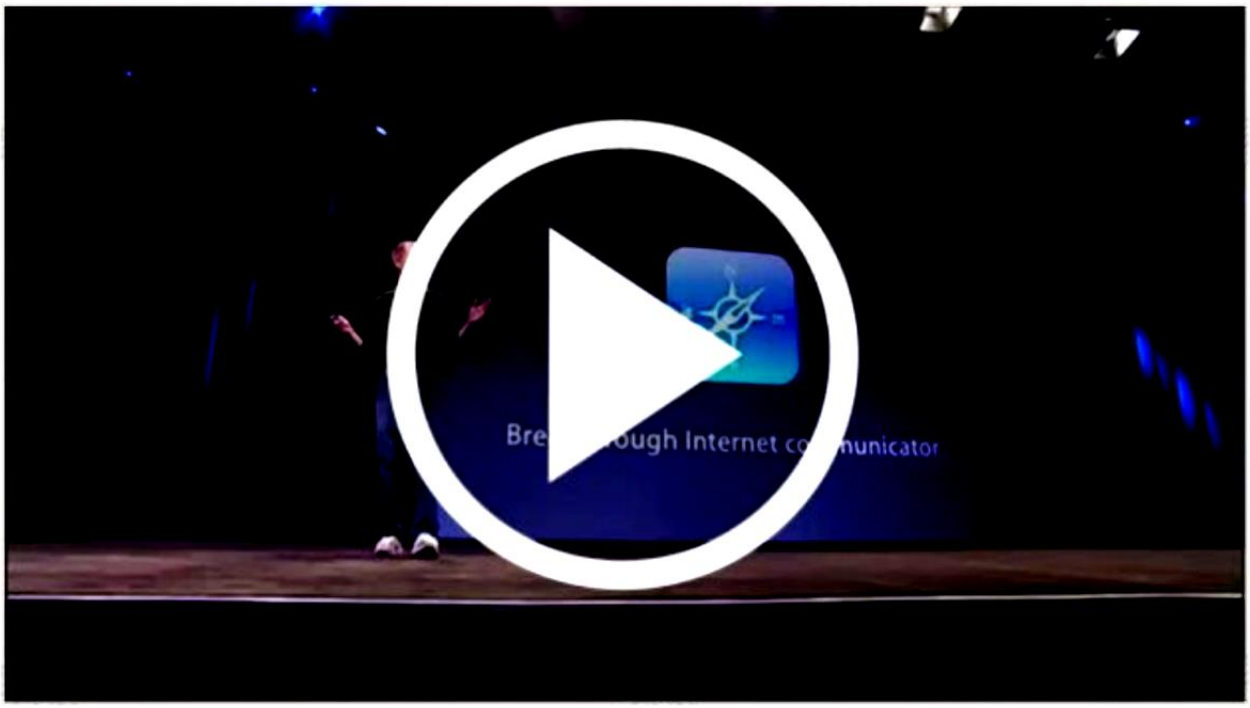
在浏览器中查看 | 在播客播放器中收听

STRATECHERY

应用程序、代理与聚合

2026年9月28日星期一

三款革命性产品——你知道那句经典台词。一台带触控的宽屏iPod、一部革命性的手机，以及一款突破性的互联网通信设备。史蒂夫·乔布斯就是这样介绍iPhone的：



今天，我们推出了这一类产品中的三款革命性产品。第一款是带触摸控制的宽屏iPod。第二款是革命性的移动电话。第三款是突破性的互联网通信设备。所以，三样东西：带触摸控制的宽屏iPod，革命性的移动电话，以及突破性的互联网

> iPod、革命性的移动电话，以及突破性的互联网通信设备。一个iPod。一部电话。还有一个互联网通信器。一个iPod。一部电话。你们明白了吗？这些不是三个独立的设备，而是一个设备，我们称之为iPhone。今天，苹果将重新发明电话。

我之前曾链接过这段片段，通常是为了指出观众并未真正理解“互联网通信设备”意味着什么，尽管这正是iPhone最具革命性的能力所在。

然而，我此刻想到的一句话是：“你明白了吗？”你明白这种信息传递方式——现在用聊天机器人，以后用自然界面——就是我们与人工智能沟通的方式吗？你明白我们不仅不会去编程计算机，甚至不会再使用它们——而是由人工智能来使用吗？你明白预构建的用户界面——一次编写，到处运行，面向所有人——已经死了吗？这些不是三个独立的预测：这就是现实，就在此刻，在2026年。未来已经到来，即使它尚未广泛普及。还是说，它已经普及了？

消息传递与待办任务

我在2014年2月18日写道，消息传递是移动设备的杀手级应用；我特别指出日期，是因为我碰巧在Facebook收购WhatsApp的前一天发表了那篇文章。我的论点是，手机最初可能只是一种便捷的信息设备，但其杀手级用例却非常人性化：

> 七年前，电脑变得可以放进口袋，但最初的应用场景主要是让信息的被动呈现不仅能在观看者方便的时间获取，还能在任何地点获取：网络如今无处不在。不过，直到最近，这个时代的杀手级应用才变得显而易见——当通信节点是智能手机时，这个应用就是即时通讯。家用电话实现了实时通信，网络实现了被动通信，而即时通讯则实现了持续通信。对话永无止境，朋友的来去节奏不由物理空间决定，而由注意力决定。而且，既然我们都是人类，渴望人际互动与情感，我们便非常乐意将大量注意力投入即时通讯，投向我们最在乎、且始终在口袋和手包里陪伴着我们的人。

Meta对WhatsApp的收购实际上在雄心壮志上是一次倒退；一年前，该公司曾推出自己的安卓界面，名为Facebook Home；当时首席执行官马克·扎克伯格对《连线》杂志表示：

> Home将你的手机变成了一部Facebook设备。即使锁屏状态下，朋友们动态的照片流也会充满屏幕。更新也会出现在你的主屏幕上。更重要的是，Home让Facebook成为你设备上的主要通信方式。该公司的消息软件与短信融合，你可以在其他应用内继续使用其“聊天头像”发消息。扎克伯格相信，社交网络在用户生活中扮演着太重要的角色，不能被浩瀚的应用海洋所淹没。“应用不是世界的中心，”他说，“人才是。”

我对此强烈反对，并在《应用、人与待办任务》中写道：

> 应用与人。据扎克伯格所言，这就是二分法。但他错了。他忽略了待完成的任务。
>
> 以下是我精心挑选的主屏幕：
>
> 
>
>
> 我的iPhone主屏幕
>
> 注意有多少图标是关于“人”的……三个屏幕上总共只有四个。人对我来说很重要，但我使用手机做的事远不止于此。
>
> 那么，如果我考虑“待完成的任务”呢？……总计：151个应用，80项待完成任务，4个以人为焦点
>
> 应用并非世界的中心……但人也不是。智能手机

> 之所以统治世界，是因为它们比人类历史上任何事物都能在更多地方为更多人完成更多任务。Facebook Home让任务更难完成，实际上将它们降级到了我第三屏的文件夹里。Facebook凭什么优先考虑我的任务？

那些截图显示了151个应用；如今我的手机上有689个。不，我很少打开其中的大多数，但嘿，如果我再次需要过去因某种原因下载的随机应用，它就在那里。更广泛地说，这些数字说明了iPhone和App Store所实现的可能：以前在移动中无法访问的各种活动的定制界面。

尽管如此，核心应用集仍然非常一致：消息应用，包括WhatsApp，社交媒体，主要是X，以及现在的ChatGPT和Claude。

为AI打造的电脑

我昨天得到了一个新应用，巧合的是，灵感来自一个WhatsApp群聊。

一段时间以来，我一直全力以赴地投入智能体。先是编码智能体，然后是我从专门的 Claude 线程中打造出的整体创意追踪智能体，最近则是一个专门针对我公司需求调校的智能体。正是这段经历让我立刻对 Muse 感到兴奋，因为它真正为人们提供了拥有自己智能体所需的基础设施。

关于智能体，关键要理解的一点是：它们不只是 AI

关于智能体，关键要理解的一点是：它们不只是 AI，而是能够访问计算机的 AI。在 ChatGPT 时代早期就已经很清楚，AI 不会取代计算机，而是会操作它们；从 2023 年的《ChatGPT 获得了一台计算机》开始，当时 ChatGPT 新增了插件，其中包括来自 Wolfram|Alpha 的一个插件：

> 这件事能如此奏效，本身就是对助理人工智能本质的一种证明：它们既不同于我们以往理解的计算机，又在“思考”和沟通方式上惊人地像人。坦率地说，那三个问题我自己也很难解决——那正是计算机该做的事！而现在，ChatGPT 有了属于自己的计算机。

这个想法并不是说概率性LLM会以某种方式变得确定性，而是说LLM能够利用确定性计算机来做计算机的事情；这正是代理所做的事，并通过其记录信息的能力得到增强。因此，要真正交付一个可供无法或不愿为其代理设置计算机的人使用的代理，你必须给人们的不仅是一个代理，还要为他们的代理提供一台计算机。

这就是为什么我在 Muse 发布中关注的不是支撑 Muse 的 Muse Spark 模型，而是 Meta 正在为美国每一位用户（想必最终也会扩展到全世界）提供一台配备 2 核处理器、8GB 内存和 8GB 存储空间的虚拟机。这是一台货真价实的计算机，相当了不起，而且也是让智能体为大多数人服务的唯一途径。

当然，问题仍然在于你实际上用这台计算机做什么，而我的新应用正是为此而生：在听说一位朋友如何用 Muse 查看他的 Instagram 收藏后，我问 Muse：

> 你能帮我整理一下我在Instagram上保存的所有食谱吗？

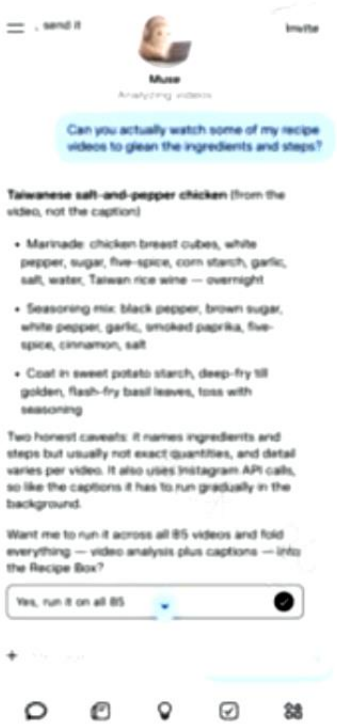
在最初交付了一份PDF后（“嗯，PDF不太实用。你不能把它们存储成更便于访问的格式吗？”），我给自己弄了一个新应用：



整个过程花了5分钟，而且整段对话都是在我遛狗时进行的。

这个应用并不完美：目前它实际上只是一种对Instagram视频进行分类的方式；我让Muse真正去读取字幕并观看视频；它正在处理这件事，但受限于

Instagram的速率限制：



我可以耐心等待；在此期间我还有689个其他应用要看。

我当然不会那么做；事实上，我查看的应用数量正在急剧减少。我理解为什么让代理访问你自己的电脑这个想法很吓人，但每个带命令行的应用早就已经能被代理使用了；从GPT-5.6 Sol开始，每个带用户界面的应用也都能被使用了，只是速度慢。Astra让它变快了。然后，就是这样：AI基本上可以使用任何我不想用的应用或网站。

坦白说，我对此并无异议。我之所以有689个应用，是因为我拥有那么多服务或游戏，至少在某个时刻，我想要或需要与之互动；它们都要求我学习如何操作，以便完成我想完成的事情。归根结底，关键在于：我不是专业的应用使用者，我只是一个想把事情完成的人。而人工智能帮我做到了。

从发现到灵感

我把我的新Muse-built Recipe Box应用标为手机上的第690个应用。但实际上，它代表着无限应用：我选择的代理可以根据命令创建任何我想要的app，即使这个app只为我一人所用。这是随着LLM兴起而变得清晰可见的另一个未来的最早体现：真正定制化的UI。摘自2024年的《The Gen AI Bridge to the Future》：

> 这就是你开始看到桥梁的地方：我所描述的其实是生成式AI的一种应用，具体来说，是针对按需UI界面的应用。这也是你能想象在现有设备上很有用的一种应用。例如，一个手表应用如果能在特定时刻直接显示你需要做出的确切选择，而不是像小型iPhone那样通过触摸导航，那将会实用得多。同样，今天我们通过确定性编程已经看到了些许端倪，但最终的应用将是通过生成式AI实现的按需应用。
>
> 当然，生成式AI也可以在手机上使用，而目前我预计生成式UI的大部分探索都将发生在手机上。我们确实看到生成式AI的广泛实验和快速发展，正如我们在PC上看到互联网的大量实验和快速发展一样。这种实验和发展不仅在PC上可用，而且还搭建了通往智能手机的桥梁；我认为生成式AI也在做同样的事情，即搭建通往可穿戴设备的桥梁——这些设备不是配件，而是本身就是通

用计算机。

上周的Meta Connect一如既往地聚焦于Meta的新设备，主要是眼镜；然而今年值得注意的一点是，Muse如何让一切都变得合理。Meta的设备并非AR或VR设备，甚至不是AI眼镜：它们是Muse的交付机制。事实上，我之前谈到的通用计算机已经存在：Meta将其提供给了每位用户，供Muse使用。

而且，正如我们在我的食谱盒应用中看到的那样，生成式UI已经到来。不，它并不是我仍然认为即将到来的即时UI，但它无疑正走在通往那里的路上：我拥有了专门为我生成的定制UI，无需与任何人共享这个应用。它实际上是一次性的，因为它是无限的。

这是网站所经历过的情景的重演。过去，出版物是稀缺的，因为它们依赖于印刷机和送货车；后来，分发变得免费，网络变得丰富。推而广之，这意味着最需要解决的最重要问题是发现；解决了发现问题的公司聚集了需求，从而获得了对供应商的权力，尤其是Meta和Google，它们因此拥有了令人难以置信的广告业务（这就是聚合理论）。

然而，真正稀缺的其实是在网页或应用上实际执行操作的能力。代理（Agents）的出现意味着执行事务的能力正变得过剩，而稀缺的是意愿。需要解决的问题不是发现，而是灵感；那些能激发灵感的公司，将获得凌驾于所有有待完成之事的实体之上的权力。一旦用户聚焦于解决某个问题，完成该问题所需的所有应用和服务，就会被抽象为实现细节，沦为纯粹的供应

商，面临与聚合器（Aggregators）时代出版物相同的命运，为从代理（Agent）那里分得残羹冷炙而挣扎——而代理，正是用户需求乃至欲望的终极守门人。

聚合平台

来自The Verge：

> 在上个月预告了其新的 Copilot“超级应用”后，微软今天正式揭晓了它。经过重新设计的 Copilot 应用将三项 AI 能力整合到一个集聊天、编码和智能体于一体的单一界面中……“代码”标签页是这款所谓“超级应用”令人意外的新增功能，而且通常不会让人联想到一款为知识工作者设计的应用。它将让任何人都能创建应用、追踪器、仪表板或自动化流程，并作为云端托管的内部应用与同事分享成果……
>
> Autopilot 是 Copilot 体验中的新成员，[微软副总裁贾里德] 斯帕塔罗将其描述为一位“数字队友”。它此前名为 Scout，最初作为桌面应用提供；Autopilot 则是与之对应的云端版本，能让个人 AI 助手在你睡觉时继续运行。与许多其他 AI 智能体一样，Autopilot 拥有自己的云端机实例，可以被指派执行诸如监视 Teams 道、运行重复性工作任务或处理后续跟进等项。



> 与其他一些AI代理相比，真正的不同之处在于微软将Autopilot打造成了企业级产品，吸引了那些依赖其Office和身份管理套件的企业。斯帕塔罗表示：“Autopilot存在于你的租户中，拥有自己的身份、记忆、计算机和工作空间，并且构建在Microsoft IQ之上，因此它了解你的组织实际如何运作。”“它出现在人们已经在工作的地方——Teams、Outlook、聊天、频道和文档——因此你可以像@同事一样@它，背后有权限、审计和治理支持。”

萨提亚·纳德拉在X上明确表示，微软的战略一如既往：做办公的操作系统。

