

AI搜索群雄逐鹿 商业化仍在探索

本报记者 曲忠芳 北京报道

最近一段时间, AI(人工智能)搜索领域好不热闹。科技大厂争相卡位, 再加上AI大模型“新

需求旺、抢入口

对于大模型企业来说, AI搜索被视为一个比较容易实现商业化的突破口。

“搜索从传统线性搜索不断跃迁至AI搜索, 其本质是用户检索信息方式的不断演进。”谦询智库创始合伙人、分析师龚斌指出, AI搜索技术不仅能够生成内容, 还更加注重对语义的理解和个性化推荐。此外, AI搜索技术还能连接不同的互联网服务场景, 构建起智能化的产品生态系统。在他看来, 无论是科技大厂, 还是大模型公司, 纷纷推出各自特色的AI搜索产品, 其背后是用户对于AI搜索范式下智能总结、主动提供智能化模式的迫切需求。

记者尝试体验了由月之暗面公司于今年10月中旬上线的“Kimi探索版”。当输入关键词或一个问题进行搜索时, 主页面显示的是AI生成的总结回答, 而页面右侧是“网页搜索”栏, 展示了包括图片、AI阅读的网页来源(如新闻报道、百科知识、社交网络等)。

另一家大模型明星企业智谱AI, 涉足AI搜索领域是以智能体(AI Agent)的形态“内嵌”于智谱清言大模型内, 输入问题搜索时, “AI搜索”结果中会注明“基于X个搜索来源”, 总结生成“答案”提供给用户, 同时用户也可以点击来源网页。

作为内容社区的知乎, 今年10月底在“知乎直答”产品内上线了“专业搜索”功能, 目前平台上支持“通用搜索”和“专业搜索”两类功能, 前者适合日常学习工作中的快速查询, 后者则瞄准需要研究特定领域、分析专业文件或

秀”的下场角逐, 给自ChatGPT诞生便热闹起来的搜索市场又新添了一把“柴”。一场AI时代的搜索新“战争”正在上演。

AI搜索能给用户带来哪些新

的体验与价值? AI搜索市场的竞争格局将走向何方? 上个时代几近成熟的结果排序算法、在线广告模式, 能否“复制”到AI搜索中? 当下大模型工具已试水的付

费订阅制, 能否扛起AI搜索的商业化“大旗”? ……针对这些业界关注的问题, 《中国经营报》记者结合国内外搜索现状进行了相关调研与采访。



谷歌、百度、Perplexity、夸克、豆包等国内外搜索巨头和新锐厂商正在不断探索AI搜索的全新形态。

视觉中国/图

文献的用户群。

而老牌搜索引擎巨头百度, 较早就在传统搜索优势积累的基础上添加甚至置顶了“AI智能回答”的结果。在11月12日举行的“百度世界大会2024”上, 百度董事长兼首席执行官李彦宏透露, 截至今年11月初, 百度文心大模型的日均调用量已超过15亿次, 相比一年前实现了约30倍的增长。这一增速“证明了AI技术的真实需求和巨大潜力”。

工信部赛迪研究院今年10月底发布的《AI搜索行业发展报告》显示, 谷歌、百度、Perplexity、夸

克、豆包等国内外搜索巨头和新锐厂商正在不断探索AI搜索的全新形态, 整体来看, 行业呈现出五个明显的趋势, 分别是: 产品力成为核心竞争力, 从信息获取工具向信息一体化处理产品形态跃迁, 搜索场景趋向垂直化、细分化和专业化, PC端产品将成为新的人口级应用, AI搜索将重塑搜索行业生态。

AI搜索市场呈现“群雄逐鹿”局面, 或许更为重要的原因是, 对于大模型企业来说, AI搜索被视为一个比较容易实现商业化的突破口。对此, 科技部国家科技专

家库专家周迪向记者分析称: 一是市场需求旺盛, 智能搜索能够提供更加精准和个性化的搜索结果, 满足用户的多样化需求; 二是技术创新, 生成式AI技术的快速发展, 推动搜索引擎向智能生产工具转变, AI搜索应运而生, 为搜索引擎行业带来新的增长空间; 三是应用场景广阔, 随着5G技术的深入应用, AI智能终端的多样化和专业化趋势愈发明显, 催生了智能手机、个人电脑、智能穿戴设备、智能家居设备、车载信息系统等新型设备形态, 为AI搜索提供了更广阔的应用场景。

AI生意的理想与现实: to B 还是 to C?

本报记者 秦泉 北京报道

AI正在成为全球最“烧钱”的一门生意。根据摩根士丹利的测算, 仅亚马逊、谷歌、Meta和微软四家公司明年的资本支出合计将达到3000亿美元左右, 其中大部分将用于AI领域, 这还不包括国内的互联网巨头和一些新兴的创业独角兽。然而, AI能为这些企业带来多少实质性的回报, 却是一个悬而未决的问题。相较于此前对AI的过度憧憬, 许多企业开始意识到, 之前预期的巨大商机并未如期而至。

多位业内人士在接受《中国经营报》记者采访时表示, 当前的AI生态链表现出严重的“失衡”, 正在走向“不可持续性”。虽然目前企业都经历了“百模大战”的洗礼, 但不能保证已经跨过了“生死线”, 现在亟须迅速发掘AI大模型的实际应用场景和商业化的途径, 来回馈过去两年的大量投资和高估值。

生态链“失衡”

谷歌、微软以及亚马逊的最新财报显示, 上个季度它们在不动产和设备上的支出, 达到了506亿美元, 同比增长66%。这些投资在美国各地催生了大量的新数据中心, 专门针对AI优化, 配备先进芯片以支持AI应用的开发与运行。谷歌2025年预计在多个园区进行吉瓦级训练, 2026年形成吉瓦级AI训练集群。微软则将凤凰城最大AI训练基地计划扩展至10座建筑, 自建24个数据中心。

相较于巨额的投资, AI带来的收益却寥寥无几, 就连被誉为“四小龙”的头部AI初创公司Anthropic、Stability AI、Inflection AI、OpenAI也在苦苦挣扎。

估值达到15亿美元的Inflection AI公司目前几乎未产生收入,

路径分歧

巨大的生存压力使得AI企业们不得不走出象牙塔, 寻找“变现”的机会。360董事长周鸿祎日前表示: “硅谷已经有人在卖大模型了, 大家都在卖产品。”而在AI大模型的商业化道路上, 逐渐分化为to B、to C两个阵营。

不过, 业内普遍认同的观点是, 创业公司更易于在面向消费者的市场(C端)发现机遇, to C应用主要集中在图像生成、AI聊天机器人、写作生成、视频生成、语音和音乐生成、AI设计、AI编程等领域。今年下半年, 部分企业也开始探索AI搜索, 如百度文心一言在9月正式升级为“文小言”, 加码AI搜索, 月之暗面也推出Kimi搜索版。

虽然企业在C端产品的花样

已经终止了其初始业务线; Stability AI公司实施了裁员措施; 同时, Anthropic公司正面临每年高达18亿美元的财务赤字, 正急切寻求解决方案。

The Information的数据显示, 2024年上半年, OpenAI的现金消耗已高达3.4亿美元, 现金仅剩10亿美元, 公司预计在2023年至2028年期间, 总亏损将高达440亿美元。而其中相当一部分支出, 预计有60%—80%将用于人工智能模型的训练和运营。

不过, 与一众科技公司陷入“烧钱”而不赚钱形成对比, 英伟达的芯片供不应求, 其业绩更是爆发。2025财年第二季度(截至2024年7月28日), 英伟达营收300亿美元, 再创历史新高, 较去年同期增长122%, 营收以成倍的

速度增长。在国内AI“六小虎”之一的零一万物的创始人兼CEO李开复看来, AI目前形成的这种模式是“不健康的”。他直言: “今天我们看到真正最赚钱的公司, 是以英伟达为主的GPU公司。这种‘倒金字塔’现象是技术行业前所未有的。过去PC和移动互联网为什么能产生这么大的价值, 是因为最终最赚钱的不是芯片商, 也不是平台提供商, 而是to B和to C的应用。而如今AI时代, 除了英伟达赚得盆满钵满外, 其他AI领域的玩家, 包括应用开发者和企业, 都挣扎在微薄的利润边缘, 整个AI生态链表现出严重的‘失衡’。”

某算力服务提供商的一位高管表示: “我们需要关注的事项, 过往的智算中心建设发展迅猛, 包含

有很多, 但不难发现, 无论是哪款应用, 营收模式都较为单一, 国内目前普遍采取的是免费商业模式。实际上, 除了订阅费之外, 其他收费模式很难推广开来。国内的AI应用也曾在收费模式上作出试探, 但效果大多不尽如人意。

李开复认为, 如果没有出现贴近用户、贴近企业的应用来创造最终价值, 用户是不会甘愿付费的, 各个层面都不会赚到真金白银。

流量和数据是至关重要的, 只有拥有足够的用户基础, 企业才能确保生存并进一步规划发展。

然而, 在业内一部分人看来, AI大模型不应该干写诗画画这些事情, 真正的应用场景应该是在to B。相较于to C, 大模型在B端的盈利模式更为多样和成熟。企业可以通过API接口调用、SaaS(软件即服务)模式、私有化部署或提供行业解决方案等方式来盈利。“但实际的情况是, 绝大多数大型模型公司仅提供应用程序接口服务, 而不提供小样本微调服务及行业特定解决方案。”上述业内人士坦言。

李开复则表示, 以大模型技术为代表的AI 2.0将重塑各行各业的生产力格局, 颠覆现有的企业组

订阅制、广告及打赏尝试

在上一轮的搜索引擎大战中, 谷歌曾凭借网页排序算法、广告模式脱颖而出, “霸占”全球搜索老大的地位达20年之久。

在当下国内外的AI搜索市场中, Perplexity是在产品商业化方面进展明显较快的一家, 这家公司创立于2022年12月就已上线。今年4月, Perplexity推出了“专业版”订阅功能, 用户每月支付20美元即可选择除了Perplexity自有大模型之外的包括GPT-4、Claude-3.5、Llama3.1等在内的第三方大模型。

事实上, 据记者了解, 就订阅会员制来说, 国内多家大模型企业在产品迭代升级中也推出了免费账户之外的会员订阅服务, 比如文心一言会员包年会员600元左右, 智谱清言包年会员约400元等。也有一些大模型工具——比如通义千问, 还没有对用户收费, 只向B端用户收取API(应用程序编程接口)调用费用。当然, API服务收费模式又包括按量付费模式、包年包月模式等。值得一提的是, 月之暗面旗下的Kimi还将社交网络中的“玩法”——打赏应用其中, 打赏不同金额可获得不同天数的高峰期优先使用权。

就在11月14日, 据TechCrunch报道, Perplexity将在其平台上试验展示广告, 以“赞助的后续问题”作为呈现形式, 放置于搜索答案旁, 并注明“赞助”字样。从订阅付费到广告, Perplexity在探索AI搜索的商业化方面走在了行业前列。

“广告计划可以帮助我们创造收入, 并与合作伙伴分享。”Perplexity官方博客中写道, “经验告诉我们, 单靠订阅无法产生足够的收入来创造可持续的收入分享计划, 而广告是确保稳定且可扩展的收入现金流的最佳方式。”

众所周知, 在上一轮的搜索

引擎大战中, 谷歌曾凭借网页排序算法、广告模式脱颖而出, “霸占”全球搜索老大的地位达20年之久。时至今日, 在线广告收入已成为搜索引擎最成熟的商业模式。随着用户行为的变化, 传统基于关键词竞价的广告模式渐趋被智能化、个性化推荐所取代, 无论是老牌搜索引擎巨头, 还是社交媒体平台, 都在利用AI技术优化广告投放, 提高转化率。

Perplexity此番在AI搜索中试图“第一个吃螃蟹”——将广告模式“移植”到AI时代的搜索引擎中, 能否成功尚待市场的检验。

针对产业界对AI搜索的商业化尝试与探索, 周迪指出, AI搜索虽然能够在短时间内找到用户需要的答案, 但提供的答案有时并不准确, 用户仍需要在AI搜索给出的结果中自行筛选有用的信息。以Perplexity为代表的国外AI搜索引擎已推出订阅服务, 国内用户普遍习惯免费的搜索服务, 能否接受付费模式还存在很大的不确定性。传统搜索引擎通过广告收入维持运营和创造巨大的现金流, 但这种模式在AI搜索中是否完全适用, 还有待观察。

龚斌表示, AI搜索正成为“兵家必争之地”, 在国内外各大巨头、新锐公司纷纷加入AI搜索的浪潮中, AI搜索或将逐渐形成平台+硬件的双轮驱动格局。AI搜索的商业化探索及产品力的比拼刚刚开始, 谁能占据市场制高点、谁能笑到最后还不好判断。但可以确定的是, 传统搜索的变革将带来海量的商业机会, 向“软件+终端”等多入口、平台型AI搜索与垂直型AI搜索的全新搜索格局演进。



在AI大模型的商业化道路上, 企业们逐渐分化为to B、to C两个阵营。

视觉中国/图

很多历史因素在内。但无论是政府、投资者还是客户, 均未预料到如此巨大的变革。实际上, 我们目前所面临的现状和问题, 首要的是算力资源利用率低下, 导致极大的浪费; 其次是硬件先行于软件的现状。我们必须首先关注应用层面, 或优先考虑软件。”

此前, 金沙江创投主管合伙人

朱啸虎曾公开表示不看好AI大模型的商业模式, 其表示: “大模型商业模式太差, 技术没有差太多的情况下, 每一代技术都要投, 现在可能3.5版本要投入几千万美元, 迭代到4版本要几亿美元, 到5版本可能又要几十亿美元, 每一代模型都要重新去投入, 而变现周期可能就只有两三年。”

织结构和规模。但是目前大模型能力并未在to B应用上释放出全面潜能, 只有进入业务核心系统、快速轻量级大批量部署才能最大程度地为企业降本增效。

不过, 在中国, to B向来是个拥挤的赛道。各家的打法各不相同, 以百度为例, 目前, 超过六成的央企和大量的民营企业, 正在联合百度智能云进行AI创新。百度智能云千帆大模型平台已经帮助客户精调了3.3万个模型、开发了77万个企业应用, 文心大模型日均调用量超过15亿次。

百度集团执行副总裁、百度智能云事业群总裁沈抖表示, AI应用正率先在B端爆发。由企业级大模型工程平台、异构算力平台组成

的新型AI基础设施, 将替代传统云计算, 为大模型应用在企业生产场景中的规模落地提供关键支撑。

零一万物选择的to B路线是AI Infra解决方案、Yi大模型开放平台、“如意”数字人解决方案以及即将正式亮相的“万视”营销短视频解决方案, 共同构成从AI Infra到模型、再到应用的面向政企客户的解决方案。

零一万物联合创始人祁瑞峰认为: “今天在大模型领域内, 大部分to B项目都集中在私有化定制模型, 如何能够真正让大模型进入客户核心业务场景, 形成一个大模型赋能下标准化、可规模复制的应用产品, 是未来to B的一大挑战。”