



Generative AI's Act Two

Generative AI's Act Two

生成式人工智能的第二序章

长期以来，科学家、历史学家和经济学家一直在研究创造创新寒武纪大生物爆发的最佳条件。在生成式人工智能领域，我们创造了一个现代奇迹，即我们这一代人的太空竞赛。

这一时刻已经酝酿了几十年。六十年的摩尔定律为我们提供了处理埃菲尔普斯数据的计算能力。四十年的互联网(由 COVID 加速)为我们提供了数万亿token (文本的基本单元) 的训练数据。二十年的移动计算和云计算让每个人都拥有了掌上超级计算机。换句话说，数十年的技术进步积累为生成式人工智能的腾飞创造了必要条件。

ChatGPT 的崛起是点燃导火索的火花，它释放出的创新密度和热情是我们多年来从未见过的——也许是自互联网诞生之初以来从未见过的。这种令人窒息的兴奋在 "智力谷 " (位于旧金山) 尤为明显。在那里，人工智能研究人员达到了摇滚明星的地位，每个周末，黑客之家都被新的自主代理人工智能和陪伴聊天机器人挤得满满当当。人工智能研究人员从众所周知的 "车库里的黑客 "转变为指挥数十亿美元计算的特种部队。arXiv 预印网站已经变得如此高产，以至于研究人员开玩笑地呼吁暂停发表新论文，以便赶上进度。

但很快，对人工智能的兴奋就变成了近乎歇斯底里的狂热。突然之间，每家公司都成了 "人工智能副驾驶"。我们的收件箱里塞满了 "AI Salesforce"、"AI Adobe "和 "AI Instagram "的无差别推销。在做出产品前的一亿美元种子轮融资又回来了。我们发现自己陷入了**难以为继的融资、人才争夺战和 GPU 采购狂潮**。

果然，裂缝开始显现。艺术家、作家和歌手对生成式人工智能**IP的合法性**提出了质疑。关于道德、监管和迫在眉睫的超级智能的争论席卷华盛顿。也许最令人担忧的是，硅谷内部开始流传一种说

法,即**生成式人工智能实际上并无用处**。产品远未达到预期,**糟糕的用户留存率**就证明了这一点。最终用户对许多应用的需求开始趋于平稳。这是否又是一个虚无缥缈的周期?

人工智能的 "不满之夏" 让批评家们高兴得手舞足蹈, 这不禁让人想起互联网的早期。1998 年, 一位著名的经济学家宣称: "到 2005 年, **互联网对经济的影响显然不会超过传真机**。"

别误会——尽管喧嚣、歇斯底里、充满不确定性和不满, 生成式人工智能的起步已经比 SaaS (软件即服务) 更为成功, 仅初创企业的收入就超过了 10 亿美元 (SaaS 市场花了数年而非数月才达到同样的规模)。一些应用程序已经家喻户晓: **ChatGPT 成为增长最快的应用**, 在学生和开发者中具有特别强的产品与市场契合度; Midjourney 成为我们共同的创作灵感, 据说**仅靠11人的团队就实现了数亿美元的收入**; Character AI 普及了人工智能娱乐和陪伴, 创造了我们最渴望的消费者 "社交" 应用——**用户在应用内平均花费两小时**。

尽管如此, 这些早期成功的迹象并不能改变这样一个现实: 很多人工智能公司根本**没有产品与市场的契合点**, 也没有可持续的竞争优势, 人工智能生态系统的整体繁荣是不可持续的。

现在尘埃落定, 我们认为这是一个合适的时机来放大并反思生成式人工智能——我们今天所处的位置, 以及我们可能的方向。

迈向第二幕

生成式人工智能走出圈的第一年——"**第一幕是从技术出发的**。我们发现了一个新的 "锤子" (基础模型), 并推出了一**大波新奇的应用程序**, 轻便地展示了很酷的新技术。

现在, 我们相信市场正在进入 "**第二幕——从用户出发 (回归用户)**。第二幕将端到端的解决人类的问题。这些应用在本质上不同于第一代应用。它们倾向于**使用基础模型作为更全面解决方案的一部分**, 而不是整个解决方案。它们引入了新的编辑界面, 使工作流程更有粘性, 输出效果更好。它们通常是**多模态的**。

市场已经开始从 "**第一幕**" 向 "**第二幕**" 过渡。进入 "第二阶段" 的公司包括: 为高端律师事务所定制法律大语言模型的 Harvey 公司; 为我们的工作空间建立索引和获取数据, 使生成式人工智能与工作更加相关的 Glean 公司; 以及正在创建数字陪伴的 Character 和 Ava 公司。

市场的两个重要趋势: **生成式人工智能从技术之锤到实际用例和价值的演变**, 以及**生成式人工智能应用日益多模态的本质**。

重新审视我们的论点

我们在2022年的文章中提出了生成式人工智能市场机遇的论点, 并假设了市场将如何发展。我们做得怎么样?

我们错在哪里?

(1) **事情发生得太快了!** 去年, 我们曾预计, 要等到近十年后, 我们才能拥有内部代码生成、好莱坞品质的视频或听起来不机械的人类品质的语音。但只要听一听 **Eleven Labs** 在 TikTok 或

Runway 的人工智能电影节上的声音, 就会清楚地意识到未来已经飞速到来。就连 3D 模型、游戏和音乐也在迅速变得优秀。

(2) **瓶颈在于供应方**。我们没有预料到最终用户的需求会在多大程度上超过 GPU 的供应。许多公司的增长瓶颈很快就不是客户需求, 而是如何获得 Nvidia 的最新 GPU。漫长的等待时间成为常态, 于是出现了一种简单的商业模式: 支付订阅费, 就可以不用排队, 获得更好的型号。

(3) **垂直分离尚未发生**。我们仍然相信, "应用层"公司和基础模型提供商之间会出现分离, 模型公司专注于规模和研究, 而应用层公司则专注于产品和用户界面。实际上, 这种分离还没有完全实现。事实上, 最成功的面向用户的应用程序都是垂直整合的。

(4) **残酷的竞争环境和从业者的迅速反应**。去年, 竞争环境中出现了一些过度拥挤的类别(尤其是图像生成和文案写作), 但总体而言, 市场还是一片空白。如今, 竞争格局的许多角落都是竞争大于机遇。从谷歌的 Duet 和 Bard 到 Adobe 的 Firefly, 现有公司的迅速反应, 以及现有公司最终愿意 "冒险" 的意愿, 都放大了竞争的热度。即使是在基础模型层, 我们也看到客户正在建立自己的基础架构, 以便在不同供应商之间互不干扰。

(5) **护城河在于客户, 而非数据**。我们曾预测, 最好的生成式人工智能公司可以通过数据飞轮产生可持续的竞争优势: 更多的使用→更多的数据→更好的模型→更多的使用。虽然这在某种程度上仍然是正确的, 尤其是在那些拥有非常专业和难以获取的数据的领域, 但 "数据护城河" 的基础并不稳固: 应用公司生成的数据并不能形成不可逾越的护城河, 而下一代基础模型很可能会抹去初创公司生成的任何数据护城河。相反, 工作流程和用户网络似乎正在创造更持久的竞争优势。

我们做对了以下几点

(1) **生成式人工智能正在改变世界**。突然间, 每个开发者都在开发生成式人工智能应用, 每个企业买家都需要生成式人工智能。市场甚至保留了 "生成式人工智能" 的名称。人才流入市场, 风险资本也纷纷涌入。生成式人工智能甚至在爆火视频中成为一种流行文化现象, 如 "哈利-波特-巴伦西亚加", 或由 Ghostwriter 创作的模仿 Drake 的歌曲 "Heart on My Sleeve", 这首歌已成为排行榜上的热门歌曲。

(2) **第一批杀手级应用出现了**。有资料显示, ChatGPT 是最快达到 1 亿 MAU 的应用, 而且只用了 6 周时间。相比之下, Instagram 用了 2.5 年, WhatsApp 用了 3.5 年, YouTube 和 Facebook 用了 4 年才达到这一用户需求水平。但 ChatGPT 并不是一个孤立的现象。Character AI 的深度参与(平均使用时长为 2 小时)、Github Copilot 的生产力优势(效率提高 55%)以及 Midjourney 的变现路径(数亿美元的收入)都表明, 第一批杀手级应用已经到来。

(3) **开发者是关键**。Stripe 或 Unity 等开发者优先型公司的核心洞察之一是, 开发者路径开启了无法想象的用例。在过去的几个季度里, 从音乐生成社区到人工智能红娘, 再到人工智能客服, 我们已经接触到了各种各样的产品。

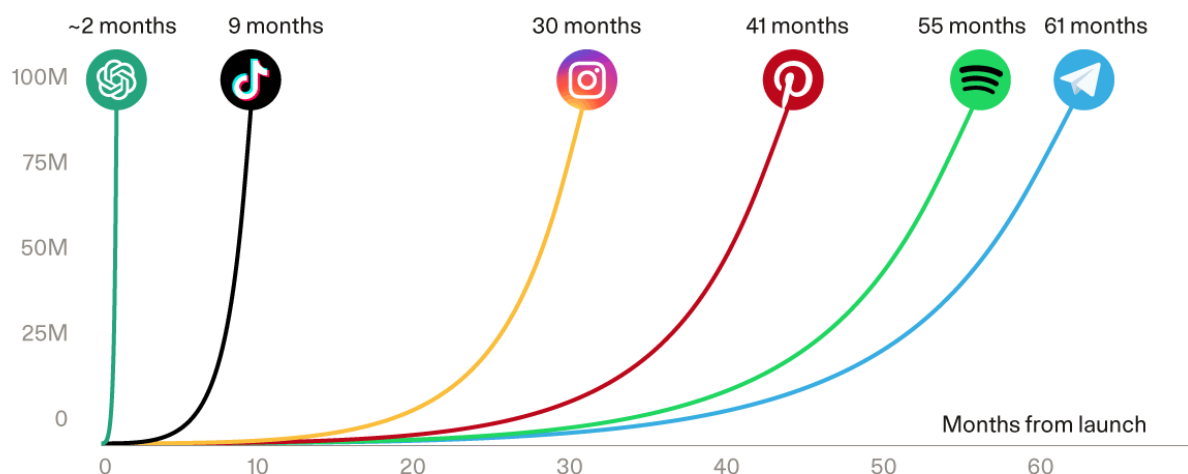
(4) **范式正在进化**。人工智能应用的最初版本主要是自动完成和初稿, 但现在这些形式因素正变得越来越复杂。Midjourney 引入的摄像头平移和填充功能很好地说明了生成式人工智能用户体验是如何变得更加丰富的。从整体上看, 范式正在从个体级生产力向系统级生产力发展, 并从人类参与决策向以执行为导向的代理系统发展。

(5) **版权、伦理和存在主义的恐惧**。关于这些热门话题的争论此起彼伏。艺术家、作家和音乐家们各执一词, 一些创作者对他人从衍生作品中获利感到义愤填膺, 而一些创作者则对新的人工智能现实表示欢迎。规则不透明: 日本宣布用于训练人工智能的内容没有知识产权, 而欧洲则提出了严厉的监管措施。

我们现在处于什么位置？生成式人工智能的价值问题

生成式人工智能并不缺乏应用案例或客户需求。用户渴望人工智能能让他们的工作变得更轻松，让他们的工作产品变得更好，这就是为什么他们以创纪录的速度涌向各种应用的原因（尽管缺乏自然分布）。

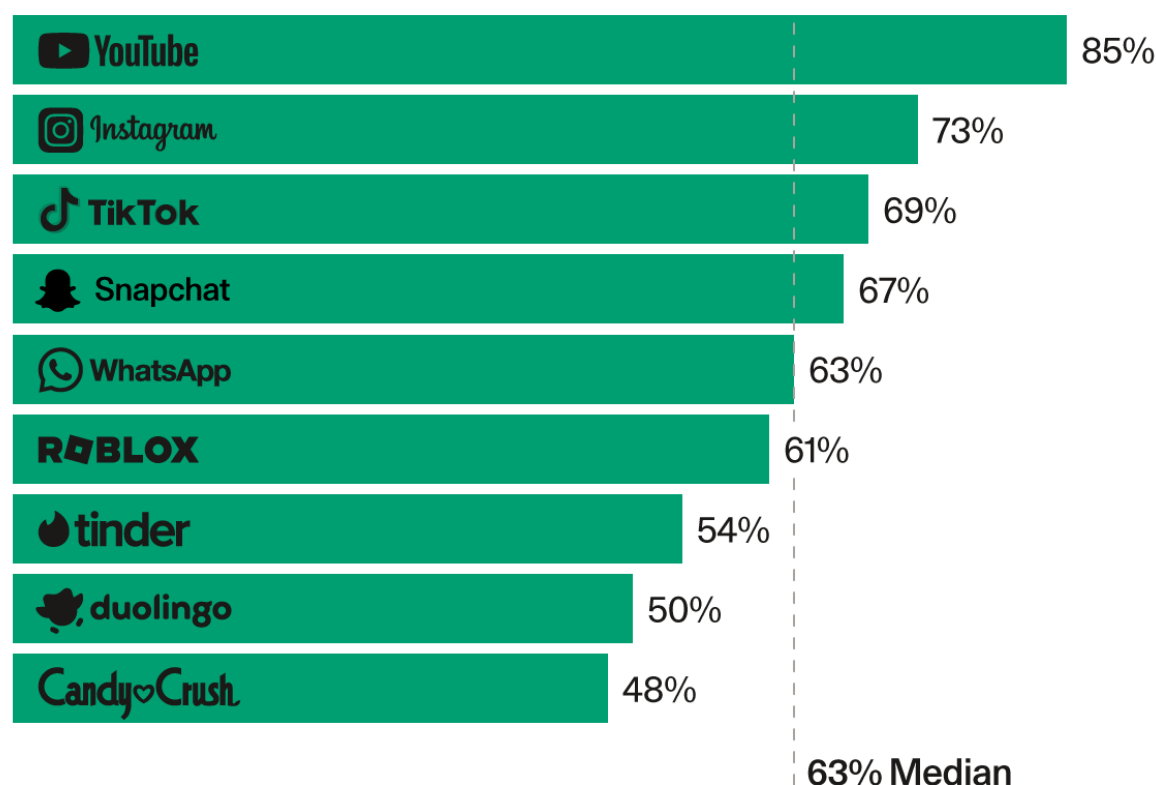
Path to 100 Million Users (stylized)



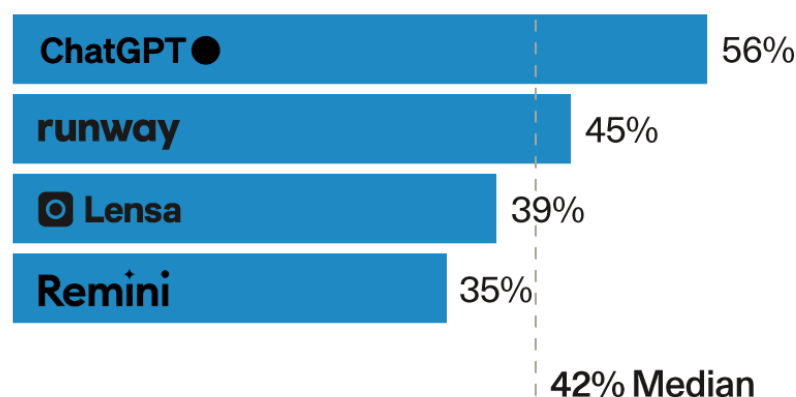
但人们会留下来吗？其实不然。下图比较了人工智能优先应用与现有公司的移动应用第一个月留存率。

One Month Retention

Incumbents



AI-First Companies

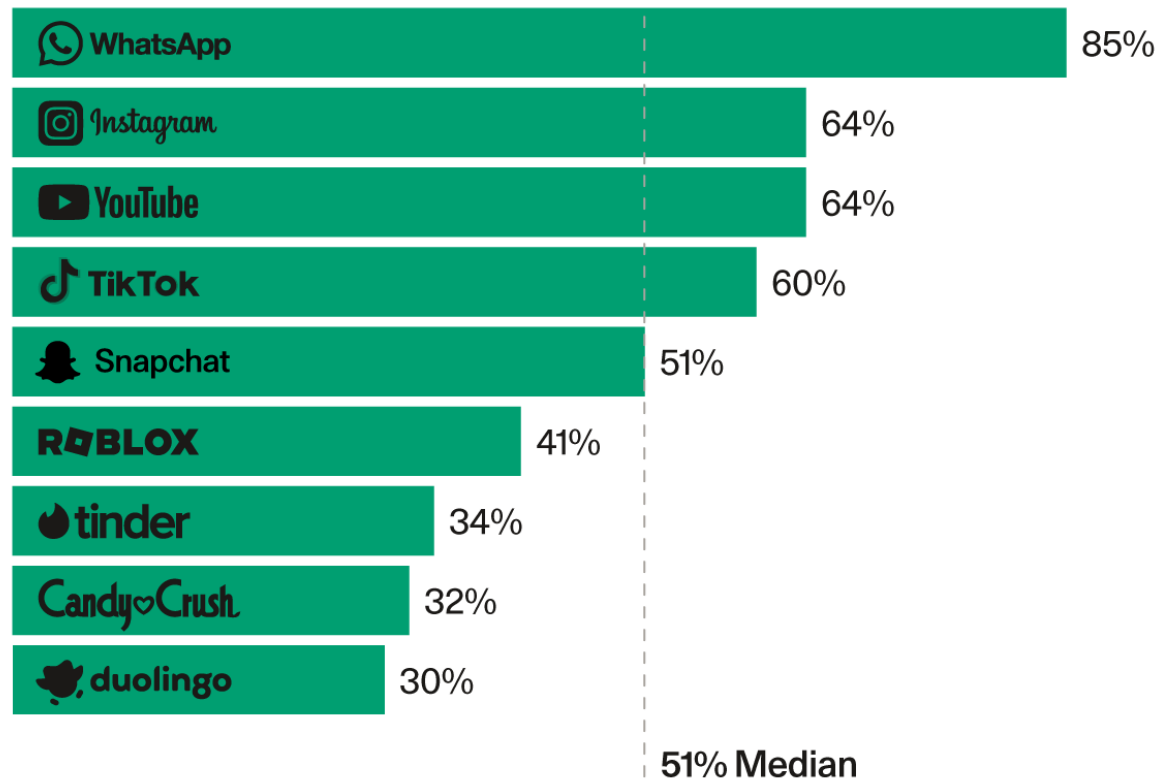


Data from mobile apps only.
Averaged over the past 12
months of cohorts in the US.

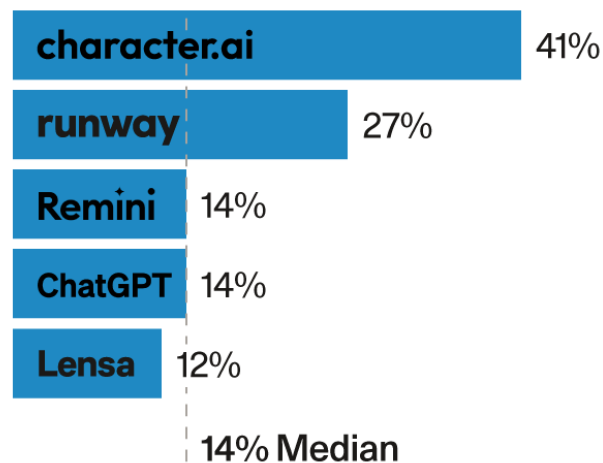
用户参与度也乏善可陈。一些优秀的消费类公司的 DAU/MAU (每日活跃用户和每月活跃用户的比例) 为 60-65%; WhatsApp 的 DAU/MAU 为 85%。相比之下, 生成式人工智能应用程序的中位数仅为 14% (Character AI 和 "人工智能陪伴" 类别是个明显的例外)。这意味着, 用户还没有发现生成式人工智能产品的足够价值, 以至于每天都要使用它们。

DAU/MAU

Incumbents



AI-First Companies



Data from mobile apps only.

简而言之，**生成式人工智能最大的问题**不是寻找用例、需求或分配，而是**证明价值**。正如我们的同事大卫-卡恩 (David Cahn) 所写的那样，"2000 亿美元的问题是：你打算用这些基础设施来做什

么？它将如何改变人们的生活？”要建立持久的业务，就必须解决用户留存问题，为用户创造足够深的价值，使他们持续使用，成为每日活跃用户。

但是，让我们不要绝望。生成式人工智能仍处于“尴尬的青少年时期”。会有曙光乍现的时刻，当产品未能达到预期时，失败往往是更可能的、可重复的和可修复的。我们任重而道远。

结束语

随着“前沿悖论”（由于我们把超越我们掌握的技术的前沿归因于人类，因此这个前沿永远是不明确的）的临近，随着变革者和扩散模型的新颖性逐渐消退，生成式人工智能市场的性质也在不断演变。炒作和噱头正在让位于真正的价值和完整的产品体验。

在红杉，我们仍然是生成式人工智能的坚定信仰者。这个市场腾飞的必要条件已经积累了几十年，市场终于到来了。杀手级应用的出现和终端用户的巨大需求加深了我们对市场的信心。

然而，Amara定律——我们往往会高估一项技术的短期效应，而低估其长期效应——正在发挥作用。我们在做出投资决策时要有耐心和判断力，仔细研究创始人是如何解决产品价值问题的。各公司正在用来推动模型性能和产品体验的发展的共同游戏规则，让我们对生成式人工智能的第二序章充满信心。

全文终。

希望红杉资本的这篇超高水准文章对大家有帮助！

如果大家有关于生成式AI的新颖的产品思路，欢迎联系我们：talktalkai.kevin@gmail.com