

我的第一位AI员工

一家一人音乐公司与ChatGPT Pro合作一年的实践报告

这不仅仅是一个关于工具的故事，这是一个关于合作、
转型以及重新定义个人企业边界的故事。

一位独奏者的多重挑战

主人公

Michael Wall，一位作曲家、创业者、父亲和丈夫。

公司

Sound for Movement，为全球舞蹈家和编舞家提供音乐。

业务规模

- 每年处理100-200个委托项目
- 每年发行5-10张新专辑
- 管理超过500首曲目的音乐库

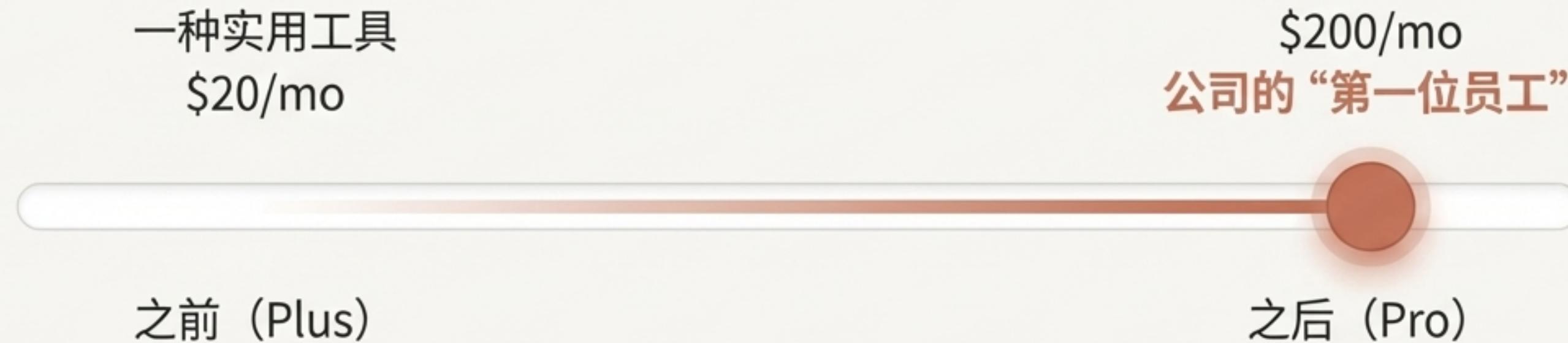
个人责任

同时需要照顾家庭，包括患有唐氏综合症的妻妹。



一项战略性投资：从工具到团队成员

将每月20美元的Plus订阅升级到每月200美元的Pro订阅，这不是成本的增加，而是对核心能力的指数级投资。



“我毫不怀疑，我从OpenAI获得的价值远远超过每月20美元，因此我在Pro套餐推出后几秒钟内就愉快地注册了。”

角色一：全栈开发者

背景：没有任何编程背景。

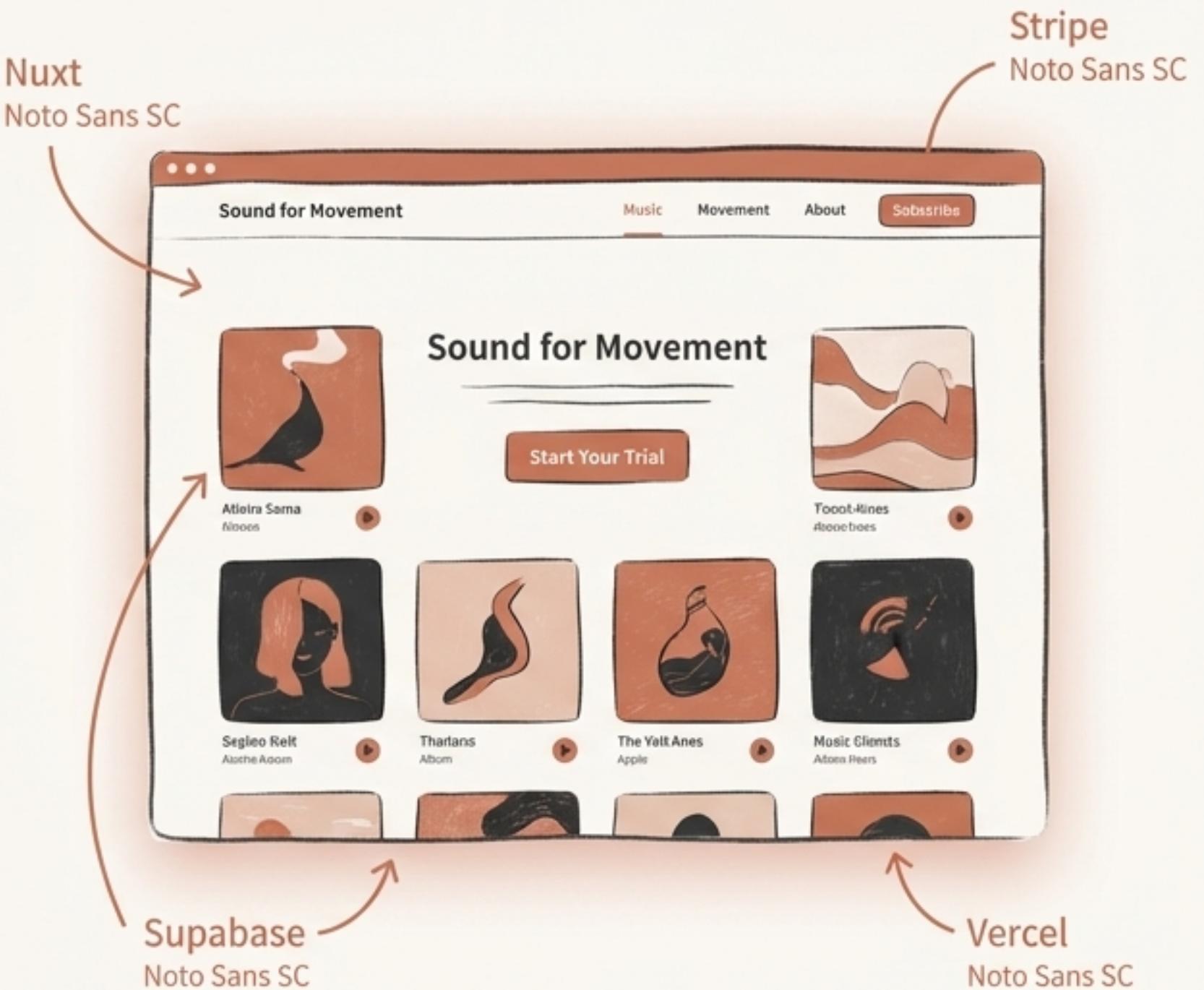
挑战：摆脱由多个SaaS产品拼凑而成的复杂系统，构建一个自有、高效的网站。

AI工作流程：

- 代码生成：**使用o1模型生成Nuxt代码文件。
- 代码应用与学习：**在Cursor中，利用其代码库感知代理应用代码，并耐心学习计算机科学基础知识。
- 迭代与上线：**经过三周和四次完全重建，新网站于2025年1月上线。

成果：

一个基于Nuxt、Stripe、Supabase和Vercel的生产级网站，已稳定运行近一年，零停机时间。



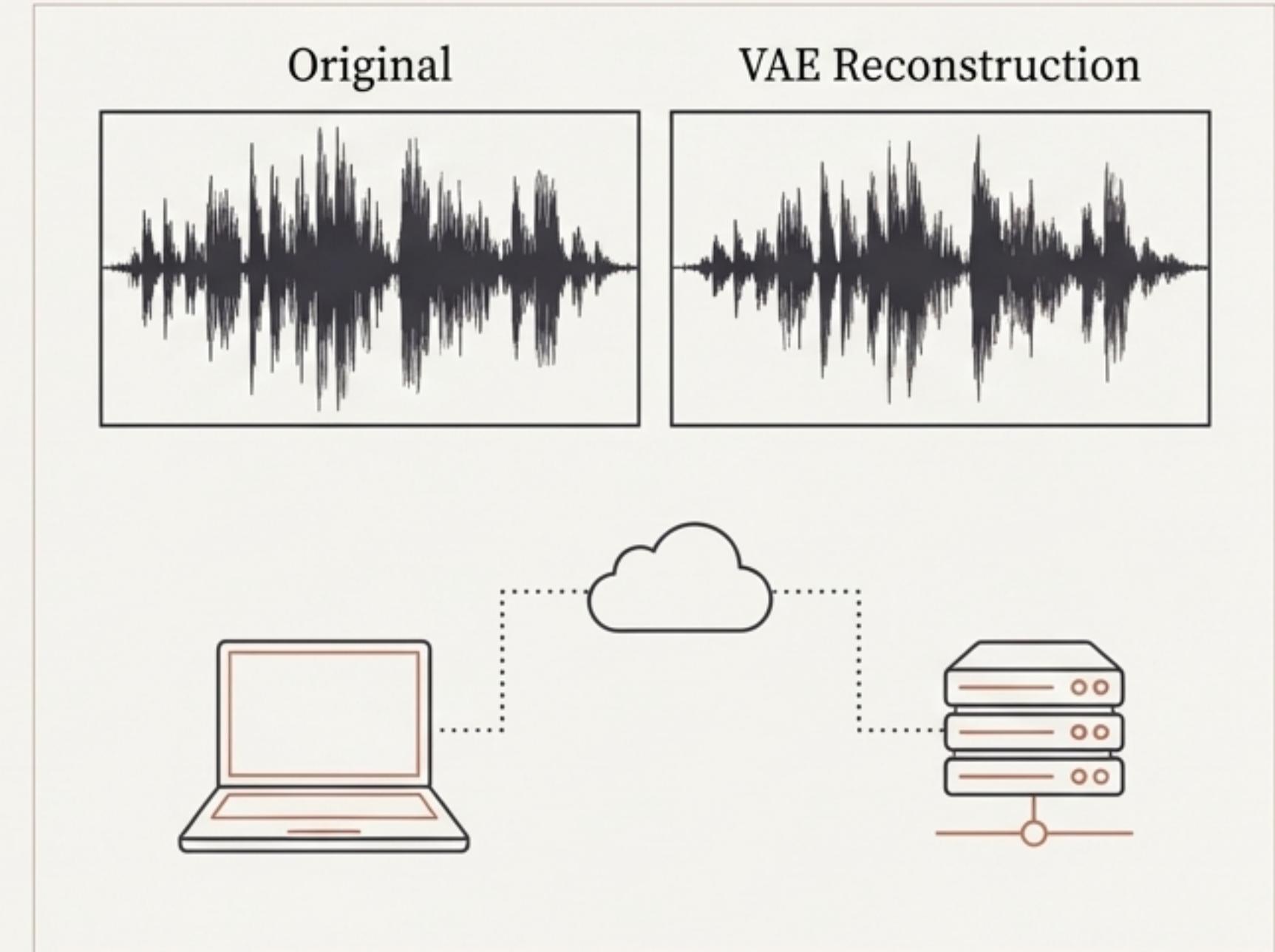
角色二：研发负责人

愿景：训练一个基于自己音乐库的生成式音频系统。

灵感来源：OpenAI的Jukebox论文。

AI协作流程：

1. **知识导师**：o1 Pro模型解释了变分自编码器(VAE)、梅尔频谱图、张量和PyTorch等概念。
2. **技术指导**：指导将60GB的音乐库转化为适用于轻量级模型的张量数据集。
3. **持续迭代**：目前已进入第三版训练流程，开始研究原始音频的条件控制和扩散模型。

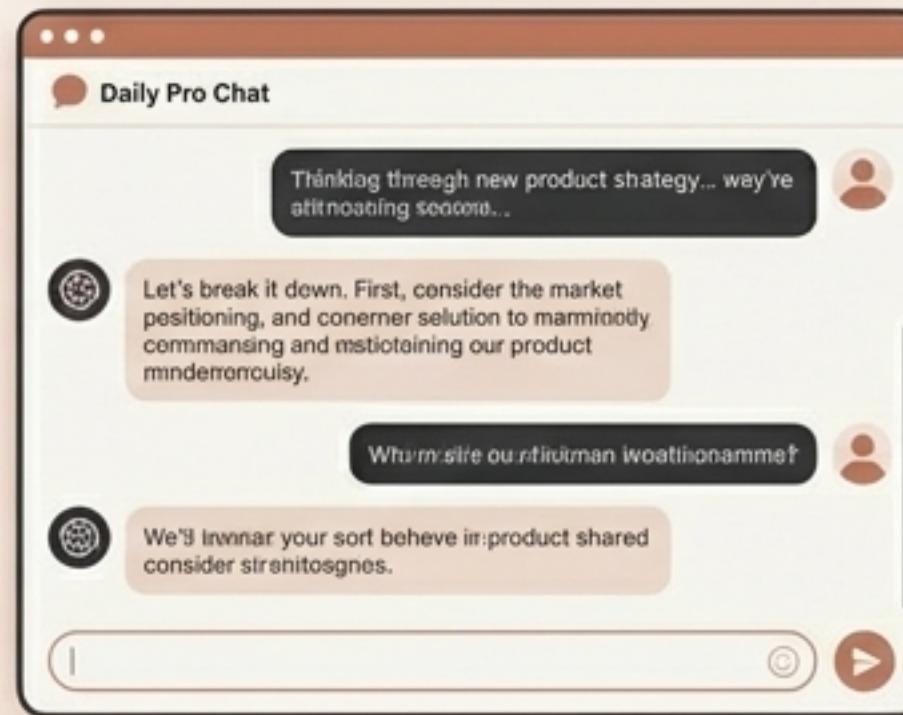


“人工智能非常擅长构建和教授人工智能，而且模型们‘让我们试试看’的态度与我自己的DIY严谨精神不谋而合。”

角色三：思考伙伴

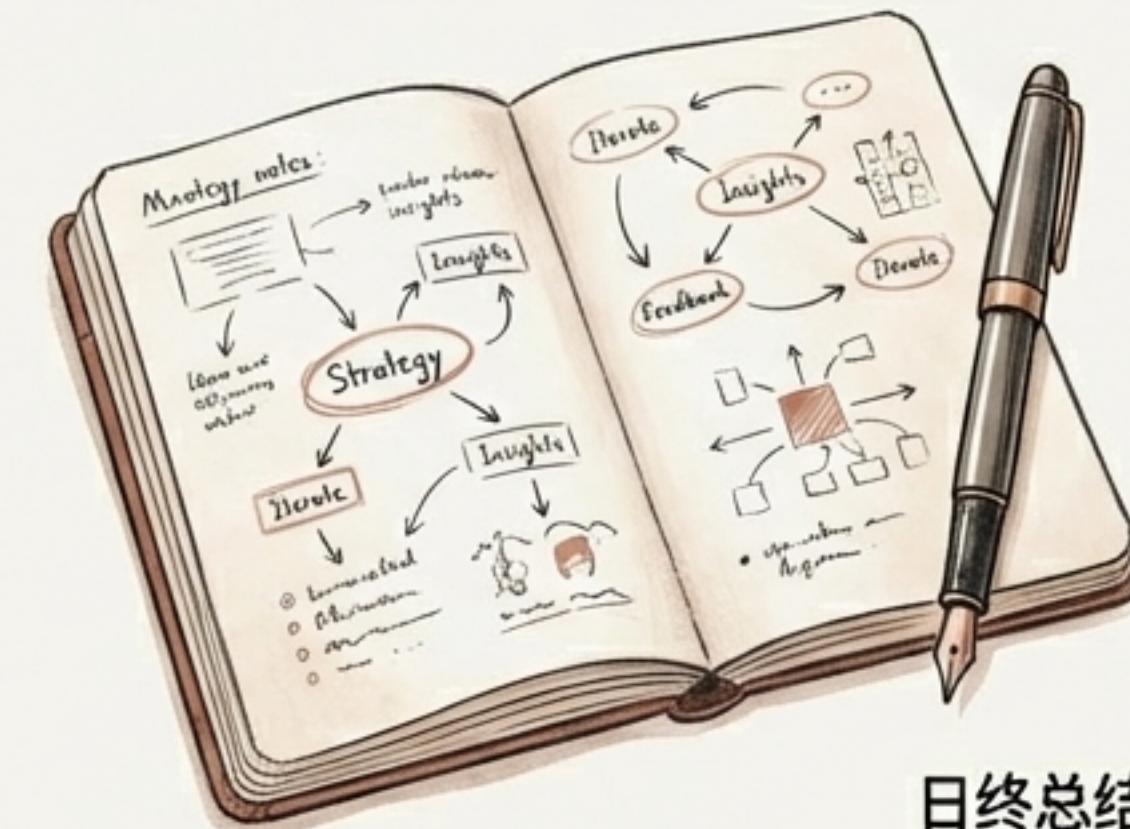
每天开启一个新的Pro聊天窗口，
作为全天的工作伙伴。

输入或口述任何想法：
业务问题、创意疑问、
实验成败、决策
决策感受。



使用原生“朗读”功
能，一边练习钢琴一
边倾听回应。

在实体笔记本
中记录灵感。



日终总结：要求模型总
结当天的聊天和笔记，
作为次日工作的开场白。

“它提供的是能够记住上下文、关心连贯性并且不厌其烦地回答下一个问题的思考存在。
.....这感觉就像一个之前混乱的系统被清理干净，让你能真正地就工作展开对话。”

高效工作流：从深度研究到语音对话

核心工具：Deep Research，一个长篇阅读和研究工具。

应用领域：机器学习论文、神经网络、音频建模、市场分析、音乐史等。



深度研究 (Deep Research)

生成比传统搜索更深入、多样化
的研究报告。

音频转换 (ElevenLabs)

将45-90分钟的文本报告转换
为音频，利用碎片时间收听。

情境对话 (Voice Mode)

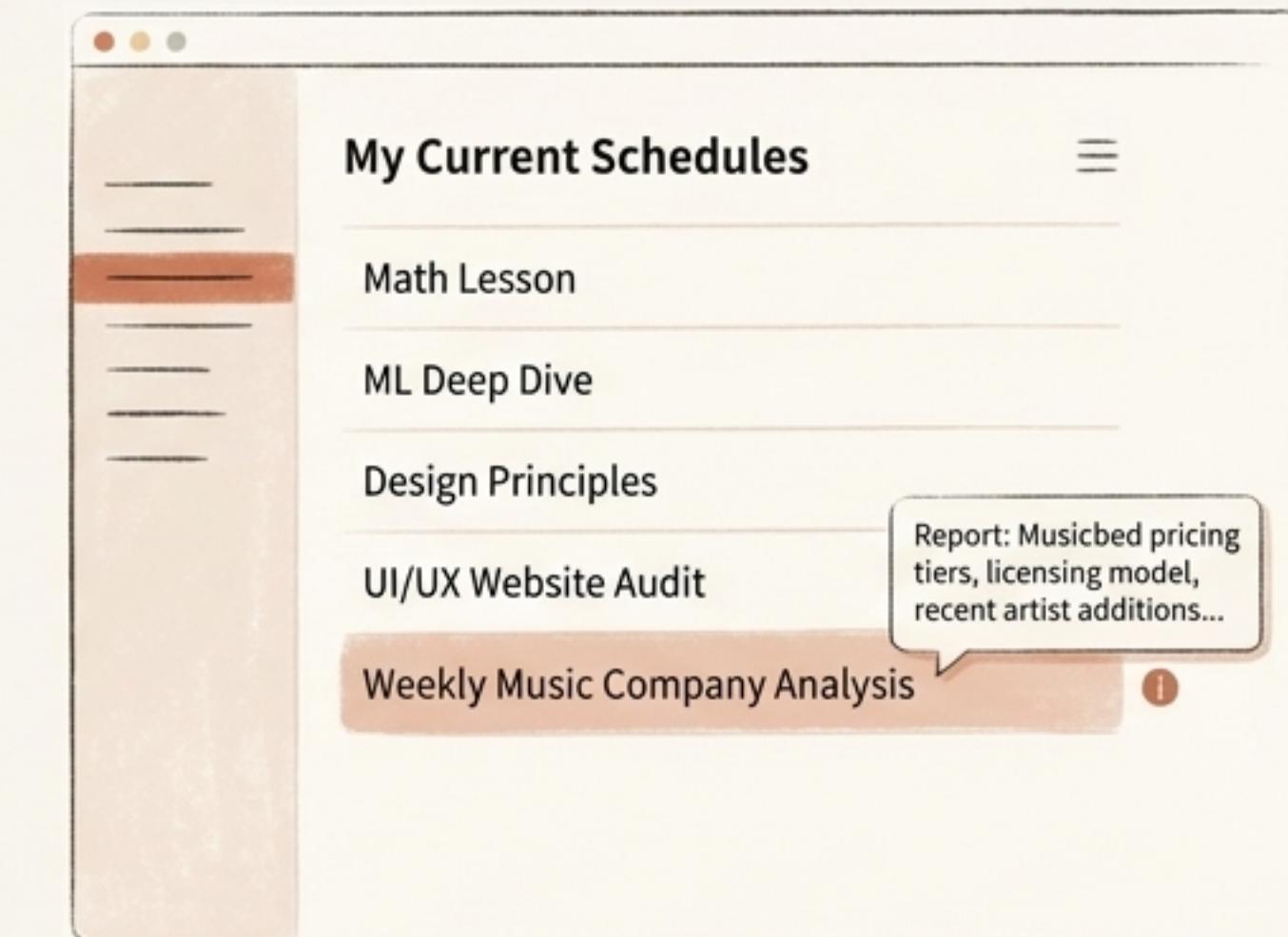
就刚刚听到的内容开启语音对
话，进行深入探讨。

成果：将被动的信息消费转变为主动的知识内化和思维碰撞。

角色四：主动的学习与组织系统

核心工具：

- **Schedules (任务)**：如同动态调整课程的教学代理。每周接收数学、机器学习、设计、市场分析等不同主题的定制化学习任务。
- **Projects (项目)**：容纳日常聊天、研究和大型实验的容器，提供静态上下文参考。
- **Canvas (画布)**：将大量文本或代码在几分钟内重组成可用形式（如提案）。



实践案例：

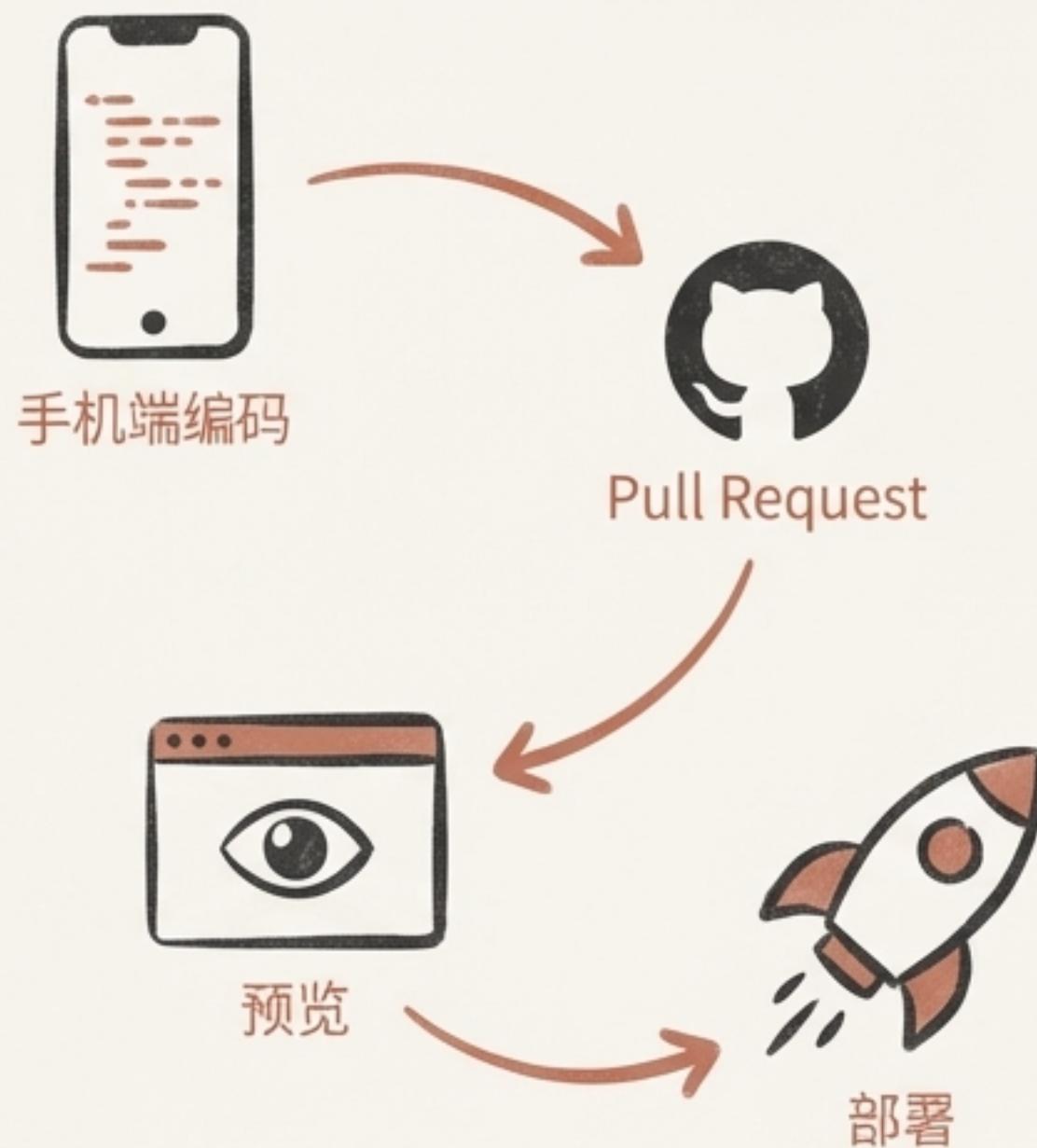
“每周音乐公司分析”任务。Schedules会自动分析Musicbed等竞争对手的网站和定价，并返回结构化的商业模式报告。

“它们不像是自动化的，因为这正是我期望公司里一位高水平的第一位员工所做的事情。”

演进：从辅助到自主代理

工具演进

- **Operator**: 早期的浏览器代理。用于执行结构化任务，如UI/UX审计，甚至在人工辅助下发布Bandcamp专辑。
- **Agent Mode**: 更强大的网页自主导航和工具调用能力。成为与Deep Research并行的首选研究工具。
- **Codex**: 每日使用的编码环境。从“制作时教学”到完整的“手机端编码 -> Pull Request -> 预览 -> 部署”流程。



关键洞察

工具的深度使用和整合，比在不同生态系统之间切换更有效率。Codex允许在已建立的ChatGPT工作流中进行深度开发。

角色五：创意助理

工具：

DALL·E 和 Sora

应用：

- **专辑封面**：创作抽象、极简风格的专辑封面。
- **博客图片**：生成文章的英雄图片。
- **Spotify Canvas**：使用Sora为单曲制作短小的无声视频。

评价：

生成的图像和视频工具取得了巨大飞跃，持续产出有价值的创意资产。Sora 2 在复杂节奏方面表现出色，带来了很多乐趣。



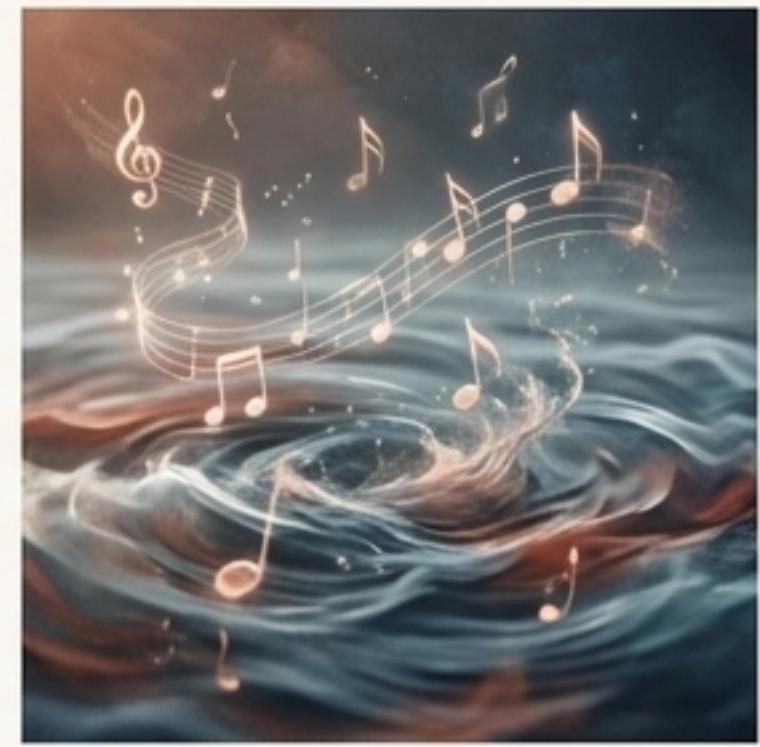
rust



blog hero

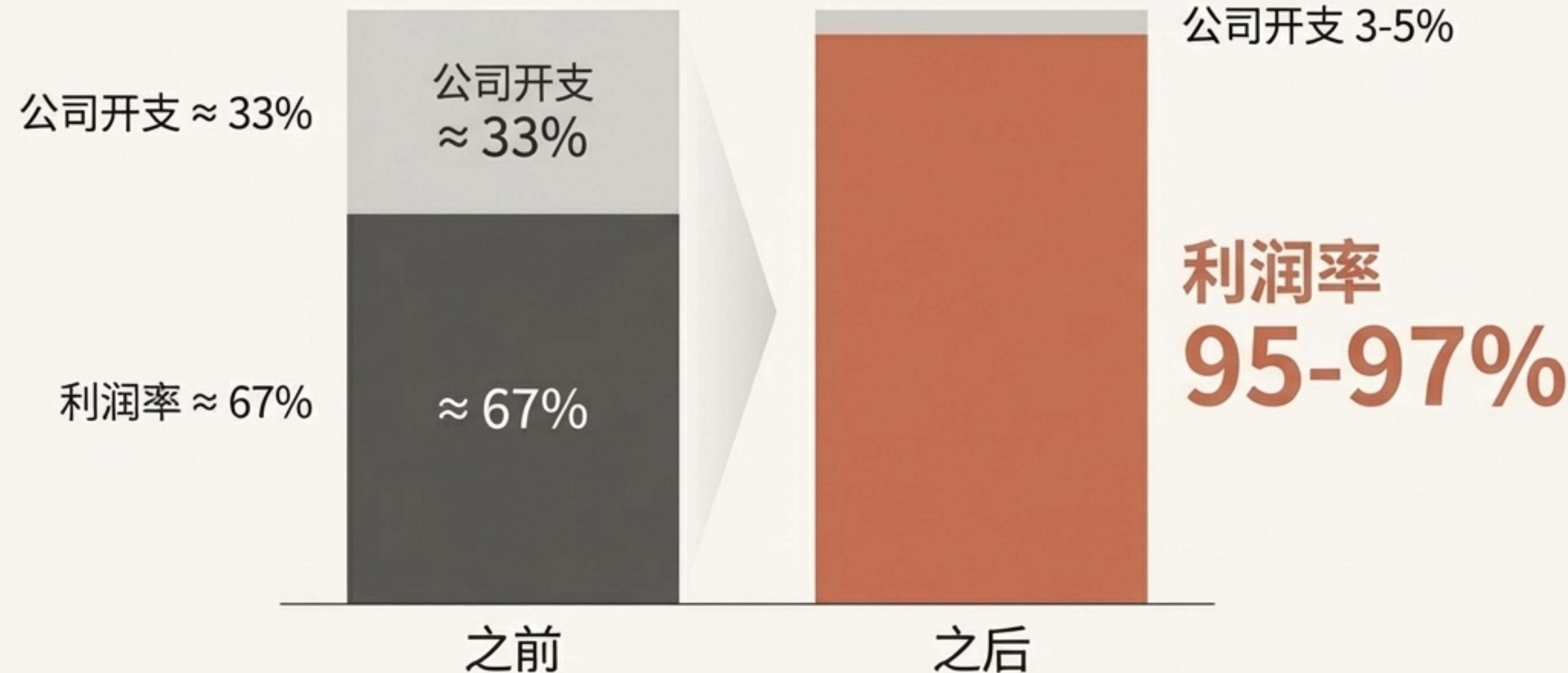


triple



Spotify Canvas

压倒性的投资回报



关键原因：摆脱了由多个SaaS产品拼凑而成的“科学怪人”系统，大幅降低了运营成本。

超越财务报表的价值

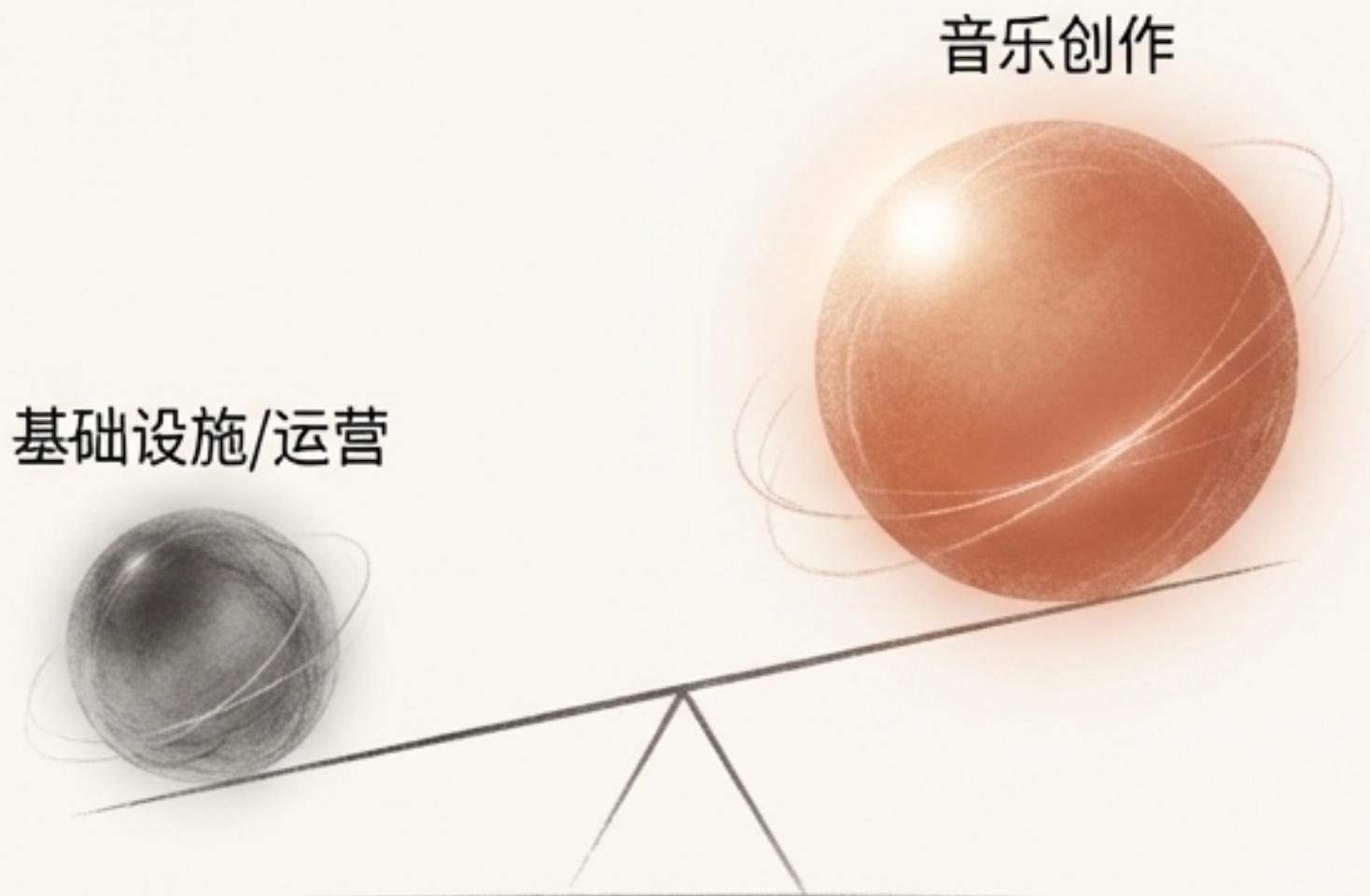
对时间的影响

- **更快的战略调整：**能够通过与模型模拟，更早地识别并放弃无效的长期策略。一个耗时数年的领悟，现在可能在几次对话中完成。
- **核心业务聚焦：**稳定和简化的基础设施意味着可以将更多精力投入到音乐创作本身。

长期目标

建立一个能大部分自主运行的公司，工作重心转向系统监控和核心创作。

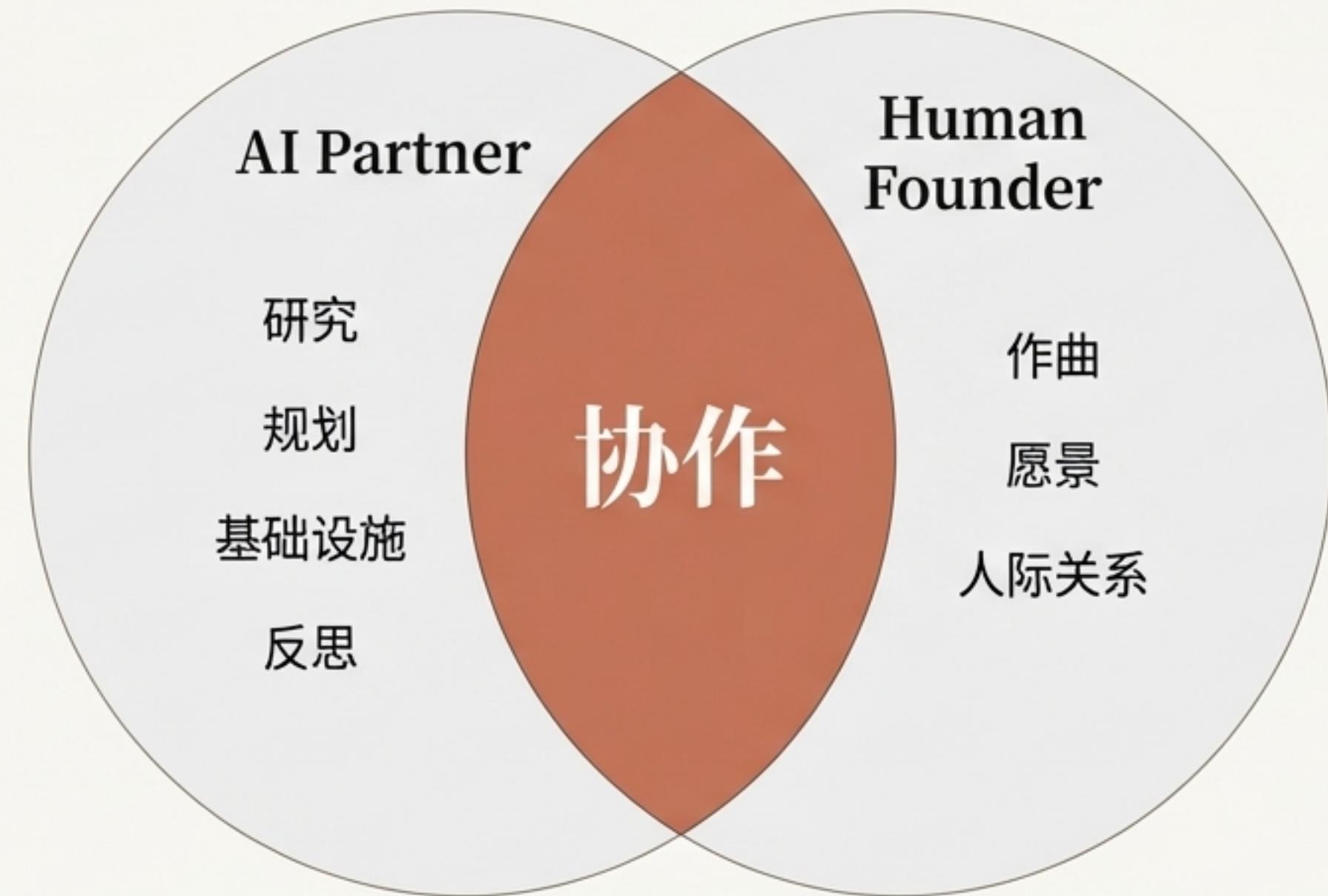
“经过这一年，这种平衡终于感觉触手可及。”



重新定义人机协作的边界

明确界限

- AI做什么：研究、规划、基础设施和反思。
- AI不做什么：它不为我创作音乐。真正的作曲过程完全由我掌控。



未来影响

- **清晰的招聘画像：**经过一年的合作，我已经非常清楚未来第一位“人类”员工需要承担的职责。
- **未来的协作模式：**当新员工加入时，我会将角色中“人性化”的部分交给他们，然后让ChatGPT Pro填补剩余的空白和新出现的任务。

合作本身就是一种乐器

决定成效的，不是模型版本或使用限制，而是你如何与它们互动。

成功的关键：

视其为同事：而非简单的搜索引擎。

提供丰富情境：上下文的质量决定了输出的质量。

提出真诚问题：并根据结果采取行动。

“高使用频率限制，就像我年轻时在琴房里长时间的练习。它们让你投入上万小时与模型协同工作，直到这种合作本身，开始感觉像一种你可以演奏的乐器。”