

# When "scary good" AI gets even better

[Ina Fried](#)

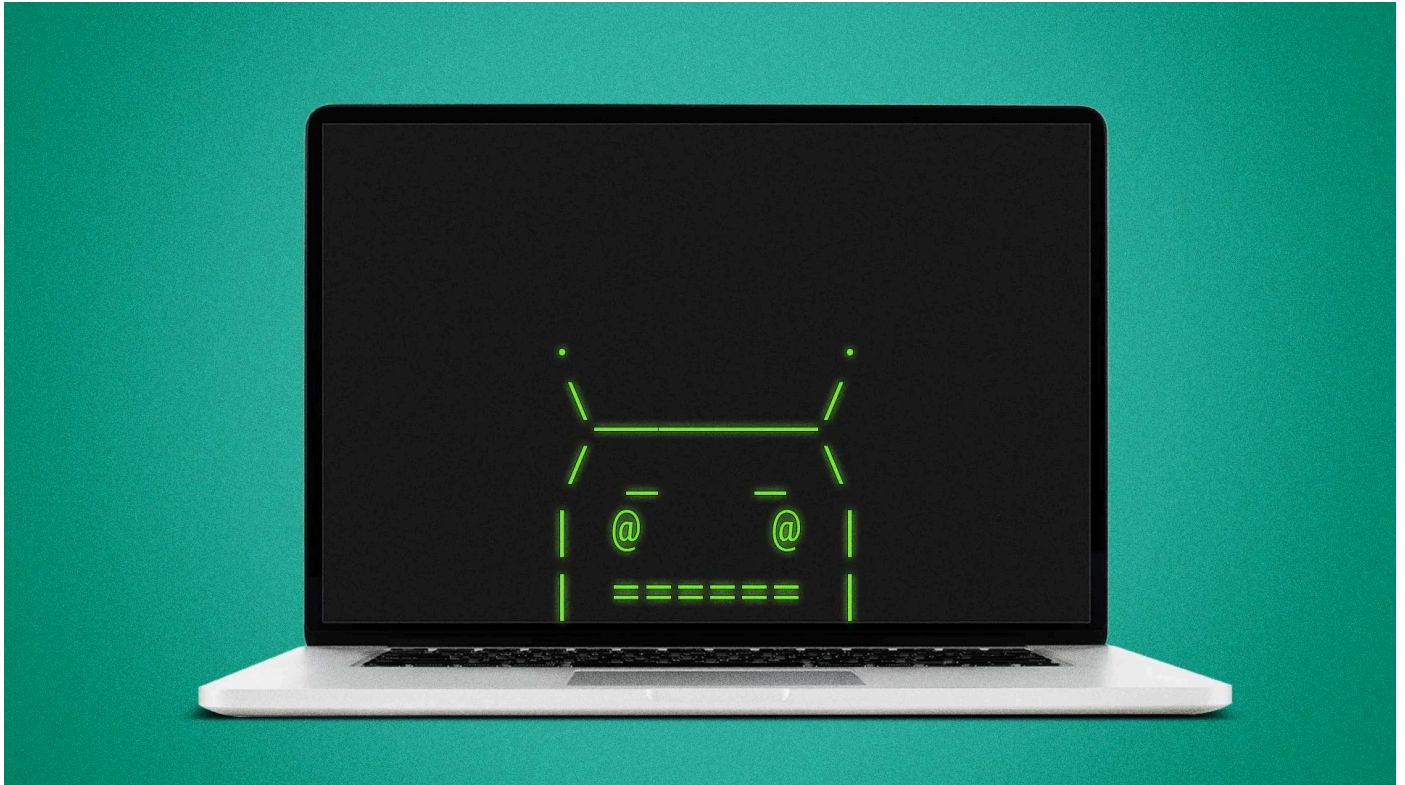


Illustration: Maura Losch/Axios

With Tuesday's [release of OpenAI's new GPT-4](#), generative AI just got a lot more powerful — and we got a fresh reminder of just how unprepared we are to deal with these new machines.

**Why it matters:** The amazing computer systems that can now ace standardized tests and maybe even do your taxes are still disturbingly prone to errors, bias and hallucinations.

**Details:** GPT-4 is an updated, significantly more powerful version of the engine that powers OpenAI's ChatGPT.

- While ChatGPT could score in the 10th percentile on the standard bar exam taken by lawyers, OpenAI says GPT-4 can score in the 90th percentile. GPT-4 is also able to pass most AP exams, OpenAI said.

**The new version** can accept and generate longer entries — up to 25,000 words. It can also generate captions and other information using an image as a starting point.

- But like its predecessor, it is only trained on information that was publicly available as of September 2021.
- On the safety side, OpenAI says GPT-4 is 82% less likely to respond when asked for content it doesn't allow and 40% more likely to produce factual responses than GPT-3.5 in internal testing.

**Microsoft [confirmed](#)** that GPT-4 has been powering its new Bing search chatbot.

- That may be your easiest route to using it right now, since OpenAI is limiting some access.
- OpenAI said GPT-4 will be available in a limited capacity to paid ChatGPT Plus subscribers via [chat.openai.com](https://chat.openai.com), and there is a waitlist for businesses and developers looking to incorporate GPT-4 via an API.

**What they're saying:** The technology is still far from perfect, OpenAI president Greg Brockman told Axios on Tuesday.

- But it's already at a stage where it can help a lot of people, he said, noting its potential to expand access to education, as well as legal and medical information. "I'm just excited to see what people build," he said.

**The big picture:** OpenAI isn't alone in debuting advances in the field. It wasn't even the only AI outfit making news Tuesday

- Anthropic, an OpenAI rival, [formally announced Claude](#), its chatbot which is being used by a range of companies including DuckDuckGo, Notion and Quora, among others.
- Google [outlined](#) how it will use generative AI to help businesses. It will offer tools that companies can use to allow generative AI engines to scour through corporate data. AI tools being added to Workspace will help summarize e-mail, craft marketing campaigns and rewrite documents.
- Microsoft, meanwhile, has scheduled an [event](#) for Thursday to talk about how it will build generative AI into its business products, including Office apps, such as Word, Excel, PowerPoint and Outlook.

**Between the lines:** AI is barreling forward even as society is still trying to come to grips with both its promise and the potential pitfalls.

- [The law has yet to catch up](#). Few laws specific to AI exist, although the EU has been working to craft a wide-ranging AI act designed to regulate use of the technology, especially in "high-risk" areas.
- Businesses are still grappling with how the technologies might augment or replace human labor. Today's generative AI can do a convincing job of crafting marketing materials, summarizing text and transforming it into new genres. At the same time it is still prone to both making up facts and committing fairly basic math errors.

**That's all leading some critics** to sound alarm bells.

- Tristan Harris, the former Googler who now runs the Center for Humane Technology, is warning against making the same mistakes with the current generation of AI that were made in the early days of social media.
- "Notice that once social media became entangled with society and its

institutions (GDP, elections, journalism, children's identity) it became impossible to regulate," Harris [tweeted](#) on Monday. "We should set guardrails for safer AI deployment and research \*before\* AI gets entangled, rather than after."

**Yes, but:** OpenAI says GPT-4 improves on its predecessors not just in its capabilities, but also with increased accuracy and more human-installed guardrails.

- At the same time, GPT-4 has gotten good enough that even its makers say it's time to start considering the impact it could have in terms of job replacement.

**Between the lines:** GPT-4's higher accuracy makes Brockman worry that people will trust it more — and let their fact-checking guard down.

- OpenAI encourages discussion about the impacts of AI, both by lawmakers and broader society, Brockman said. "We believe there should be more AI regulation," he said, but added that's not sufficient: "You really need an educated public."