MEMO *Published September 19, 2024 · 7 minute read*

America's Next Top AI Models

Mike Sexton, Senior Policy Advisor for Artificial Intelligence and Digital Technology

The AI Runway—Prepare to be wowed! Fact: The best AI models in the world are being built in America. *Plural.* ChatGPT is not a one-off. American tech titans like Nvidia, Meta, Amazon, Google, and Microsoft are developing new, groundbreaking generative AI models for a variety of uses. Smaller startups like Reka (US), [1] Cohere (Canada), [2] and Mistral (France) [3] have their designs on breaking in as well. The US is not just outperforming China: it's outperforming everyone.

The AI runway right now is competitive, though there are early signs that the AI industry is slouching towards consolidation. [4] That is because training frontier AI models becomes exponentially more complex over time—technologically, financially, legally, and politically. This complexity has led to copyright infringement lawsuits, [5] public apologies, [6] office [7] and boardroom [8] drama, and even synthetic child pornography. [9]

There is a spectrum of opinion on whether AI's advancement and adoption will benefit society. However, a diversity of continuously improving and competing frontier AI models presents key

advantages to an industry dominated by a small handful of players. From Abu Dhabi to Nairobi, [10] Toronto to Tel Aviv, [11] and Shenzhen [12] to San Francisco, there is little doubt that the advancement of frontier AI models approaching the theoretical ideal of "artificial general intelligence" (AGI) [13] is inevitable.

What is in doubt is *how* the technology will advance. Who will develop, control, and benefit from these models? A sprawling mosaic of ambitious competitors, or a select few companies and players?

AI Competition is Real

This product aims to introduce you to AI competition and to the top American AI models. Beyond their inspirational novelty, they are useful in ways we are only beginning to understand. If you haven't tried on any of these AIs yet, you should take a look, and if you do, you should accessorize and try more. Learn where they excel and fall short; explore how they can sate your curiosity, organize your thinking, and manifest your imagination. As policymakers, think about what excites you from a public policy standpoint and what gives you pause.

When more people use more AI models, we collectively slouch away from AI monoculture [14] —a world where one AI model becomes dominant and evolves into a knowledge chokepoint.

We include six chatbots and two image generators, briefly discuss what distinguishes them from each other, and link to useful external resources with more comprehensive information.

So without further ado, here are America's next top AI models, generated using Midjourney:

Read more: What is AGI? (Amazon Web Services)

ChatGPT

ChatGPT

➤ Parameters: 1.76 trillion (GPT-4, estimate)
➤ Users: over 100 million
➤ Multimodal: text, image, speech

Released by OpenAI in 2022, ChatGPT is first down the catwalk and instantly became a household name and the model to beat. OpenAI was founded in 2015 to build AGI for the public good, and its flagship model has dominated the scene like Naomi Campbell. OpenAI has signed an exclusive licensing agreement with Microsoft, making GPT the foundation for apps like Copilot for the foreseeable future.

- Parameters: 1.76 trillion (GPT-4, estimate) [15] (*Parameters are the guts of an LLM that comprise its capacity to reason. While inexact, their total number is a decent proxy for a model's breadth of knowledge.*)

- Users: over 100 million [16]

- Multimodal: text, image, speech

Read more: Introducing ChatGPT, Hello GPT-4o, But what is a GPT? Visual intro to transformers (3Blue1Brown)

Claude

Anthropic was founded in 2021 by former OpenAI employees who supported a stricter approach to developing AGI for the public good (aka "the alignment problem."). While AI models with superhuman intelligence can be immensely helpful for humans, the alignment problem posits that they can be dangerous unless programmed carefully to share our values like respect for human life and difference of opinion. Fad or fashion icon? Its latest premium model, Claude 3 Opus, has previously outperformed GPT-4 [17] on benchmarks for math, code, and textual reasoning. [18]

- Parameters: unknown

- Users: 'millions' according to CEO Dario Amodei [19]

- Multimodal: text & image, not speech

Read more: AI Startup Anthropic Debuts Claude Chatbot as an iPhone App

Pi

Florals for spring? Pi won't judge. Inflection AI developed Pi with an eye not just for factual but *emotional* intelligence. While it makes a fine research assistant, its sense of humor and breezy communication style make it seem like an ideal life coach for navigating social or personal issues.

- Parameters: 13 billion [20]

- Users: over one million [21]

- Multimodal: text & speech, not images

Read more: <u>Inflection AI's friendly chatbot tops 1 million daily users</u>, <u>Microsoft to Pay Inflection AI $650 Million After Scooping Up Most of Staff</u>

LLaMa

LLaMa

➤ Parameters: 70 billion
➤ Users: unknown
➤ Multimodal: can be made multimodal

LLaMa may not be a household name to everyone, but it is to AI experts. Meta's groundbreaking decision to open-source its large language model (LLM, hence LLaMa) in 2023 means anyone can use it as a foundation for their own program. Its impact on the AI ecosystem is hard to overstate, with websites like Ollama and LlamaIndex dedicated to supporting its developer community.
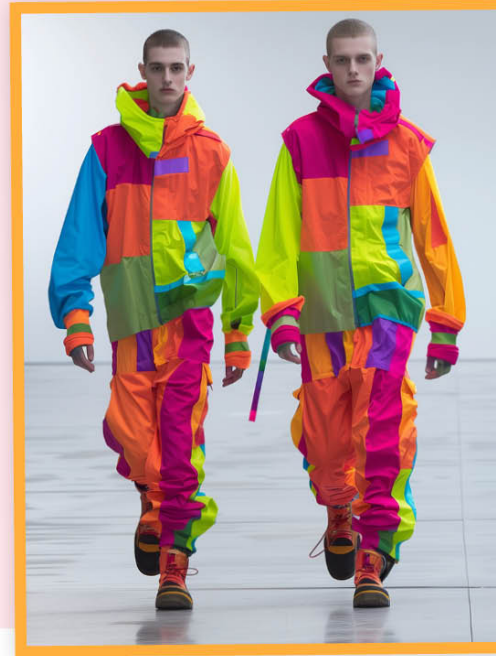
- Parameters: 70 billion [22]

- Users: unknown

- Multimodal: can be made multimodal

Read more: <u>Meet Your New Assistant: Meta AI, Built With Llama 3</u>, <u>Llama 3 Establishes Meta as the Leader in "Open" AI</u>, <u>What is "Open" AI?</u>

Gemini

Gemini is Google's AI chatbot. Whereas chatbots typically only "know" what's in their training data, Gemini integrates with industry-leading Google Search, grounding its results in real-time with supporting sources. Lightweight, open-source versions named Gemma are available to download and tinker with, too—just like LLaMa.

- Parameters: unknown (2B-7B for Gemma) [23]

- Users: unknown

- Multimodal: text, image, speech

Read more: These 3 Gemini upgrades will help you get more out of the AI assistant, How Prados Beauty is using Gemini to grow their business

Perplexity

When ChatGPT went public, it transfixed the public like a bona fide oracle, but its hallucinations and lack of citations were early shortcomings in its immediate incarnation. Enter Perplexity: a chatbot-meets-search-engine that generates answers with ChatGPT and other leading LLMs and automatically identifies and cites sources for their answers.

- Parameters: not applicable (*Perplexity integrates other companies' AI models like GPT, Llama, and Claude with source-citing capabilities*)

- Users: 10 million [24]

- Multimodal: not multimodal

Read more: Does Perplexity's "answer engine" threaten Google?, AI Search Startup Perplexity Valued at $1 Billion in Funding Round

Character.ai

Character.ai is a marvel and an enigma—a service for building personalized AI chatbots that can embody a wide range of characters, from occupations (a travel agent, a psychologist specializing in cognitive behavioral therapy, a polyglot language tutor) to celebrities living (Elon Musk) and dead (Albert Einstein). [25] While ChatGPT has an online "shop" of personalized GPTs, too, [26] this is Character.ai's *raison d'etre*, built on top of a proprietary LLM the company has shared few details about.

- Parameters: unknown

- Users: 20 million [27]

- Multimodal: text and voice

Read more: [Meet My AI Friends](#), [Ex-Google engineers who founded Character.AI rejoin company with new AI partnership](#)

DALL-E

(Image generated by DALL-E)

DALL-E is an image-generator developed by OpenAI, the makers of ChatGPT. First released less than two years ago, [28] its name is a combination of the Disney robot WALL-E and surrealist Salvador Dali. Built with GPT-3, [29] DALL-E's transformer architecture make it more effective at grasping semantics, actions, and the relationships between objects in a prompt than other models.
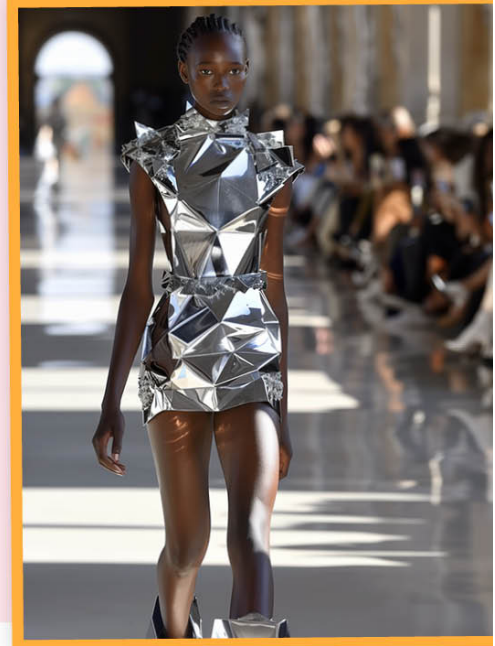
- Users: 1.5 million (2022) [30]

Read more: <u>Meet DALL-E, the A.I. That Draws Anything at Your Command</u>, Paid ChatGPT customers can now use AI to edit DALL-E images

Midjourney

Midjourney is another image-generation AI. Unlike DALL-E, which can rephrase prompts and add valuable details, Midjourney apparently lacks the deep conceptual understanding of its GPT-based competitor. What it lacks, though, it makes up for with stylization, photorealism, higher resolution, and powerful customization options that make it ideal for eye-catching visuals.

- Users: 20 million [31]

Read more: How To Create AI Images on Midjourney, How a tiny company with few rules is making fake images go mainstream

ENDNOTES

1. "Reka AI – Products, Competitors, Financials, Employees, Headquarters Locations." *CB Insights*, https://www.cbinsights.com/company/reka. Accessed 22 May 2024.

2. "Cohere – Products, Competitors, Financials, Employees, Headquarters Locations." *CB Insights*, https://www.cbinsights.com/company/cohere-3. Accessed 22 May 2024.

3. "Mistral AI – Products, Competitors, Financials, Employees, Headquarters Locations." *CB Insights*, https://www.cbinsights.com/company/mistral-ai. Accessed 22 May 2024.

4. Newton, Casey. "The Indie AI Companies Are Falling Apart." *Platformer*, 27 Mar. 2024, https://www.platformer.news/indie-ai-collapse-stability-ai-mostaque-inflection-suleyman/. Accessed 11 April 2024.

5. Grynbaum, Michael M., and Ryan Mac. "The Times Sues OpenAI and Microsoft Over A.I. Use of Copyrighted Work." *The New York Times*, 27 Dec. 2023. *NYTimes.com*, https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html. Accessed 11 April 2024.

6. Robertson, Adi. "Google Apologizes for 'Missing the Mark' after Gemini Generated Racially Diverse Nazis – The Verge." *The Verge*, 21 Feb. 2024, https://www.theverge.com/2024/2/21/24079371/google-ai-gemini-generative-inaccurate-historical. Accessed 11 April 2024.

7. Knight, Will, and Steven Levy. "OpenAI Staff Threaten to Quit Unless Board Resigns | WIRED." *Wired*, 20 Nov. 2023, https://www.wired.com/story/openai-staff-walk-protest-sam-altman/. Accessed 11 April 2024.

8. Paul, Kari. "OpenAI Reinstates CEO Sam Altman to Board after Firing and Rehiring." *The Guardian*, 9 Mar. 2024. *The Guardian*, https://www.theguardian.com/technology/2024/mar/08/openai-sam-altman-reinstated. Accessed 11 April 2024.

9.   Farivar, Cyrus. "Suspect Charged After Allegedly Using AI To Create Images Of Child Sexual Abuse." *Forbes*, 26 Mar. 2024, https://www.forbes.com/sites/cyrusfarivar/2024/03/26/stable-diffusion-15-cited-in-case-involving-ai-generated-child-sexual-abuse-images/. Accessed 11 April 2024.

10.   Source, Microsoft. "Microsoft and G42 Announce $1 Billion Comprehensive Digital Ecosystem Initiative for Kenya." *Stories*, 22 May 2024, https://news.microsoft.com/2024/05/22/microsoft-and-g42-announce-1-billion-comprehensive-digital-ecosystem-initiative-for-kenya/. Accessed 22 May 2024.

11.   Wiggers, Kyle. "AI21 Labs' New AI Model Can Handle More Context than Most." *TechCrunch*, 28 Mar. 2024, https://techcrunch.com/2024/03/28/ai21-labs-new-text-generating-ai-model-is-more-efficient-than-most/. Accessed 22 May 2024.

12.   "Baidu Says AI Chatbot 'Ernie Bot' Has Attracted 200 Million Users." *Reuters*, 16 Apr. 2024. *www.reuters.com*, https://www.reuters.com/technology/baidu-says-ai-chatbot-ernie-bot-has-amassed-200-million-users-2024-04-16/. Accessed 22 May 2024.

13.   "What Is AGI? - Artificial General Intelligence Explained - AWS." *Amazon Web Services, Inc.*, https://aws.amazon.com/what-is/artificial-general-intelligence/. Accessed 22 May 2024.

14.   Leffer, Lauren. "Too Much Trust in AI Poses Unexpected Threats to the Scientific Process." *Scientific American*, 18 Mar. 2024, https://www.scientificamerican.com/article/trust-ai-science-risks/. Accessed 11 April 2024.

15.   Bastian, Matthias. "GPT-4 Has More than a Trillion Parameters - Report." *THE DECODER*, 25 Mar. 2023, https://the-decoder.com/gpt-4-has-a-trillion-parameters/. Accessed 11 April 2024.

16.   Hu, Krystal. "ChatGPT Sets Record for Fastest-Growing User Base - Analyst Note | Reuters." *Reuters*, 2 Feb. 2023, https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/. Accessed 11 April 2024.

17. Lance Whitney. "Claude 3 Overtakes GPT-4 in the Duel of the AI Bots. Here's How to Get in on the Action." *ZDNet*, 28 Mar. 2024, https://www.zdnet.com/article/claude-3-overtakes-gpt-4-in-the-duel-of-the-ai-bots-heres-how-to-get-in-on-the-action/. Accessed 21 May 2024.

18. "Gpt4 Comparison to Anthropic Opus on Benchmarks - Community." *OpenAI Developer Forum*, 24 Apr. 2024, https://community.openai.com/t/gpt4-comparison-to-anthropic-opus-on-benchmarks/726147. Accessed 21 May 2024.

19. "Google-Backed Anthropic Releases Claude Chatbot Across Europe." US News & World Report, 13 May 2024, //www.usnews.com/news/technology/articles/2024-05-13/google-backed-anthropic-releases-claude-chatbot-across-europe.

20. Sr, Kain Sosa. "Inflection AI Presents Pi: The Groundbreaking Personal Intelligence." *Medium*, 5 July 2023, https://kainsosa.medium.com/inflection-ai-presents-pi-the-groundbreaking-personal-intelligence-a7afd2bd98d. Accessed 11 April 2024.

21. "Inflection AI's Chatbot Pi Surpasses 1 Million Daily Active Users | Reuters." *Reuters*, 7 Mar. 2024, https://www.reuters.com/technology/inflection-ais-chatbot-pi-surpasses-1-million-daily-active-users-2024-03-07/. Accessed 11 April 2024.

22. "Llama 2." *Llama*, https://llama.meta.com/llama2/. Accessed 11 Apr. 2024.

23. "Gemma: Introducing New State-of-the-Art Open Models." *Google*, 21 Feb. 2024, https://blog.google/technology/developers/gemma-open-models/. Accessed 11 April 2024.

24. "The Latest Perplexity AI Stats (2024)." *Exploding Topics*, 19 Mar. 2024, https://explodingtopics.com/blog/perplexity-ai-stats. Accessed 2 May 2024.

25. *Character.Ai | Personalized AI for Every Moment of Your Day*. https://character.ai/. Accessed 22 May 2024.

26. "Explore GPTs." *ChatGPT*, https://chatgpt.com/gpts. Accessed 22 May 2024.

27. Shewale, Rohit. "Character AI Statistics 2024 - Active Users & Revenue." *DemandSage*, 27 Mar. 2024, https://www.demandsage.com/character-ai-statistics/. Accessed 22 May 2024.

28. Bobby Allyn. "Surreal or Too Real? Breathtaking AI Tool DALL-E Takes Its Images to a Bigger Stage." *NPR*, 20 July 2022. *NPR*, https://www.npr.org/2022/07/20/1112331013/dall-e-ai-art-beta-test. Accessed 2 May 2024.

29. "DALL·E: Creating Images from Text." *OpenAI*, https://openai.com/index/dall-e. Accessed 2 May 2024.

30. DALL·E Now Available without Waitlist." OpenAI, 28 Sept. 2022, https://openai.com/index/dall-e-now-available-without-waitlist/.

31. Midjourney. "Join the Midjourney Discord Server!" Discord, 16 Sept. 2024, https://discord.com/invite/midjourney.