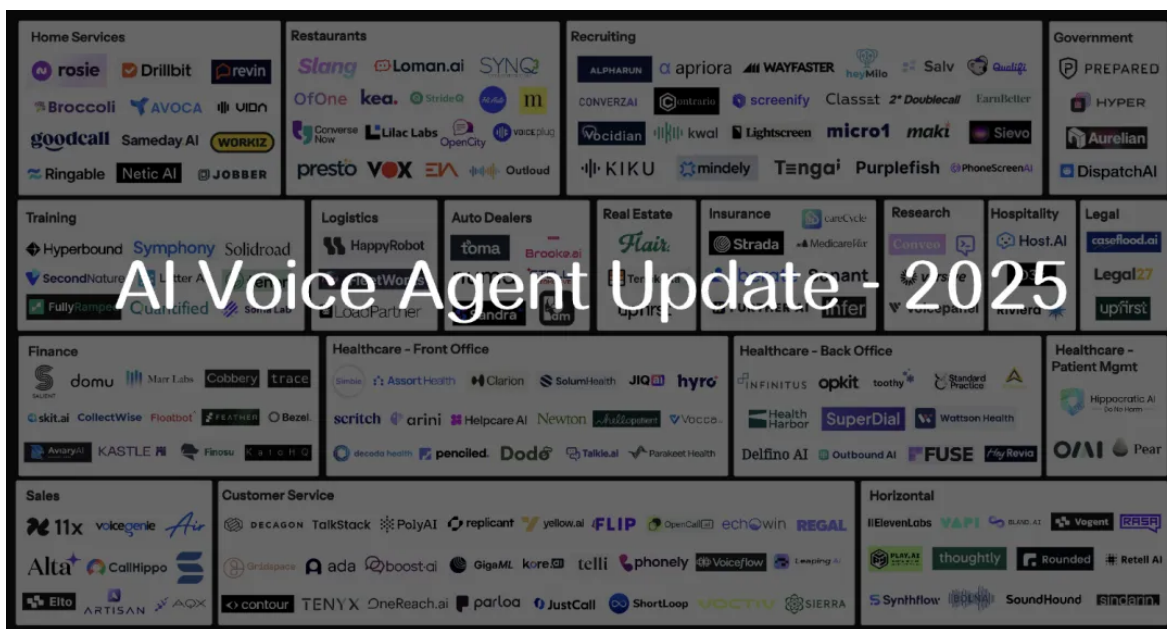


2025 最新 AI Voice Agent 研究与图谱

宇宙编辑部 特工宇宙 2025年02月04日 23:06 浙江



特别感谢特工宇宙战略顾问 @庄明浩 老师的推荐。该研究由 a16z 最近发布，原文稍有晦涩，再保证对齐原意的基础上，经 DeepSeek 生活化处理翻译，宇宙编辑部审校修改如下。



为什么要研究 Voice Agent?

语音技术是 AI 应用的关键钥匙。语音是人类最常用的沟通方式，信息量也最大。现在有了 AI，我们第一次能真正让电脑理解并处理语音。这对企业和服务普通用户都有大用处。

对企业来说，用 AI 接电话比雇人划算多了——便宜、反应快、不会出错，有些时候甚至比真人做得更好。比如餐馆订位、银行客服这些需要接电话的工作，AI 语音助手可以 24 小时在线。再也不用担心时差问题（就像下午三点从美国西海岸打给东海岸的银行，照样有人接）。

对普通人来说，以后和 AI 打交道，说话应该是最自然的方式。可能是随时能聊天的智能助手，也可能是人人都用得起的语言老师——以前请不起真人外教的人，现在用手机就能练口语。

现在的 AI 语音技术就像智能手机刚出来那会儿，基础功能都有了，真正好用的 App 才刚开始出现。随着技术升级，语音不会单独存在，而是会成为打开各种服务的“万能钥匙”。我们特别看好那些先用语音功能吸引用户，再慢慢做成综合服务平台的新公司。



2024 AI 语音发展时间线

五月

- ▶ OpenAI 放出大招：GPT-4o 语音版能实时唠嗑了
- ▶ Cartesia 公司推出真人级语音工具 Sonic

六月

- ▶ Character 软件试水语音通话（三百万用户狂打两千万通电话）
- ▶ 苹果把 ChatGPT 塞进 Siri——现在叫“苹果智能”

七月

- ▶ OpenAI 开始测试高级语音模式
- ▶ Speechmatics 家的 Flow 模型能实时对话了

八月

- ▶ 亚马逊给 Alexa 装上 Claude 大脑
- ▶ Meta 找明星录声音做 AI 朋友（想象用霉霉声音叫你起床）

九月

- ▶ NotebookLM 的语音总结功能突然爆火
- ▶ PlayHT 和 Hume 两家都升级了 2.0 版模型

十月

- ▶ OpenAI 开放实时语音 API（程序员狂喜）
- ▶ Kyutai 公司搞出能抢话聊的 Moshi 模型

十一月

- ▶ ElevenLabs 发布聊天专用 AI
- ▶ 英伟达出了能改声音的 Fugatto 工具
- ▶ 谷歌 Gemini 语音版 APP 上线

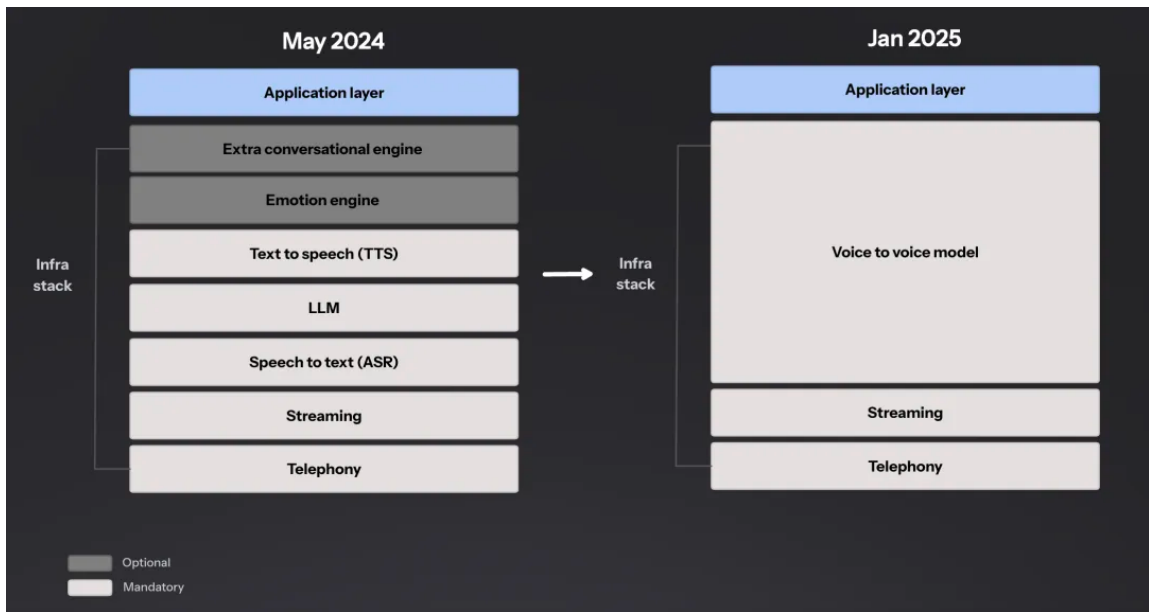
十二月

- ▶ ChatGPT 语音模式终于能上网查资料
- ▶ 1-800-CHATGPT 电话服务开通（没想到被美国人打爆）

有什么新变化？

1. 技术升级后模块越来越简单

现在做 AI 语音就像给手机换模块——底层技术越做越精简，语音助手延迟更低、表现更猛（这波质变 80% 都是最近半年搞定的）。



2. 价格越来越亲民

2024 年 12 月 OpenAI 搞价格大跳水：GPT-4o 实时 API 输入费用降 60%（现在每百万 tokens 40 刀），输出费用砍 87.5%（每百万 tokens 只要 2.5 刀）。连 GPT-4o mini 现在也能实时聊天了，这价格快跟朋友发微信差不多划算。



现在发展到哪一步了？

1. 对话质量堪比真人

延迟、抢话、语气这些技术难题基本搞定，AI 语音客服现在和真人呼叫中心水平持平甚至更好。这得感谢 OpenAI 的实时语音模型和 ElevenLabs 的聊天专用 AI 这些黑科技。

2. 市场攻略像坐火箭

用技术换人力的模式太好复制了，小公司和世界 500 强都能快速铺开（比如给快餐店装自动接单系统）。不过传统大企业可能卡在数据安全这些老问题上，这时候拼的就是谁动作快、后续功能跟得上。

3. 赚钱套路在变天

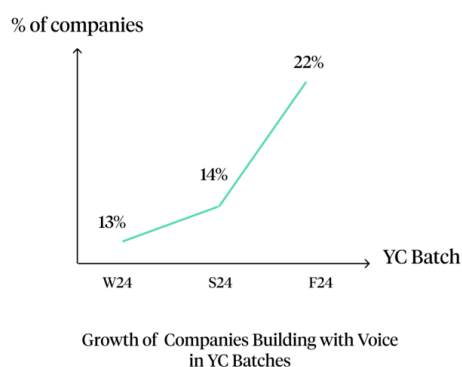
之前都是按通话分钟数收费（毕竟 AI 烧钱），现在模型成本暴跌，价格战打得飞起。以后估计要玩组合拳：收平台月费 + 按用量计费，就像充视频网站会员还得买点播券。

4. 竞争对手分三大派系

- **程序员派**：给公司 IT 部门用的开发工具（自己造轮子）
- **小白派**：不懂代码也能部署的傻瓜平台
- **垂直派**：专攻医疗、教育这些细分领域

市场进化趋势

2024 下半年语音助手市场像坐火箭——据 Cartesia 统计，现在 YC Batches 中每 5 家新创公司就有 1 家在搞语音 AI。



玩家大致分三大门派：

- **搞底层技术的**：ElevenLabs、Cartesia 这种模型大佬
- **通用平台型的**：Vapi、Bland 这种小白也能用的平台
- **深入垂直行业**：HappyRobot（物流）、Wayfaster（招聘面试）这种专攻某个领域的

大厂攻略秘籍

想直接让 AI 接管所有电话？大公司才不吃这套！聪明团队都先找突破口：

1. 捡漏非高峰来电

比如下班后的订位电话、爆线时的溢出来电。AI 至少能记下客户需求安排回拨，如果连着订位系统，说不定当场就把生意做成了。

2. 狂打骚扰电话

以前人工打推销电话亏本，现在用 AI 打就是纯赚！给个电话号码列表就能开工，适合激活老客户、开发新客户。

3. 处理后台脏活累活

诊所给药店打电话确认库存、公司给供应商催货——这些无聊到爆的电话最适合 AI 干。哪怕只是帮真人转接电话（比如按 1 转采购部），也能省下大把时间。



2024 AI 语音融资大战

▣ 模型公司（造引擎的）

- ElevenLabs - 2024.1 获 a16z 等 8000 万美元 B 轮
- Hume - 2024.3 EQT 投 5000 万美元 B 轮
- PlayAI - 2024.11 种子轮 2100 万 (Kindred Ventures)
- Cartesia - 2024.12 Index 种子轮 2700 万
- WaveForms AI - 2024.12 a16z 种子轮 4000 万

▣ 通用平台（搭舞台的）

- Kore - 2024.1 融 1.5 亿 C 轮 (FTV Capital + NVIDIA)
- Rasa - 2024.2 PayPal 参投 3000 万 C 轮
- Parloa - 2024.4 Altimeter 投 6600 万 B 轮
- PolyAI - 2024.5 获 5000 万 C 轮 (Zendesk 跟投)
- Synthflow - 2024.6 种子轮 740 万 (无代码神器)
- Thoughtly - 2024.6 Greycroft 投 300 万种子
- Bland - 2024.8 Scale 投 2200 万 A 轮
- Retell AI - 2024.8 种子轮 460 万
- 11x - 2024.9 A 轮 2400 万 / 10 月 B 轮 5000 万
- Decagon - 2024.6 A 轮 3500 万 / 10 月 B 轮 6500 万
- Sierra - 2024.10 Greenoaks 投 1.75 亿 B 轮
- Artisan - 2024.10 种子轮 1150 万
- Vapi - 2024.12 Bessemer 投 2000 万 A 轮

▣ 垂直应用（专精特长生）

- Hippocratic AI - 2024.3 A 轮 5300 万 / 2025.1 B 轮 1.41 亿

- Assort Health - 2024.3 种子轮 350 万
- Parakeet Health - 2024.10 种子轮 300 万
- HelloPatient - 2024.10 种子轮 630 万
- Hyro - 2024.12 B 轮 3500 万
- Numa - 2024.10 B 轮 3200 万 (4S 店智能销售)
- HappyRobot - 2024.12 a16z 投 1560 万 A 轮
- Prepared911 - 2024.10 B 轮 2700 万 (急救调度)
- Voicepanel - 2024.5 种子轮 240 万 (客户回访)
- Letter - 2024.6 种子轮 270 万 (销售话术教练)
- Solidroad - 2024.7 种子轮 120 万 (客服培训)
- Apriora - 2024.8 种子轮 280 万 (AI 面试官)
- Wayfaster - 2024.9 种子轮 250 万 (招聘专线)
- FullyRamped - 2024.10 种子轮 230 万 (销售培训)
- Maki - 2025.1 A 轮 2860 万 (HR 猎头)
- Tenor - 2024.12 种子轮 540 万 (领导力培训)



语音 AI 的主战场在哪里？

简单说：先找那些接电话接到手软的行业。

□ 首选战场：烧钱大户

这些行业本来就在呼叫中心/BPO 外包上砸大钱：

- 金融服务（比如讨债公司天天打电话）
- 保险（客服和后台都要打电话）
- 政府机构（想想社保热线）
- 医疗（诊所天天给保险公司扯皮）

□ 别碰这些坑

如果公司员工只是偶尔接电话（比如 IT 小哥顺带处理客户问题）：

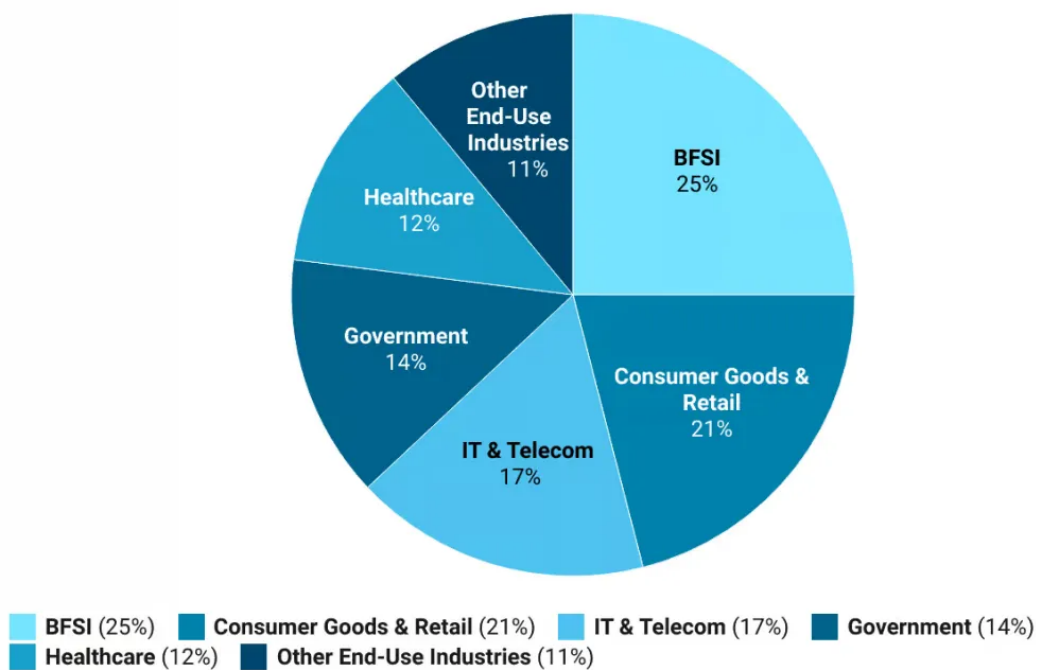
- 老板根本不舍得买 AI——除非专门养了 10 个话务员
- 省了多少钱根本算不清（财务小姐姐会撕了你的报价单）

除了接电话，现在土豪公司也愿意为“给高薪岗位当 AI 陪练”买单：

- 销售模拟器（练到能把客户聊哭）
- 急诊医生特训（用 AI 病人练手）
- 空中管制员实战演练（比请真人教练便宜，比看教学视频管用）

Global Contact Center as a Service Market Share – By End-use Industry

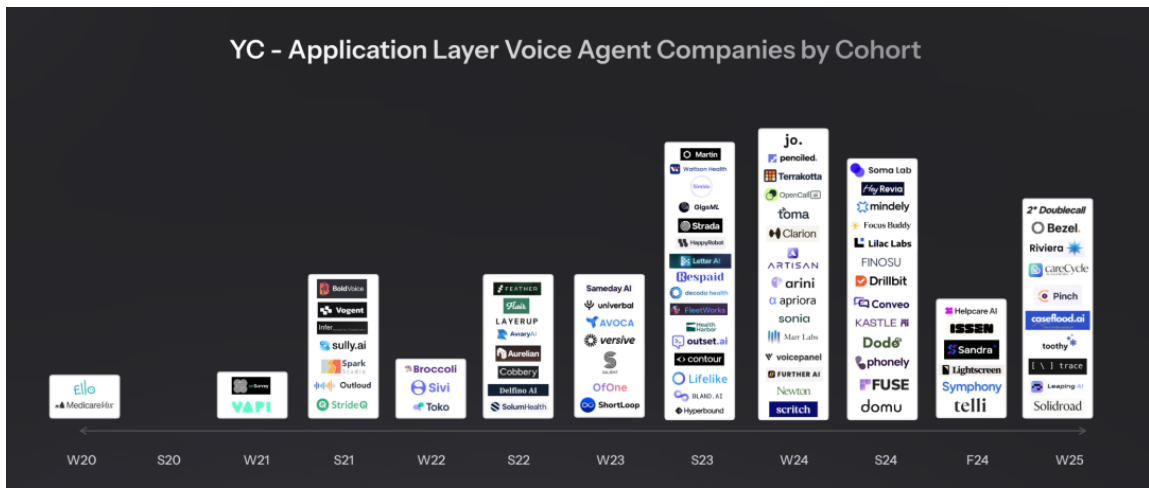
Market Share in Percentage



YC 的语音 AI 版图

用硅谷最火孵化器的数据看未来趋势。

2020 年至今，YC 孵化了 90 家语音 AI 公司。最近一期 W25（还没完全公布）就有 10 家，像滚雪球一样越滚越大。2023 年前的很多公司，其实都是中途转型杀进来的。



行业分布三大山头

□ B2B 企业服务（占 69% 大头）

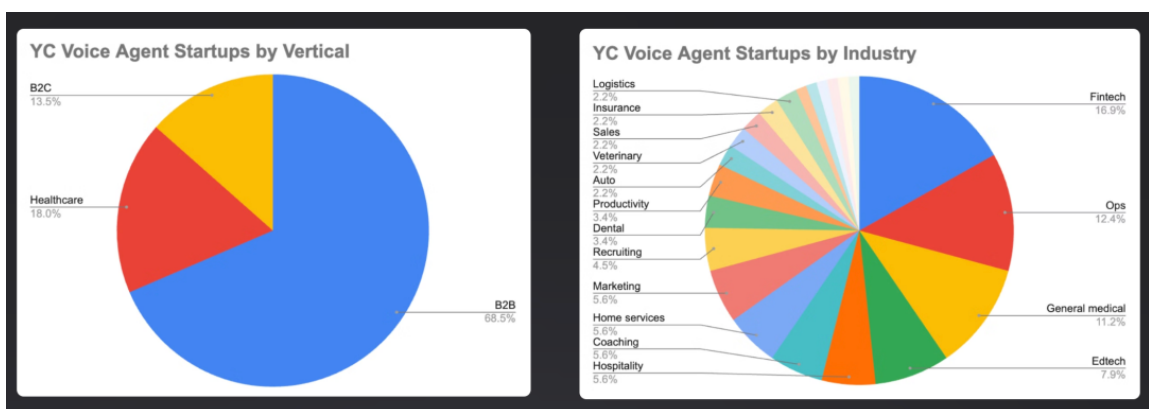
- **金融科技（16.9%）**：比如自动催债机器人
- **客服系统（12.4%）**：7x24 小时接投诉电话

□ 医疗健康（18% 闷声发大财）

- **前台服务**：智能导诊助手（11.2%）
- **后台苦力**：药房库存催货 AI（3.4%牙科诊所专用 / 2.2%宠物医院版 / 1.1%康复理疗师助手）

□ 消费者领域（13% 玩跨界）

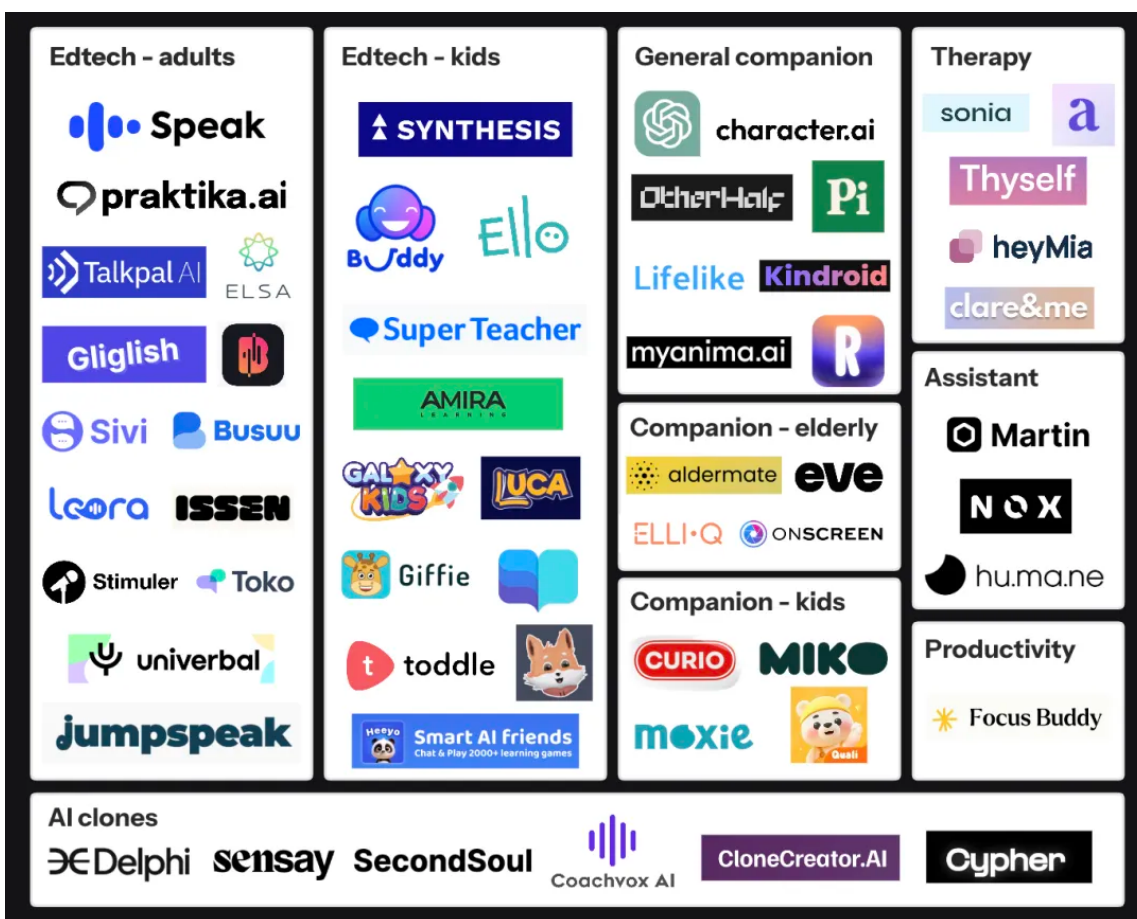
比如用霉霉声音叫你起床的闹钟 App



Voice Agents 市场图谱 - B2B



Voice Agents 市场图谱 - B2C



特工宇宙 Agent Universe



特工少女 内容运营

关注 AI+泛娱乐、AI+教育领域
欢迎交流，备注来意，加群请私信



Agent Universe, 这里有关于Agent的一切。



扫码加入特工宇宙
社群矩阵👉



Agent细糠 · 目录 ≡

〈 上一篇

23 年爆火的大模型综述，如今出纸质书了 |
文末送书

下一篇 >

MCP 重构 Agent 生态，深入探讨其现状与未
来

修改于2025年02月19日