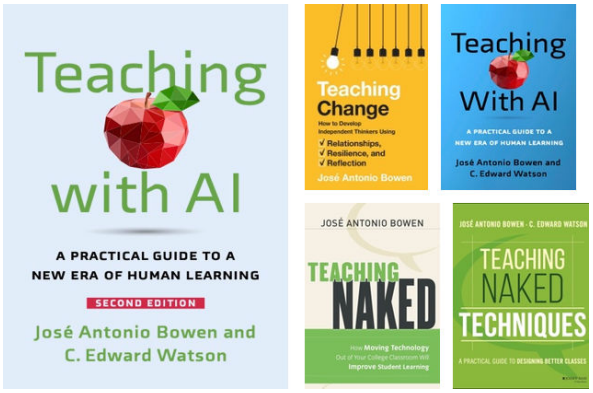
Are We Asking the Right Questions about Human Thinking and AI?



SLIDES & CITATIONS

José Antonio Bowen

Get the 2nd Edition of Teaching with AI.
Discount code HTAI25 = 30% off at JHUP.
WebSite: https://WeTeachWithAI.com/
This doc was prepared by Claude using the images, text and notes from my ppt deck.

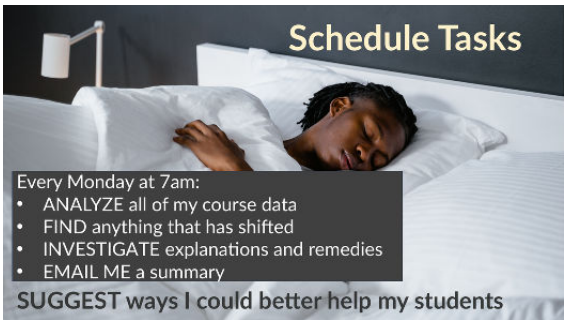
Where's Rigoletto?

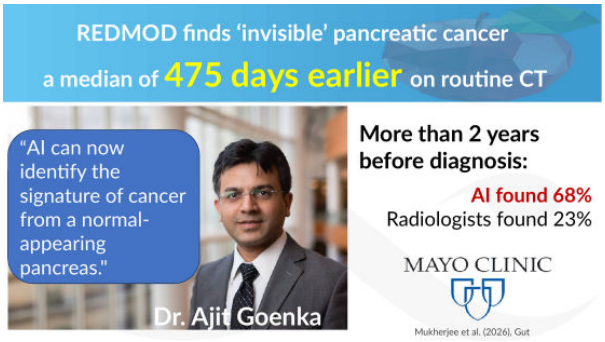
(That you can figure out on your own.)

WE ARE ALREADY ASKING A LOT OF GOOD QUESTIONS (slide 3)

How should we regulate AI?
Can we reduce the climate impact?
What (if any) AI use is ethical?
How do we manage privacy?
Do we need to provide access?
What about students who object?

HOW IS AI CHANGING WHAT IS POSSIBLE?





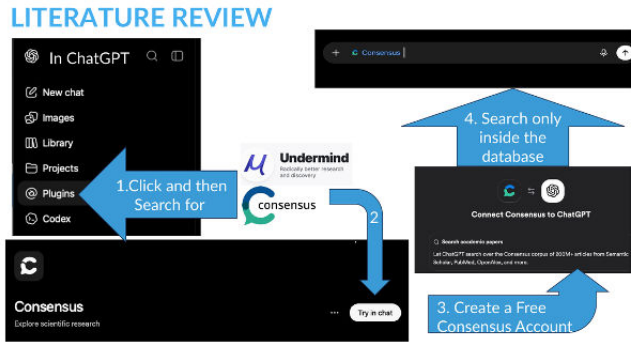
Mukherjee, S., Antony, A., Patnam, N. G., et al. (2026). Next-generation AI for visually occult pancreatic cancer detection in a low-prevalence setting with longitudinal stability and multi-institutional generalisability. Gut. [published online 28 April 2026] https://doi.org/10.1136/gutjnl-2025-337266

AGA News. (2026, June). AI model flags pancreatic cancer on CT more than a year before diagnosis. [overall test-set sensitivity 73.0% vs 38.9% for radiologists; specificity 81.1%] https://news.gastro.org/issues/2026/june-2026/ai-model-flags-pancreatic-cancer-on-ct-more-than-a-year-before-diagnosis/

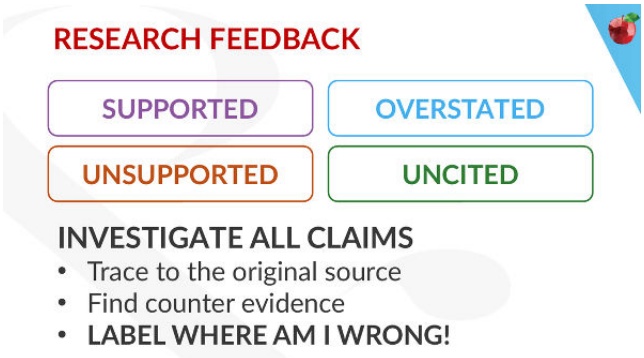


- MatrAIx

Li, X., Hao, Y., Hou, J., et al. (2026, August 4). MatrAIx: Simulating the world with 8.3 billion persona agents. arXiv [preprint]. https://arxiv.org/abs/2608.04205



- Consensus
- Undermind



More on Research with AI workflow here:

https://weteachwithai.com/ai-research-workflow/

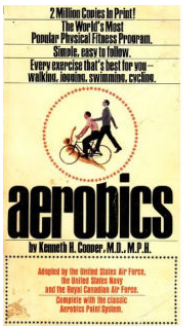


- Devil's Advocate (Duke Office of Information Technology)
- Getting started with Devil's Advocate

AI Opportunities in Health Care

- Automation of Tasks and Workflow
  - Documentation and Notes
  - Cleveland Clinic Identify deteriorating Conditions
  - Extract Data from Pharmacy Benefit Management (PBM) contracts
- Enhance Diagnostic Accuracy
  - Butterfly iQ+ Ultrasound or PathAI
- Improve Patient Care
- Faster Detection
  - Freenome Early Cancer Detection
- Personalized Treatment Plans
  - MD Anderson Cancer
- Predictive Analytics

- Butterfly iQ+ (Butterfly Network)
- PathAI
- Freenome



Dr. Kenneth H. Cooper (1968)

"The streets would be littered with dead joggers"

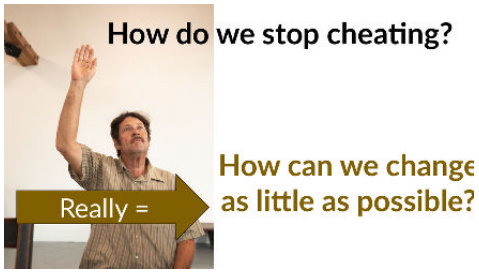
Cooper, K. H. (1968). Aerobics. M. Evans.

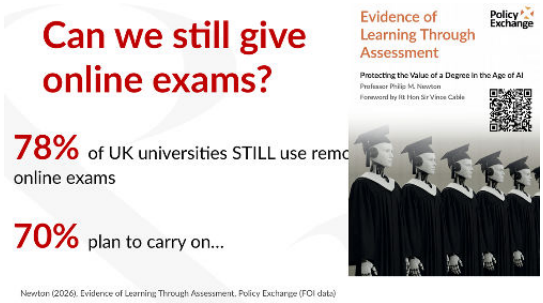
Texas Tech University Health Sciences Center. (2026, Winter). Father of fitness. Pulse. [Cooper recounts the warning that "the streets would be littered with dead joggers"] https://pulse.ttuhsc.edu/issue/winter-2026/father-of-fitness/

WHAT DOES THIS MEAN FOR HUMAN COGNITIVE FITNESS?

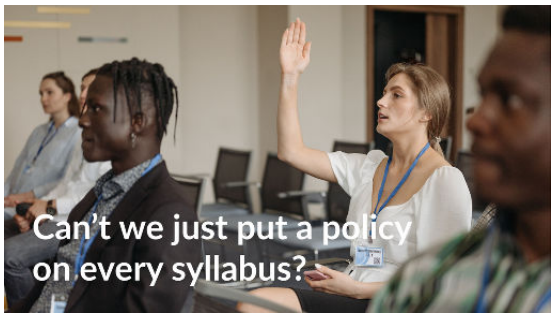
EIGHT QUESTIONS (slide 14)

What are our principles and our practice?
How can we design better learning?
How will AI alter our disciplinary learning goals?
How might we amplify and raise standards?
Who will verify what AI produces and how?
How can we build more supportive relationships?
What will we do with the extra capacity?
What does AI mean for our curriculum?





Newton, P. M. (2026, August 17). Evidence of learning through assessment: Protecting the value of a degree in the age of AI. Policy Exchange. [think-tank report; Freedom of Information requests to UK universities] https://policyexchange.org.uk/publication/evidence-of-learning-through-assessment/

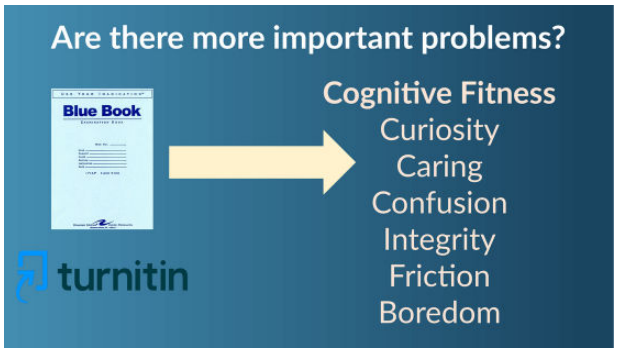








▸ Caltech Undergraduate Admissions: Research Papers & Publications

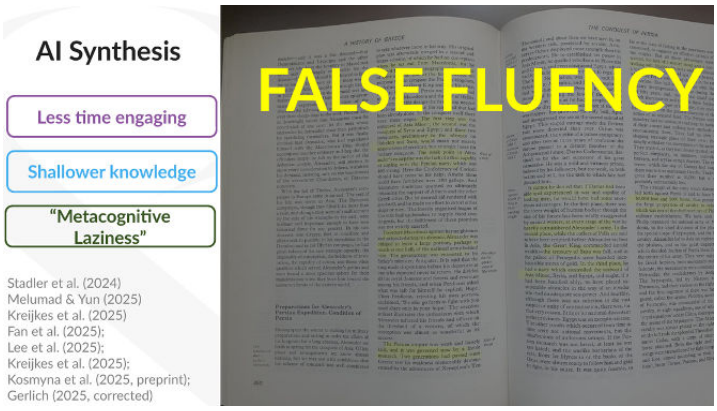




Marken, S. (2026, April 1). AI is routine for college students, despite campus limits. Gallup. [Lumina Foundation–Gallup 2026 State of Higher Education study; 3,801 students; opt-in online panel] https://news.gallup.com/poll/704090/routine-college-students-despite-campus-limits.aspx

Chirikov, I., Smirnov, I., & Kizilcec, R. F. (2026). Generative AI use and misuse call for assessment reform in higher education. Science, 392(6800), 818–820. [95,513 students at 20 US public research universities; substantial heterogeneity in GenAI use and misuse across disciplines and student groups; misuse estimated by indirect questioning, not self-report] https://doi.org/10.1126/science.aec5115

Stephenson, R., & Armstrong, C. (2026). Student generative AI survey 2026 (HEPI Report 199). Higher Education Policy Institute. [1,054 UK undergraduates; 12% put AI-generated text directly into assessed work, up from 8% in 2025 and 3% in 2024] https://www.hepi.ac.uk/reports/student-generative-ai-survey-2026/



Melumad, S., & Yun, J. H. (2025). Experimental evidence of the effects of large language models versus web search on depth of learning. PNAS Nexus, 4(10), pgaf316. https://doi.org/10.1093/pnasnexus/pgaf316

Stadler, M., Bannert, M., & Sailer, M. (2024). Cognitive ease at a cost: LLMs reduce mental effort but compromise depth in student scientific inquiry. Computers in Human Behavior, 160, 108386. https://doi.org/10.1016/j.chb.2024.108386

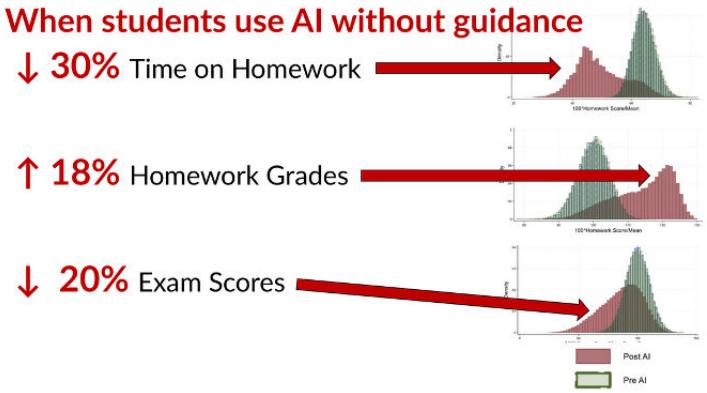
Kreijkes, P., et al. (2026). Effects of LLM use and note-taking on reading comprehension and memory: A randomised experiment in secondary schools. Computers & Education, 243, Article 105514. https://doi.org/10.1016/j.compedu.2025.105514

Fan, Y., et al. (2025). Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. British Journal of Educational Technology, 56(2), 489–530. [RCT, 117 students: essay scores up, no gain in knowledge or transfer] https://doi.org/10.1111/bjet.13544

Lee, H.-P., et al. (2025). The impact of generative AI on critical thinking. CHI '25. [confidence in AI ↔ less critical thinking] https://doi.org/10.1145/3706598.3713778

Kosmyna, N., et al. (2025). Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing. arXiv [preprint; small sample — cite as a cultural phenomenon, not as evidence]. https://arxiv.org/abs/2506.08872

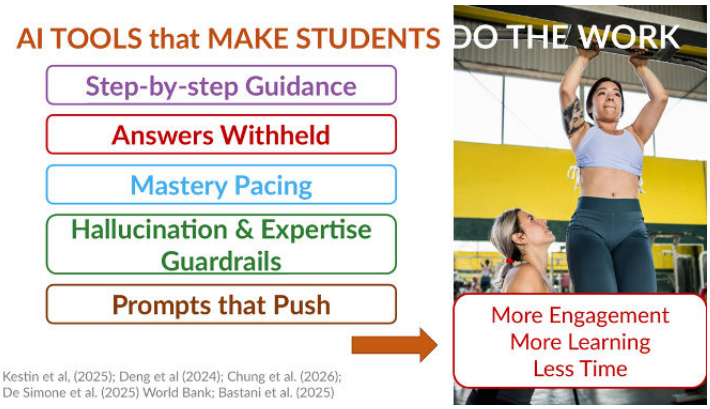
Gerlich, M. (2025). AI tools in society. Societies, 15(1), 6. [self-report perception data, 666 participants; formally corrected September 2025] https://doi.org/10.3390/soc15010006



Strömberg, D., Lei, V., & Wu, Y. (2026). The generative AI learning penalty: Evidence from Chinese secondary education (CEPR Discussion Paper No. 21577). Centre for Economic Policy Research. [working paper] https://cepr.org/publications/dp21577

Wecks, J. O., Voshaar, J., Plate, B. J., & Zimmermann, J. (2025). Generative AI usage and exam performance. ECIS 2025 Proceedings. https://arxiv.org/abs/2404.19699

Rismanchian, S., et al. (2026). Faster completion, less learning: Generative AI reduced study time on math problems and the knowledge they build. arXiv [preprint]. https://arxiv.org/abs/2605.21629



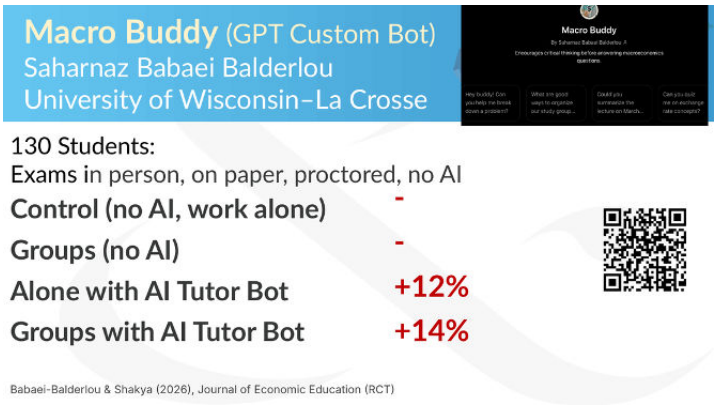
Kestin, G., Miller, K., Klales, A., Milbourne, T., & Ponti, G. (2025). AI tutoring outperforms in-class active learning. Scientific Reports, 15, 17458. [Harvard RCT, N = 194] https://doi.org/10.1038/s41598-025-97652-6

Deng, R., Jiang, M., Yu, X., Lu, Y., & Liu, S. (2024). Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies. Computers & Education, 227, 105224. [69 studies; higher-order gains largely from subjective student reports] https://doi.org/10.1016/j.compedu.2024.105224

Chung, A. T.-H., Zhang, B., Kung, L.-C., Bastani, H., & Bastani, O. (2026). Effective personalized AI tutors via LLM-guided reinforcement learning. SSRN [working paper; +0.15 SD unassisted exam gain ≈ 6–9 months of schooling] https://papers.ssrn.com/sol3/papers.cfm?abstract_id=6423358

De Simone, M. E., et al. (2025). From chalkboards to chatbots: Evaluating the impact of generative AI on learning outcomes in Nigeria. World Bank [working paper; 800 students; girls gained most] http://documents.worldbank.org/curated/en/099548105192529324

Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö., & Mariman, R. (2025). Generative AI without guardrails can harm learning: Evidence from high school mathematics. PNAS, 122, e2422633122. [RCT, n ≈ 1,000] https://doi.org/10.1073/pnas.2422633122

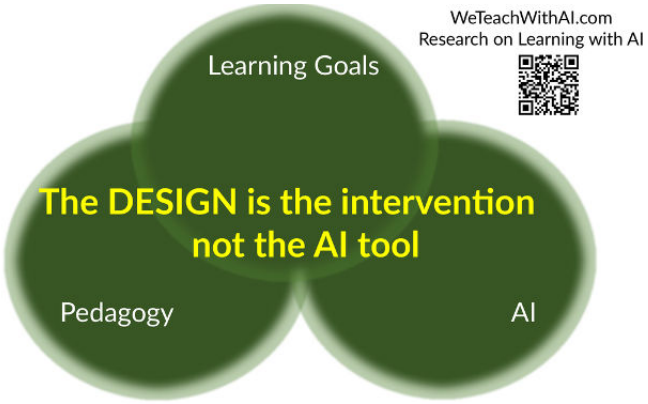


Saharnaz Babaei-Balderlou, University of Wisconsin–La Crosse.

- Macro Buddy (GPT custom bot)

Babaei-Balderlou, S., & Shakya, S. (2026). Can structured AI tutoring and study groups improve learning outcomes? Evidence from a randomized controlled trial. The Journal of Economic Education, 1–18. [RCT, 4 sections, 130 students; the first randomized evidence of AI-enhanced learning in economics education] https://doi.org/10.1080/00220485.2026.2695361

Babaei-Balderlou, S., & Shakya, S. (2025, July 25). The invisible hand of Gen-AI: Can AI-enhanced study groups improve learning outcomes? SSRN. [working-paper version, the one cited on the slide] https://doi.org/10.2139/ssrn.5696622



▸ Designing Learning with AI: findings, principles and examples

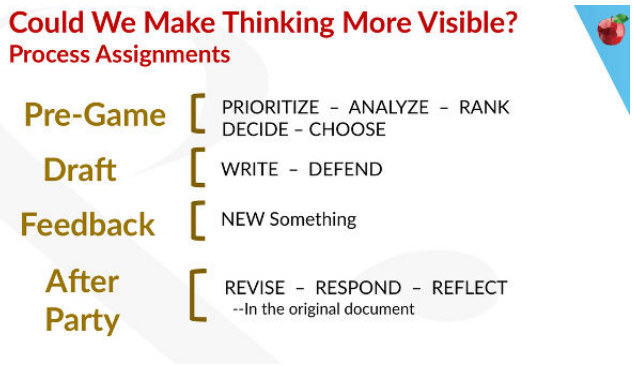
Li, C., Cui, H., & Hagedorn, L. S. (2026). The cognitive impact of ChatGPT in higher education: A systematic review of critical and creative thinking outcomes. Computers and Education: Artificial Intelligence, 10, Article 100571. [PRISMA review of 67 empirical studies, 2022–2025 — the strongest synthesis to date] https://doi.org/10.1016/j.caeai.2026.100571

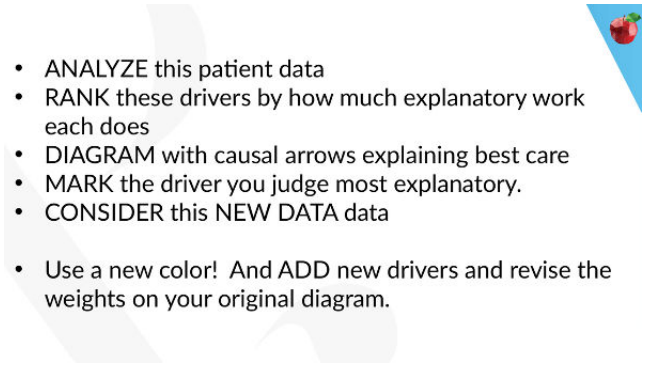
Wang, J., & Fan, W. (2025). RETRACTED (22 April 2026): The effect of ChatGPT on students' learning performance… Humanities and Social Sciences Communications, 12, 621. [RETRACTED — never cite as evidence] https://doi.org/10.1057/s41599-025-04787-y

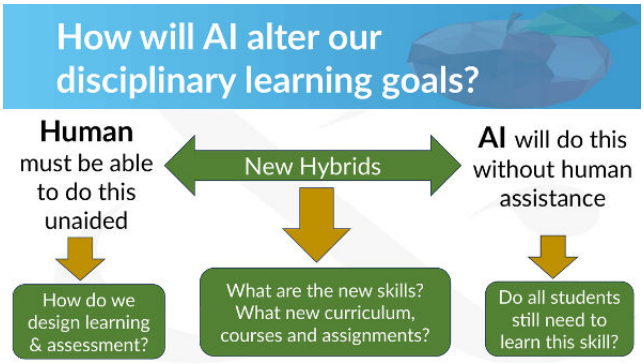
Bartoš, F., Martinková, P., & Wagenmakers, E. (2025, May 12). Adjusting for publication bias reveals no evidence for the effect of ChatGPT on students' learning performance. [preprint reanalysis behind the retraction] https://doi.org/10.31234/osf.io/8vs32_v1

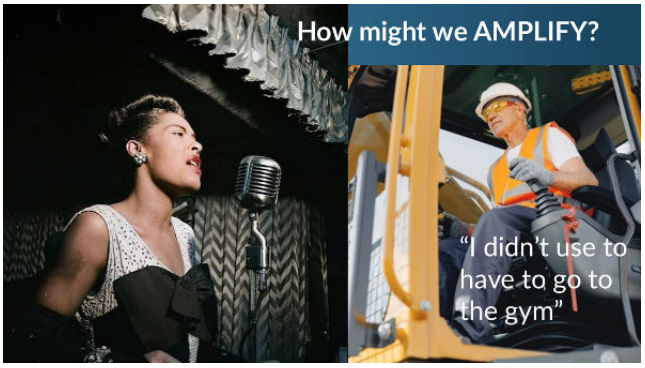
LIMITS. The strongest evidence of harm comes from secondary students outside the United States (China, Turkey, Germany, the U.K.); the strongest evidence of benefit comes from elite, high-structure settings such as Harvard. The typical university classroom sits in the understudied middle, where age, selectivity and design are confounded. Most of these studies measure exam scores; we still lack long-term evidence about durable, transferable learning.

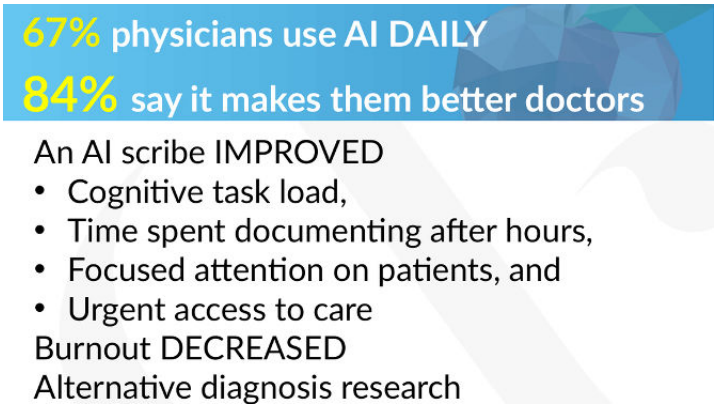






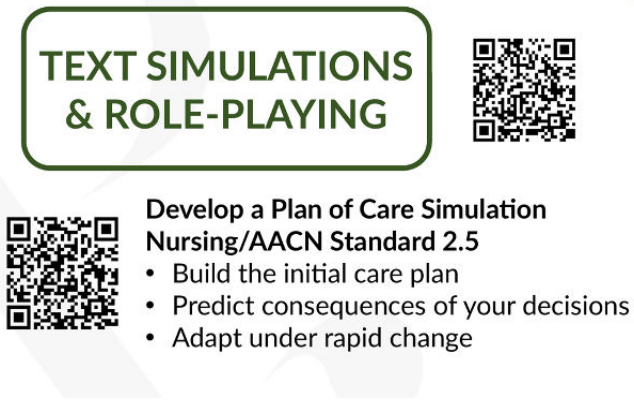






Offcall. (2025). The 2025 physicians AI report. [survey of 1,000+ U.S. physicians across 106 specialties; physician-founded company] https://2025-physicians-ai-report.offcall.com/

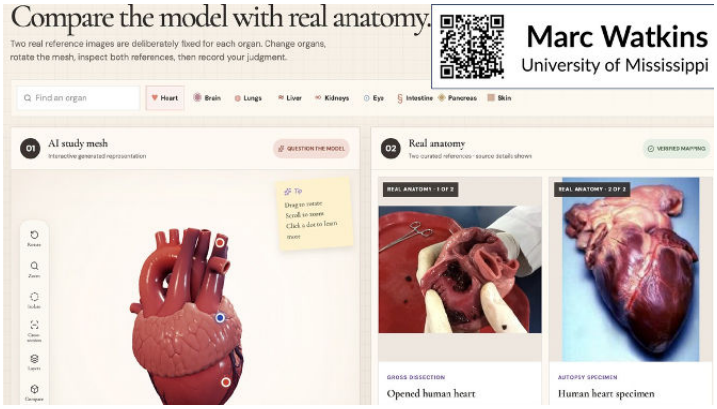
Olson, K. D., et al. (2025). Use of ambient AI scribes to reduce administrative burden and professional burnout. JAMA Network Open, 8(10), e2534976. [quality improvement study, pre/post surveys, 263 clinicians in 6 health systems, 30 days; burnout 51.9% to 38.8%] https://doi.org/10.1001/jamanetworkopen.2025.34976



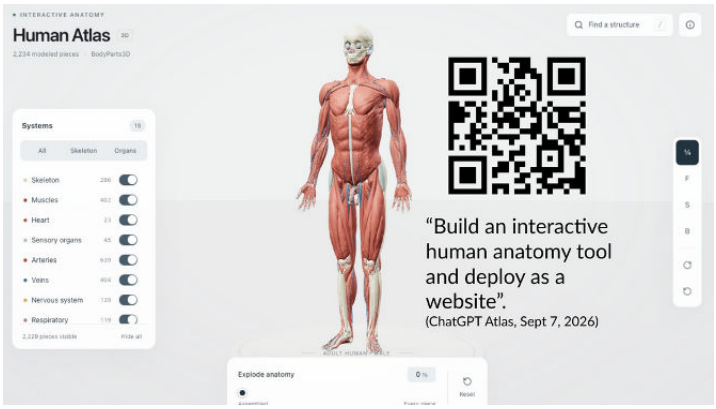
- Plan of Care simulation (Claude artifact)
- Simulations, games & role-playing



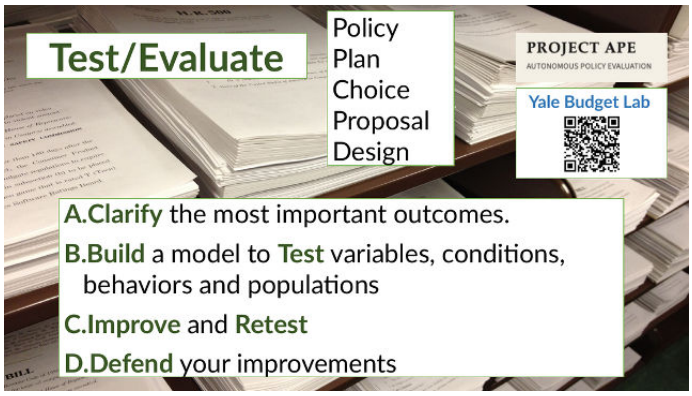
- Vibe coding examples



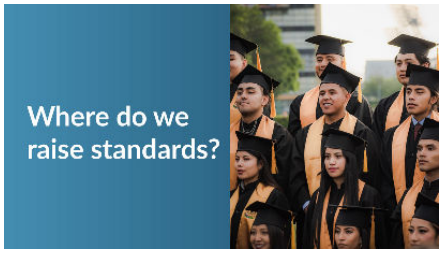
- Anatomy Studio — compare the AI model with real anatomy

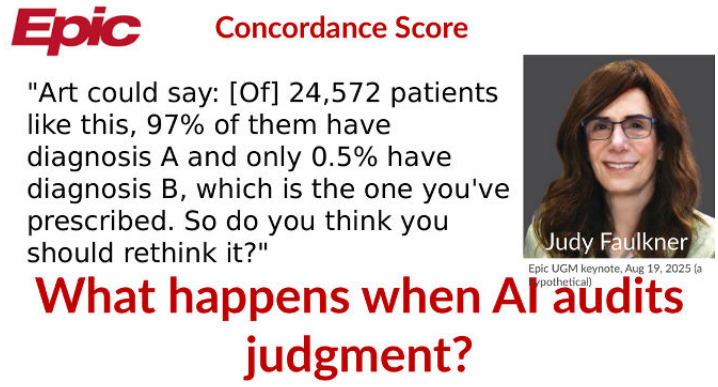


- Human Atlas



- Yale Budget Lab
- Project APE: Autonomous Policy Evaluation (Social Catalyst Lab)





Becker's Hospital Review. (2025, August). Judy Faulkner unveils 'healthcare intelligence.' [Epic UGM keynote, Verona, WI, August 19, 2025; reports "0.5%"] https://www.beckershospitalreview.com/healthcare-information-technology/ehrs/judy-faulkner-unveils-healthcare-intelligence/

InBusiness. (2025, August). Epic UGM: Futuristic Faulkner sets the AI tone. [prints "5.5%"] https://ibmadison.com/epic-ugm-futuristic-faulkner-sets-the-ai-tone/



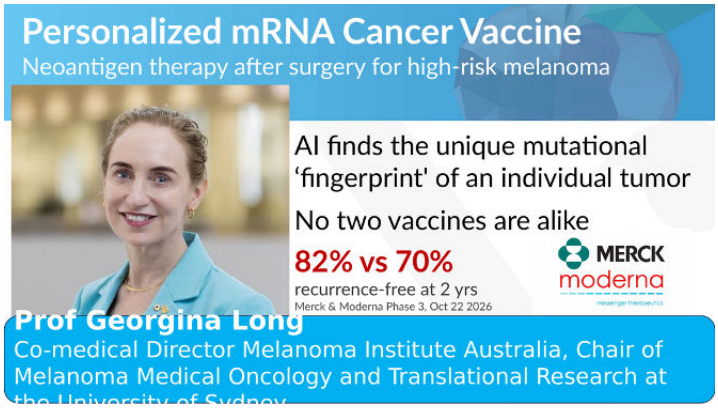
- Refine.ink
- Stanford Agentic Reviewer (paperreview.ai)

Roberts, S. (2026, October 8). 'Breathtaking,' 'Devastating': Mathematics reels after new OpenAI release. The New York Times. https://www.nytimes.com/2026/10/08/science/mathematicians-respond-openai-release.html

Gallegos, A. (2026, October 8). OpenAI withdraws three preprints a day after releasing 722 manuscripts on unsolved math problems. Retraction Watch. [quotes the Association for Human Mathematics statement of October 7, 2026] https://retractionwatch.com/2026/10/08/openai-withdraws-preprints-722-manuscripts-unsolved-math-problems/

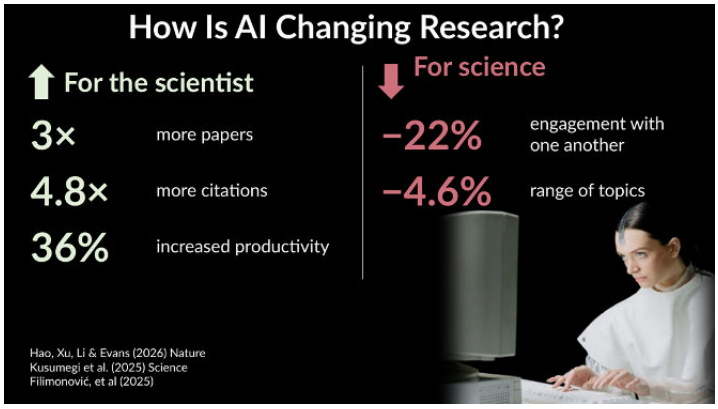
Association for Human Mathematics, Communications Working Group. (2026, October 7). AHM statement on OpenAI's October 6 release of mathematical documents. [guest post on Terence Tao's blog] https://terrytao.wordpress.com/2026/10/07/ahm-statement-on-openais-october-6-release-of-mathematical-documents/

Refine. (2026, August 6). Refine partners with the American Economic Association and the Econometric Society. [company blog] https://www.refine.ink/blog/economics-journal-partnerships



Weber, J. S., Carlino, M. S., Khattak, A., et al. (2024). Individualised neoantigen therapy mRNA-4157 (V940) plus pembrolizumab versus pembrolizumab monotherapy in resected melanoma (KEYNOTE-942): A randomised, phase 2b study. The Lancet, 403(10427), 632–644. [randomised phase 2b; Georgina Long is a senior author] https://doi.org/10.1016/S0140-6736(23)02268-7

Merck & Moderna. Phase 3 INTerpath-001 trial of intismeran autogene plus Keytruda met endpoints of recurrence-free survival and distant metastasis-free survival. [company press release] https://www.merck.com/news/merck-and-moderna-announce-phase-3-interpath-001-trial-of-intismeran-autogene-plus-keytruda-met-endpoints-of-recurrence-free-survival-rfs-and-distant-metastasis-free-survival-dmfs-in-patient/



Hao, Q., Xu, F., Li, Y., & Evans, J. (2026). Artificial intelligence tools expand scientists' impact but contract science's focus. Nature, 649(8099), 1237–1243. [41.3M papers, 1980–2025; machine-learning methods, mostly pre-LLM] https://doi.org/10.1038/s41586-025-09922-y

Kusumegi, K., et al. (2025). Scientific production in the era of large language models. Science, 390, 1240–1243. https://doi.org/10.1126/science.adw3000

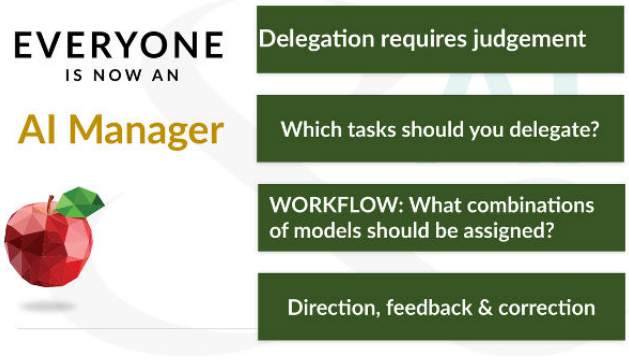
Filimonović, D., Rutzer, C., & Wunsch, C. (2025, October). Can GenAI improve academic performance? Evidence from the social and behavioral sciences. arXiv [working paper; see p. 14]. https://arxiv.org/abs/2510.02408

Women at the Keys: a searchable database of women composers of solo piano music, built by a coding agent from a five-step prompt.



- Women at the Keys
- Vibe coding examples

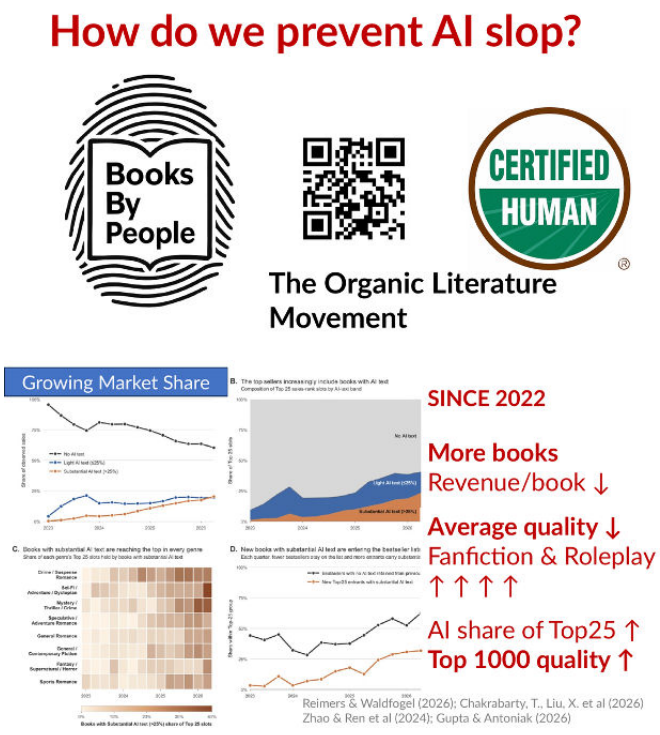






Korst, J., Puntoni, S., & Tambe, P. (2026, October). The adaptive advantage: Who's winning with AI and how (Year Four report). Wharton Human-AI Research & GBK Collective. [online survey, about 1,400 U.S. employees at firms with 1,000+ staff, fielded July 2026; share agreeing "My focus has shifted from creating outputs to managing AI-generated outputs"] https://knowledge.wharton.upenn.edu/wp-content/uploads/2026/10/Wharton-GBK-AI-Adoption-Report-2026.pdf





- Books By People

The fanfiction and roleplay line comes from the WildChat corpus: in more than 500,000 consented, English-language ChatGPT conversations, over a third involve generating fiction, especially fanfiction and erotica which generate what the authors call a "solipsistic reader-writer," “who both generates and consumes fiction within a closed conversational loop.”

Reimers, I., & Waldfogel, J. (2026). AI and the quantity and quality of creative products: Have LLMs boosted creation of valuable books? NBER Working Paper 34777. [working paper] https://doi.org/10.3386/w34777

Chakrabarty, T., Liu, X., et al. (2026, July). Generative AI floods and dilutes the market for books. arXiv [preprint]. https://arxiv.org/pdf/2607.20349

Gupta, N., Antoniak, M., & Walsh, M. (2026, June 23). AI fiction in the wild. arXiv [preprint; provisionally forthcoming in MFS: Modern Fiction Studies]. https://arxiv.org/abs/2606.22748

Zhao, W., Ren, X., Hessel, J., Cardie, C., Choi, Y., & Deng, Y. (2024). WildChat: 1M ChatGPT interaction logs in the wild. ICLR 2024. [consented corpus of 1M conversations; the dataset behind Gupta et al.] https://arxiv.org/abs/2405.01470



Schironi, F. (2018). The best of the grammarians: Aristarchus of Samothrace on the Iliad. University of Michigan Press.

Kato (2022). Plus or minus? Adamantius, 28.

Fölster & Staack (2021). [title not verified]

Elseadawy (2023). Freie Universität Berlin. [title not verified]

Seidensticker (2020). Journal of Islamic Manuscripts, 11. [title not verified]

Hirschler (2014). Al-ʿUṣūr al-Wusṭā, 22. [title not verified]

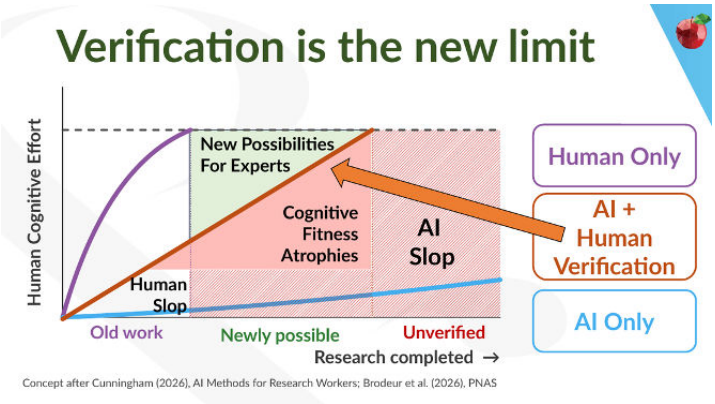
Grafton, A. (2011). The culture of correction in Renaissance Europe. British Library.

Hornschuch, H. (1608). Orthotypographia. Leipzig.

Dickey, C. (2019). The rise and fall of facts. Columbia Journalism Review.

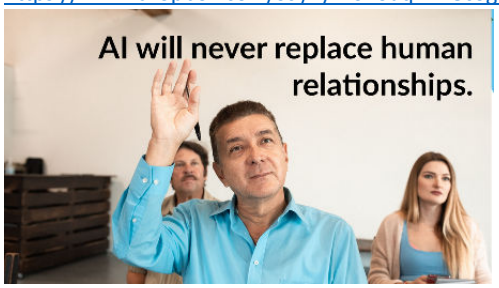
Rossner, M. (2008). A false sense of security. Journal of Cell Biology, 183.

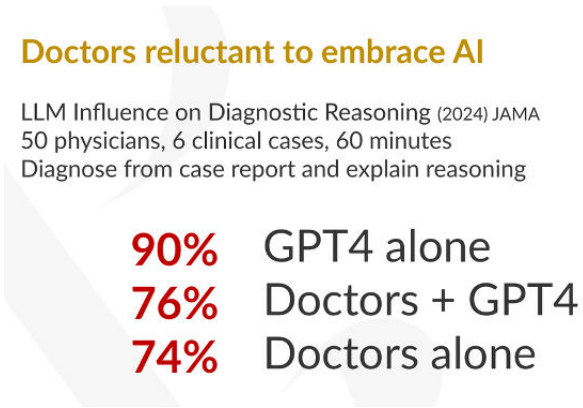
Rossner, M., & Yamada, K. M. (2004). What’s in a picture? The temptation of image manipulation. Journal of Cell Biology, 166.



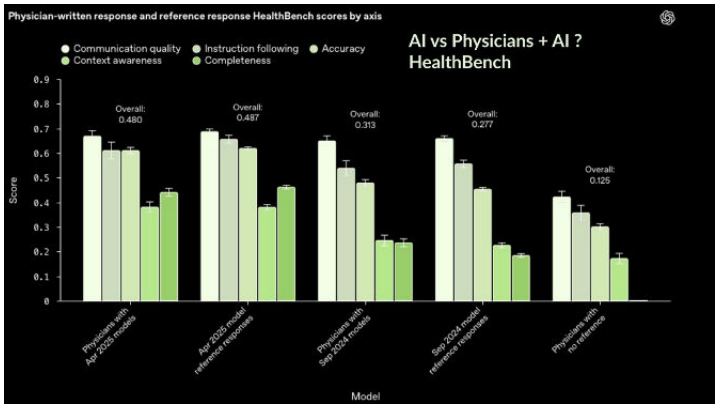
Brodeur, A., Valenta, D., Marcoci, A., Aparicio, J. P., Mikola, D., Barbarioli, B., et al. (2026). AI-assisted teams outperform AI-led teams but not human-only teams in assessing research reproducibility in quantitative social science. Proceedings of the National Academy of Sciences, 123(22), e2524747123. [experiment, 103 teams]
https://doi.org/10.1073/pnas.2524747123
See also Cunningham, S. (2026, October 2). AI methods for research workers [Workshop slides]. Texas A&M University.

https://www.dropbox.com/scl/fi/xro10uqr7w8c0g2k6u00p/tamu_deck.pdf?rlkey=93xgrh4oc2deunptbxif05vg8&dl=0

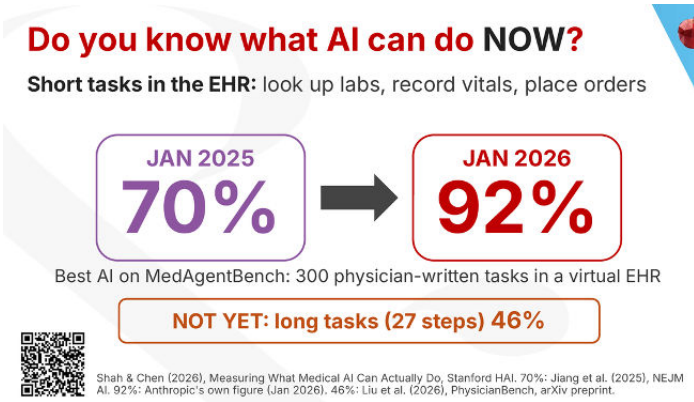




Goh, E., Gallo, R., Hom, J., et al. (2024). Large language model influence on diagnostic reasoning: A randomized clinical trial. JAMA Network Open, 7(10), e2440969. [RCT, 50 physicians]
https://doi.org/10.1001/jamanetworkopen.2024.40969



Arora, R. K., Wei, J., Hicks, R. S., et al. (2025). HealthBench: Evaluating large language models towards improved human health. arXiv [preprint; OpenAI-authored benchmark, 5,000 conversations, 262 physicians].
https://arxiv.org/abs/2505.08775

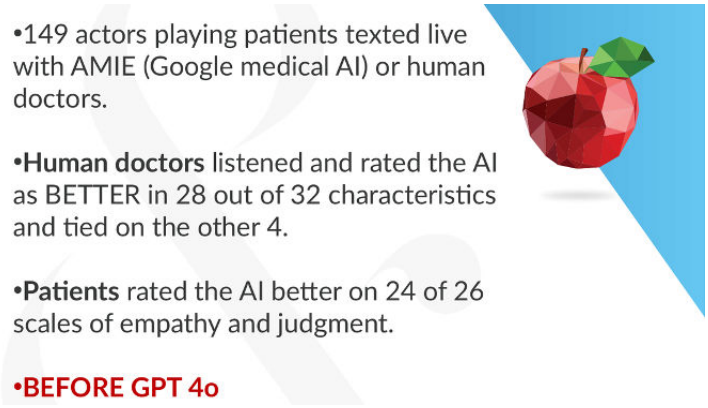


Shah, N., & Chen, J. (2026, October 19). Measuring what medical AI can actually do. Stanford HAI. https://hai.stanford.edu/news/measuring-what-medical-ai-can-actually-do

Jiang, Y., Black, K. C., Geng, G., Park, D., Zou, J., Ng, A. Y., & Chen, J. H. (2025). MedAgentBench: A virtual EHR environment to benchmark medical LLM agents. NEJM AI. https://doi.org/10.1056/AIdbp2500144

Anthropic. (2026, January). Advancing Claude in healthcare and the life sciences [vendor-reported]. https://www.anthropic.com/news/healthcare-life-sciences

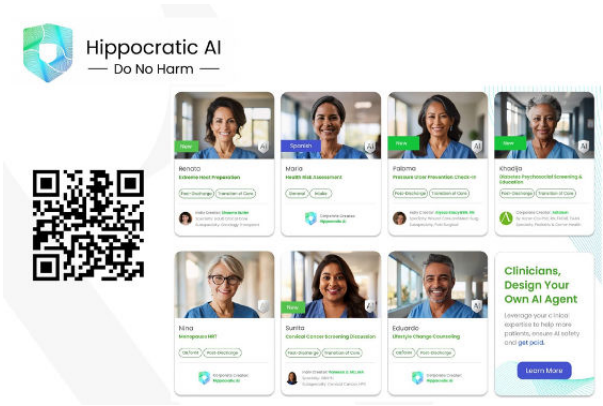
Liu, R., Mohiuddin, I. Q., Schoeffler, A. J., et al. (2026). PhysicianBench: Evaluating LLM agents in real-world EHR environments. arXiv [preprint]. https://arxiv.org/abs/2605.02240



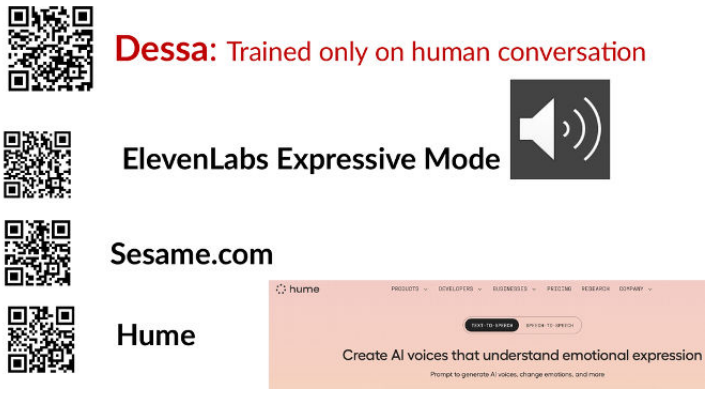
- AMIE (Google Research)

Tu, T., Schaekermann, M., Palepu, A., et al. (2025). Towards conversational diagnostic artificial intelligence. Nature, 642(8067), 442–450. [randomized, double-blind crossover; text-chat consultations with patient-actors; Google-authored] https://doi.org/10.1038/s41586-025-08866-7

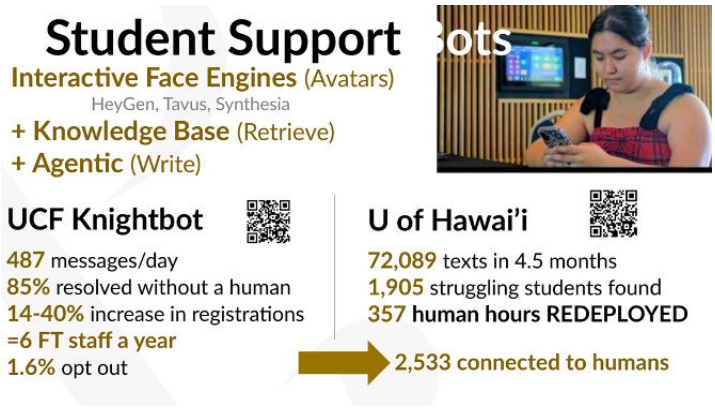
Tu, T., et al. (2024). Towards conversational diagnostic AI. arXiv [preprint; source of the 149 scenarios, 28 of 32 and 24 of 26 axes; the Nature version reports 159 scenarios, 30 of 32 and 25 of 26] https://arxiv.org/abs/2401.05654



- Hippocratic AI

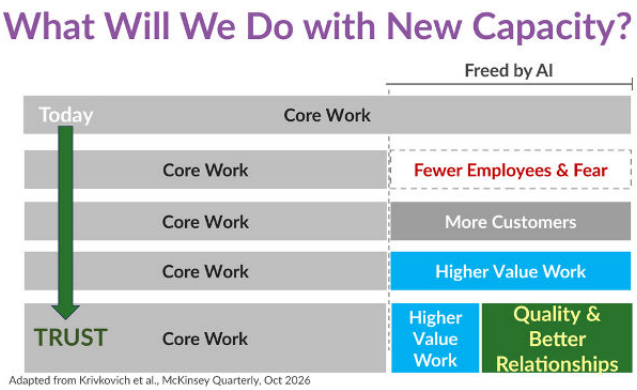


- Dessa (Block): conversation demos

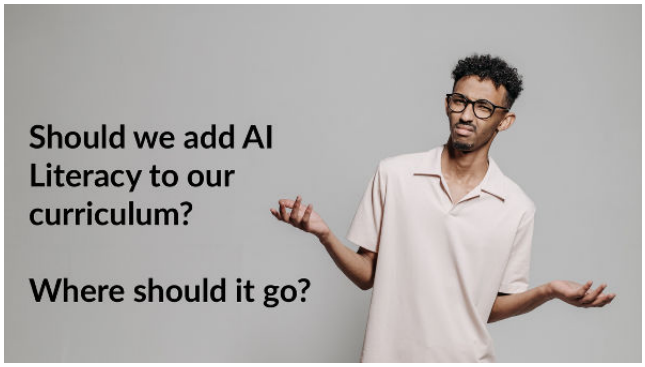


- UCF Knightbot
- University of Hawai'i (UH News)
- HeyGen
- Tavus
- Synthesia

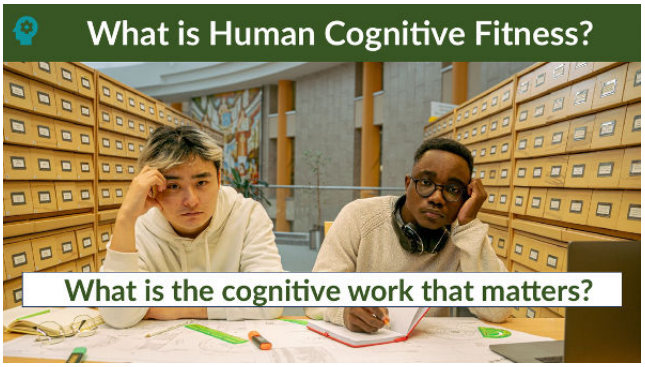




Krivkovich, A., Weddle, B., Price, H., & Sohoni, V. (2026, October 5). AI is changing work. Now it has to change the organization. McKinsey Quarterly. [adapted from Exhibit 1; shares are illustrative, not measured] https://www.mckinsey.com/capabilities/people-and-organization/our-insights/ai-is-changing-work-now-it-has-to-change-the-organization







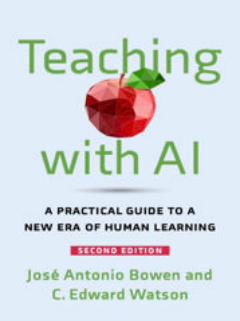
What new HUMAN question will you bring to your department? (slide 69)





CITATIONS & RESOURCES

- WeTeachWithAI.com contains many many more resources, prompts, assignments and links. It is updated weekly if not daily.
- Buy the book at 30% off!
- Teaching with AI at Johns Hopkins University Press — code HTAI25 for 30% off
- Teaching with AI on Amazon



- If you don't want to use AI, but want students to want to do things the hard way, then read Teaching Change: How to Develop Independent Thinkers Using Relationships, Resilience and Reflection.
- Teaching Change at Johns Hopkins University Press — code HTWN for 30% off
- Teaching Change on Amazon (often 25% off)

