“Hey AI, Can you take this next consult?”

What can Generative AI do for heme/onc fellows in March 2026?

…and what should you let it do?

Jonathan Berry, MD | BIDMC / Harvard Medical School
Clinical Informatics Lead, BIDMC Hematology
Associate Program Director, BIDMC Internal Medicine Residency
Tech Enthusiast

Disclosures

No relevant financial relationships to disclose.

Note: GenAI tools (Gemini, NotebookLM, Claude, and Claude Code) were used for generation of teaching materials, including outlines, slides, web activities, and graphics.

Don't worry – I'll show you how.

I have reviewed and vouch for all information presented today.

Additional Note: Institutional guidelines and policies are changing rapidly in this space. I have done my best to understand the landscape at BIDMC, HMS, MGB, and DFCI, but policies around various use cases can and will continue to change

Framing & Orientation

What is AI, why it matters for heme/onc, and some ground rules

Who's in the Room

Pre-Survey · n = 21 Heme-Onc Fellows

GenAI Usage Level

2

Non-users

10% of group

15

Casual Users

71% of group

4

Power Users

19% of group

LLM Knowledge Level

5% No understanding

38% Basic

43% Intermediate

14% Advanced

Bias Familiarity

48% Not at all familiar

47% Moderately familiar

5% Extremely familiar

How You're Currently Using AI

TOOLS

- **General Foundation LLMs** — 65%
  ChatGPT · Claude · Gemini
- **Medical-Focus LLMs** — 65%
  Open Evidence · UpToDate AI
- **Image / Slide Generation** — 30%
  ChatGPT Images · Gamma
- **Ambient AI Scribes** — 15%
  HEIDI · Abridge · DAX
- **Epic-Integrated Tools** — 15%
  Patient Insights · ART

TOP TASKS

Clinical

- Evidence & data lookup
- Targeted drug / dosing Q&A
- Clinical reasoning support
- Documentation & appeal letters

Administrative

- Drafting & editing emails
- Letters of Recommendation

Research

- Literature review — by a wide margin

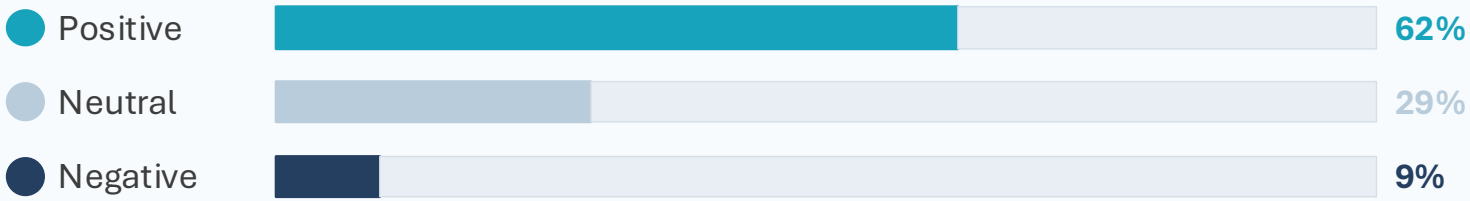
Attitudes: Hopeful but Hedged

PATIENT CARE

62%

expect positive impact

29% uncertain · only 9% expect harm

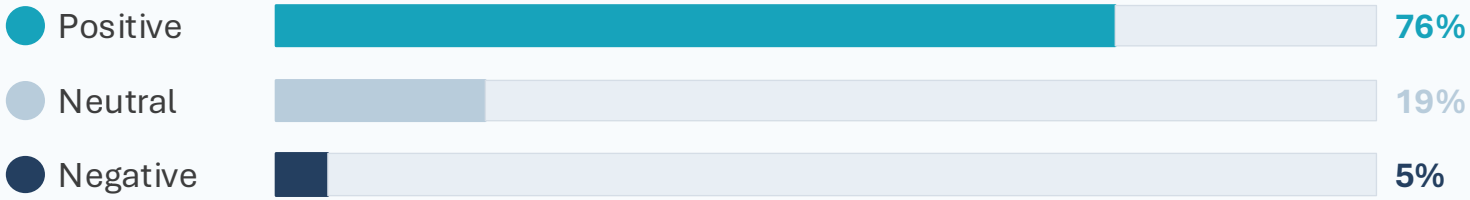


WORKFLOW EFFICIENCY

76%

expect time savings

The strongest consensus

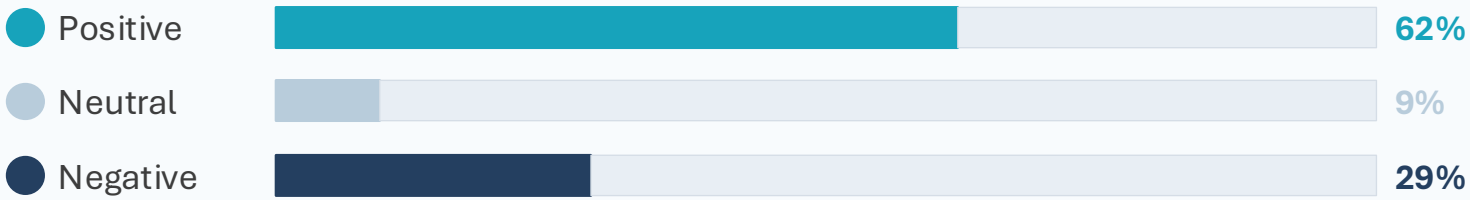


REASONING SKILLS

62% vs 29%

hopeful · worried

Defining tension: gain vs. erosion



What's Holding You Back

Top barriers reported by fellows

"I don't know enough to use it well"

Most respondents aren't aware of what's even possible beyond basic prompting

HIPAA & institutional policy uncertainty

"Not sure what's legal or what my hospital allows" — a near-universal hesitation

Accuracy & hallucination concerns

Distrust of clinical outputs without a clear framework for verification

Fear of skill erosion

"I don't want to stop knowing how to research things myself"

Practical friction

Epic integration gaps, clunky ambient scribe setup, no time to learn

What You're Here For

What fellows asked to get out of today

What to do — and what NOT to do

Clear guardrails: when to trust AI and when to push back

Practical, implementable clinical examples

Real workflows you can walk out and use this week in clinic

Understanding limitations & trust calibration

How to verify outputs and build appropriate skepticism

Epic, documentation & education integration

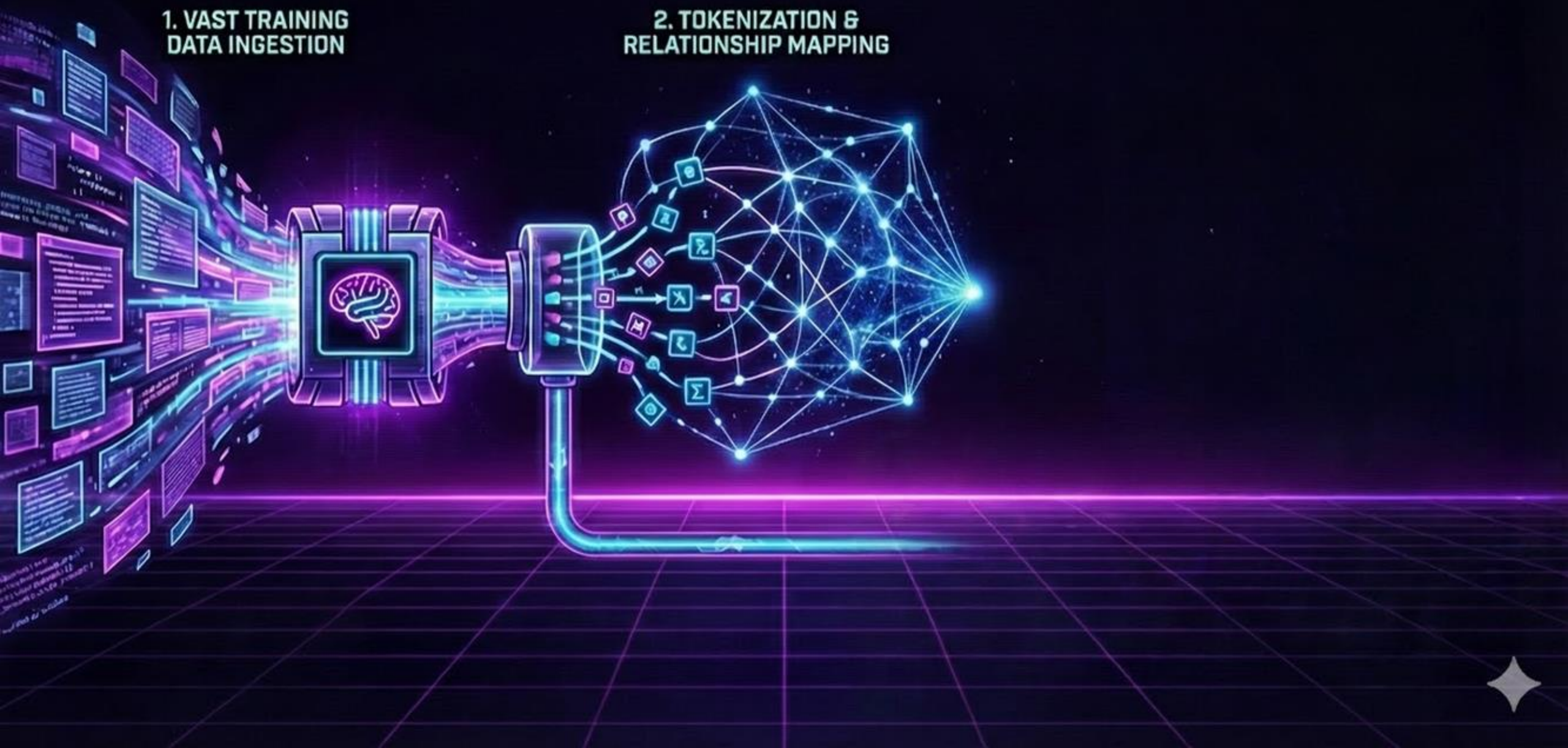
AI in the chart, the scribe, and the teaching session

Beyond ChatGPT — what else is out there?

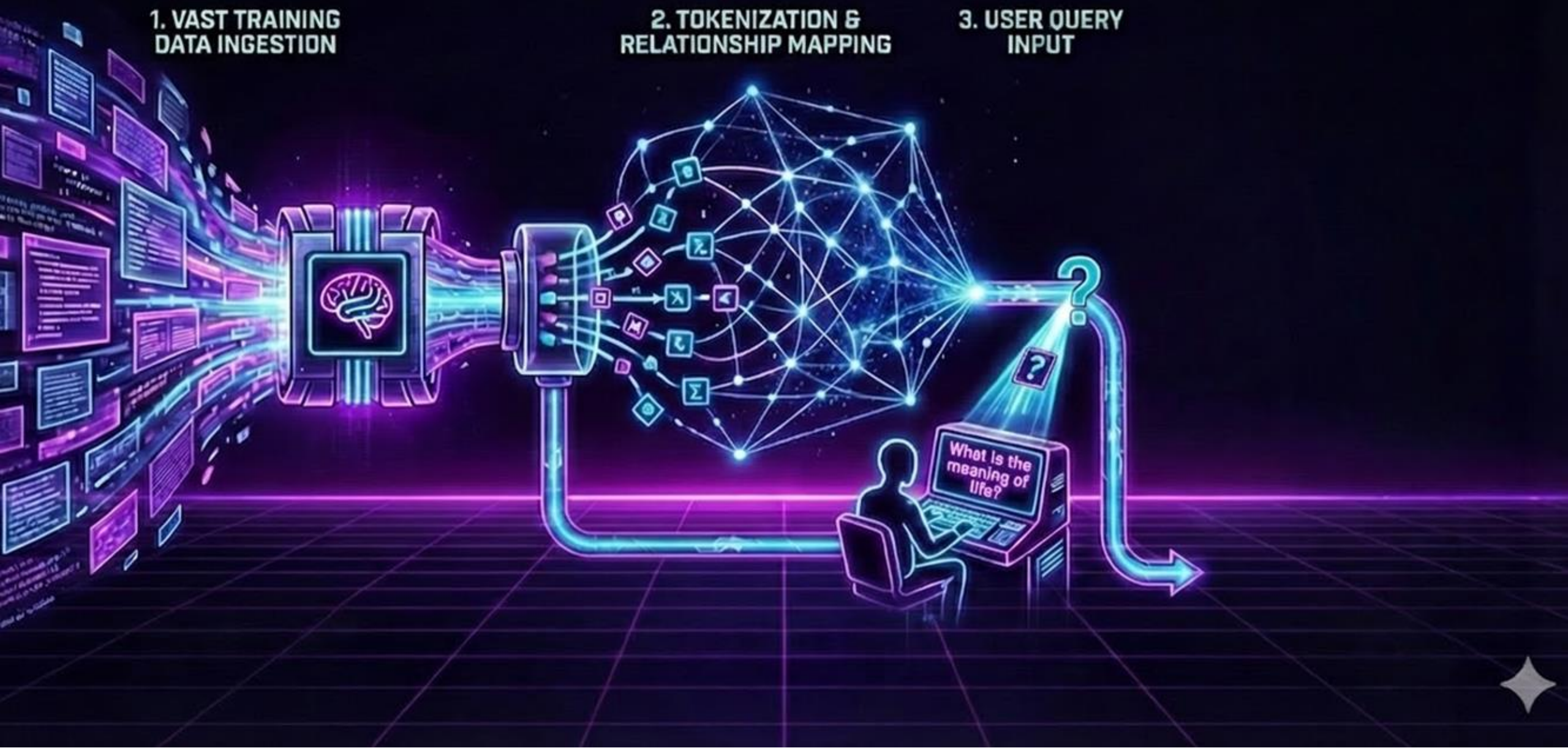
Curated tools and how to make them work in your style

LLMs & GenAI

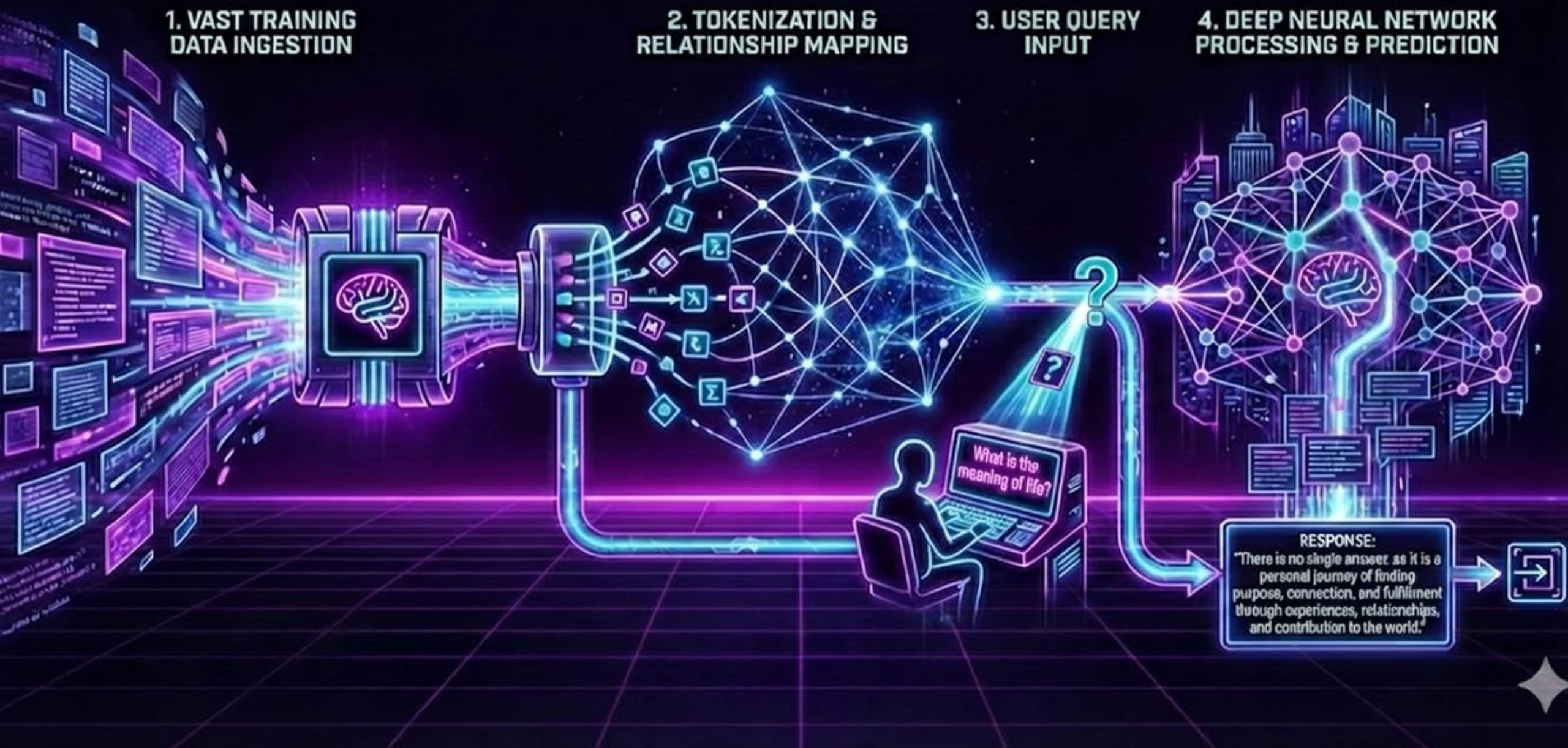
LLMs & GenAI


1. VAST TRAINING
DATA INGESTION
2. TOKENIZATION &
RELATIONSHIP MAPPING

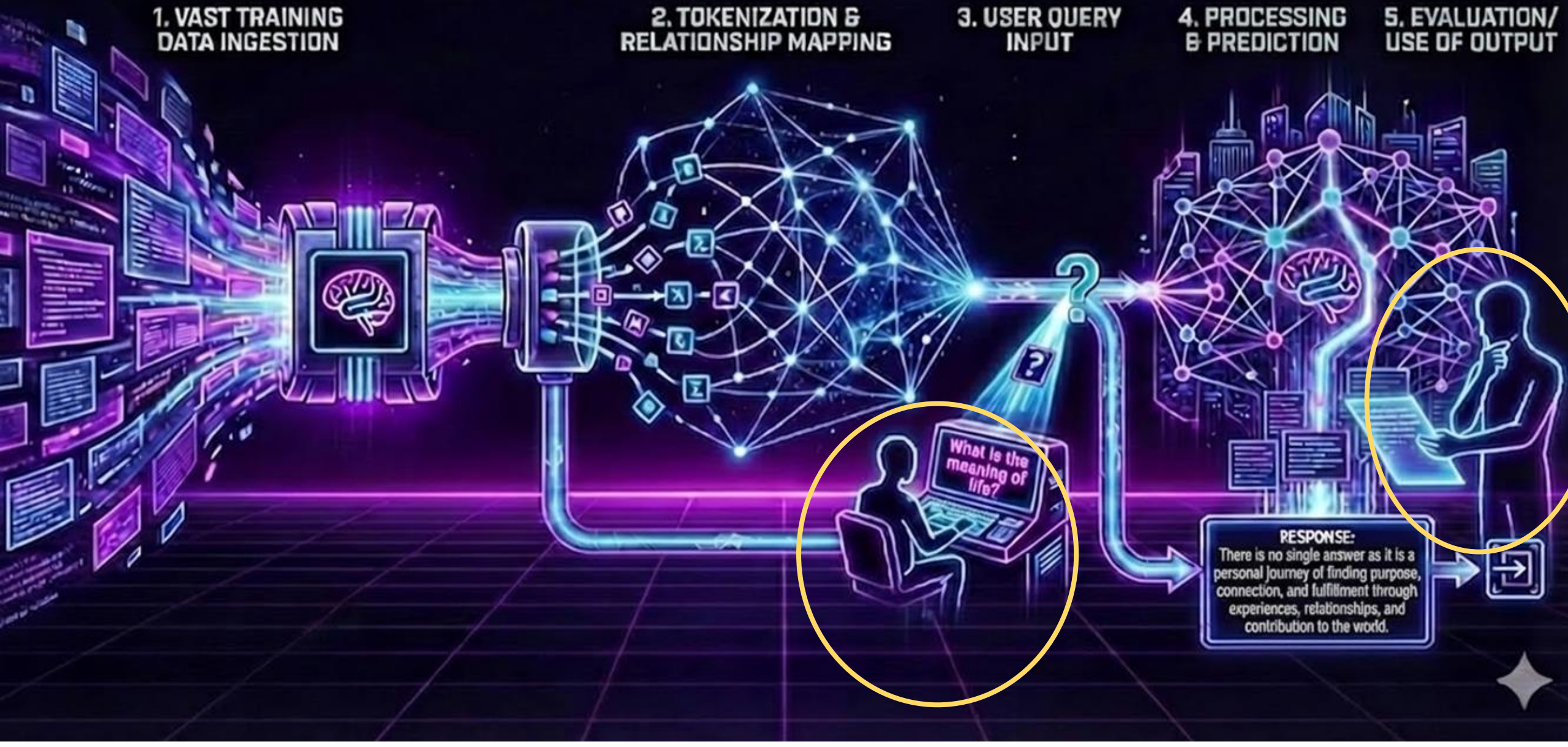
LLMs & GenAI


1. VAST TRAINING DATA INGESTION
2. TOKENIZATION & RELATIONSHIP MAPPING
3. USER QUERY INPUT
What is the meaning of life?

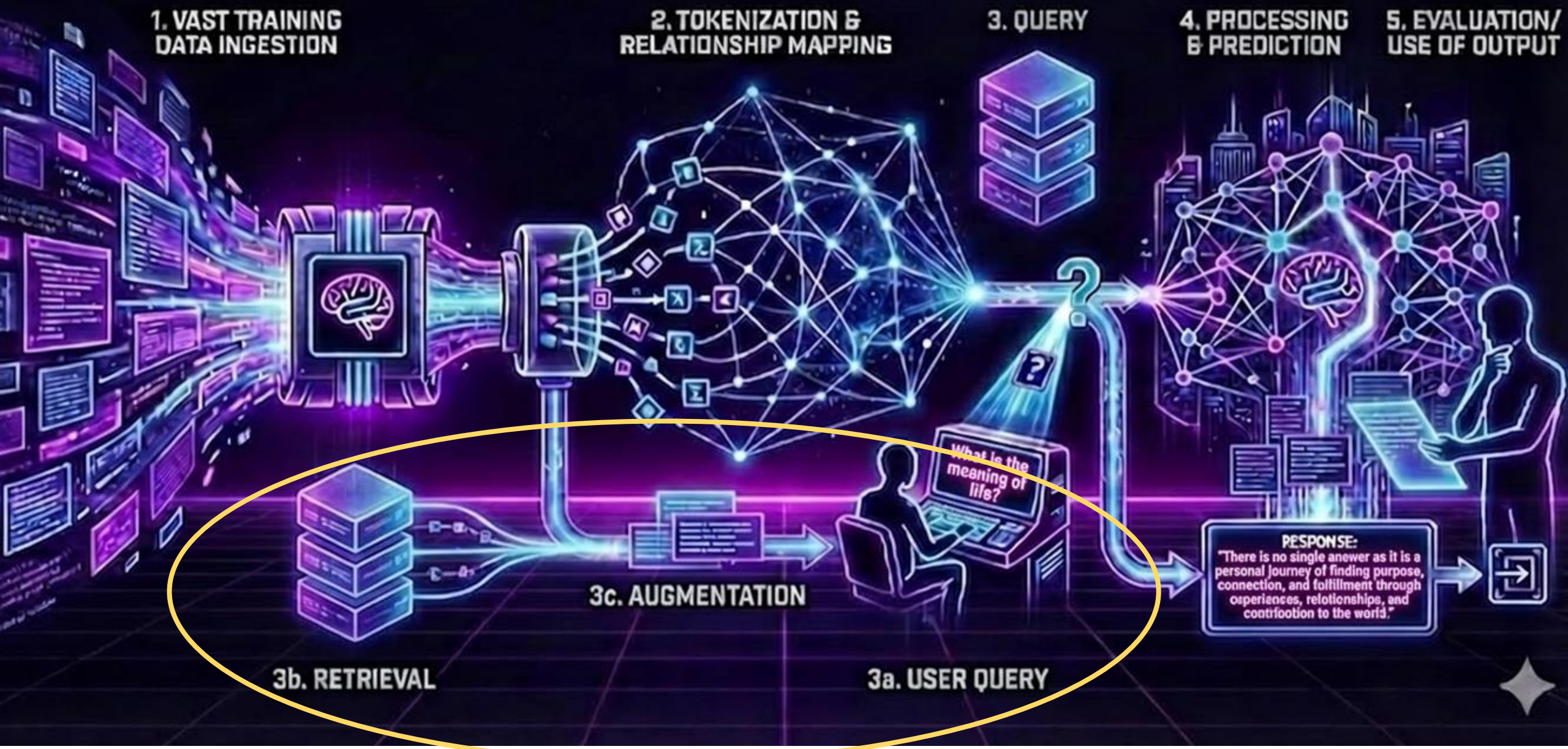
LLMs & GenAI


1. VAST TRAINING DATA INGESTION
2. TOKENIZATION & RELATIONSHIP MAPPING
3. USER QUERY INPUT
4. DEEP NEURAL NETWORK PROCESSING & PREDICTION
What is the meaning of life?
RESPONSE:
"There is no single answer, as it is a personal journey of finding purpose, connection, and fulfillment through experiences, relationships, and contribution to the world."

LLMs & GenAI


1. VAST TRAINING DATA INGESTION
2. TOKENIZATION & RELATIONSHIP MAPPING
3. USER QUERY INPUT
4. PROCESSING & PREDICTION
5. EVALUATION/ USE OF OUTPUT
What is the meaning of life?
RESPONSE:
There is no single answer as it is a personal journey of finding purpose, connection, and fulfillment through experiences, relationships, and contribution to the world.

Retrieval Augmented Generation



Platforms / Foundation Models

ChatGPT

Gemini

Claude

AI Assistants (Chatbots)

ChatGPT → Microsoft 365 Copilot Chat

Gemini

Claude

RAGs

OpenEvidence

perplexity

consensus

ChatPPGD
AI powered chat interface for BIDMC PPGD

AI Tools

Heidi

Nuance® DAX™ Copilot

ABRIDGE

Gamma

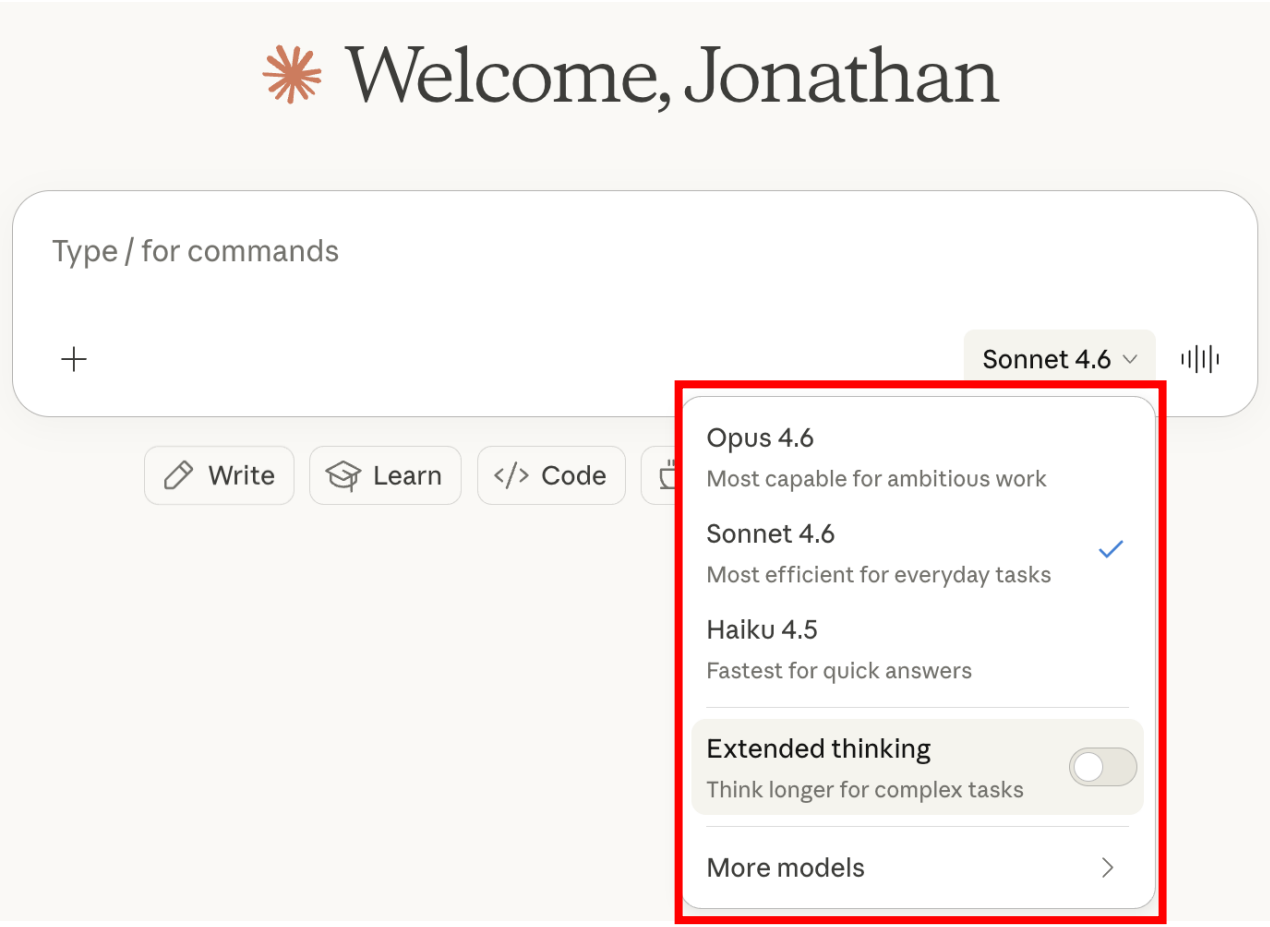
NotebookLM

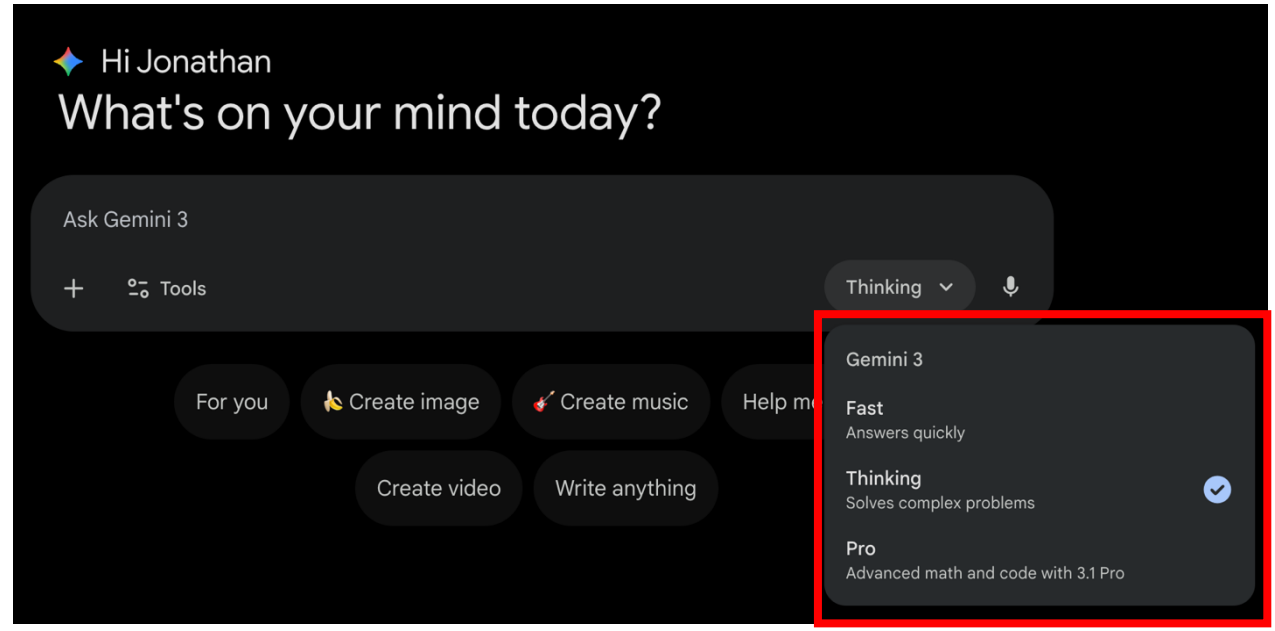
Agentic AI

Claude Code v2.0.24
Sonnet 4.5 · Claude

CLAUDE COWORK

perplexity


Welcome, Jonathan
Type / for commands
Sonnet 4.6
Write
Learn
Code
Opus 4.6
Most capable for ambitious work
Sonnet 4.6
Most efficient for everyday tasks
Haiku 4.5
Fastest for quick answers
Extended thinking
Think longer for complex tasks
More models


Hi Jonathan
What's on your mind today?
Ask Gemini 3
Tools
Thinking
For you
Create image
Create music
Create video
Write anything
Gemini 3
Fast
Answers quickly
Thinking
Solves complex problems
Pro
Advanced math and code with 3.1 Pro

Why AI Is Particularly Relevant to Heme/Onc

Massive, nuanced data landscape

Molecular markers, cytogenetics, prognostic scoring — more complexity than almost any other specialty

Complex, highly-invested patients

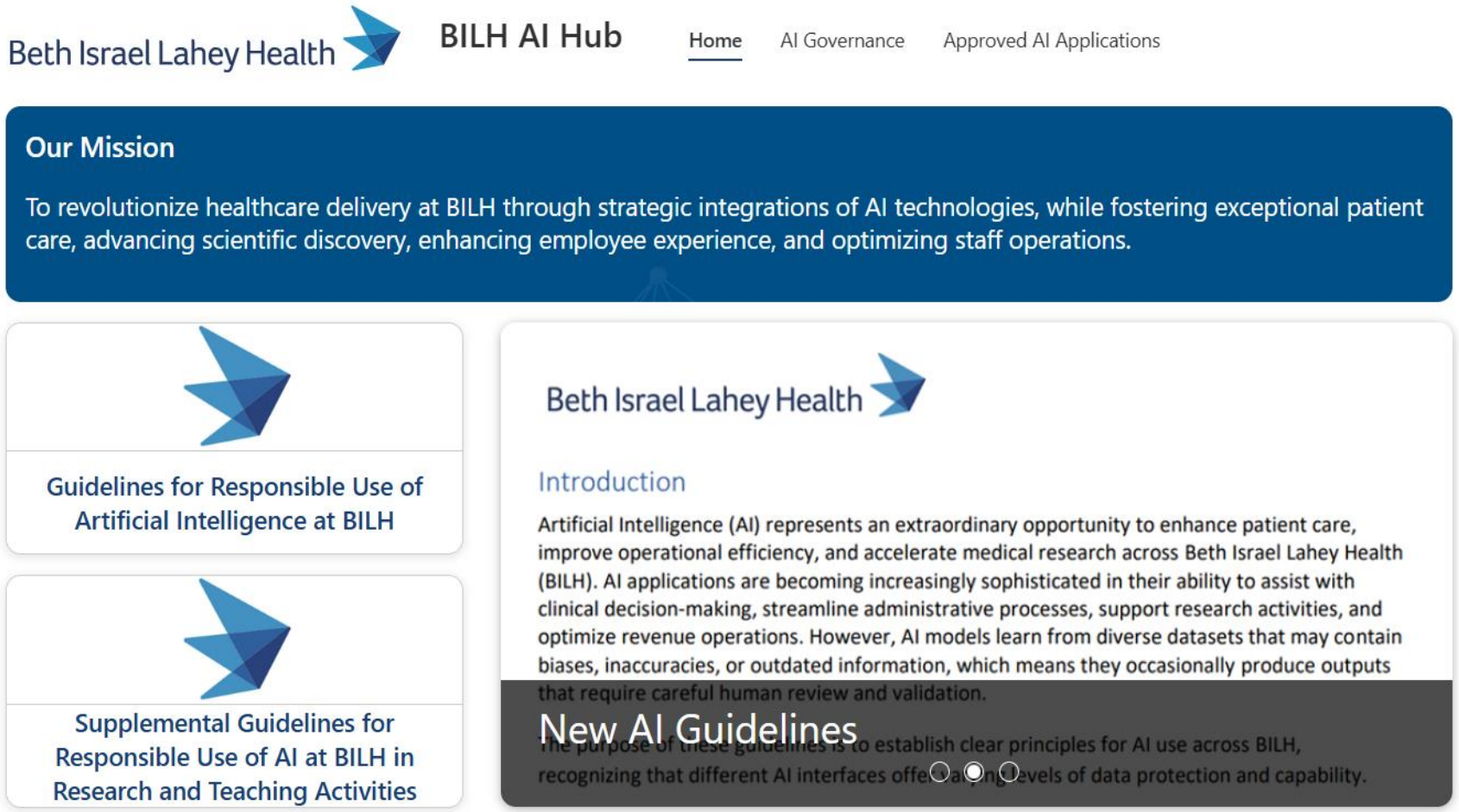
High patient education demand; patients actively researching their own disease and treatment options

High-stakes drug regimens

Prior auths, appeals, toxicity education, monitoring — all time-consuming and highly AI-amenable tasks

HIPAA, Ethics & Hospital Policy: PHI

⚠ Hospital policies vary and will change



BILH HIPAA + AI De-Identification

HIPAA, Ethics & Hospital Policy: PHI

⚠ Hospital policies vary and will change

1. No PHI in free-text prompts to consumer-grade AI tools
2. PHI only in hospital-approved, HIPAA-compliant tools
3. At BIDMC: Copilot Chat, HEIDI, and Epic-integrated tools are approved for PHI
4. De-identified cases: fair game — but be thoughtful about what "de-identified" truly means

AI & Research: Disclosure Framework

ICMJE Consensus Framework (most journals follow this)

✗ AI cannot be listed as an author

✓ Total human accountability for all content — the human author is responsible

✓ Mandatory disclosure in Acknowledgments and/or Methods section

⚠ Check the specific journal — many have their own AI policy requirements beyond ICMJE

AI Disclosure: Clinical & Professional Use

With Patients

- Audio/video recording (e.g. Ambient scribes) requires patient disclosure
- Other AI use: no current legal mandate, but be prepared to discuss
- Never deny AI use if a patient asks

With Colleagues

- Educational materials: disclose AI assistance
- Policy documents: disclose
- Professional communications: use your judgment
- Institutional policy is actively evolving

Hands-On: Diagnostic Reasoning

Let's think about some stuff

1 Case Scenario #1

55-year-old with remote Hodgkin lymphoma (1980s) — s/p chemo + XRT

Presents with 6 months of: weight loss (15 lb), dry cough, exertional dyspnea
CXR: widened mediastinum, large anterior mediastinal mass

Key labs: Hb 10.5 | LDH 320 | Ca 10.3 | SpO_2 94% RA

tinyurl.com/55e4ubt8

The task: Generate a differential diagnosis with your AI tool of choice

tinyurl.com/mr2vjzz7
Landing Page

2 Case Scenario #2

32-year-old G2P1 at 28 weeks gestation — previously uncomplicated pregnancy
Persistent frontal headache, fatigue, nausea × 5 days

BP 148/94 (baseline 110/70), RUQ discomfort, dark urine

Key labs: Plt 62k | LDH 800 | Hapto <10 | Hb 8.6 | AST 150 | TBili 3.4

Smear: rare schistocytes (1–2/HPF)

tinyurl.com/33ubzvxy

The task: Try at least 2 different prompting approaches to try and understand what is most likely going on with this patient and what tests you may want to perform. Compare output with your neighbor.

tinyurl.com/mr2vjzz7
Landing Page

Prompting Approaches

1 **Model Choice**

Fast vs. Thinking vs. Pro — pick based on task complexity

2 **Zero-Shot**

Just ask the question cold — the baseline approach

3 **Chain-of-Thought**

"Think through this step by step before answering"

4 **Assign Expertise**

"You are an expert hematologist. Your patient..."

5 **Give Flexibility**

"Before answering, ask me any clarifying questions"

6 **Iterative Disagreement**

"Now argue against your recommendation and explain why it might be wrong"

⚠ Pitfall — Priming Bias: lead with a diagnosis → model anchors to it even when better alternatives exist

We will reconvene at *

tinyurl.com/33ubzvxy

Debrief — What Did You Find?

Next: understanding WHY it sometimes gets things wrong…

Limitations: Bias & Accuracy

Why AI outputs can be wrong — and what to do about it

Context Drift

Models re-read the entire conversation with every prompt

Practical tip: Start new chats often — especially when topics shift. You lose some context but gain better attention.

Context Windows (in tokens)

200k

Free: 16k
Plus: 32-256k

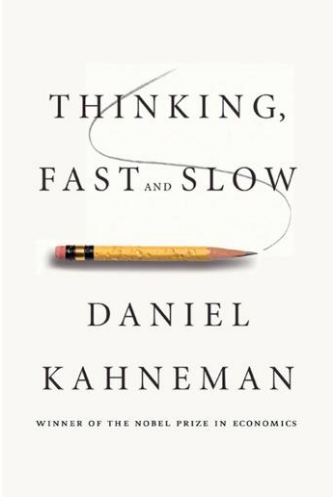
1 million

Claude

ChatGPT

Gemini

200k tokens ~ 150k words ~ a 500 page book



Training Data Cutoffs

Your AI doesn't know what it hasn't read

General chatbots:

ChatGPT / Claude / Gemini all have optional web search — but check if yours is enabled

BILH Copilot Chat:

No web search (intentional — keeps PHI contained within the enterprise environment)

If your AI can't look for the latest information, all it knows is what is in its training data!

Hallucination: Confidently Wrong

Definition: Model generates plausible-sounding but fabricated content

Good news: Much better than 2022 — "thinking" and reasoning models significantly reduce hallucination rates

⚠ **Bad news: Hallucination rate is not (and likely never will be) 0**

Mitigation: You must read output before using in any capacity. If you're using AI for information, ask for citations — then actually follow the links.

Bias: Human Bias Is AI Bias

Training Data	Label Bias	Outcome Bias
Underrepresentation of certain populations in the medical literature and trial data	Who gets labeled as 'high risk' or 'good candidate for treatment' in the data the model learned from	Who receives aggressive treatment in the literature — and whose outcomes get studied

Limiting Errors & Mitigating Bias

1. Ask for citations — and follow them
2. Start fresh chats often
3. Know your own expertise (and verify more in areas you don't)
4. Ask about populations: "Are there groups for whom this might differ?"
5. Know your tool's web access — and whether it has current data

Hands-On: Text Generation Workflows

Let’s make some stuff

A Option A: Patient Message Response

"I read that venetoclax + azacitidine is not as good for younger patients. Should I be getting intensive chemo instead?" — Patient with newly diagnosed AML

Tasks:

- Draft a patient-facing response using AI
- Then: "Rewrite this at a 6th grade reading level"
- Adjust tone, remove jargon, check for accuracy

💡 Better yet — bring a REAL patient question from your own inbox. No PHI needed in the prompt.

B Option B: Prior Auth Appeal Letter

Patient needs VIPOR for relapsed/refractory DLBCL. Insurer denied citing "insufficient evidence for clinical benefit."

Tasks:

- Use AI to draft an appeal letter — add your own de-identified clinical details
- Ask it to include relevant trial data and why it makes sense for your patient

💡 No PHI needed — describe the clinical scenario in general terms. AI does the heavy lifting on structure and citations.

C Option C: Patient Education Handout

New consult: patient with recently diagnosed CLL, no indication for treatment yet. Needs to understand their diagnosis.

Tasks:

- Generate a one-page patient handout on CLL — written at 8th grade level
- Then: ask the model to critique it for medical accuracy
- Can you turn it into an infographic? (Hint, use NotebookLM or Gemini)

💡 Ask the model to check itself for accuracy and flag anything uncertain — a useful self-audit step.

D Option D: Teaching Case Generator

You're leading a small group on oncologic emergencies for IM residents. Build a structured patient case that you can use for teaching.

Tasks:

- Choose: malignant bowel obstruction, SVC syndrome, or TLS
- Generate a realistic patient case with staged disclosure (H&P → labs → imaging)
- Then separately: generate 3–5 discussion questions for resident teaching
- Compare: having it do all at once vs. in stages — which is better?

💡 Try generating the case and the teaching questions in separate chats — you often get better results each way.

Choose Your Own Adventure — Pick One or Two

A: Patient Message

Draft a response to a patient question about their AML treatment options

B: Prior Auth Appeal

Write an insurance appeal letter for a denied DLBCL treatment

C: Patient Education

Generate a CLL patient education handout — then critique it

D: Teaching Case

Build a teaching case on oncologic emergencies for residents

Better yet — bring a REAL task from your own work!

tinyurl.com/mr2vjzz7
Landing Page

We will reconvene at *

tinyurl.com/4s29urxu
Straight to activity

Debrief — What Did You Create?

1. What worked well? What didn't?

2. Where did it save you time vs. create extra work?

3. How much editing was needed to make it clinically usable?

Governance, Liability & the Bigger Picture

Who's responsible, and what does AI mean for your training?

Governance & Liability

Who is responsible when AI gives bad advice that harms a patient?

- No clear medicolegal precedent yet — but the standard remains physician accountability

What Does AI Mean for Your Training?

Deskilling is a real concern — be intentional about what you outsource

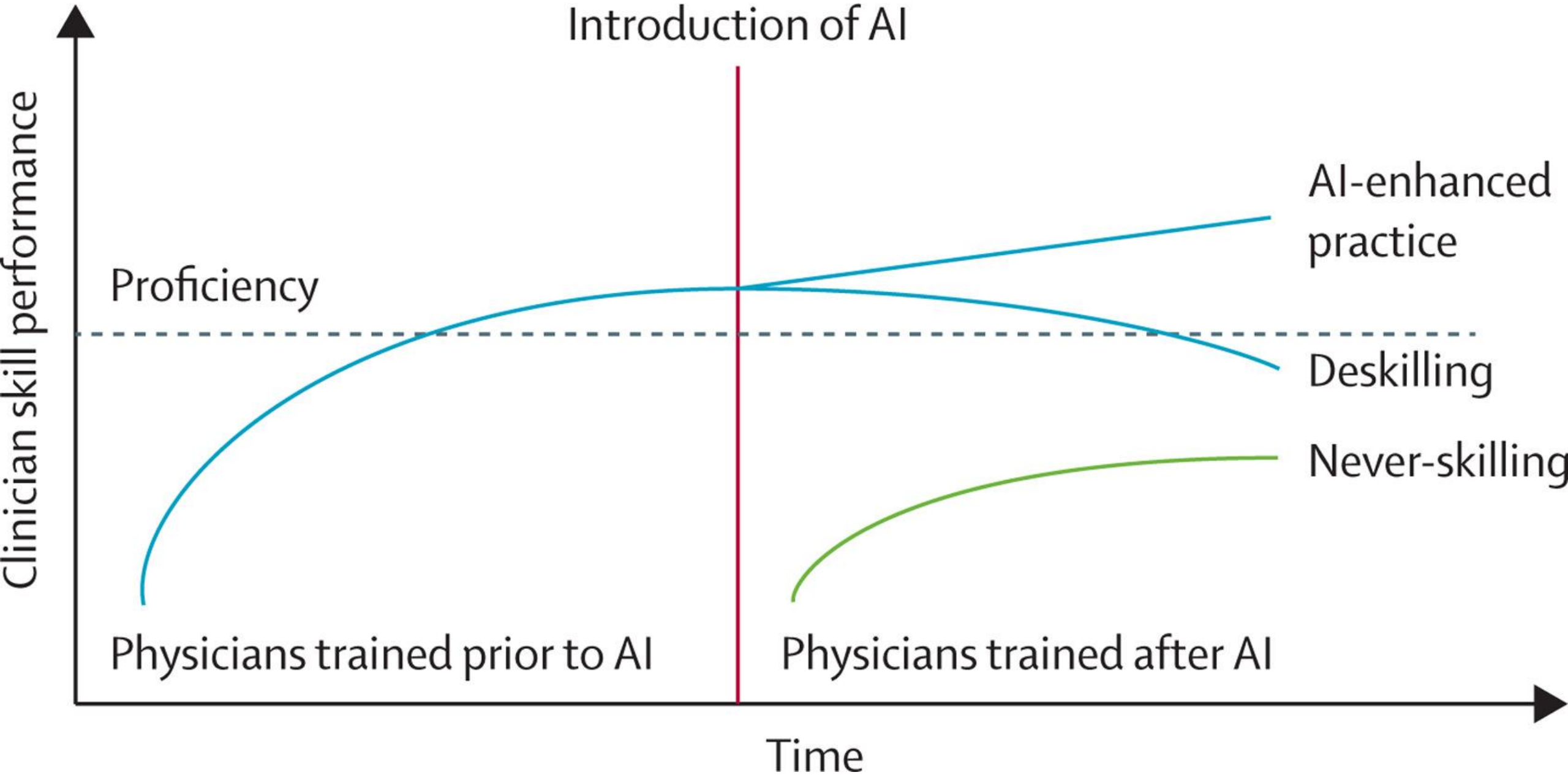
ARTICLES · Volume 10, Issue 10, P896-903, October 2025 Download Full Issue

Endoscopist deskilling risk after exposure to artificial intelligence in colonoscopy: a multicentre, observational study

Krzysztof Budzyń, MD [a,b] · Marcin Romańczyk, MD [a,b] · Diana Kitala, PhD [c] · Paweł Kołodziej, MD [d] · Marek Bugajski, MD [e] · Hans O Adami, MD [f,g] · et al. Show more

Affiliations & Notes · Article Info · Linked Articles (7)

Get Access · Cite · Share · Set Alert · Get Rights · Reprints



Preserving clinical skills in the age of AI assistance; Berzin, Tyler M et al.; The Lancet, Volume 406, Issue 10513, 1719

Computers in Human Behavior 175 (2026) 108779

Contents lists available at ScienceDirect

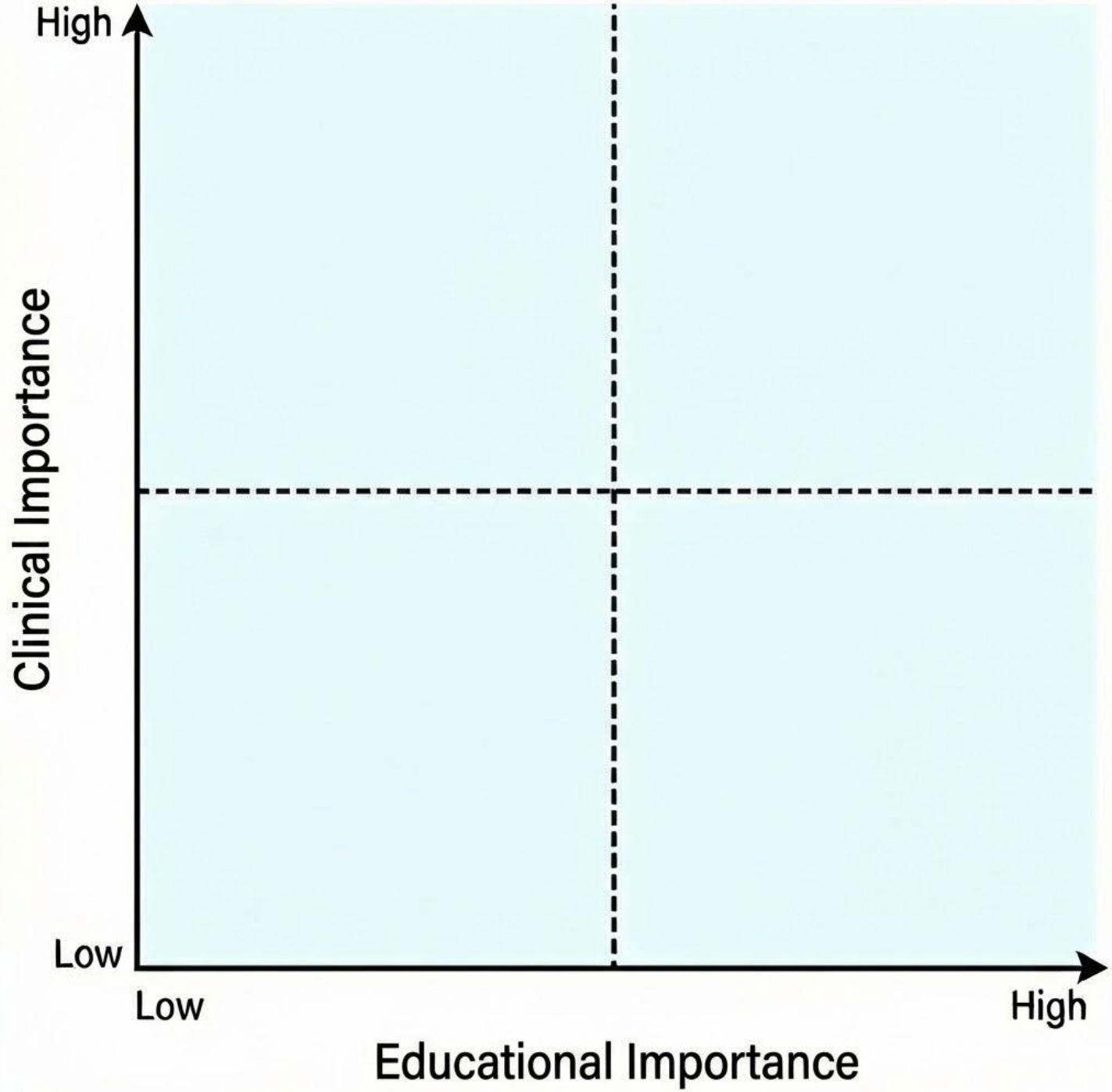
Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

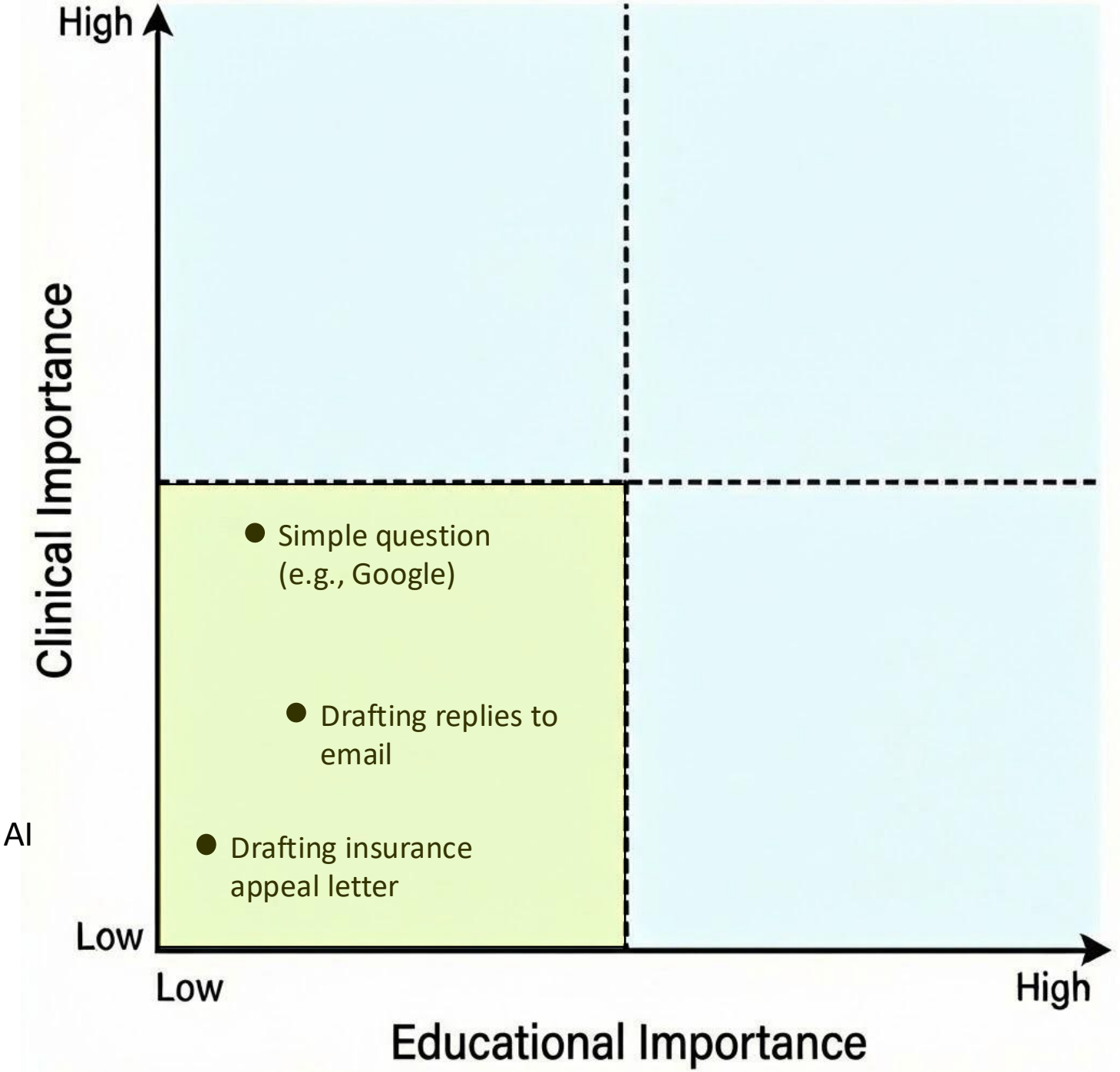
Full length article

AI makes you smarter but none the wiser: The disconnect between performance and metacognition

Daniela Fernandes [a,*], Steeven Villa [b], Salla Nicholls [a], Otso Haavisto [a], Daniel Buschek [c], Albrecht Schmidt [b], Thomas Kosch [d], Chenxinran Shen [e], Robin Welsch [a]

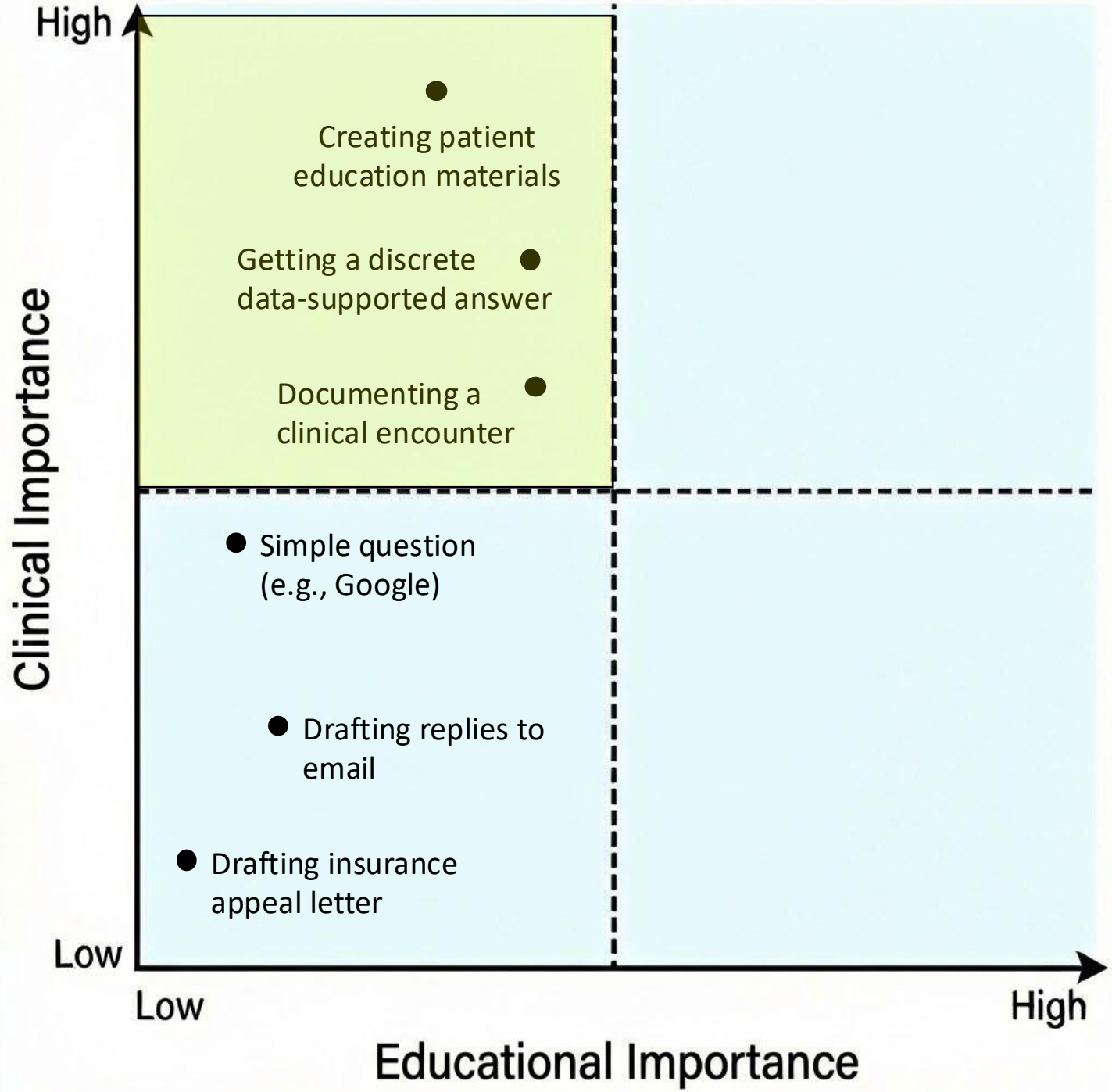

High
Clinical Importance
Low
Low
High
Educational Importance

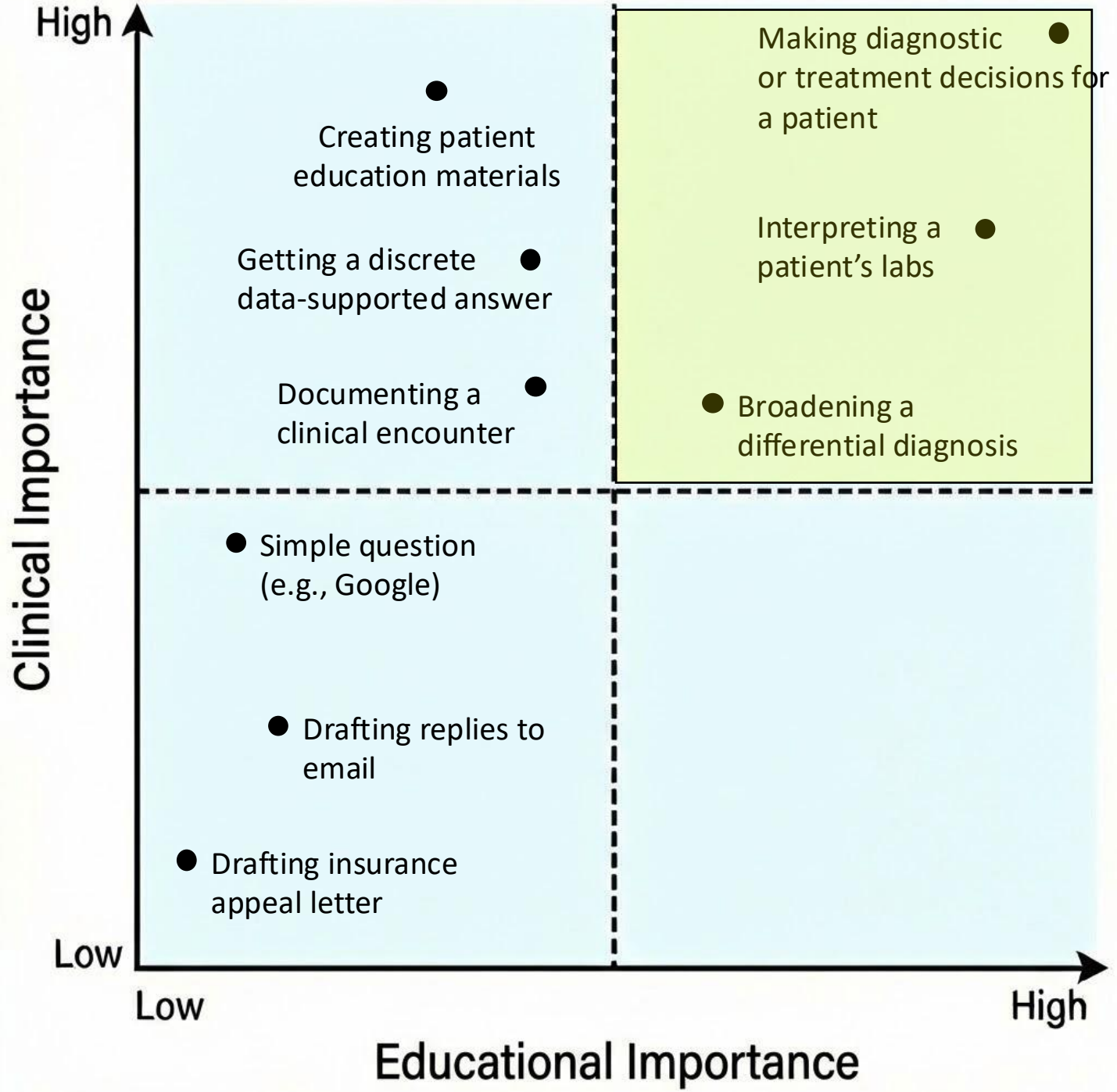
Low/Low = Administrative
Great place to deploy AI



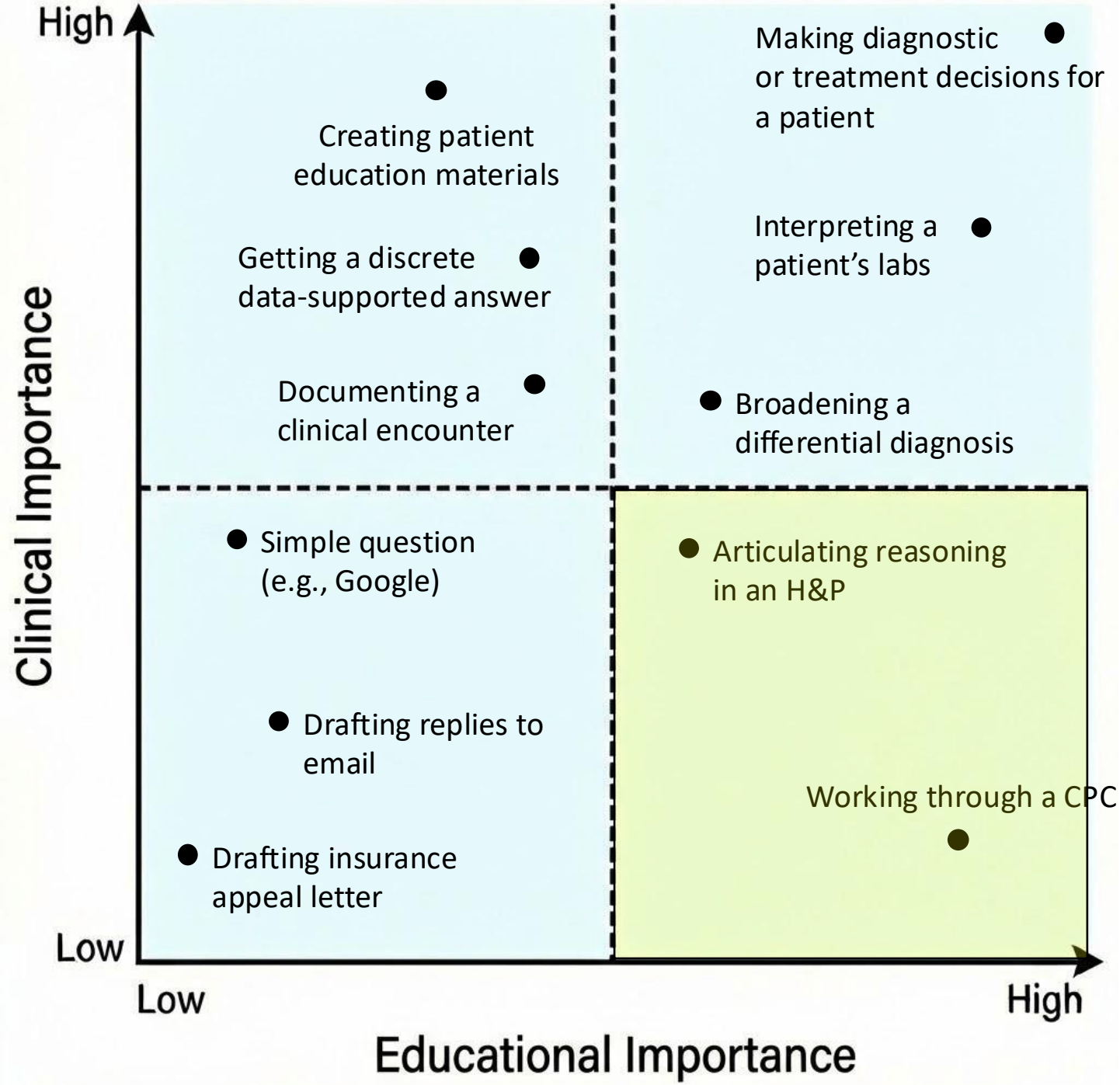
High/Low = Verify

Ok to use AI here, but important to understand its limitations and verify sources/output





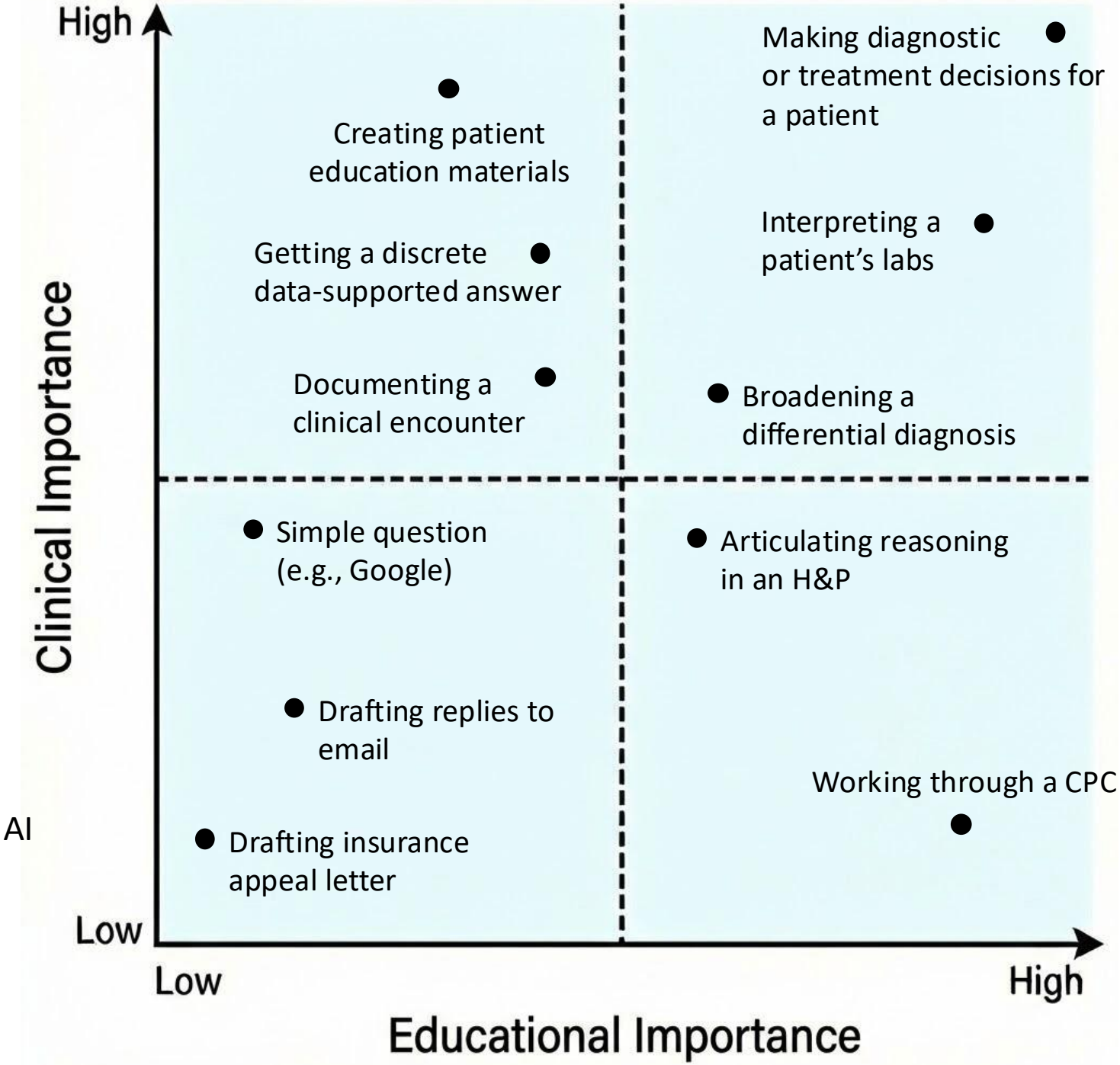
High/High = Education + Clinical
Tasks are highly educational and important for accuracy. Requires careful use and careful verification



Low/High = Pure Education

These tasks are primarily educational; there is really no role to outsource them

High/Low =
Verify
Ok to use AI here, but important to understand its limitations and verify sources/output



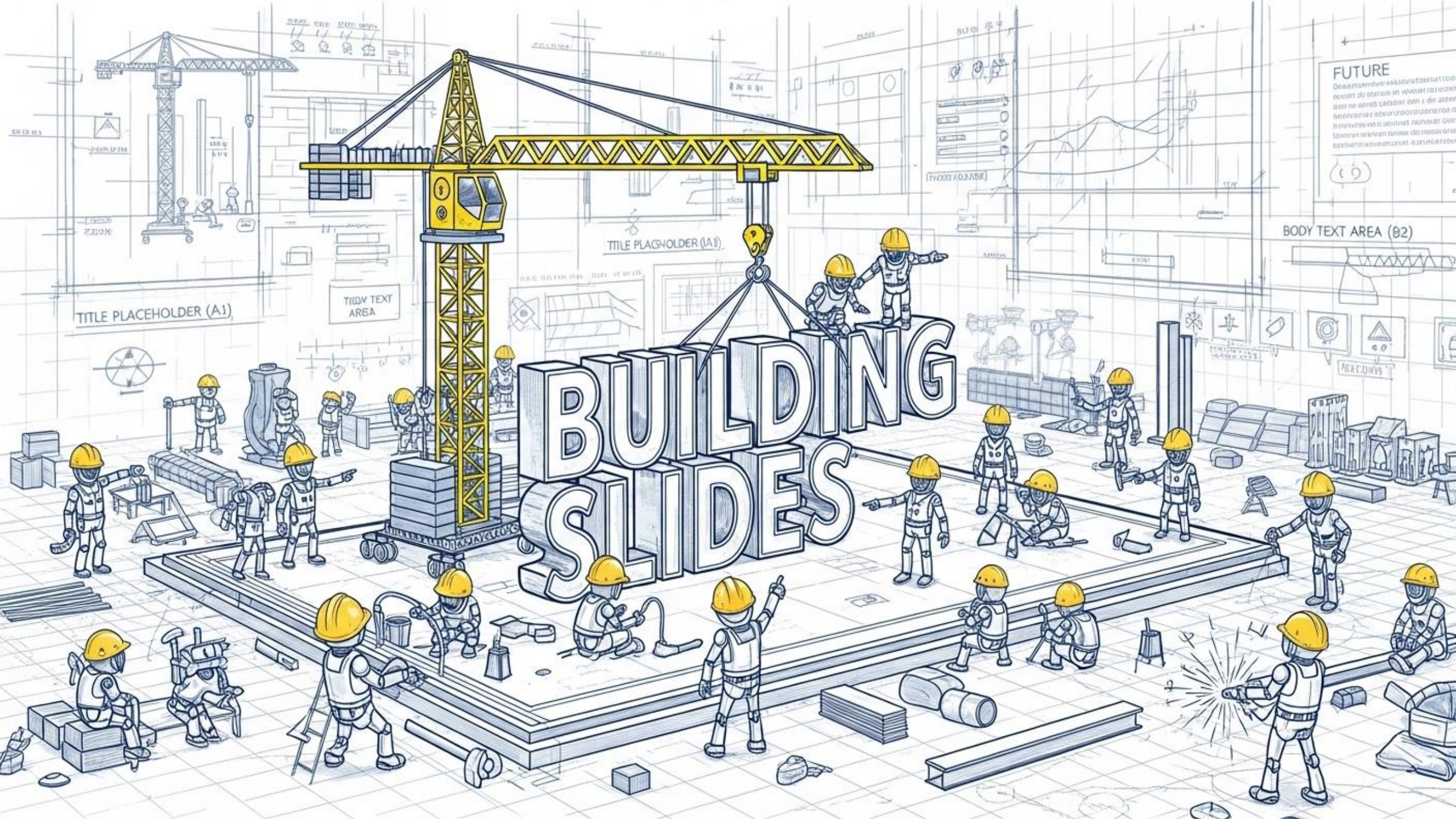
High/High =
Education + Clinical
Tasks are highly educational and important for accuracy. Requires careful use and careful verification

Low/Low =
Administrative
Great place to deploy AI

Low/High =
Pure Education
These tasks are primarily educational; there is really no role to outsource them

Beyond the Clinic

A few non-clinical use cases to consider


BUILDING SLIDES
TITLE PLACEHOLDER (A1)
BODY TEXT AREA
FUTURE

Less slog; more conversation

tinyurl.com/5hxur6c2


NotebookLM

Cognitive Load in Medical Education

+ Create notebook | Share | Settings

Sources

+ Add sources

Search the web for new sources

Web | Fast Research

Select all sources

- 2410.03703v2.pdf
- 9 ways to reduce cognitive load in ...
- AI & cognitive load.pdf
- Ambient AI, cognitive load.pdf
- Brit J Educational Tech - 2024 - Fa...
- Graphic Design for Clinical Teacher...
- Managing the Load: AI and Cogniti...
- Medical Education - 2009 - Van M...
- Pharmacology multimedia design.p...

Chat

- **Promote Schema Building:** Use structural aids like flowcharts, diagrams, and metaphors to help learners connect new information to existing knowledge 36 .
- **Include Active Prompts:** Integrate **self-explanation prompts** or case studies at the end of sections to force learners to actively process and apply information 37 ... .
- **Employ Fading Guidance:** Start with fully worked examples, move to completion tasks with partial solutions, and end with unassisted problems to transition the learner toward expertise 29 ... .
- **Act as a "Coach," not a Repository:** Design content that guides users through structured reasoning frameworks (e.g., listing associated words or strategies) rather than providing direct answers to prevent **metacognitive laziness** and cognitive offloading 42 ... .

Save to note

How can AI act as a coach for convergent thinking?

What are the risks of AI causing homogenization of ideas?

How does the modality effect help prevent cognitive overload?

Start typing... 13 sources

NotebookLM can be inaccurate; please double check its responses.

Studio

Audio Overview | Video Overview

Mind Map | Reports

Flashcards | Quiz

Infographic BETA | Slide Deck BETA

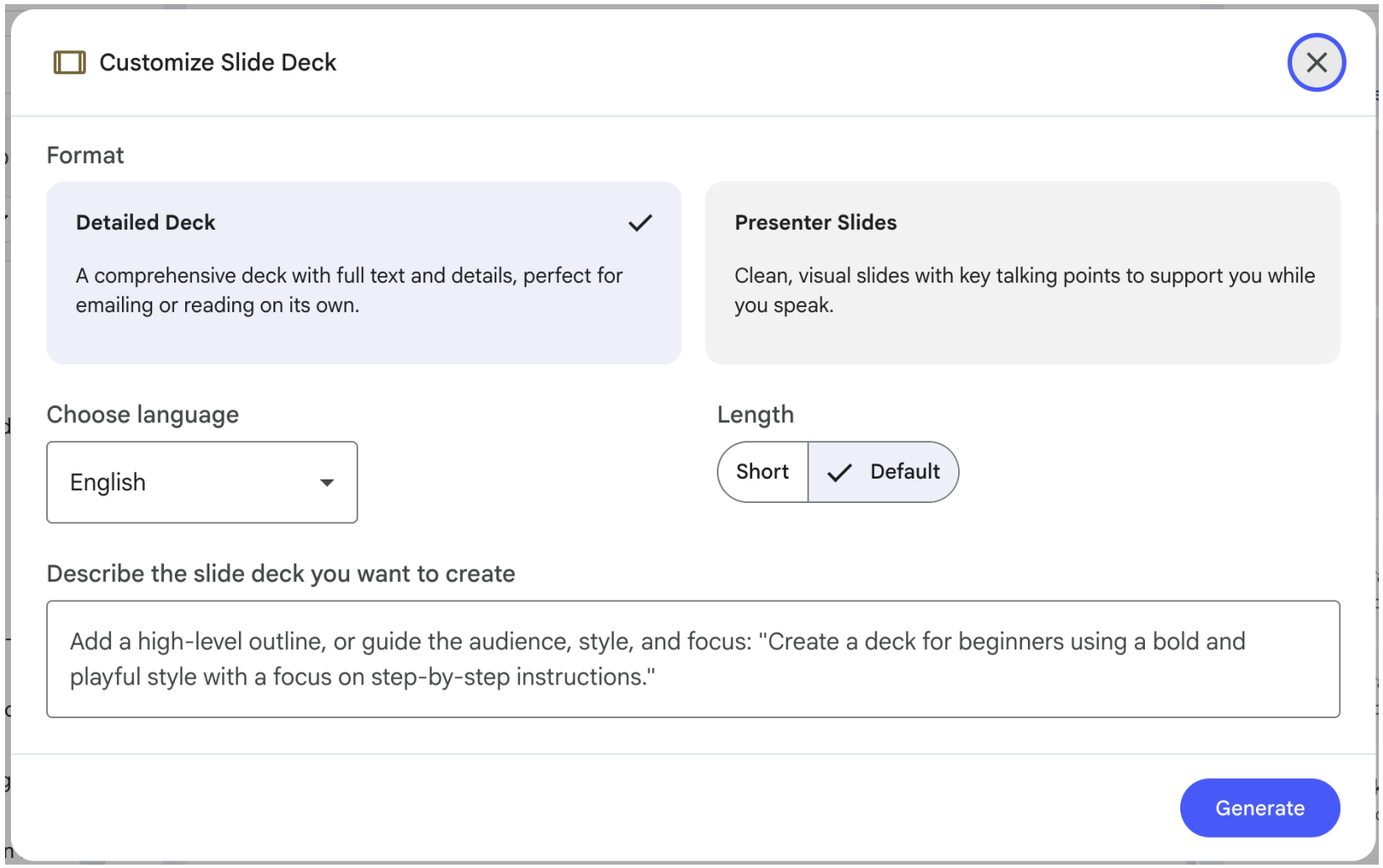
Data Table

Generating Slide Deck...
based on 13 sources

Generating Slide Deck...
based on 13 sources

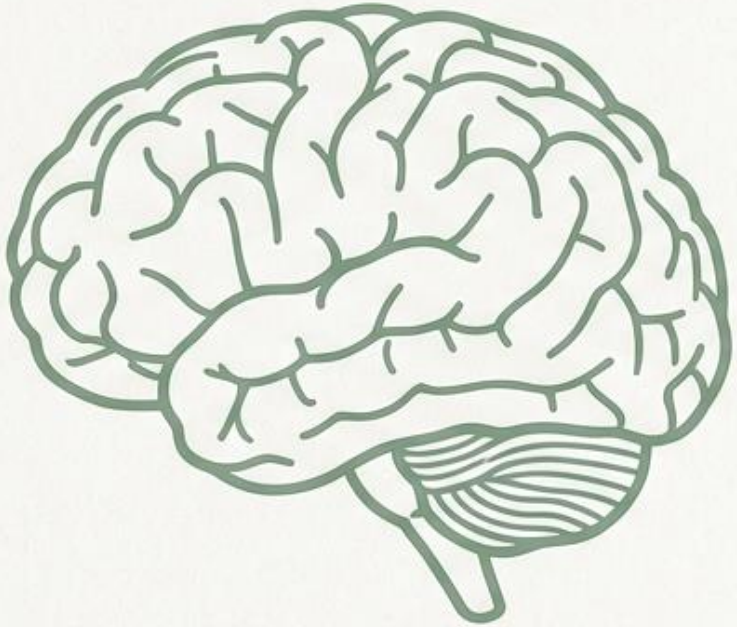
Unlocking Deep Learning:...
13 sources · 12d ago

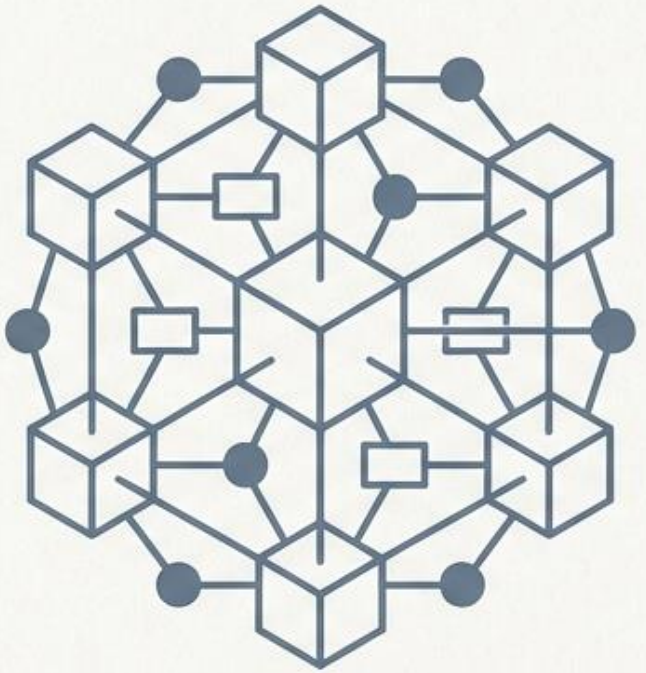
Add note


Customize Slide Deck
Format
Detailed Deck
A comprehensive deck with full text and details, perfect for emailing or reading on its own.
Presenter Slides
Clean, visual slides with key talking points to support you while you speak.
Choose language
English
Length
Short
Default
Describe the slide deck you want to create
Add a high-level outline, or guide the audience, style, and focus: "Create a deck for beginners using a bold and playful style with a focus on step-by-step instructions."
Generate

The Cognitive Paradox

Navigating cognitive load, creative homogenization, and the future of assisted learning.





The Double-Edged Sword of Cognitive Assistance

The Promise

Unprecedented productivity, rapid ideation, personalized tutoring, and complex synthesis.

GPT-4 scored **88.1%** on **Medical State Exams** vs. students' **74.6%**.

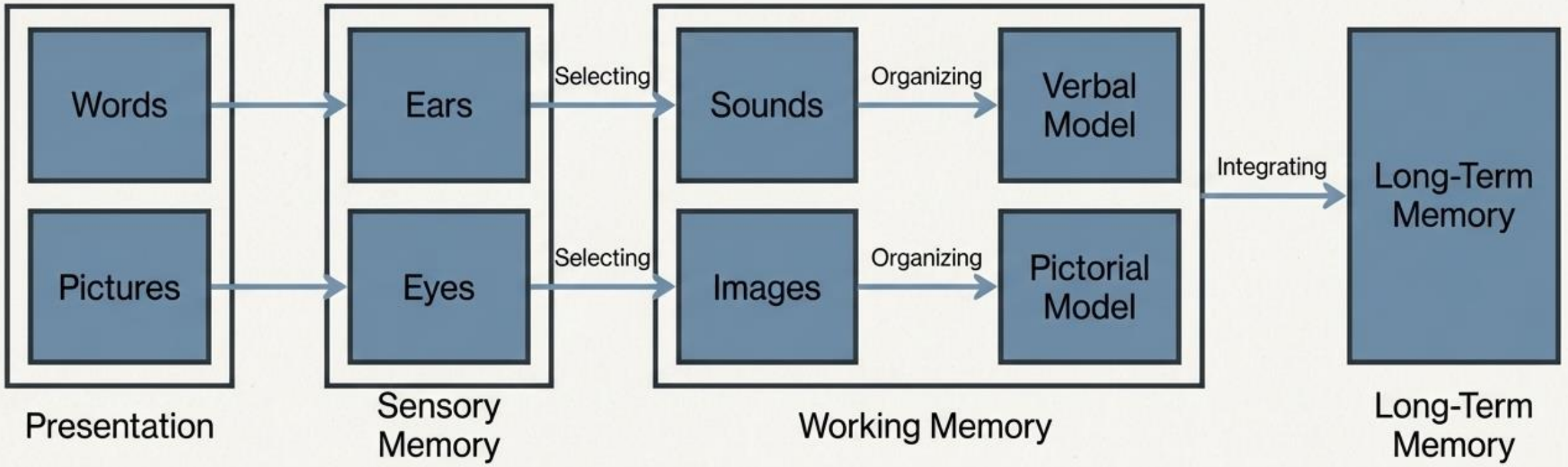
The Peril

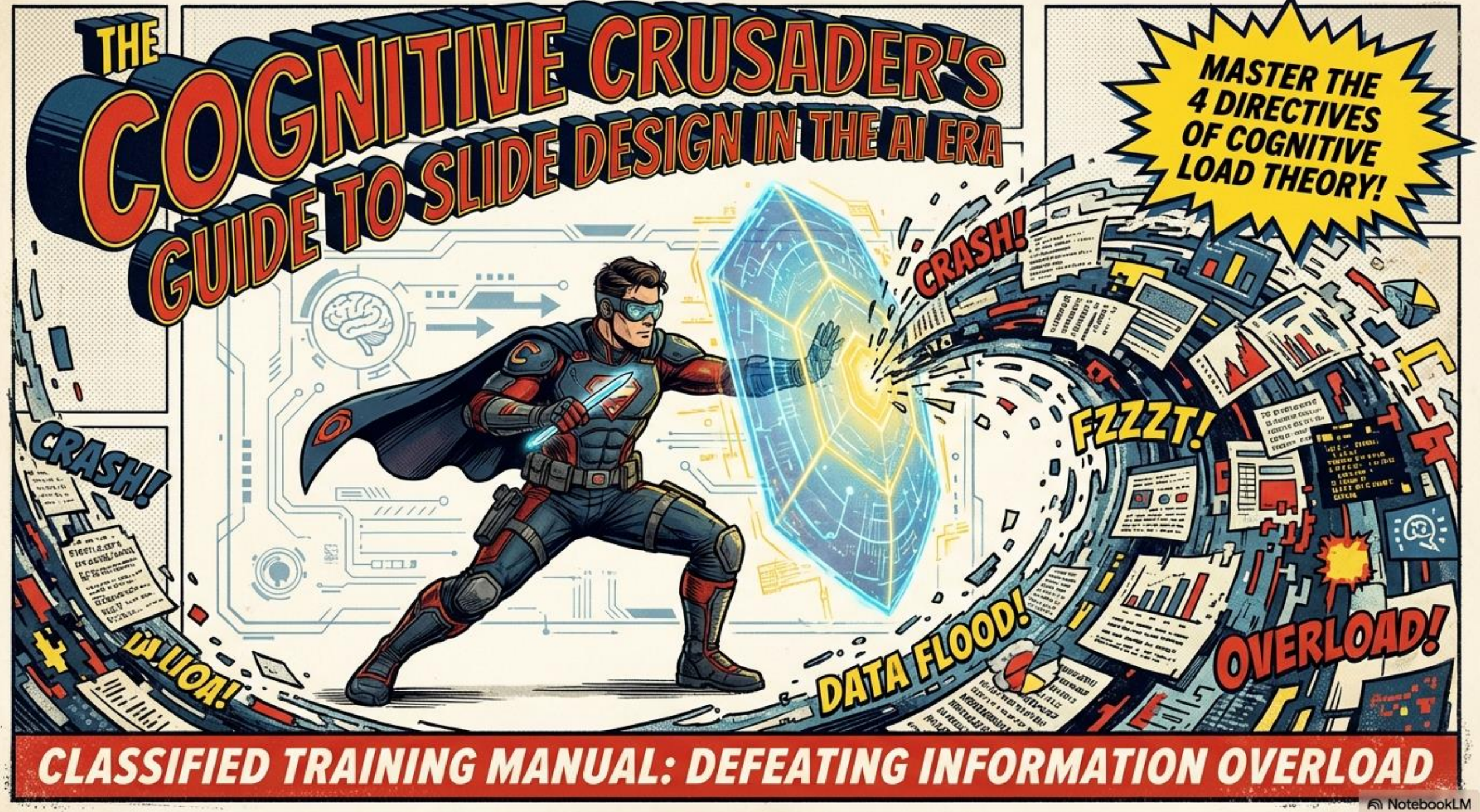
Cognitive overload, "offloading" of analytical tasks, creative homogenization, and the degradation of independent reasoning.

Overreliance correlates with **reduced problem-solving ability** (d = **-0.41**).

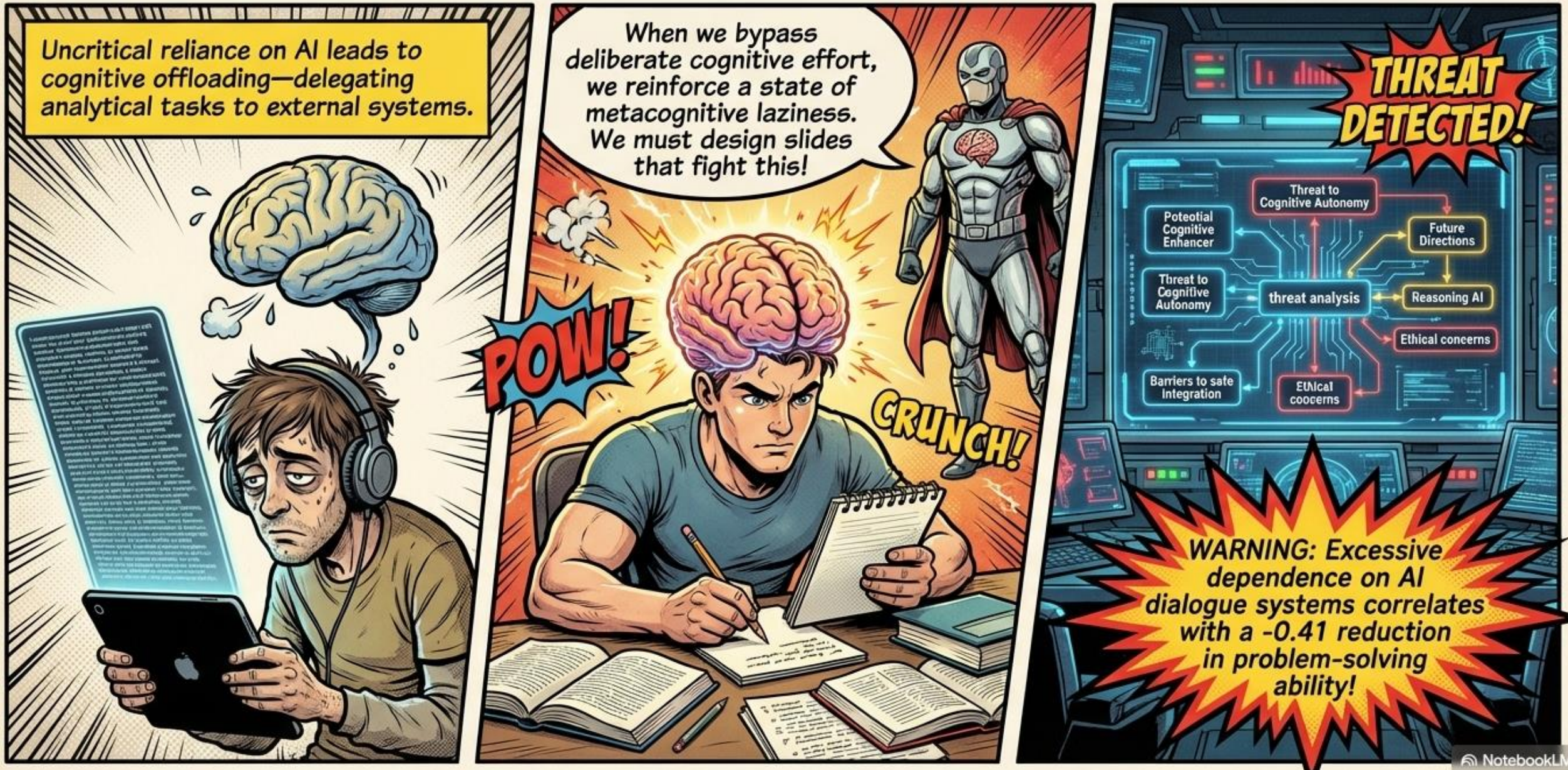
The Three Pillars of Cognitive Processing

- **Dual Channels:** Separate pathways for visual/pictorial and auditory/verbal data.
- **Limited Capacity:** Working memory bottlenecks.
- **Active Processing:** Meaningful learning requires selecting, organizing, and integrating.



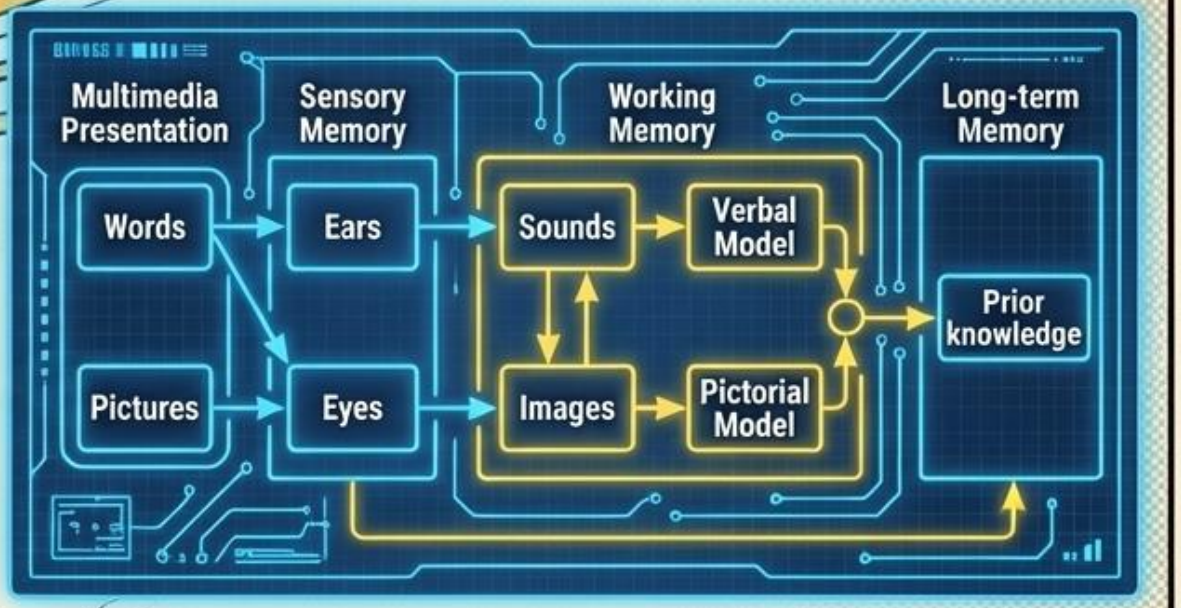

THE COGNITIVE CRUSADER'S GUIDE TO SLIDE DESIGN IN THE AI ERA
MASTER THE 4 DIRECTIVES OF COGNITIVE LOAD THEORY!
CRASH!
CRASH!
FZZZT!
DATA FLOOD!
OVERLOAD!
CLASSIFIED TRAINING MANUAL: DEFEATING INFORMATION OVERLOAD

GENERATIVE AI THREATENS TO TRIGGER METACOGNITIVE LAZINESS AND COGNITIVE OFFLOADING



DESIGN MUST ENGAGE BOTH THE AUDITORY AND VISUAL CHANNELS SIMULTANEOUSLY.

The Dual-Channel Assumption:
Humans possess separate systems for processing pictorial and verbal material.



TACTIC: Maximize processing capacity by distributing information across both channels, avoiding a jam in the visual processor.

Each channel has a strictly limited capacity. Processing bottlenecks lead to catastrophic cognitive overload.

Choosing the Right Tool

A quick reference for which AI tool fits which clinical task

Choosing the Right Tool

Use Case	Best Tool(s)	Notes
Differential diagnosis, clinical reasoning	Gemini, Claude, ChatGPT	Use Thinking or Pro models
Quick targeted clinical questions	Open Evidence	Cites sources reliably
Literature review / evidence synthesis	Consensus, Open Evidence, Perplexity	Better citation reliability
Patient messages, letters, appeals	Claude, ChatGPT, Doximity	Often no PHI needed; Claude writes most naturally
Summarizing a complex chart	Epic-Integrated Tools	Claude may be coming to Epic soon
MCQ / teaching case generation	ChatGPT, Claude, Gemini	Great for self-study and teaching prep
Image / slide generation	Gemini, Claude (Code)	Gemini is by far the best image model Claude is quite good at coding slides
Data cleaning / file conversion	ChatGPT, Claude, Gemini	Agents (Claude Code) especially strong here

Know the task → pick the tool. General LLMs handle most things; specialized tools add value for evidence-based queries.

Next Steps

Beginner

- Think about some repetitive task you do regularly. Could you use AI to help you with that? Open an AI assistant and talk with it to try and brainstorm!
- The next time you are on consults, intentionally practice incorporating AI into your workflow. Reflect on where it is helpful and where it isn't, and how it changes your learning.

Intermediate

- Think about an area where you need to process or sift through large amounts of information; try using an AI tool to help you with this.
- Try to build a workflow that allows you to quickly create patient education material either from an ambient scribe or another process

Advanced

- I would strongly encourage starting to explore Agentic AI. It has blown my mind the ways that it can simplify and streamline productivity systems, allow you to create things that you never would have dreamed possible, and assist in complex projects over multiple sessions
- I've been using Claude Code, but if the terminal scares you Claude Cowork is quite good as well. Other options (that I have no experience with) include ChatGPT Codex, Perplexity Computer, Manus.IM, and Openclaw

Challenge: What's ONE thing you'll actually try this week?

Learning More – With Whom Am I In Conversation?

www.oneusefulthing.org

hannahstulberg.substack.com

theconvivialsociety.substack.com

fitzyhistory.substack.com

tinyurl.com/3rnssdyd

Thank You

Questions?

Jonathan Berry, MD | BIDMC / Harvard Medical School
Jlberry@bidmc.harvard.edu

[Slides & resources: tinyurl.com/5znszs8h]

tinyurl.com/mr2vjzz7

Landing Page