Distribution

OpenAI still sets the agenda for new models, mostly. It has 8-900m weekly active users, it has the mind-share, and it has the consumer brand. But none of that is based on any fundamental, structural, competitive advantage. Half a dozen to a dozen companies regularly ship SOTA models now, and there are no network effects in LLMs, at least not yet. Only 5% of those 800m users pay, and 80% post less than three messages per day. That usage is a mile wide and an inch deep - which is to say, people can move very easily.

So, how can it turn that position into something durable? It has to make itself a daily habit, and it has to close the 'capability gap' - the fact that most people who use ChatGPT don't find it very useful. Anthropic has gone for the developers and the enterprise API market, but Sam Altman wants OpenAI to be the new Google, Microsoft, and Apple all rolled into one.

That makes ads an interesting choice. On one hand, ChatGPT has marginal cost, and most users aren't paying yet, so ads are a way to close that gap on those less engaged users (and perhaps tempt them to upgrade). It also makes it easier to give free users more features that have more marginal cost.

Indeed, contextual ads in a chatbot seem like continuity with search ads - everything else that's happening with AI and the web is a fundamental change, but asking a question and getting an ad in the answer is the same model we got used to 25 years ago. There are much bigger ways that AI will change search, ads, and links than this.

More strategically, if ads are part of the future of chatbots, does that mean it's better to start building that business soon rather than later, and especially before Google runs away with it? Well, perhaps. But Google has decided the opposite: it won't annoy Gemini users with ads, for now, concentrating on improving the product and the experience. After all, Google can afford it - this is a loss-leader for now.

But Google is posing much bigger questions than that for ChatGPT. It's spent the last six months pushing Gemini to all the distribution points that it already has, taking on ChatGPT directly, and it's also looking at use cases by adding it to products where OpenAI has nothing to say. Yes, I could set up ChatGPT to look at my email and my calendar and tell me what's going on tomorrow, but Google already has that context ready to go, and can see it from the inside out. Meta can do the same in Instagram: I can ask ChatGPT what I'd like, but Meta can see what I've already liked and looked at.

The Apple Gemini deal, to power 'Siri 2.0', covers all of these stories. It's cashflow for Google, and a vote of confidence, but not *exactly* distribution: the personal assistance that's integrated with your own data on your iPhone or Mac will be branded 'Siri', though more 'world-model' questions, that today Siri 1.0 can send to ChatGPT, might go to Gemini instead (we'll have to wait until the release to see exactly how Apple's implementing this). But this is really Apple, like Google, proposing a set of use cases where an LLM is a technology to power a feature, not a new tool in its own right.

I think this is a much more general question for OpenAI. Today a general purpose LLM is a commodity, as experienced by a normal consumer for normal, general purpose use cases. If you're a developer or spend hours in these things every day they feel different, but if you're posting three prompts a day they're all the same, and we don't know how that would change. So how do you compete? You can extend the chatbot itself to enable use cases and drive adoption (how?), or use it to power new features for your existing products (which OpenAI doesn't have), or try to invent new, dedicated vertical use-cases, unbundling the raw chatbot and abstracting it into tools and products - which means competing with every startup in the world.