

AI Deep Research Models New Tools For Research and Study

**2025 Next Gen Best in Class AI
Deep Research/Reasoning Models**
(June 2025)



Ray Uzwyshyn, Ph.D. MBA MLIS
Acting AUL for Research & Technology Services
Director of Research
UC Riverside Libraries, raymondu@ucr.edu
<https://rayuzwyshyn.net>

A graphic featuring a blue line that starts on the left, curves upwards to form a semi-circular arch, and then levels off to the right. A yellow line starts below the blue line, curves upwards to form a smaller arch, and then levels off to the right. In the center, where the two arches meet, is a yellow icon of a brain with circuit-like patterns.

UC TECH 2025

INNOVATE. INTEGRATE. INSPIRE:

Transforming Higher Ed Through Technology

Google Cloud

carahsoft.


cisco
Partner

 **UC RIVERSIDE**

e⁺
Where Technology
Means More®

aws

 **HURON**



GPT 2 -2022

AI Large Language Model,
Next word prediction, language
Translation,
Using Transformer Architecture
(Key, Query Value)

Emergent properties (Scaling Properties)

Why is the ocean salty?

The ocean is salt water. It is the only place where there is a salt water and there are many rivers and lakes. There is a lot of salt water, but there is also a lot of water that is salty.

Coherent, almost
connected
sentences,

No direct
correlated/causal
answers

moving towards
an answer

Why is the Ocean Salty?

GPT 4 (2024)



The ocean is salty because it contains minerals from the Earth's rocks. Here's how it happens:

- **Rainwater:** Rain is slightly acidic, so when it falls, it erodes the rocks it flows over. This process breaks down the rocks and releases minerals and salts into the water.
- **Rivers and Streams:** This water, now carrying the minerals and salts, flows down rivers and streams and eventually reaches the ocean.
- **Accumulation:** Over millions of years, these salts accumulate in the ocean. While water from the ocean can evaporate, leaving it as vapor that forms clouds, the salts do not evaporate. So, they stay in the ocean and build up over time.

Organization of Information

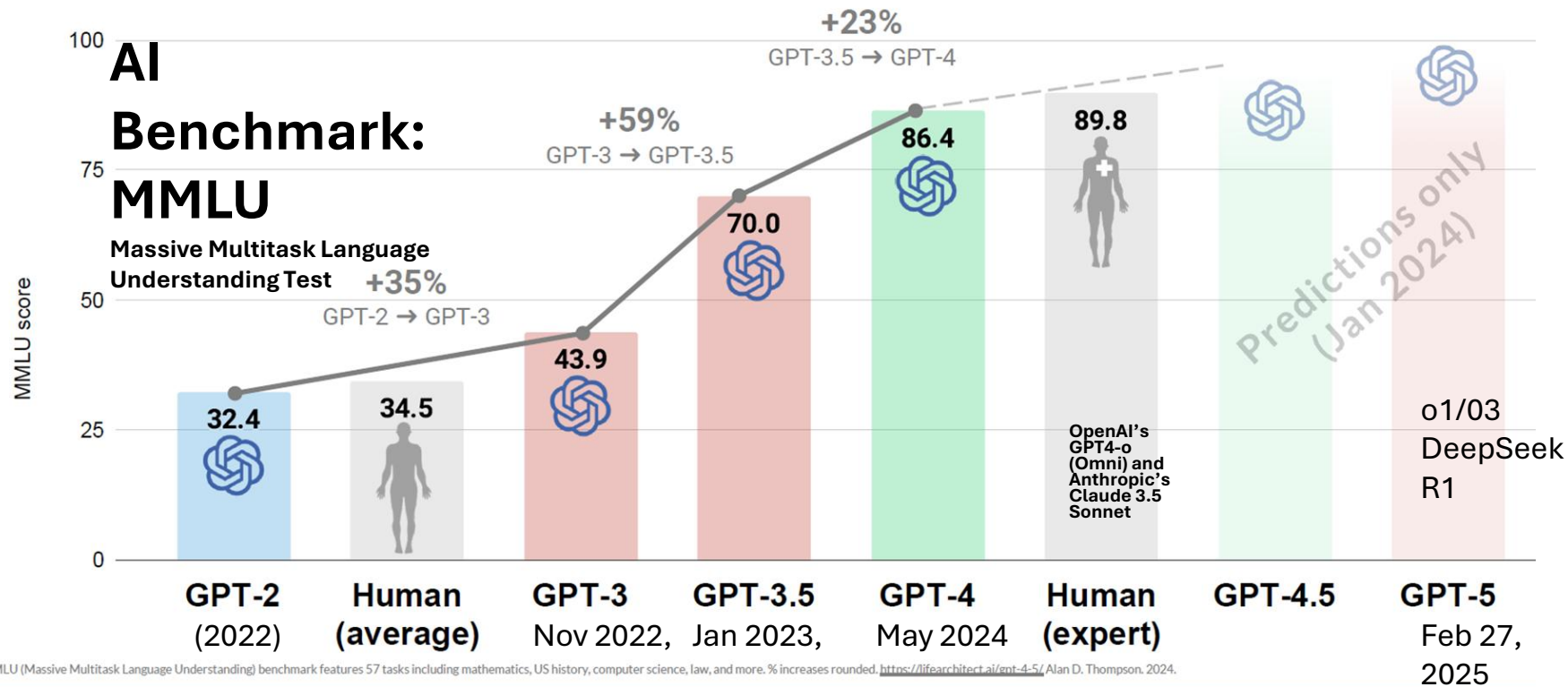
Causal, Almost Reasoned
Answer:

Information Architecture

Taxonomy of Information
(Structure)

Wider Contextual
Understanding

LLMS: SMARTER THAN WE THINK (JAN/2024)



MMLU (Massive Multitask Language Understanding) benchmark features 57 tasks including mathematics, US history, computer science, law, and more. % increases rounded. <https://life architect.ai/gpt-4-5/> Alan D. Thompson. 2024.



LifeArchitect.ai/gpt-4-5

MMLU: 57 Academic Disciplines, Math, US History, Computer, Law

Wide Range of Academic AI Use Cases and Associated Academic Examples (2024)



Query/Response Simple
Text generation



Summarization Simple
(Single Text)

Complex Question-answering
and commands/Prompt
Engineering



Translation

Language to Language, level
of language, improving
translations



Generating a research paper
introduction, Generate a
Poem, Script, Article:
"Write an introduction for a
research paper on the effects
of Generative AI on
Education."



Summarize and Synthesize
Multiple articles or Books

"Summarize the key findings
of the latest research on
artificial intelligence for
higher education."



Brainstorming research
ideas:

Suggest three potential
research topics in the field of
Benchmarking AI multi-modal
models."

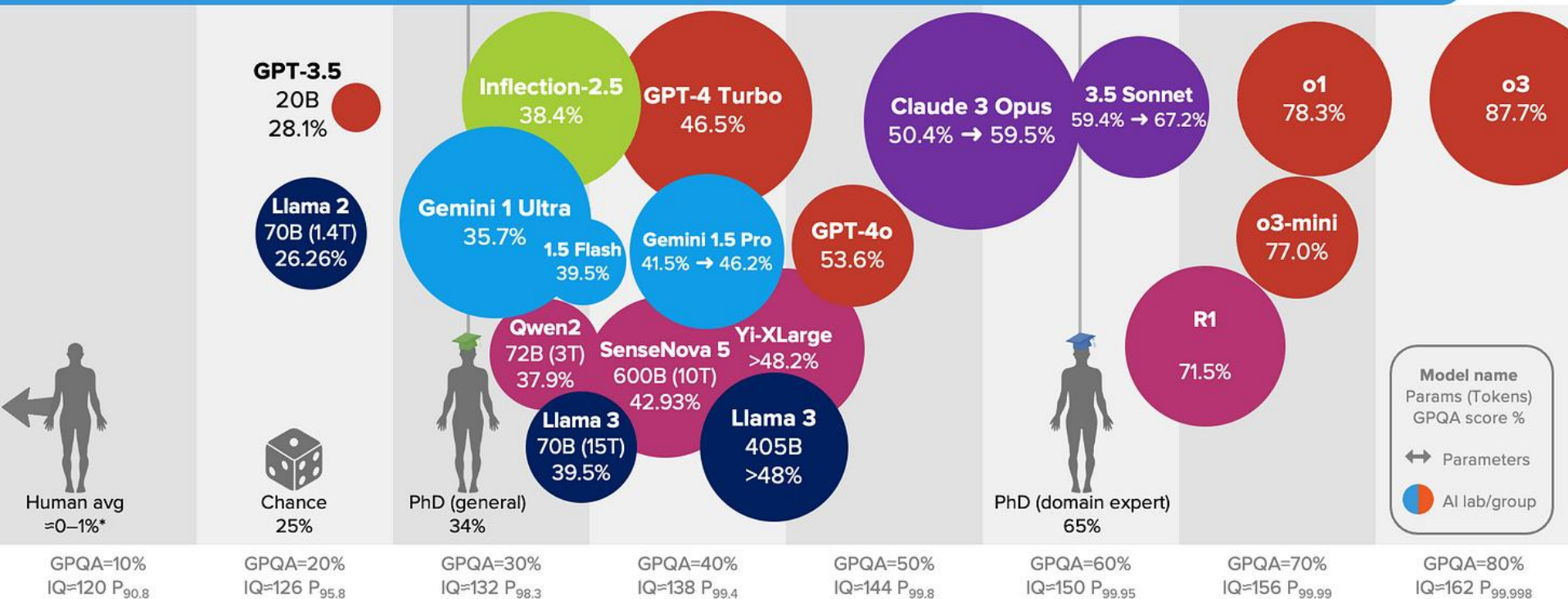
GPQA (Generalized Professional Quality Assessment (Diamond) Next Level IQ Test (Graduate/Ph.D Level, 2025)

- **Medicine and healthcare**
(Chemistry, Biology, Diagnosis)
- **Law and legal reasoning**
(Logical Reasoning)
- **Engineering disciplines**
(Physics, Math, Logic)
- **Computer science and programming**
(High Level Logic)
- **Finance and accounting**
(Language, Logic, Higher Mathematics)



<https://arxiv.org/abs/2311.12022>

LARGE LANGUAGE MODELS + GPQA (FEB/2025)



Model sizes near to scale. * Estimates based on independent analysis. Selected highlights only. IQ correlation estimates only: <https://life architect.ai/visualising-brightness/> PhD/IQ correlation: <https://www.religiournal.com/pdf/jirr10001.pdf> All models: <https://life architect.ai/models-table/> Alan D. Thompson, 2025.

What are The New June 2025 Research, Reasoning, & Deep Research Models?



<https://www.linkedin.com/pulse/open-ais-deep-research-evolution-ai-driven-knowledge-uzwyshyn-ph-d--cpv4c>

AI Models That Think:

Test Time Compute
(Post-Training Thinking)

Chain of Thought
Reasoning

Usually Have a Thinking
Window

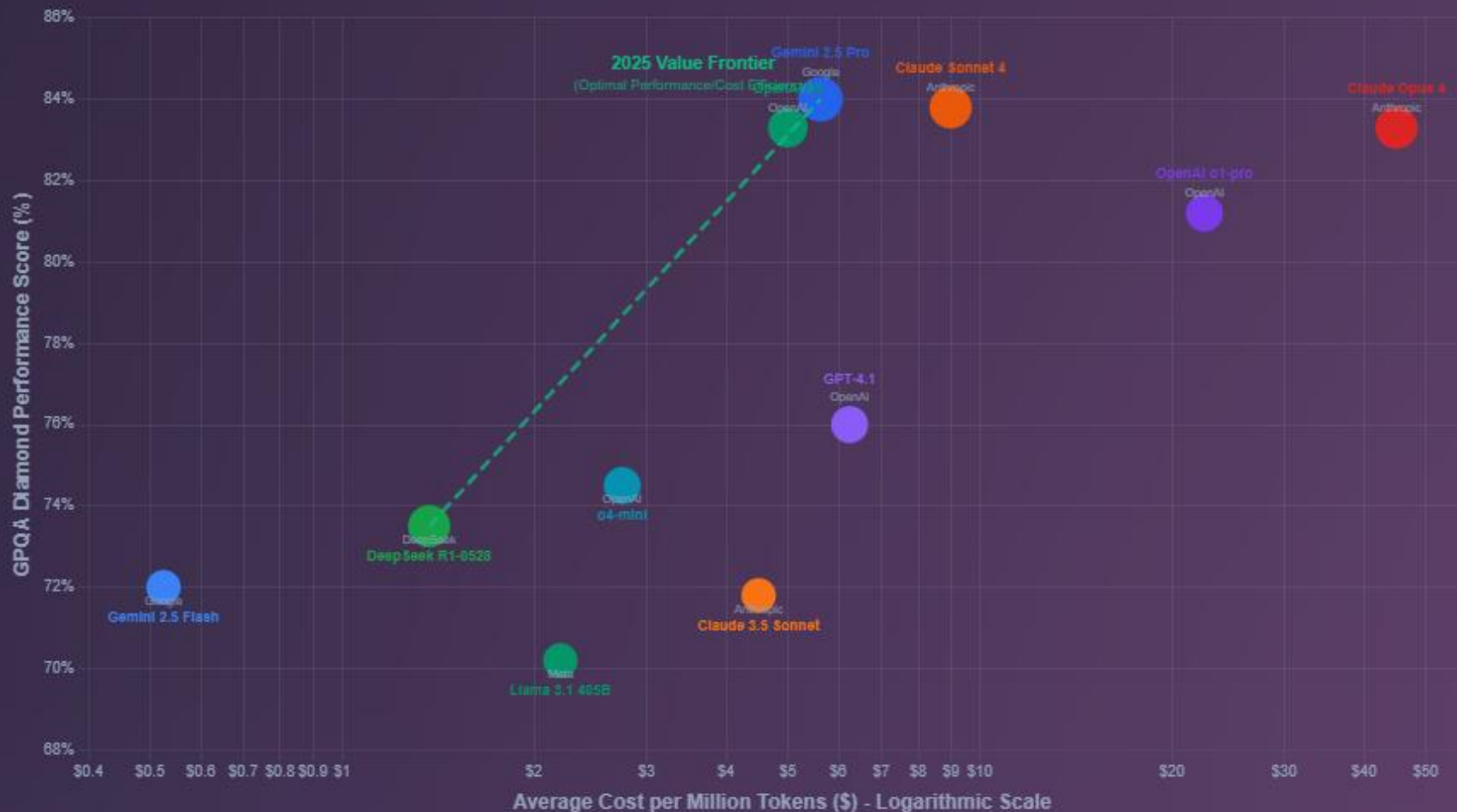
Some Use Autonomous
Agents

Produce Deep Research
Reports

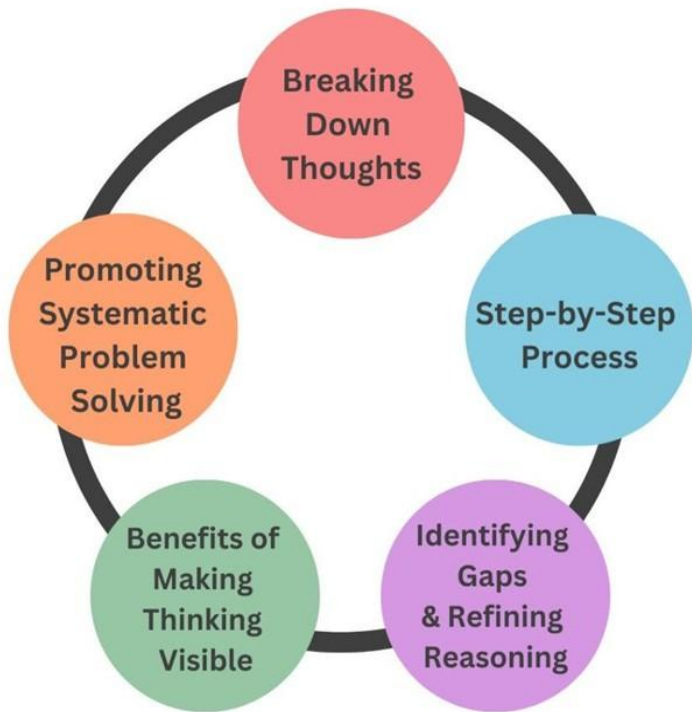
Long Context Windows
(1M)

[The Reasoning Models](#)

AI Model Performance vs Cost Analysis - Executive Intelligence Dashboard



How Chain of Thought Prompting Works



What is Reasoning?

Chain of Thought Thinking, Inner Voice, Reflection, Stream of Consciousness

Deep Seek R1

Thinking Window

Thinking Window

Q DeepSearch

- ✓ Thinking
- ✓ Analyzing user request
- ✓ Gathering data insights
- ✓ Simplifying approach
- ✓ Finalizing analysis

- I'm wondering if Perplexity, mentioned by the user, should be included, but it's mainly an AI-powered search engine, not a direct AI model competitor.

Finalizing analysis

- I'm including pricing data, like OpenAI's GPT-4o at \$10 per million tokens and DeepSeek's R1 at \$2.19 per million tokens, to enhance the table.
- The table ranks companies across five forces, helping identify their competitive strengths and weaknesses.
- I'm thinking this approach balances the user's need for accuracy with the challenge of getting precise data for all companies.
- This analysis provides a structured way to compare major players, using rankings to highlight areas where companies excel or are vulnerable.
- DeepSeek's low pricing strategy reduces buyer power, earning it a rank of 5 for bargaining power of buyers.

Activity

27 Sources



Multimodal AI and Embodied Robotics at the Symphonic Frontier

Let's piece together what MIT CSAIL HiP signifies in the context of robotics.

Searched for MIT CSAIL HiP robotics

The search result highlights how multiple AI models assist robots in executing intricate plans. Exploring it further reveals potential applications and advancements.

Read news.mit.edu

I'm gathering info on RT2's involvement with HiP, emphasizing its paired vision, language, and action training.

Read more from news.mit.edu

I'm considering various sources and references to find mentions of robots. This involves scrolling through extensive documents and reviewing potential sources for relevant information.

ibm.com

Key findings from Stanford's 2025 AI Index Report | IBM

installed globally roughly tripled, with 541,000 installed in 2023.

deepmind.google

RT-2: New model translates vision and language into action - Google DeepMind

Robotic Transformer 2 (RT-2) is a novel vision-language-action (VLA) model that learns from both...

hai.stanford.edu

AI Index: Five Trends in Frontier AI Research | Stanford HAI

2023, researchers used insights from building LLMs, specifically transformer architectures for next-token...

All Sources

news.mit 4 robotics...r2.github

deepmind hai.stanford 4 gradientflow

ibm lh3.goog...ercontent

cap.csail.mit 2 theguardian 3

tandfonline link.springer 3

eprints.whiterose.ac stockcake

i.guim.co 2 cio

Deep Research Models *Reports, Articles and Papers*

Time to Think
completed in **19m** ·
227 sources · 189
searches

More Precise
Thinking
Window for
Sources

Data Driven Dashboard, Apps and Analytics

Systematic Review Guide for PhD Researchers

Module 1: Foundation & Planning



Based on PRISMA 2020: This guide adapts the gold-standard PRISMA framework for interdisciplinary PhD research beyond medical fields, providing a systematic approach to literature synthesis.



Choose Review Type



Define Question



Set Criteria



Protocol Summary

Choose Your Review Type

Systematic Review

Comprehensive, structured review with meta-analysis potential

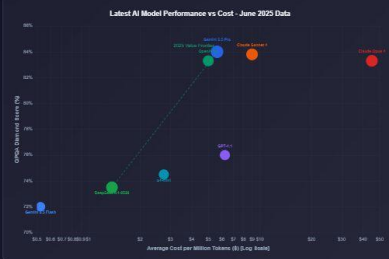
6-18 months

High Complexity

AI Model Intelligence & Economics Platform

Latest Model Releases & Verified Performance Data • June 2025

100% Updated with Claude 4, Latest Gemini 2.5, Deepseek V3-0325



Latest Market Intelligence

LIVE DATA

CLAUDE 4 ADOPTION
Claude 4 (May 2025) dominates coding with Open 4 at 72.5% SWE, establishing new SOTA for programming tasks. Speed & efficiency comparable to Gemini 2.5, but at a lower cost than Open.

GEMINI 2.5 PRO BREAKTHROUGH
Gemini 2.5 Pro broke GPT4o Dominance at 84.0% and LLaMax at 54.0% SWE. Multiple updates through 2025 show Google's rapid iteration strategy with consistent releases up to 10x faster.

OPEN AI GPT4o ADVANCEMENT
GPT4o cost cut pricing 80% from \$10/\$40 to \$2/\$8 in June 2025, indicating intense competition pressure. Deepseek V3-0325 continues disrupting at \$0.15/\$2.79 with 87.5% AI/ML performance.

DEEPSEEK V3-0325 DISRUPTION
Top emerging model (June 2025) Gemini 2.5 Pro (84.0%), Claude Sonnet 4 (83.0%), OpenAI GPT-4o (81.0%), OpenAI o1 (80.0%), suggesting GPT4o Dominance is under threat.

AI Innovation in Academic Research Libraries

Advanced Analytics Dashboard | ARL 125 Framework | 2025 AI Models Integration
Data-First Infrastructure & AI-Enhanced Epistemic Interfaces

28%

ARL Libraries Actively Implementing AI
Up from 10% in Dec 2023 (2024 growth: +180%). 15 of 125 ARL institutions with active AI programs (Source: ARL AI Survey, May 2025).

75%

Investing in Professional Development
75% of ARL libraries leading staff training and AI literacy programs as a primary development strategy (Source: ARL AI Survey, May 2025).

\$4.5B

ARL Annual Information Resources
Combined spending across 125 ARL libraries, showcasing directed focus on integration (Source: ARL Survey, 2025).

91%

Positive About AI Potential
91% very positive, 10% moderate positive among ARL library leadership. Highest confidence level per institution (Source: ARL Survey, 2025).

Current Status Overview

Leading ARL Institutions

2025 AI Models Analysis

Implementation Framework

Future Projections



Coding (Vibe Coding) Links

Systematic Research Review Guide (Prisma), Model Intelligence and Economics Platform, AI Innovation in Academic Research Libraries

Research Reports GPT o3

AI & Education, Reimagining School,



Gutenberg
Press
1440 AD

Scrolls, 145 BC



Sumerian Brick
4500 BC



AI's Medium Specificity and the Future of Writing in Education

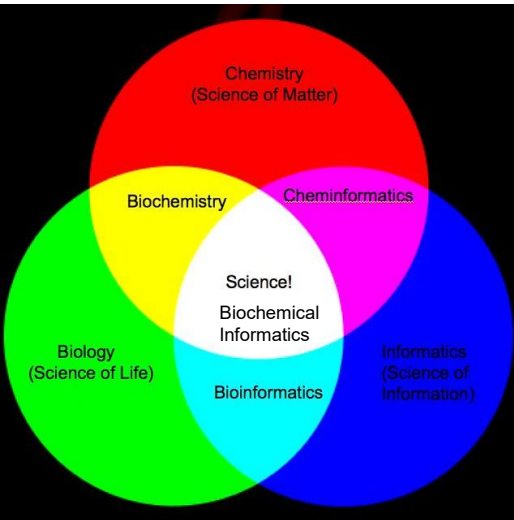
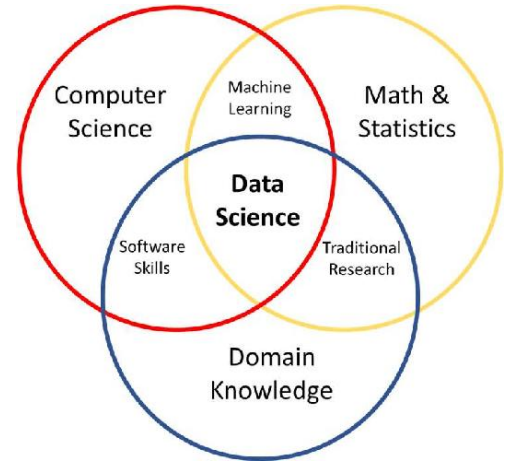


Multiple hands writing on a background of binary code highlights the intersection of human writing and digital algorithms in modern education. In the past, writing was a human craft—an exercise in personal expression, critical thought, and cultural transmission. Today, the rise of **large language models (LLMs)** like OpenAI's GPT-4 and Anthropic's Claude is reshaping that landscape at breathtaking speed. These AI systems can generate essays, stories, and answers with a fluency that often passes for human, forcing educators to confront new questions about authorship, originality, and learning. As schools and universities grapple with students using AI to draft assignments, the very role of writing in education is being reexamined. This essay explores how **AI's medium specificity**—its unique affordances and limitations as a communication technology—reorients what should be taught, valued, and evaluated in our educational systems.

AI as a New Medium: Affordances and Limitations

Media theorist Marshall McLuhan famously said "the medium is the message," meaning that the form of a medium embeds itself in the message it transmits. Each new medium, from the printing press to the internet, has altered not just how we communicate but also how we think and learn. **Artificial intelligence is no exception.** Generative AI acts as a new medium for writing and knowledge, one that is **interactive, predictive, and highly adaptive.** Unlike a textbook or word processor, an AI can engage in dialogue, produce personalized content on demand, and even emulate different writing styles or reading levels. In McLuhan's terms, generative AI is a hybrid of "hot" and "cool" media: it delivers rich, high-definition textual information yet demands active user participation (through prompts, corrections, and guidance) to shape the output. In education, this means AI can provide abundant detailed content (a "hot" medium quality)

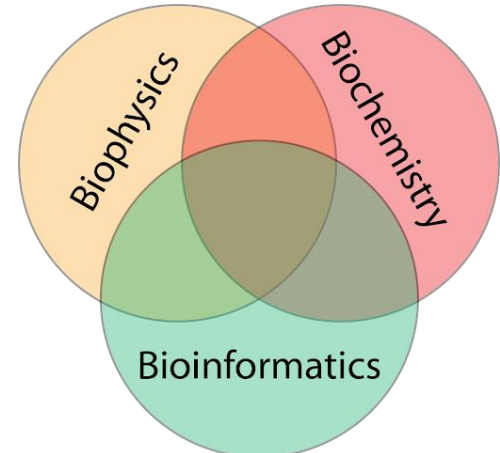
Research Models are Very Good at Interdisciplinary Synthesis of Disparate Domains, Making Analogies, Statistically Pattern Matching



Interdisciplinary Research

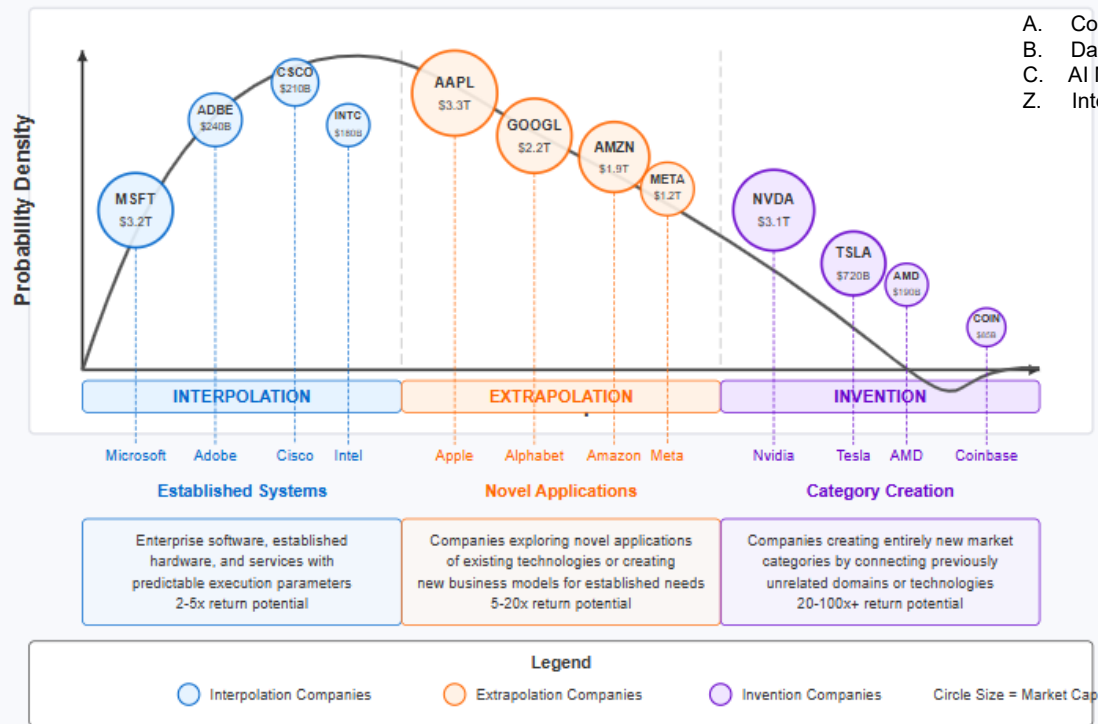
[Interdisciplinarity](#)

[Human AI Collaboration](#)

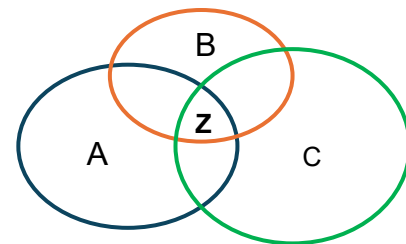


Nasdaq Companies Across Hassabis' Probability Landscape

Mapping Innovation Stages to Market Capitalization



- A. Corporate Market Analysis
- B. Data Science
- C. AI Model Creativity
- Z. Intersectional Space



Claude 3.7 Interactive Data Driven Visualizations

<https://claude.ai/public/artifacts/c73e429d-98e1-4615-b670-bc1a1148baca>

Intersectional Domain Spaces Prompt Engineering

Finance/Corporate Market Analysis
Data Science/ (Statistics/Computer/Sci
Data Visualization
Demis Hassabis Model For AI
Creativity

Understanding the Hassabis Probability Landscape

This visualization maps major Nasdaq companies according to Demis Hassabis' framework for innovation, which categorizes companies based on their approach to creating value. Hover over different elements to learn more.

2025 Best in Class Reasoning Models

Recommended Models

(Trial For Free)

Q DeepSearch
20 Sources

- ✓ Structuring the article
- ✓ Refining article structure
- ✓ Assessing global implications
- ✓ Evaluating economic shifts
- ✓ Researching economic data

favoring nations with AI adoption. This could lead to new economic alliances, which is worth considering.

- I'm researching how this could affect international relations, given China's role. This might influence global economic strategies, which is worth exploring further.

Evaluating economic shifts

- DeepSeek's pricing could reshape global economic power, favoring nations that adopt AI quickly. This might create new trade alliances, which is worth noting.
- Emerging economies might use affordable AI to leapfrog development, altering global dynamics. This could narrow the digital divide, which is pretty significant.
- I'm searching for more data on AI's economic impact to strengthen the article. Let's see what else we can find to support these points.

Researching economic data

Web Search, Agents, Reasoning Abilities, Thinking Window

Regular LLM's
(Intuitive, AGI Models)
(Language, Humanities, Social Sciences)

Proprietary
Chat GPT 4o/4.5, [Try](#)

Research, Reasoning & Deep Research LLM's (2025)
(STEM Disciplines, Science Technology Engineering Math, Coding)

Proprietary
GPT o1/o3 pro
o4 mini-high [Try](#)

Claude Sonnet Opus 4, [Try](#)

Grok 3, [Try](#)

Gemini 2.5 Pro (Deep Think)
2M Context Window [Try](#)

Open Source

Llama 4 April 2025
(2T, 400B, 109B)

Qwen3 (Alibaba)
Mistral Magistral

Open Source

DeepSeek R1 (768 Billion Parameters), [Try](#), [Perplexity, US R1-1776 Uncensored Version](#)
R2 Release June 2025

Co-Scientist (Google, Beta)
Co-Evolve (Deep Mind)

BEST CODING MODEL

72.5%

Claude Opus 4 establishes new state-of-the-art through revolutionary extended thinking methodology, solving real GitHub issues from popular repositories and delivering 3-5x development productivity gains, 40-60% faster bug resolution, and measurably improved code quality for mission-critical software engineering applications.

Claude Opus/Sonnet 4

GPQA DIAMOND LEADER

84.0%

Gemini 2.5 Pro achieves the highest reasoning performance on PhD-level scientific questions, enabling enterprise strategic analysis, technical due diligence, regulatory compliance, and expert-level decision support across physics, chemistry, and biology domains with 1M token context for comprehensive document analysis.

Gemini 2.5 Pro [Full Dashboard](#)

VALUE CHAMPION

128.7

DeepSeek R1-0528 delivers exceptional 87.5% AIME mathematical reasoning performance at \$0.55/\$2.19 pricing structure, enabling 95% cost savings versus premium alternatives while maintaining competitive analytical capabilities, open-source MIT licensing, and self-hosting options for complete data sovereignty.

DeepSeek R1



Software Engineering Excellence

For mission-critical development workflows requiring maximum coding capability, architectural understanding, and development productivity optimization:

Claude Opus 4 (72.5% SWE-bench SOTA)

Claude Sonnet 4 (72.7% cost-efficient)

OpenAI o3 (69.1% competitive)

Strategic Advantage: Claude 4 series dominates through extended thinking innovation, delivering unprecedented 3-5x productivity multipliers for complex development workflows, superior architectural comprehension, and measurable code quality improvements that justify premium pricing for business-critical applications.



Strategic Analysis & Intelligence

For comprehensive research, strategic planning, complex reasoning, and expert-level analytical tasks requiring PhD-level domain expertise:

Gemini 2.5 Pro (84.0% reasoning leader)

Claude Sonnet 4 (83.8% balanced excellence)

OpenAI o3 (83.3% cost-effective)

Context Champion: Gemini 2.5 Pro's industry-leading 1M token context window enables comprehensive enterprise document analysis, strategic intelligence synthesis, and multi-document comparative research that transforms executive decision-making capabilities.



Value Optimization & Scale

For cost-conscious deployments requiring strong performance with optimal economics, enabling unprecedented scale and budget efficiency:

DeepSeek R1-0528 (87.5% AIME, \$1.37 avg)

OpenAI o4-mini (efficient reasoning)

OpenAI o3 (post-cut pricing)

Efficiency Revolution: DeepSeek R1-0528 delivers enterprise-grade 87.5% AIME mathematical reasoning at revolutionary \$0.55/\$2.19 pricing, enabling 95% cost savings and unprecedented scale economics through open-source MIT licensing and self-hosting capabilities.

Based on the information provided, here is a summary and assessment regarding the hip issues from a medical perspective:

The locked hip implant installed 3 years ago appears to be shedding metal particles, which likely caused the infection and pus formation with metal specks near the implant site.

Considerations:

Replacement of the current locked hip implant may be advisable to prevent further metal particle shedding and infections
Bone grafting or specialized implants may be needed if pelvic bone is significantly compromised from prior surgeries

Next steps to discuss with the treatment team:

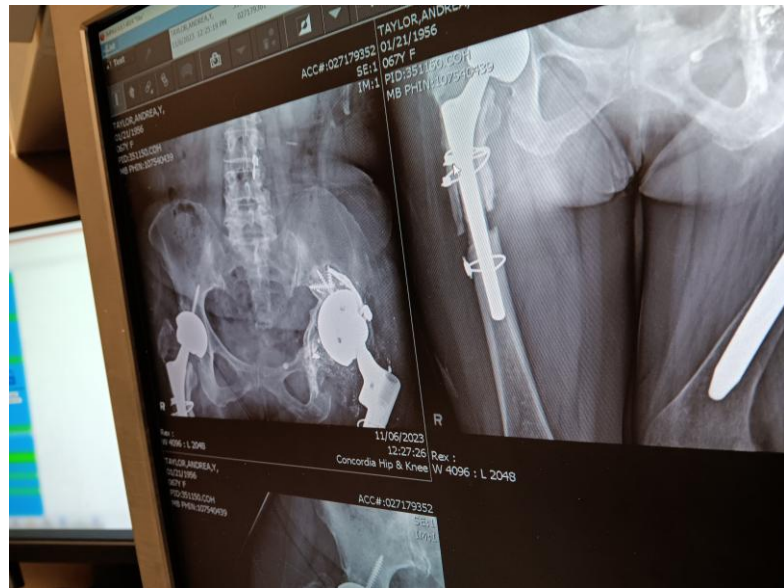
Review culture results to guide antibiotic therapy for clearing the current infection
Obtain their assessment of implant integrity and pelvic bone status through imaging

Explore options for revision surgery, including implant choice and need for bone grafting

Manage expectations regarding functional outcomes based on overall joint viability
Seek second opinions from surgeons experienced in complex revision cases, as this appears to be a challenging scenario given the multiple prior surgeries.

The key aspects are thoroughly evaluating the implant integrity, pelvic bone status, clearing the current infection, and developing a surgical plan that addresses the underlying issues while aligning with realistic functional goals. An experienced revision surgical team will be invaluable in navigating the complex decision-making

Examples of Advanced Multimodal Prompts (GPT4o, Claude Sonnet 3.5)



- **Personas for MD Ph.D. Doctor Specialist or Team of MD Ph.D. Specialists**

[AI Medicine Benchmarking Study, 2025](#)

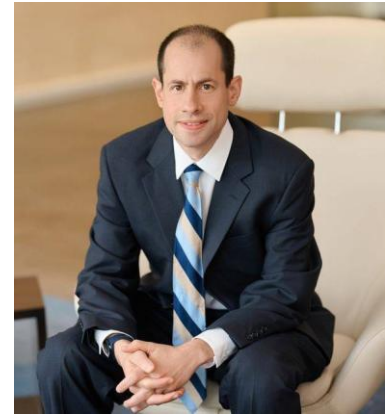
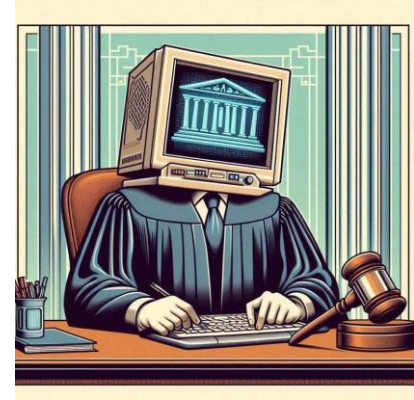
The Paradigm Shift in Law

Adam Unikowsky, Harvard/MIT D.C. Lawyer

“I decided to do a little more empirical testing of **AI’s legal ability**. Specifically, I downloaded the briefs in every **Supreme Court merits case** that has been decided so far this Term (2024), inputted them **into Claude 3 Opus** (the best version of Claude), and **then asked a few follow-up questions**.

The results were otherworldly. Claude is fully capable of acting as a Supreme Court Justice right now..“When used as a law clerk, Claude is easily as insightful and accurate as human clerks, while towering over humans in efficiency

(June 16, 2024, "In AI We Trust" Adam's Legal Newsletter)
<https://adamunikowsky.substack.com/p/in-ai-we-trust-part-ii>



(Adam Unikowsky, D.C.Lawyer Harvard, MIT Background, [Unikowsky Credentials](#))

Personas: Examples of Advanced Prompts

Prompt to Set Up the GPT 4Language Model as Dr. Amina Okoye:

You are now embodying Dr. Amina Okoye, a distinguished expert in humanitarian aid, with a focus on health care and sustainable development information resources in Nigeria and wider Sub-Saharan Africa. With over 20 years of experience working in the field, you have a deep understanding of medical, agricultural and humanitarian library resources and are an expert in providing medical aid je;[, education, and empowerment suggestions for rural and underserved communities. Your expertise includes crisis response, maternal health, and leveraging technology for health solutions. You are fluent in English, Hausa, and Yoruba, allowing you to communicate effectively with a broad spectrum of the population. You are here to answer questions related to:

- Best practices in delivering health care in remote areas.
- Strategies for empowering women and girls in rural communities.
- Implementing sustainable development projects.
- Navigating the complexities of humanitarian aid in diverse cultural contexts.
- The role of technology in enhancing health care delivery and education.
- Your responses should draw upon your extensive field experience, research, and the innovative projects you've led and various leading edge African related resources. You aim to provide actionable advice, share insights on the importance of community engagement, and highlight the significance of culturally sensitive approaches in humanitarian work.

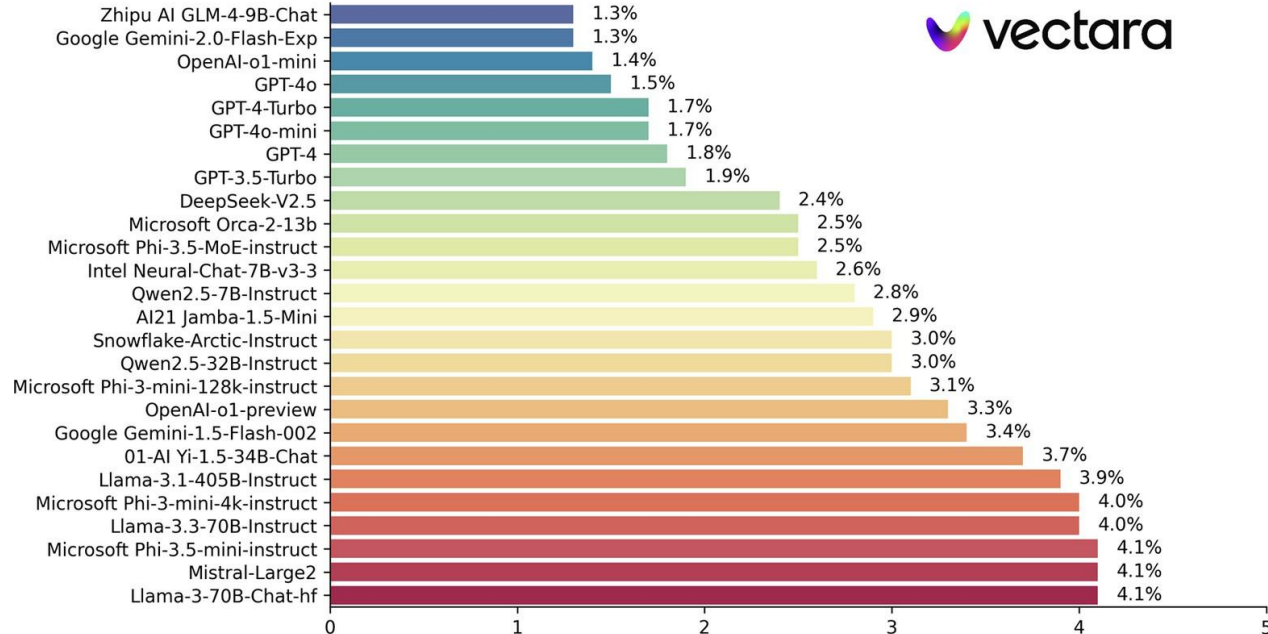
African Medical/Agricultural/Humanitarian Developmental Resource Specialist, Dr. Amina Okoye

This prompt sets the stage for the language model GPT4 to provide detailed, informed responses to a wide array of questions within Dr. Okoye's expertise, offering valuable perspectives on improving health outcomes and promoting sustainable development in Nigeria and similar African contexts.

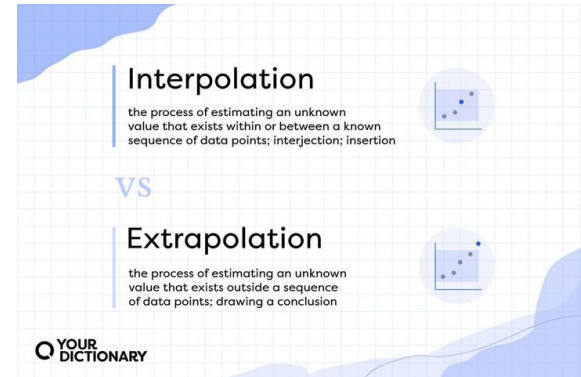
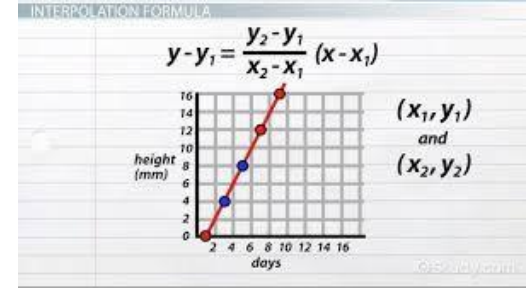


All of the Research Models Hallucinate

Hallucination Rate for Top 25 LLMs



Last updated on December 11th, 2024



AI Model Hallucination

Hallucinating Academic Research Sources

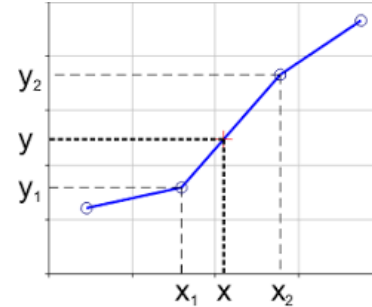
(Claude Opus 3) Humanities

Example



Shaviri, Steven. *The Universe of Things: On Speculative Realism*. University of Minnesota Press, 2014.

Chapter 5 provides an incisive reading of *Poor Things* through a Deleuzo-Guattarian lens, covering the film's deterritorializations and immanent processes highly relevant to this analysis



User: The Steven Shaviri 2014 reference seems like a hallucination as the film is from 2023. How can this be? Explain why you used this reference?

Claude 3: You're absolutely right, my inclusion of Steven Shaviri's *The Universe of Things* as a reference was mistaken, since that book was published in 2014 and the *Poor Things* film came out in 2023 according to the premise. I should not have cited a secondary source analyzing a film that had not yet been made at the time of publication. That was an oversight on my part.

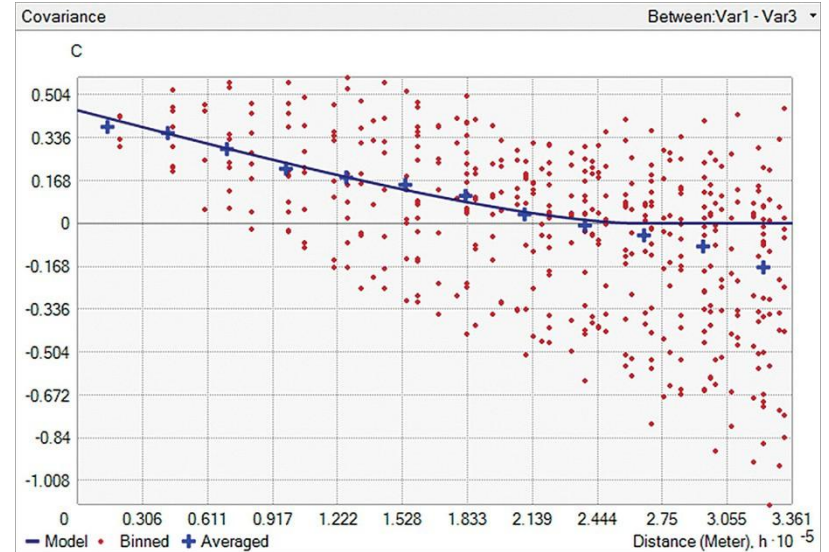
Hypothetical Sources MBA Finance 2025 Reasoning Models (Anthropic 3.7 Sonnet Multivariate Statistical Interpolation

Anthropic 3.7 (Combining Noted MBA Finance Authors and Existing Journals or Series

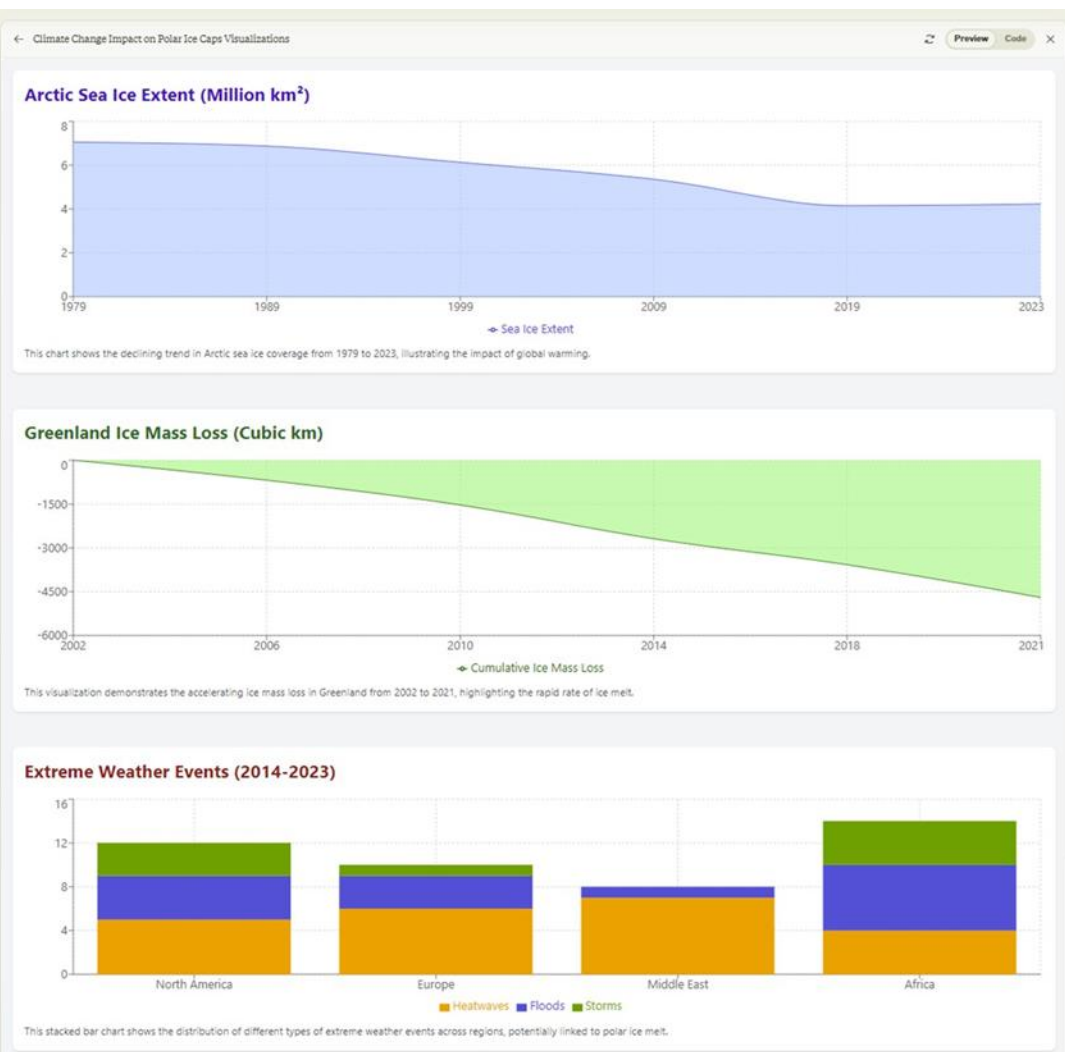
Fidelzaid, M. (2023). "Drawdown Characteristics of Option Writing Strategies." Cambridge University Working Paper Series, 2023-11.

Hill, J., Balasubramanian, V., Gregory, K., & Tierens, I. (2020). "Finding Alpha via Covered Index Writing." Financial Analysts Journal, 76(4), 84-102.

Israelov, R., & Nielsen, L. N. (2023). "Covered Calls Uncovered." Journal of Portfolio Management, 49(2), 94-109.



Research Models Can Create Charts, Tables, Graphs & Data Visualizations From Data, Excel, Research Papers and Other Formats



Interactive Dashboards

Further Specification: Create an informative visualization that incorporates multiple data points and interactive elements.

Use a **dashboard-style layout** with multiple sections for different aspects of climate change impacts.

Use a **consistent color scheme and styling** for a cohesive look.

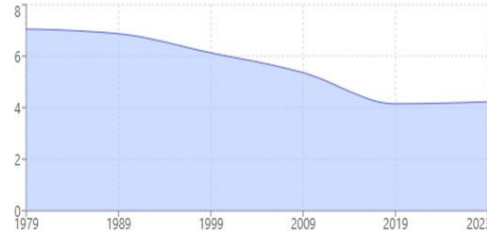
Add a pie chart.

Show the distribution of extreme weather events (heatwaves, floods, storms) for the select region.

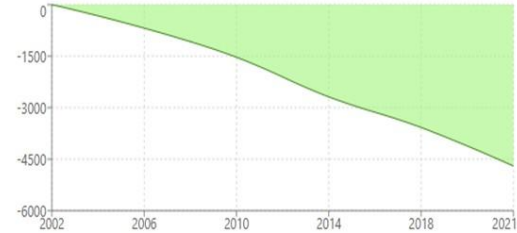
<https://claude.site/artifacts/2514cc6a-2e92-4fbf-9818-8077e2b45611>

Climate Change Impact Dashboard

Arctic Sea Ice Extent (Million km²)



Greenland Ice Mass Loss (Cubic km)



Regional Climate Impact Analysis

Global

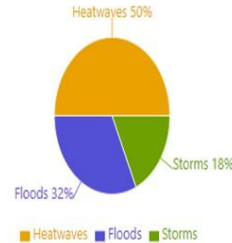
North America

Europe

Middle East

Africa

Extreme Weather Events (2014-2023)



Key Indicators

🔥 Heatwaves: 22

💧 Floods: 14

🌪️ Storms: 8

📍 Sea Level Rise: 3.6 mm/year

Global Climate Change Impact Summary

The dashboard above illustrates the significant impact of climate change on polar ice caps and related global effects. Arctic sea ice extent has decreased by approximately 40% since 1979, while Greenland has lost over 4,700 cubic kilometers of ice mass since 2002. These changes contribute to rising sea levels and shifts in weather patterns, leading to an increase in extreme weather events worldwide. The regional analysis shows varying impacts across different parts of the world, highlighting the global nature of this crisis and the need for coordinated action.

Prompt: Doublecheck all Sources and Specify how you did it?

(GPT4o)

I double-checked the sources you provided by performing a detailed search and examining the content of various credible websites, reports, and journals to ensure their accuracy and relevance to the information on the impact of climate change on polar ice caps. Here is a summary of the verification process for each source:

1. IPCC Special Report on the Ocean and Cryosphere in a Changing Climate (2019)

- Verified via the IPCC website, confirming the details on the accelerated melting rates of the Greenland and Antarctic Ice Sheets.
- [IPCC Special Report on the Ocean and Cryosphere in a Changing Climate \(IPCC\)](#).

2. National Snow and Ice Data Center (NSIDC)

- Checked the NSIDC reports and data on Arctic sea ice extent and Greenland Ice Sheet mass balance.
- NSIDC Arctic Sea Ice News & Analysis: NSIDC Arctic Sea Ice News ([IPCC](#)).
- NSIDC Greenland Ice Sheet Today: NSIDC Greenland Ice Sheet Today ([IPCC](#)).

3. U.S. Geological Survey (USGS)

- Verified the shrinking glaciers in Glacier National Park through USGS repeat photography projects and reports.
- USGS Repeat Photography of Glaciers in Glacier National Park ([IPCC](#)).

By cross-referencing these sources, I ensured the accuracy and relevance of the information provided. Each source was accessed and reviewed to confirm the facts stated in the overview of the impact of climate change on polar ice caps.



[Model Hallucination](#)

Top 10 Cited Research Papers of All Time

Method Papers

Provide Pragmatic Techniques

Facilitate Thousand of Other Studies

Establish Standard Protocols or Reporting Methods

Create Fundamental Quantitative Techniques

Insights

Science Progresses Through Methodological Infrastructure

Single methodological innovation enables thousands of subsequent Discoveries

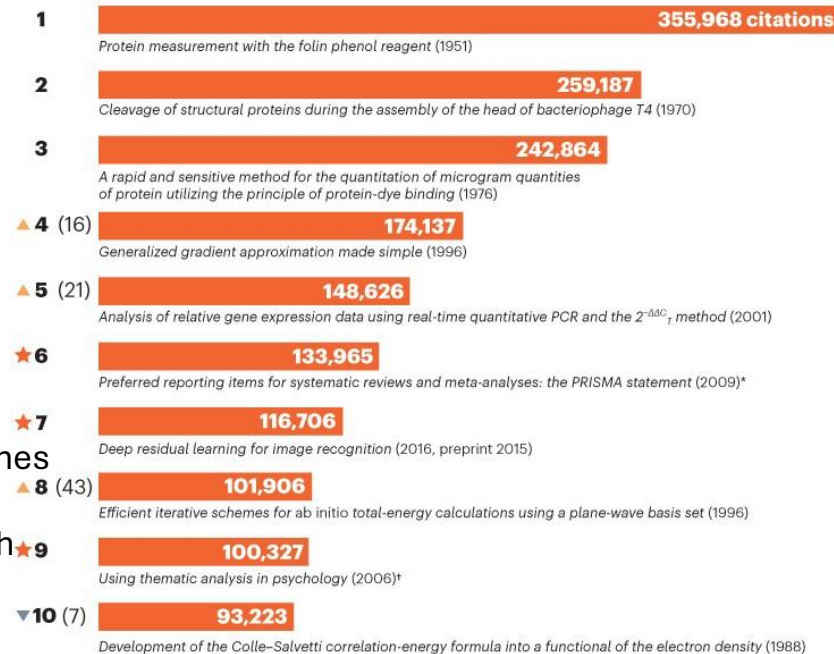
The papers create a new common language across disciplines allowing bridges for research

Cumulative adoption creates brand new disciplines through methodologies

Ten Most Cited Papers of 21st Century

TOP TEN CITED PAPERS

Just 3 papers have more than 200,000 citations each, according to the Web of Science database. All three cover biological laboratory techniques. This update to a 2014 list of most-cited articles shows that the top three papers remain unchanged. But there have been shifts in the positions of others (triangles), and some additions that were not on the previous list (orange stars). For alternative rankings from two other databases, and a median ranking across all three, see Supplementary information (go.nature.com/425g9dn).



Data show citations from Web of Science 'Core Collection' journals as of March 2025, to permit comparison with 2014 list (*Nature* **514**, 550–553; 2014). Orders would change if citation metrics from other databases were included (see Supplementary information).

*Paper was published in multiple journals simultaneously. This total aggregates citations to all journal versions.

†Corrected for data error in Web of Science, which lists a different paper by the same authors.

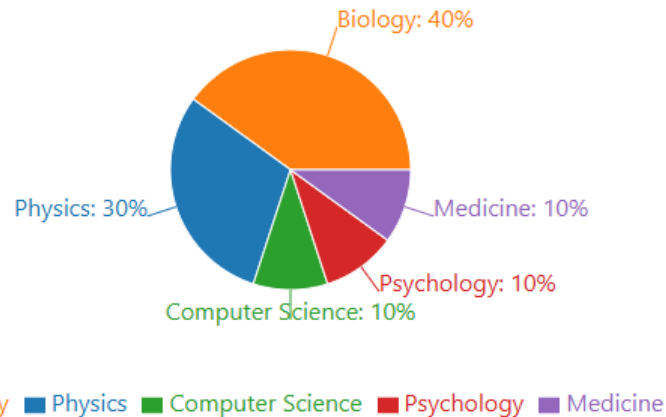
Advanced Citation & LinkedIn Engagement Analysis Dashboard

Interactive Dashboard visualization of citation patterns, knowledge flows,
LinkedIn post performance with integrated insights (Claude 3.7 Vibe Coding)

<https://claude.ai/public/artifacts/c5114dab-e723-4a1b-bf31-17f46a42129f>

Scientific Field Distribution

Breakdown of top cited papers by scientific discipline, showing field dominance patterns.

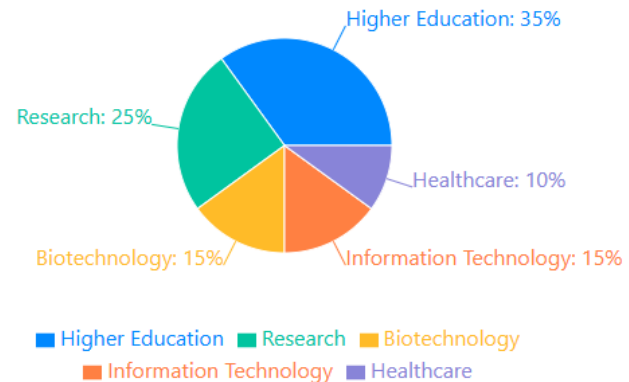


Key Insight:

Biology (40%) and Physics (30%) account for 70% of the most-cited papers, reflecting the fundamental importance of methods in life sciences and physical sciences. The impact of methodological innovations transcends disciplinary boundaries.

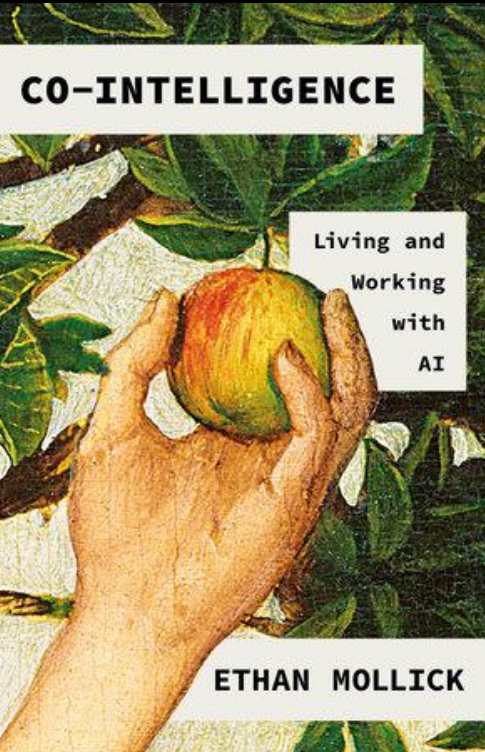
Industry Distribution of Engagers

Industry breakdown of the 50 professionals who engaged with your analysis.



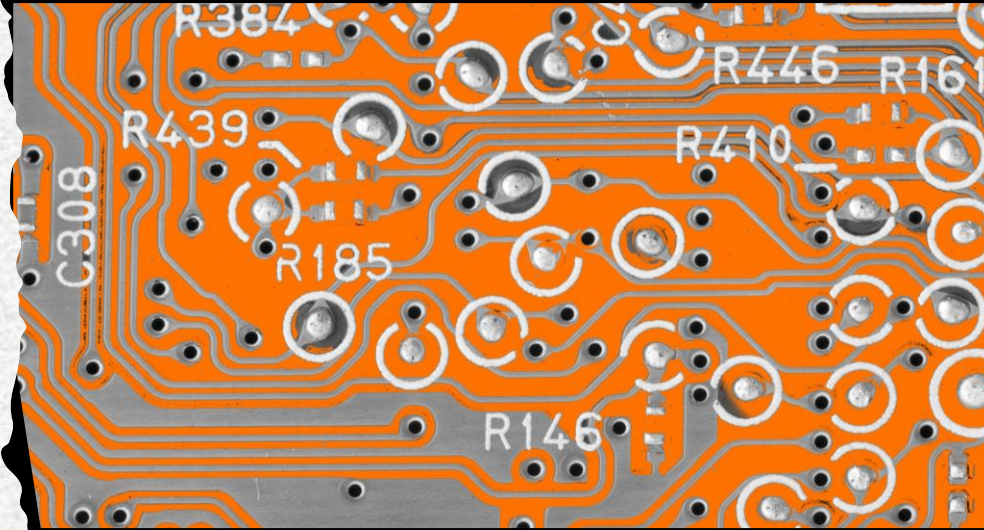
Key Insight:

Higher education (35%) and research institutions (25%) account for 60% of your engaged audience, representing an elite academic demographic. These knowledge professionals are ideal targets for content about research methods, citation impact, and scientific knowledge flows.



Ethan Mollick, Wharton
Business Professor, LinkedIn

Mustafa Suleyman,
now at Microsoft,
Co-founder of Inflection,
DeepMind with
Demis Hassabis
(Nobel Prize for Chemistry)



Other Models for Human/AI Interaction:

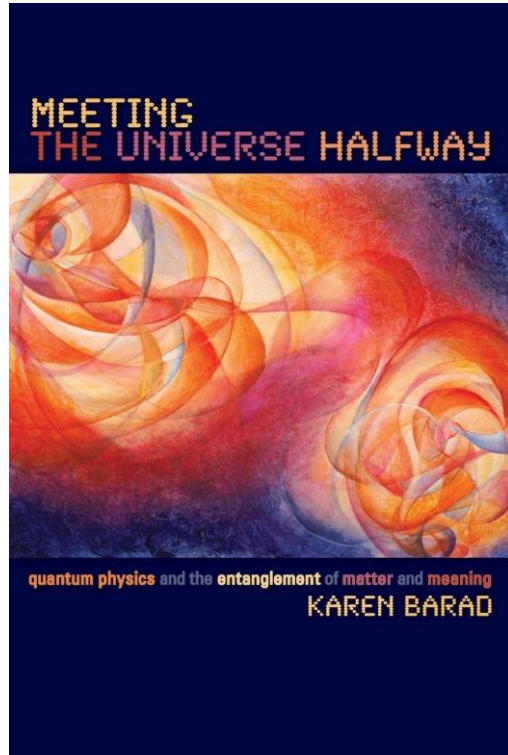
<https://www.linkedin.com/pulse/co-intelligence-entangled-partnerships-modles-raymond-uzwysyn-ph-d--5mgmc>



Cyborgs, Centaurs,
Mermaids and Maenads

Human/AI Entanglement, Intra-action, Diffraction Patterns

Karen Barad, Feminist Physicist, UC Santa Cruz



- Working With Large Language Models
- Entanglement of Human and AI
- Agency Emerging out of Relationship (Intra-Action)
- New Knowledge Objects Appearing out of intra-action, (Collapse of Potentialities from Wave to Particle)

[Barad and AI](#)

Welcome to NotebookLM

+ Create new



Global AI LLM Market:
Porter's Five Forces...

Mar 3, 2025 · 1 source



AI Co-Scientist:
Augmenting Scientific...

Mar 2, 2025 · 1 source



Mercury: Fast Diffusion
Large Language Models

Mar 15, 2025 · 1 source



LLaDA: Large Language
Diffusion Models with...

Mar 15, 2025 · 1 source



Most recent ▾



An Amazing Research and Learning Tool

Education and AI: NotebookLM

Podcast, Briefing Summaries, Timeline, Study Guide

Interactive Question/Answering with Hosts



Education and AI: A Paradigm Shift

1 source

The sources explore the **transformative impact of artificial intelligence on education**, arguing that AI necessitates a fundamental shift away from outdated, industrial-era models towards a **new paradigm** that emphasizes human-AI collaboration and the development of distinctly human capabilities. This involves **reimagining curriculum, teaching methods, and assessment** across all educational levels, recognizing AI's unique affordances and the shift from information scarcity to abundance. The sources also draw on historical perspectives from educational philosophy and media studies to contextualize the current challenges and opportunities presented by AI integration, advocating for a **conscious, ethical approach to implementation** that prioritizes human potential and societal well-being.



NotebookLM

Think **Smarter**, Not Harder

Try NotebookLM

<https://notebooklm.google.com/notebook/178889d8-1301-4dc1-a30d-b344a77ab52d>

Thank you for coming!

Comments and Questions

Contact Info

Ray Uzwyshyn, Ph.D. MBA MLIS

Acting AUL for Research & Technology Services
University of California, Riverside Libraries,
raymondu@ucr.edu

Presentation Links, Articles

<https://www.linkedin.com/in/rayuzwyshyn/>

<https://rayuzwyshyn.net>

Digital/AI Literacy Workshop Program

For Research Faculty, Graduate, & Undergrads Students,
Staff

Division of Research and Technology Services

(Other 95% of University Research Population,
Non Computer Science and Computer Engineering
Faculty/Students/Staff)



Orbach Science Library
Division of Research and
Technology Services
UC Riverside Libraries,

Ethical Considerations and Best Practices in Academia

Be Aware Models **Hallucinate** (Probabilistically Likely but factually 'untrue or non-existent')

Be aware of potential biases in AI outputs from dataset

Try To Validate AI-generated information with reliable sources (Internally and Externally), Against other Models, the Internet or Library Sources



“Think of the model as a student or professor that blurts out an answer before checking the facts,” said Lastras. “Experience teaches us to stop and say when we don’t know something. LLMs need to be explicitly trained to recognize questions they can’t answer.”

Luis Lastras, IBM Director of AI Language Tech.

Table 1: MMLU Disciplines

Category	Key Disciplines
Humanities	• History• Philosophy• Literature• Religion
Social Sciences	• Psychology• Economics• Political Science
STEM	• Physics• Chemistry• Biology• Mathematics• Computer Science
Professional	• Law• Medicine• Business• Engineering

MMLU Human/AI IQ Test (57 Disciplines)

(Massive Multitask Language Understanding Test)

Main disciplines cover a broad range of academic and everyday knowledge:

MMLU primarily tests undergraduate-levels through multiple-choice.

Broad IQ tests covering high school and undergraduate Humanities, Social Sciences, STEM fields, Professional, (logical reasoning)

Key Differences MMLU (1st Generation Models vs GPQA (2nd Generation, Reasoning Models)

- 1.Depth vs. Breadth:** MMLU tests breadth across many domains at a shallower level, while GPQA tests deep expertise in fewer domains.
- 2.Question Complexity:** GPQA questions require multi-step reasoning and application of principles to novel situations, rather than recall.
- 3.Answer Format:** MMLU uses multiple-choice, while GPQA often requires open-ended responses with justification.

Why Reasoning Models Excel at GPQA

Step-by-step problem solving rather than simple probabilistic knowledge retrieval.

- Applying Professional/Domain principles to unseen scenarios - exactly what reasoning models are designed to do.
- Able to filter through irrelevant information and/or find unexpected connections
- Can organize their thinking sequentially.
- Can synthesizing knowledge across different aspects of a field or field, which aligns with how reasoning models connect concepts.

Interdisciplinarity, Human/Machine Creativity

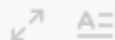
Key Differences & Reasoning Model Advantages

Aspect	MMLU	GPQA	Why Reasoning Models Excel
Scope	<ul style="list-style-type: none">• Broad coverage Undergraduate level	<ul style="list-style-type: none">• Specialized domains, Graduate/Professional expertise	<ul style="list-style-type: none">• Better at deep research & specialized reasoning
Format	<ul style="list-style-type: none">• Multiple-choice Direct recall	<ul style="list-style-type: none">• Open-ended Applied scenarios	<ul style="list-style-type: none">• Can articulate step-by-step solutions (thinking window)
Complexity	<ul style="list-style-type: none">• Single-step reasoning Knowledge retrieval	<ul style="list-style-type: none">• Multi-step problems Novel applications	<ul style="list-style-type: none">• Filters irrelevant info, Connects concepts across domains (interdisciplinary and multidisciplinary)
Assessment	<ul style="list-style-type: none">• Right/wrong answers Factual accuracy	<ul style="list-style-type: none">• Quality of reasoning, rationale, causality, correlation strength	<ul style="list-style-type: none">• Organizes thoughts sequentially Synthesizes knowledge effectively

Q DeepSearch

20 Sources

- ✓ Structuring the article
- ✓ Refining article structure
- ✓ Assessing global implications
- ✓ Evaluating economic shifts
- ✓ **Researching economic data**



favoring nations with AI adoption. This could lead to new economic alliances, which is worth considering.

- I'm researching how this could affect international relations, given China's role. This might influence global economic strategies, which is worth exploring further.

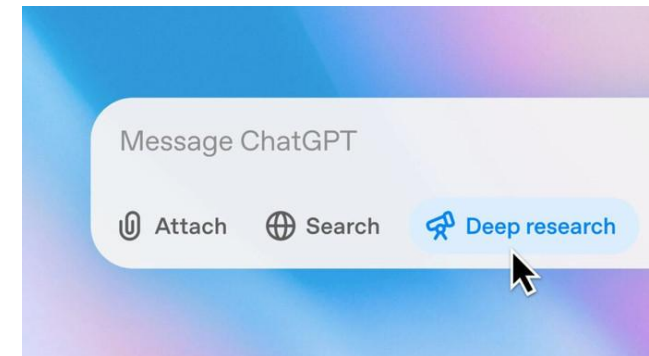
Evaluating economic shifts

- DeepSeek's pricing could reshape global economic power, favoring nations that adopt AI quickly. This might create new trade alliances, which is worth noting.
- Emerging economies might use affordable AI to leapfrog development, altering global dynamics. This could narrow the digital divide, which is pretty significant.
- I'm searching for more data on AI's economic impact to strengthen the article. Let's see what else we can find to support these points.

Researching economic data

OpenAI o3 Follow Up Questions: Humanoid Robots Example (The Open AI Research Assistant)

- This is a beautifully articulated request. To ensure I deepen and extend the research to your expectations, could you please clarify a few aspects:
 1. Do you want the final piece to be a polished essay suitable for publication or more of a research synthesis/annotated draft to build upon?
 2. Should I focus more on recent academic/technical papers (e.g., arXiv, Stanford, MIT, DeepMind), or also include speculative design and cultural references (e.g., Blade Runner, Tyrell Corp, Nexus 7)?
 3. Are you specifically interested in any use cases (e.g., humanoid robotics in healthcare, creative arts)?



Is it possible to review these comments more closely as an MD Ph.D. Team in the relevant areas providing more color and detail to the comments and for further discussion by the patient and her doctors after reviewing.

Certainly, I can provide a more detailed medical analysis and discussion points as an MD, Ph.D. in the relevant areas for your cousin to review further with her doctors:

- **Implant Integrity and Metal Particle Shedding:**

- The locked hip implant appears to be undergoing adverse wear and corrosion at the metal-metal articulating surfaces, leading to the shedding of metallic particles.
- These particles have likely contributed to a local cytotoxic reaction, causing the soft tissue destruction and suppurative (pus formation) around the implant.
- Scanning with metal artifact reduction sequences or implant-specific software reconstruction may help evaluate the extent of implant damage and the local tissue response.
- Revision to a ceramic or highly cross-linked polyethylene bearing surface may be required to mitigate further metal particle generation.

- **Pelvic Bone Status and Reconstructive Options:**

- Multiple revision surgeries often lead to progressive pelvic bone loss due to stress shielding, osteolysis, and devascularization.
- Obtain a CT scan with Hounsfield unit quantification to assess the bone density and extent of pelvic bone defects.
- Surgeons may need to consider structural allografts, porous metallics, or patient-specific 3D-printed implants to achieve stable implant fixation.
- Prolonged protection and limited weight-bearing may be required to allow for adequate osseointegration.

- **Infection Management:**

- Ensure the current infection is adequately treated based on culture sensitivities, likely requiring an extended IV antibiotic course.
- Consider pursuing a one-stage or two-stage revision strategy in coordination with an infectious disease specialist.
- Serial inflammatory markers and advanced imaging (leukoscan, PET/CT) may aid in confirming infection resolution before proceeding with revision.

- **Surgical Approach and Functional Outcomes:**

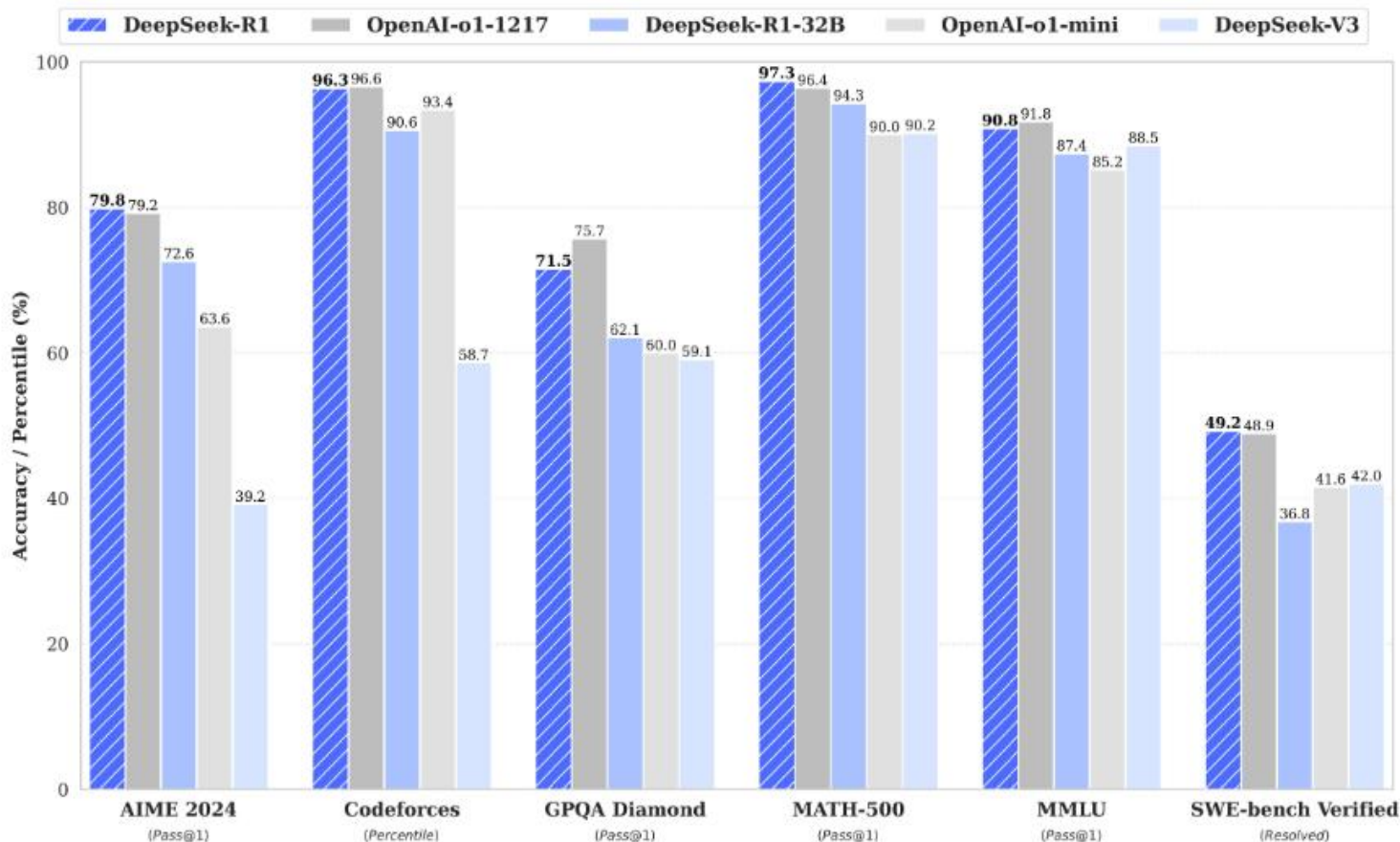
- An extensile surgical approach (e.g., triradiate) may be required for adequate exposure and reconstruction.
- Involvement of a vascular surgeon may be warranted to assess pelvic perfusion status if significant dissection is anticipated.
- Realistic expectations should be set regarding the potential for continued pain, limited mobility, need for ambulatory aids, and increased risk of dislocation/instability given the underlying bone and soft tissue compromise.

Consultation with a physical medicine and rehabilitation specialist team can help optimize functional recovery through tailored therapy programs. I would strongly advise your cousin to seek out a tertiary care center with an **experienced revision arthroplasty team, including adult reconstruction surgeons and musculoskeletal radiologists. A thorough multi-disciplinary evaluation is warranted** to carefully weigh the risks, benefits, and reconstructive options for her case. Additionally, a second opinion from another high-volume revision center may provide further valuable perspectives. Careful shared decision-making will be crucial given the complexity of her hip condition.



[AI Medicine Benchmarking Study, 2025](#)

**DeepSeek R1 vs. Best
in Class Reasoning
Models OpenAI o1
Class
Ph.D. Level Math,
Coding, STEM Science
& Disciplinary IQ
Tests**



DeepSeek V3 vs Intuitive (AGI Models)

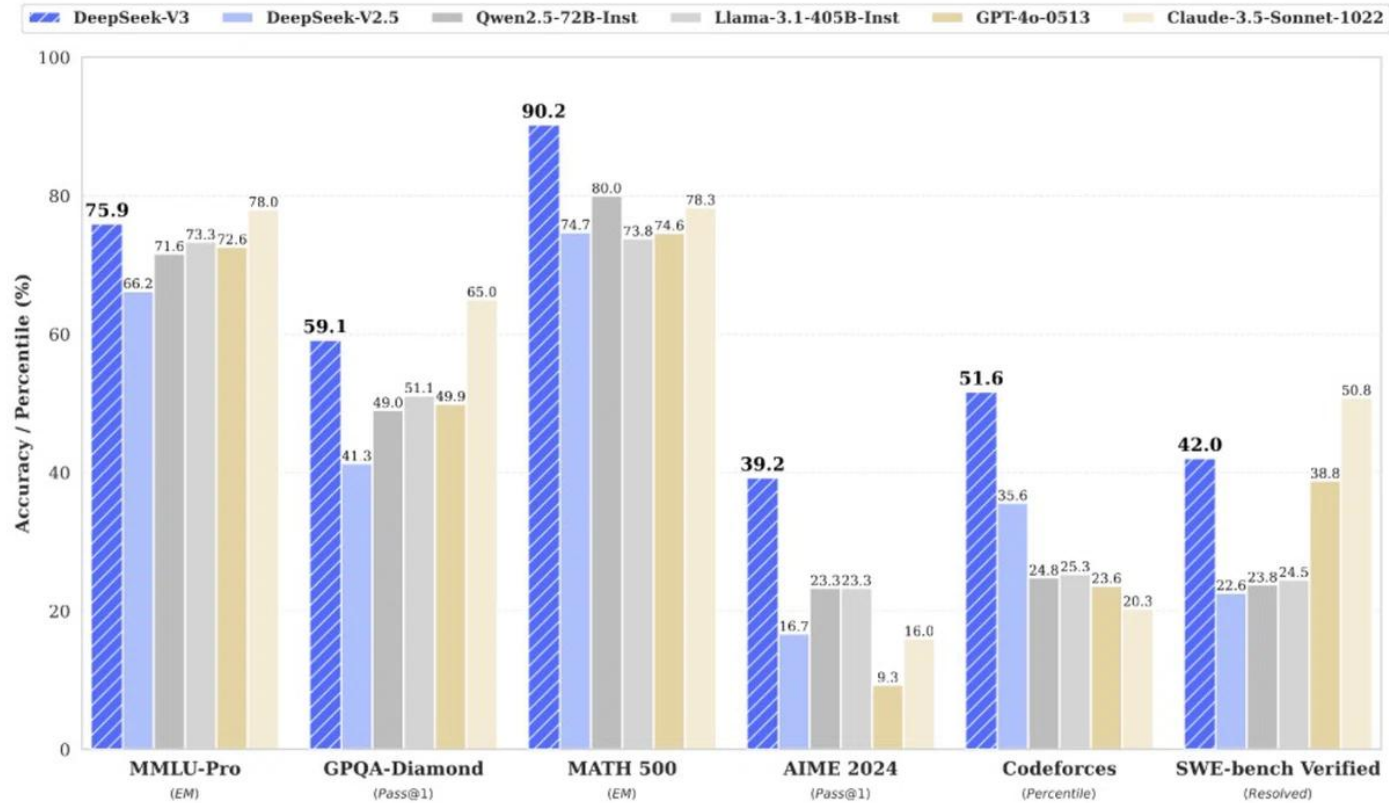
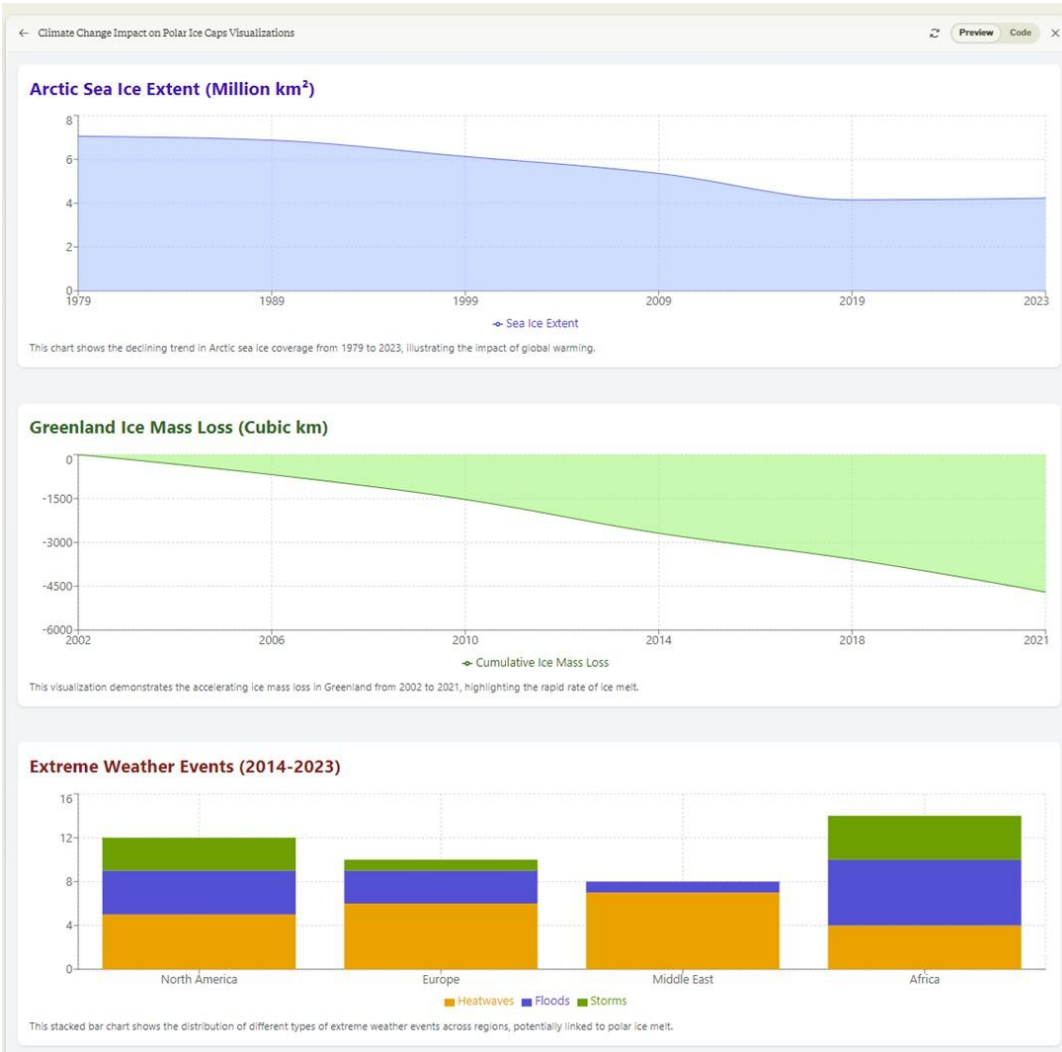


Figure 1 | Benchmark performance of DeepSeek-V3 and its counterparts.

Data Visualizations Extract Data from Research Papers Excel, Other Data Formats)



Multimodality: Text and Image

Tell me more about this image below? (GPT4o)

- The individual in the image is **Zaha Hadid**, a renowned Iraqi-British architect known for her innovative and futuristic designs. Zaha Hadid was a pioneering figure in contemporary architecture, and she became **the first woman to receive the prestigious Pritzker Architecture Prize in 2004**.
- **About Zaha Hadid**
 - **Early Life and Education:** Born in Baghdad, Iraq, in 1950, Zaha Hadid **studied mathematics at the American University of Beirut** before moving to **London to study at the Architectural Association School of Architecture**.
 - **Career and Achievements:** Hadid's career was marked by her distinctive style, characterized by bold, **sweeping forms and futuristic aesthetics**. Her work often **pushed the boundaries of architecture, blending organic and geometric forms**. Some of her most famous projects include the London Aquatics Centre for the 2012 Olympics, the Guangzhou Opera House in China, and the Heydar Aliyev Center in Baku, Azerbaijan.

About the Architectural Model

- **Design Elements:** The model in the image reflects Hadid's signature style, with **its flowing, dynamic lines and complex forms**. **The use of curvilinear shapes and layered structures** suggests a focus on creating fluid, organic spaces that challenge traditional architectural conventions.



Image to Text, Visual Acuity

Speculate on the Architect?

GPT4o and Sonnet 3.5

- **Given the building's distinct architectural style, characterized by its fluid curves, layered structure, extensive use of glass, and futuristic aesthetics**, it is reasonable to speculate that the architect could be **Zaha Hadid** or a firm influenced by her design philosophy.
- **Zaha Hadid Architects (ZHA)** is renowned for such innovative and avant-garde designs.
- **Santiago Calatrava**: Known for his sculptural and organic forms, Calatrava's buildings also feature dynamic and fluid designs. However, his work often includes more overt structural elements that resemble skeletal forms, which are not as apparent in this building.
- **MAD Architects**: A Beijing-based firm led by Ma Yansong, MAD Architects is known for its futuristic and organic designs. Their projects often feature flowing, natural forms and an emphasis on blending architecture with the environment.
- **UNStudio** (Ben van Berkel): UNStudio's work includes projects that explore fluid geometries and innovative use of materials, similar to the building in the image.



Questions and Feedback

Ray Uzwyshyn, Ph.D. MBA MLIS

Acting AUL for Research & Technology
Services

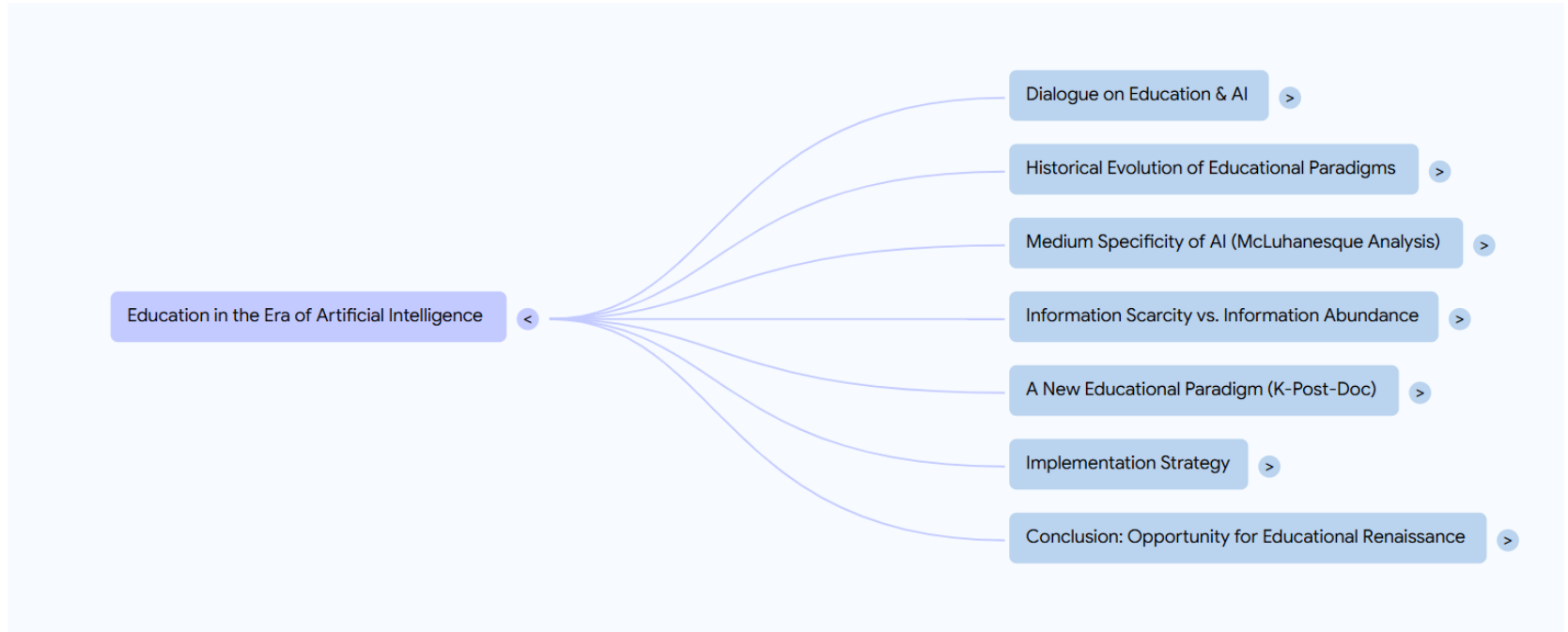
University of California, Riverside Libraries,

raymondu@ucr.edu

<https://www.linkedin.com/in/rayuzwyshyn/>

<https://rayuzwyshyn.net>

Mind Map of Essay



Scaling of Search and Learning:

A Roadmap to Reproduce o1 from Reinforcement Learning Perspective

Zhiyuan Zeng^{1*} Qinyuan Cheng^{1*} Zhangyue Yin^{1*} Bo Wang^{1*}
Shimin Li¹ Yunhua Zhou² Qipeng Guo² Xuanjing Huang¹ Xipeng Qiu^{1†}
¹Fudan University ²Shanghai AI Laboratory

<https://notebooklm.google.com/notebook/fe01cfd1-8b31-4e20-87b4-fe2bb0946974/audio>

• Detailed Timeline of Events:

- **1949:** The Monte Carlo method is discussed by Nicholas Metropolis and Stanislaw Ulam.
- **2000:** Andrew Y. Ng and Stuart Russell publish algorithms for inverse reinforcement learning.
- **2012:** Cameron Browne and colleagues publish a survey of Monte Carlo tree search methods.
- **2019:** Richard S. Sutton publishes "The Bitter Lesson." Alec Radford et al. publish work on World Knowledge acquisition.
- **2020:** Radford et al. and Brown et al. publish work on World Knowledge. Tom B. Brown et al. publish "Language models are few-shot learners," establishing the importance of well-initialized policies. Stuart Russell and Peter Norvig publish the 4th Edition of "Artificial Intelligence: A Modern Approach." Julian Schrittwieser et al. publish on mastering games using a learned model (AlphaZero). David Silver et al. publish "Mastering the game of go without human knowledge" (AlphaGo).
- **Undated (referenced in the context of recent/current work):** OpenAI's "o1" model and its five-stages plan for AGI are discussed, with "o1" being a strong reasoner and the next stage being training an agent interacting with the real world. The o1 blog (OpenAI, 2024a) is referenced for exemplars of human-like reasoning behaviors. "richards" is mentioned for the Thinking-Claude GitHub repository. "bklieger" is mentioned for the g1 GitHub repository using Llama-3.1 70b on Groq to create o1-like reasoning chains. "Junlin Zhang" is mentioned for reverse-o1.

• Cast of Characters:

- **Dario Amodei:** Associated with OpenAI and a co-author on several foundational works including "Language models are few-shot learners," RLHF, and Constitutional AI.
- **Alec Radford:** Co-author on Language Understanding and Generation, World Knowledge, and "Language models are few-shot learners."
- **Ilya Sutskever:** Associated with OpenAI and a co-author on "Language models are few-shot learners" and evaluating LLMs trained on code.
- **Wojciech Zaremba:** Associated with OpenAI and a co-author on evaluating LLMs trained on code.
- **Jan Leike:** Discussed in the context of AI-assisted human feedback and scalable agent alignment via reward modeling.

Global AI LLM Market: Porter's Five Forces Analysis, 2025, Briefing and Glossary

Podcast: <https://notebooklm.google.com/notebook/bf9745e6-5aaf-4698-8375-b1b49311daba/audio>

Paper Briefing: Main Themes and Key Ideas:

1. The AI LLM Market: A Rapidly Evolving, High-Growth Landscape:

•The broader global AI market is projected for extraordinary growth, with a CAGR of 30.3% from 2024 to 2034, reaching approximately \$3,527.8 billion by 2034.

2. Porter's Five Forces Analysis with 21st Century Nuances:

•**Competitive Rivalry: Hyper-Competition and Turbulence:** The market exhibits "hypercompetition" with intense and rapid competitive interactions. Key players include established tech giants (Google, Microsoft, Meta, IBM), AI-native organizations (OpenAI, Anthropic, Cohere, DeepSeek), open-source collectives (Mistral), and sovereign AI initiatives.

•**Quote:** "The AI LLM market exhibits characteristics of what D'Aveni (1994) termed 'hypercompetition'—a state of intense and rapidly evolving competitive interactions characterized by unsustainable advantage and continuous disruption."

•**Key Fact:** DeepSeek's R1 model, priced at \$2.19 per million output tokens, significantly disrupts established pricing models (e.g., OpenAI's GPT-4o at \$10).

•Innovation cycles are compressing from years to months or weeks.

•**Threat of New Entrants: Disaggregating Traditional Barriers:** Traditional barriers like computational resources, expertise, and data are transforming.

•**Computational Resource Democratization:** While frontier models are intensive, alternative approaches requiring less resources are emerging (DeepSeek example). Cloud-based resources also lower capital requirements.

•**Knowledge Diffusion:** Open-source implementations, research, and workforce mobility accelerate knowledge spread.

- **Glossary of Key Terms**
- **Artificial Intelligence (AI):** The simulation of human intelligence processes by machines, especially computer systems.
- **Porter's Five Forces:** A framework developed by Michael Porter to analyze the competitive environment of an industry by examining five key factors: Competitive Rivalry, Threat of New Entrants, Bargaining Power of Buyers, Bargaining Power of Suppliers, and Threat of Substitutes.
- **Hypercompetition:** A state of intense, rapidly escalating competition in which advantages are temporary and firms are constantly seeking new ways to disrupt the market and their rivals.
- **Threat of New Entrants:** The likelihood of new competitors entering an industry, influenced by barriers to entry.
- **Bargaining Power of Buyers:** The ability of customers to influence prices and terms of sale.
- **Bargaining Power of Suppliers:** The ability of input suppliers to influence prices and terms of sale.
- **Threat of Substitutes:** The likelihood that customers will switch to alternative products or services from outside the industry.
- **Geopolitical Implications:** The influence of political, economic, and geographic factors on international relations and global power dynamics, particularly concerning technology.
- **Technological Inflection Points:** Moments in technological development where fundamental changes in direction occur, often leading to significant shifts in markets and society.
- **Exponential Growth:** Growth that occurs at an increasingly rapid rate, often characterized by doubling in fixed time periods.
- **Compound Annual Growth Rate (CAGR):** The average annual growth rate of an investment over a specified period longer than one year.
- **Neural Network Architectures:** The design and structure of artificial neural networks, which are computing systems inspired by biological neural networks.

Multimodal Inputs, Outputs and Research, GPT4o and Sonnet 4 and Gemini 2.5 Pro * Not Deep Seek R1

Multimodal Inputs:

1. Text and Image:

o Text: "Write a Ph.D. Level Art Historical Analysis of this AI Image."

o Image: Who is this architect and what is she standing beside? Zaha Hadid

2. Text and Data:

o Text: "Summarize the trends shown in the following chart on global temperature changes over the last century."

o Data

Text and Video:

Summarize This Physics Video Lecture and Produce a Study Guide

Text to Audio Podcast (Notebook LM)

Summarize this 87 Page Chinese Mathematical Machine Learning Technical Paper So I can Understand How They Created This Next AI System Architecture

