Sam Altman: OpenAI may reach AGI this year

Sam Altman：OpenAI 今年可能实现 AGI

What I learned during two weeks inside the company as it slowed frontier training and previewed its next major model, Astra.

在公司放缓前沿模型训练并预演其下一个重大模型 Astra 之际，我在公司内部进行了为期两周的深入调研，以下是我的所见所闻。



ALEX HEATH

AUG 28, 2026

Listen

Yesterday, TIME and I published a cover story called **"Inside OpenAI's Reboot."** I spent two weeks at OpenAI's San Francisco headquarters and spoke with more than 20 executives, investors, customers, and rivals.

昨天，《时代》周刊和我共同发布了一篇名为《OpenAI 重启内幕》（Inside OpenAI's Reboot）的封面文章。我在 OpenAI 旧金山总部待了两周，采访了 20 多位高管、投资者、客户和竞争对手。

I went into the reporting process wanting to understand

OpenAI's last year, what allowed Anthropic to gain so much ground, and the internal mood around Astra, OpenAI's next major family of models that is coming soon.

在开始报道时，我希望了解 OpenAI 过去一年的情况、是什么让 Anthropic 赢得了如此多的市场份额，以及公司内部对即将推出的下一代重大模型系列 Astra 的氛围。

I happened to be in the office the week OpenAI began slowing some of its frontier reinforcement-learning work due to alignment concerns. I had multiple conversations with **Sam Altman** as the decision trickled through the company. I saw genuine alarm among the researchers over the capabilities they were seeing.

我恰好在 OpenAI 开始因对齐担忧而放缓部分前沿强化学习工作的那个星期身处公司办公室。随着这一决定在公司内部传达，我与 Sam Altman 进行了多次交谈。我看到研究人员对他们所见识到的模型能力表现出了真正的警惕。

At times, it was jarring to observe that caution con[illegible]h the frenetic pace at which OpenAI is operating on commercial side of the house, which is largely over[illegible] **Greg Brockman**. I can confidently say that people OpenAI are working harder than the company's Big Tech



incumbents. You can feel the urgency when you're in the building.

有时，这种谨慎态度与 OpenAI 在商业端（主要由 Greg Brockman 负责）运作的疯狂节奏形成了鲜明对比，这让人感到有些违和。我可以自信地说，OpenAI 内部员工的工作强度超过了那些大型科技巨头。当你身处大楼之中，你能感受到那种紧迫感。

There are several scoops in the over 4,000+ word feature that was published yesterday. I'm breaking out what stood out to me below, along with some reporting that didn't make it past the cutting room floor for the main story:

在昨天发布的这篇超过 4000 字的专题报道中，包含了多项独家内幕。我在下方列出了令我印象深刻的内容，以及一些未能进入正文的报道素材：

Astra is built for continuous agents

Astra 为连续型智能体而生

At an early August closed preview for VIP customers that I attended, OpenAI showed off Astra, which I expect to be released soon. Altman told the room:

在八月初我参加的一场面向 VIP 客户的闭门预览会上，OpenAI 展示了 Astra，我预计它很快就会发布。奥特曼（Altman）在现场表示：

> "I expect this will be the first model where the model actually invents new things in a way that matters."
>
> "我预计这将是第一个能以真正重要的方式发明新事物的模型。"

OpenAI researchers showed Astra coordinating multiple agents on a math proof, tearing through desktop software at what Altman called a "superhuman, very fast" pace, and producing slides, financial reviews, and analysis of messy data. It is designed to work for days or weeks, remember corrections, collaborate with other agents and people, and act across software tools. (I left wondering what the implications will be for all of the enterprise software that OpenAI clearly trained Astra to operate faster than a human.)

OpenAI 的研究人员展示了 Astra 如何协调多个智能体完成数学证明，以奥特曼所说的"超越人类、极快"的速度处理桌面软件，并生成幻灯片、财务审查以及对杂乱数据的分析。它旨在持续工作数天或数周，记住修正意见，与其他智能体及人类协作，并跨软件工具执行任务。（我不禁开始思考，对于那些 OpenAI 显然为了让

Astra 能以超越人类的速度进行操作而训练过的企业软件来说，这将意味着什么。）

Altman said OpenAI wants to offer a version that "runs forever" in ChatGPT and through its API, though it will be expensive at first. Customers pressed the company on memory, permissions, audit trails, duplicate agents, and whether an agent should act under an employee's identity or its own.

Altman 表示，OpenAI 希望在 ChatGPT 及其 API 中提供一个可以"永久运行"的版本，尽管初期价格会很昂贵。客户就内存、权限、审计追踪、重复智能体，以及智能体应该以员工身份还是以自身身份运行等问题向公司施压。

Thibault Sottiaux, who runs OpenAI's combined product organization, told me a new interface for continuous agents is coming "very soon."

负责 OpenAI 综合产品部门的 Thibault Sottiaux 告诉我，一种用于持续运行智能体的新界面"很快"就会推出。

"There isn't really an app," he said. It will be accessible from almost anywhere and will "definitely" feel like a new thing rather than an upgrade.

“它其实算不上是一个应用程序，”他说道。它将几乎随处可用，并且“绝对”会让人感觉是一个全新的事物，而非一次简单的升级。

Leaders think they basically have AGI

领导层认为他们基本上已经实现了 AGI

Altman told me OpenAI was “very close,” based on its latest internal progress. He put his estimate at the end of this year:

Altman 告诉我，根据 OpenAI 最新的内部进展，公司已经“非常接近”了。他给出的预估时间是今年年底：

> “I would guess that by the end of the year, we would have an internal system where I would say, ‘Okay, fine, it’s an AGI.’”
>
> **“我猜到今年年底，我们内部会拥有一个让我感叹‘好吧，行了，这就是 AGI’的系统。”**

Chief research officer **Mark Chen** estimated that OpenAI is “80% of the way” there. **Greg Brockman** told me people looking back two years from now may remember this as the period when AGI was created.

首席研究官 Mark Chen 估计 OpenAI 已经完成了“80%

的进度”。Greg Brockman 则告诉我，两年后人们回首往事时，可能会将这段时期视为 AGI 诞生的时刻。

OpenAI's charter still defines AGI as "highly autonomous systems that outperform humans at most economically valuable work." But it's clear the term has lost its meaning over time, despite it still being plastered all over the walls of OpenAI HQ as I walked between meetings. Still, it feels significant for company leadership to admit the milestone has practically been reached, and to do so with what amounts to a shrug.

OpenAI 的章程仍将 AGI 定义为“在大多数具有经济价值的工作中表现优于人类的高度自主系统”。但显而易见的是，尽管在我穿梭于会议室之间时，这个词依然贴满了 OpenAI 总部的墙壁，但随着时间的推移，它已经失去了原有的意义。尽管如此，公司领导层承认这一里程碑已基本实现，且表现出一种近乎“耸耸肩”般云淡风轻的态度，这依然显得意义重大。

A key AI researcher milestone has been hit

一项关键的 AI 研究人员里程碑已经达成

Chief scientist **Jakub Pachocki** told me OpenAI has met its internal benchmark for automating an entry-level AI

researcher.

首席科学家 Jakub Pachocki 告诉我，OpenAI 已经达到了其内部基准，实现了初级 AI 研究人员工作的自动化。

Given an experimental idea, Pachocki said Astra can implement it in OpenAI's codebase, run the experiment, and return the results. He said it can also take a paper and perform work that would have taken a human researcher about a week.

Pachocki 表示，给定一个实验想法，Astra 可以在 OpenAI 的代码库中实现它，运行实验并返回结果。他说，它还可以处理一篇论文，并完成人类研究人员大约需要一周才能完成的工作。

Pachocki described day-scale execution, not a system capable of choosing an entire research program or deciding what its results mean. The benchmark is internal and has not been independently replicated. By his account, OpenAI's models are already performing some of the experimental work used to build their successors. I'm still seeing a lot of AI experts talk about recursive self-improvement as if it's coming soon, but my takeaway from spending time inside OpenAI is that it's already here.

Pachocki 描述的是天级别的执行能力，而非一个能够

选择整个研究项目或决定其结果意义的系统。该基准是内部的，尚未经过独立复制。根据他的说法，OpenAI 的模型已经在执行一些用于构建其后继版本的实验工作。我仍然看到许多 AI 专家在谈论递归自我改进，仿佛它即将到来，但我通过在 OpenAI 内部观察得出的结论是：它已经实现了。

OpenAI's slowdown of frontier research is unprecedented

OpenAI 放缓前沿研究的举措是前所未有的

OpenAI kept its clusters running, and ordinary pretraining continued. The company delayed or froze some frontier reinforcement-learning runs in which models interact with computers and tools.

OpenAI 保持了集群的运行，普通的预训练也在继续。但公司延迟或冻结了一些前沿的强化学习运行，在这些运行中，模型会与计算机和工具进行交互。

Altman said OpenAI had never paused training runs this way before. There was no single smoking gun. The decision followed a mix of concerning behavior, fast capability gains, and a belief that safety, alignment, and security had fallen

behind. Researchers who did not normally work on alignment volunteered to help, and the compute shifted toward monitoring and safeguards.

Altman 表示，OpenAI 此前从未以这种方式暂停过训练运行。这并非由单一的突发事件引起，而是源于一系列令人担忧的行为、能力的快速提升，以及一种认为安全、对齐和保障工作已经滞后的信念。平时不从事对齐工作的研究人员也自愿提供帮助，计算资源也转向了监控和防护措施。

Pachocki said confidence in alignment and safety has quickly become as limiting to OpenAI's progress as access to computing power. That's quite a statement and a sign of what's coming in terms of model capability.

Pachocki 表示，对对齐和安全的信心已迅速成为限制 OpenAI 进展的因素，其影响程度不亚于计算能力的获取。这是一个相当有分量的表态，也预示着模型能力即将迎来突破。

The former leader of ChatGPT is starting an enterprise product team

ChatGPT 前负责人正组建企业产品团队

Nick Turley, who ran ChatGPT from the beginning, now

leads a new enterprise-product division.

从 ChatGPT 创立之初就负责该产品的 Nick Turley，目前正领导一个新的企业产品部门。

He described an applied-product effort organized around industry verticals. It will target large sectors of the economy that rely on computer-based knowledge work, then adapt OpenAI's models, connectors, interfaces, and "whole product stack" to each domain, often in partnership with startups.

他描述了一项围绕垂直行业组织的垂直应用产品计划。该计划将针对依赖计算机知识工作的各大经济领域，通过与初创公司合作，将 OpenAI 的模型、连接器、界面以及"整个产品栈"适配到各个特定领域。

OpenAI is going to build a humanoid

OpenAI 将制造人形机器人

OpenAI's early research work trained a robot hand in simulation. It later disbanded that team to focus on language models and began rebuilding its robotics effort in 2024.

OpenAI 早期的研究工作曾在模拟环境中训练过机械手。后来，该公司解散了该团队以专注于语言模型，并于 2024 年开始重建其机器人研发工作。

Altman told me the company plans to eventually build a humanoid. He cited data-center construction and robots that build other robots as earlier uses than a personal robot. It sounds like this is still many years away, but Altman was very declarative that it's coming.

Altman 告诉我，公司计划最终制造出人形机器人。他提到，数据中心建设和制造其他机器人的机器人将比个人机器人更早投入应用。虽然这听起来仍需多年时间，但 Altman 非常肯定地表示这一天终将到来。

The first consumer device with Jony Ive will sit on a table

与 Jony Ive 合作的首款消费级设备将放置在桌面上

Altman said the first device is coming "soonish" and described a small handful of form factors: one for a table, another for a pocket, and another to be worn on the body. He said he does not like glasses because he finds it uncomfortable talking to someone wearing a camera and a light.

Altman 表示第一款设备将"很快"面世，并描述了几种形态：一种放在桌上，一种装在口袋里，还有一种佩戴在身上。他说自己不喜欢眼镜，因为他觉得与戴着摄像头和指示灯的人交谈会感到不适。

关和指示灯的人交谈会感到不适。

The products OpenAI is designing with **Jony Ive** are proactive, ambient computers. Altman acknowledged the potential privacy problem. He said OpenAI plans to introduce new privacy controls as the launch gets closer.

OpenAI 与 Jony Ive 共同设计的这些产品是主动式的环境计算设备。Altman 承认了潜在的隐私问题。他表示，随着发布日期的临近，OpenAI 计划推出新的隐私控制措施。

There's a debate about whether to sell compute

关于是否出售算力的争论

There's an unresolved debate within OpenAI about whether to eventually sell compute capacity. "Not anytime soon. We just need the compute," Altman told me. He left the door open if OpenAI's chip, robotics, supply-chain, and data-center programs eventually let it build infrastructure more cheaply and effectively than anyone else.

OpenAI 内部对于最终是否出售算力容量存在尚未解决的争论。"短期内不会。我们自己就需要这些算力，" Altman 告诉我。但他留有余地，如果 OpenAI 的芯

片、机器人、供应链和数据中心项目最终能让其比任何人都更廉价、更高效地构建基础设施，情况可能会有所改变。

Sachin Katti, who directly oversees OpenAI's compute efforts, described two possibilities: monetizing older chips near the end of five-year contracts that frontier models may outgrow but smaller open-source models could still use, or leveraging OpenAI's infrastructure expertise to build for others. "Nothing has been decided," he said.

直接负责 OpenAI 计算业务的 Sachin Katti 描述了两种可能性：一是将接近五年合同期满的旧芯片货币化，虽然前沿模型可能不再适用，但较小的开源模型仍可使用；二是利用 OpenAI 的基础设施专业知识为他人进行建设。他表示："目前尚未做出任何决定。"

"We just don't have enough." **Uday Ruddarraju**, who oversees OpenAI's end-to-end compute stack, told me he negotiates scarce capacity among research, consumer products, enterprise customers, and longer-term bets. "If we had more compute, we'd have definitely made more revenue last year."

"我们就是没有足够的资源。"负责 OpenAI 端到端计算堆栈的 Uday Ruddarraju 告诉我，他在研究、消费者产品、企业客户以及长期押注之间协调稀缺的算力容量。"如果我们有更多算力，去年的收入肯定会更高。"