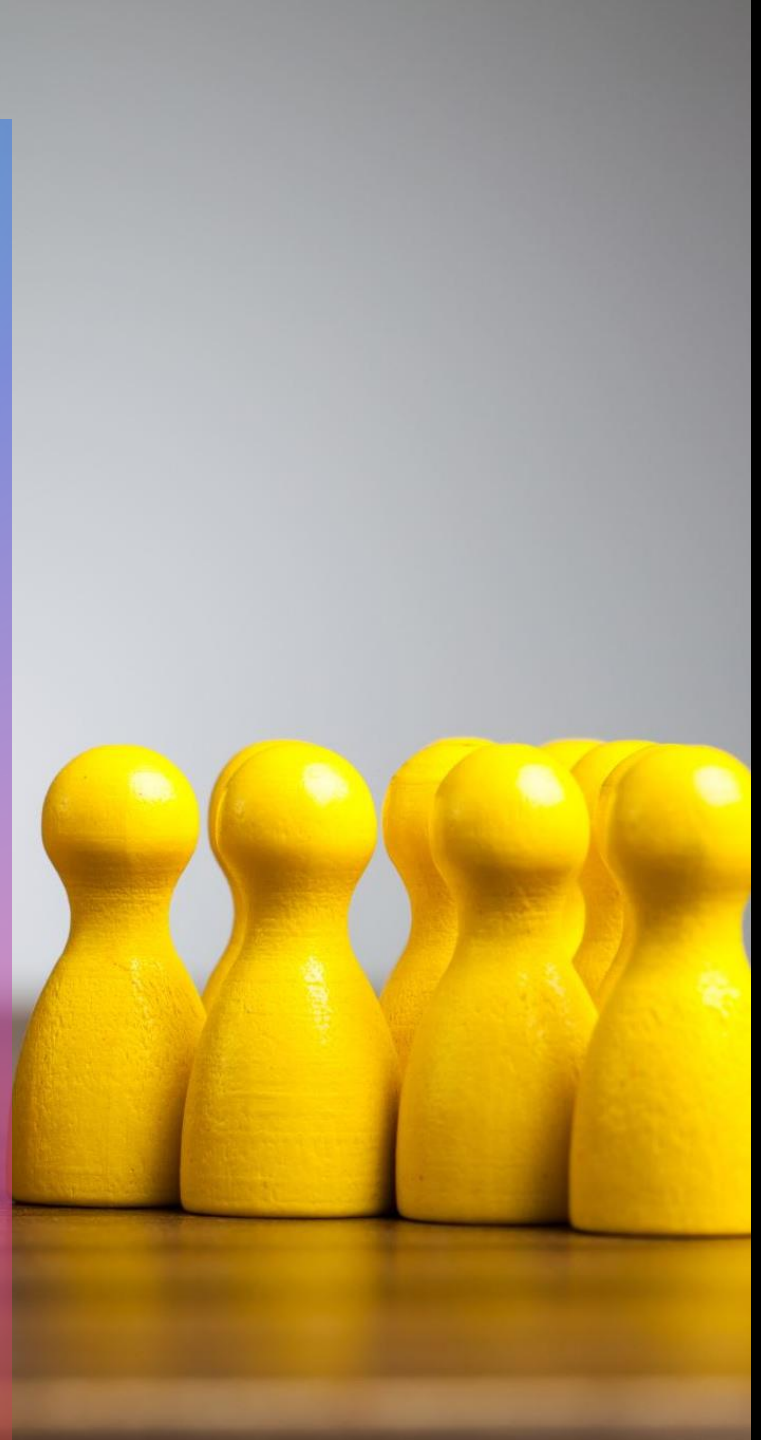
If I want conservative, guideline-aligned answers → favor physician-focused tools. (OpenEvidence or DoxGPT)

You’re not sure which platform? → try both and compare.



Medical versus general platform

“When accuracy matters, narrow the model; when thinking matters, widen it.”



3. Advanced tips & tricks: Power without Pain

- If you have the file, upload it
- Reverse prompting: let AI challenge your thinking
- Projects: pick up where you left off
- Use personal AI notebooks
- Chain your tasks: rough thought to polished output

Stop explaining the problem: Show the source

Traditional Approach

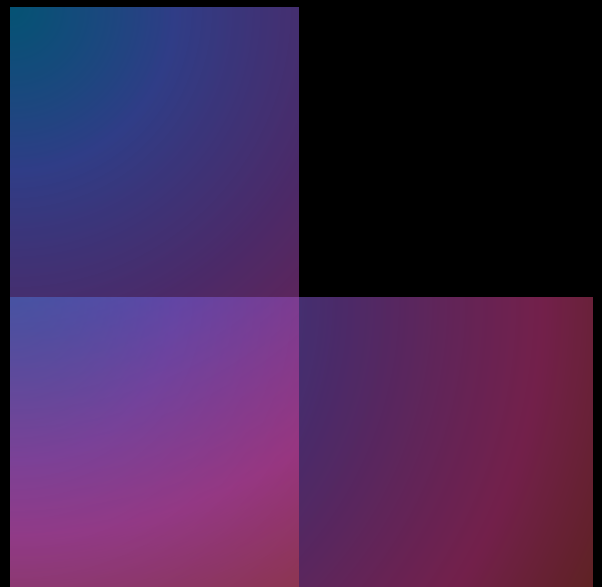
- Physician summarizes information
- Problems:
  - Risk of missing details
  - Time spent summarizing
  - AI only sees your summary

AI-assisted approach

- Upload the document:
  - Research paper
  - Clinical guidelines
  - Hospital policy
  - Data export
- Then request action

If the information already exists in a file, don't paraphrase it – upload it

Uploading files: letting the model analyze the source, not your summary



If the information exists in a document, upload it rather than summarizing it yourself.

Works with guidelines, papers, spreadsheets, policy documents, and data exports.

Images, screenshots

The quality of AI output is proportional to the quality of its input

Reverse prompting

- *Ask AI to critique, stress-test, or find weaknesses in your own clinical reasoning or written work* — which is a meaningfully different cognitive move than asking it to synthesize information for you.
- Metacognition

Projects: Using AI as a Persistent Workspace

Without Projects

Scattered chats

- Repeating queries or uploads each session
- Lost decisions and drift
- Repeated setup work
- "Didn't we already decide this?"

With a Project

One ongoing workspace

- Goal stays constant
- Prior decisions remembered
- Documents accumulate
- Work resumes, not restarts

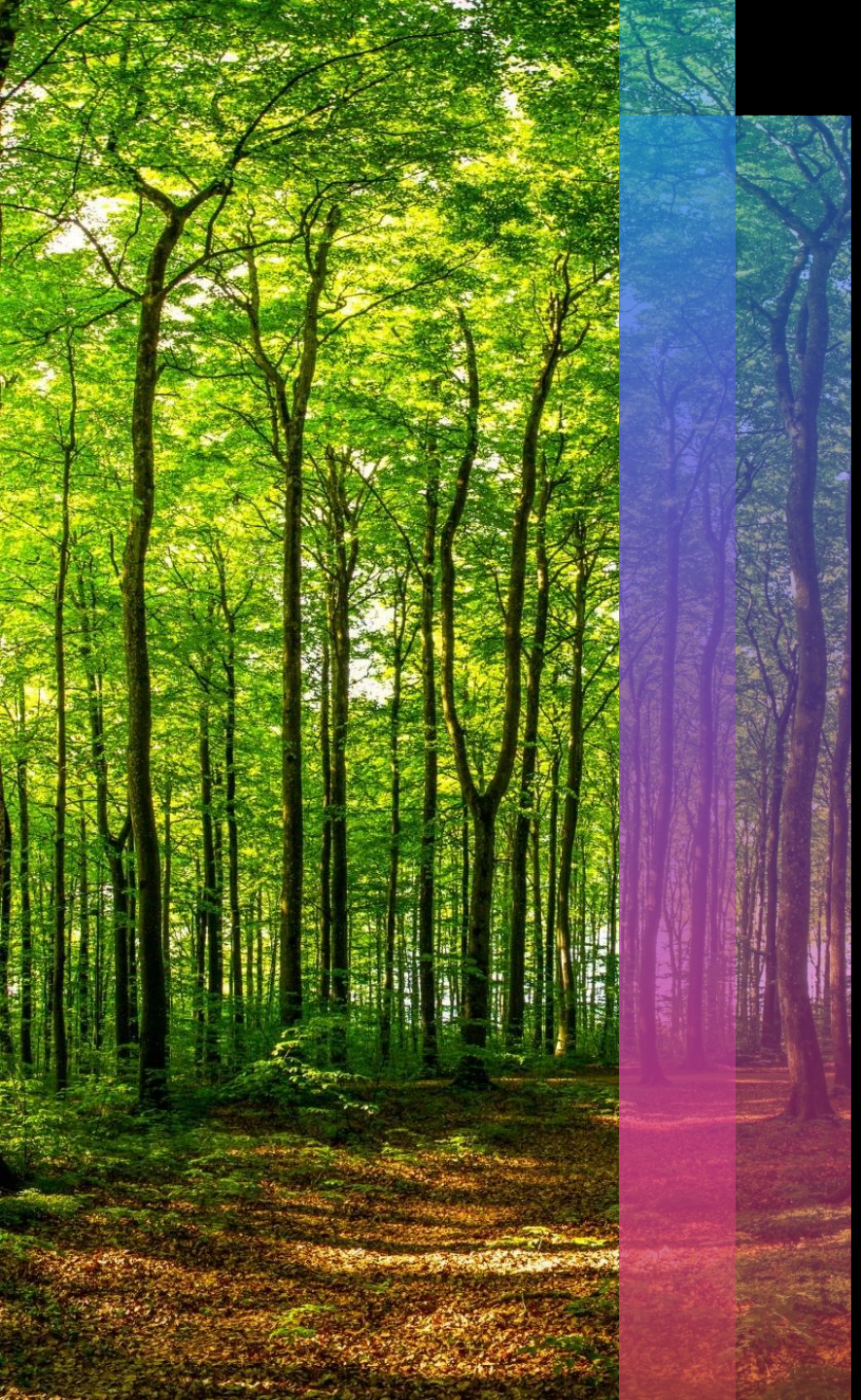
Projects turn AI from a conversation into a workspace.

Use personal AI notebooks

- Upload a curated collection of guidelines, papers, or CME materials and query them as a unified knowledge base. Turns a static PDF library into an interactive reference system.
- NotebookLM

Chaining tasks

- Use AI sequentially — outline → draft → refine → distill.
- Each pass improves the previous one. This is where advanced users get disproportionate leverage.



Parting advice: Keeping up with AI

- *Panta rei (Heraclitus)*
- Focus on the forest, not the trees.
  - Trends matter, details not so much
- Find a one or two reliable sources and stick with them
  - There is a cognitive cost to changing platforms
- Let experience be your guide, and be open to new approaches
  - You can learn by doing. The barriers are low.
- Ask the AI to help you

"These teams reached insightful conclusions that neither a human nor a machine could have produced on its own. They were the only group to consistently rival the prediction market's accuracy. On certain questions, they even outperformed it.

"It's not that these people were more intelligent than the others in the study. But they demonstrated two important qualities: perspective-taking and ***intellectual humility***."

Is AI Smarter Than People? It's Complicated.

As a neuroscientist, I conducted research into artificial versus human intelligence. The results surprised me—and suggest we've been worrying over the wrong things.

By Vivienne Ming

Who's smarter, the human or the machine?

In the 30 years I've worked in artificial intelligence, that's been the question driving the conversation.

We've also been sold a story about AI that goes something like this: It will handle the tedious, routine work—the research, the first draft, the number-crunching—while we focus on the interesting parts: creativity, judgment, the human touch.

My research suggests we've been asking the wrong question and drawing

mation had come across their feeds that morning. The large AI models—ChatGPT and Gemini, in this case—performed considerably better, though still short of the market itself.

But when we combined AI with humans, things got more interesting.

Most hybrid teams used AI for the answer and submitted it as their own, performing no better than the AI alone. Others fed their own predictions into AI and asked it to come up with supporting evidence. These "validators" had stumbled into a classic confirmation bias-loop: the sycophancy that leads chatbots to tell you what you want to hear, even if it isn't true. They ended up performing worse than an AI working solo.

Largely unasked: Is the product building human capacity or

intelligent than the others in the study. But they demonstrated two important qualities: perspective-taking and intellectual humility.

Perspective-taking is the ability to genuinely inhabit another point of view. Not to debate it, not to tolerate it, but to actually inhabit it. Intellectual humility is the ability to recognize the edge of your own knowledge and sit with that discomfort rather than trying to rush to fill it.

Both of these qualities are, at root, emotional skills. Perspective-taking requires genuine curiosity about minds other than your own. Intellectual humility requires a kind of emotional courage: the willingness to feel uncer-

missing?" rather than default to "Great, that's done." To disagree with something that sounds authoritative and to trust your instinct enough to follow it.

We don't build these capacities by avoiding discomfort. We build them by choosing it, repeatedly, in small ways: the student who struggles through a problem before checking the answer; the person who asks a follow-up question in a conversation; the reader who sits with a difficult idea long enough for it to actually change one's mind. Most AI chatbots today default to easy answers, which is hurting our ability to think critically.

I call this the Information-Exploration Paradox. As the cost of information approaches zero, human exploration collapses. We see it in students who perform better on AI-assisted tasks and worse on everything afterward. We see it in developers shipping more code and understanding it less. We are, in ways that feel like progress, slowly optimizing ourselves out of the loop.

This is the divergence I worry about. Not the dramatic science-fiction scenario of AI replacing humans wholesale, but the quieter process of people gradually outsourcing their judgment in increments too small to notice.

Over time, this produces two kinds of people: Those who use AI as a genuine intellectual partner—whose thinking actually gets sharper through the friction of the collaboration—and those who get better at securing quick answers and worse at knowing what questions to ask.

So what can any of us actually do about it?

Start with the reframe: The goal of working with AI isn't to get the answer faster. It's to find out what you're missing. Don't deploy AI minions to "do the boring work" for you, as so many sales pitches argue; use it as a savant collaborator to explore uncertainty.

In practice, that means before you accept an AI's answer, ask it for the strongest argument against itself. When it hedges or qualifies, pay attention—that's usually where the real uncertainty lives. Treat it like a brilliant colleague who has read everything and understands nothing—useful precisely because ~~they're different from you, not~~ because they'll agree with you.

For the AI industry, a key design question has gone largely unasked: Is the product building human capacity or consuming it? Nearly all AI benchmarks measure what AI agents can do alone. We desperately need benchmarks for hybrid intelligence. Errors are signals our brains use to trigger learning. An AI that eliminates friction entirely is often eliminating the learning along with it.