

编者按:随着以ChatGPT为代表的生成式人工智能技术融合发展,搜索引擎正从简单的信息获取工具转变为智能生产工具。当前,国内外科技企业纷纷布局人工智能(AI)搜索市场。与国外同类产品相比,我国的AI搜索产品在技术成熟度、用户接受度、盈利能力、算力供给等方面仍存在进一步的提升空间。为推动搜索行业向“新”而行,《中国电子报》围绕AI搜索发展现状、面临的挑战和未来机遇进行深度调研,并形成报告。

人工智能引发搜索行业变局

本报记者 宋婧

不同于传统搜索主要依赖关键词匹配的模式,AI搜索拥有生成内容的能力,并且更加注重语义理解和个性化推荐。同时,AI搜索还具备连接不同互联网服务场景的能力,从而形成智能化的产品生态。业内人士普遍认为,AI技术将带动搜索行业迈向更高级别的智能化。

AI搜索赛道按下“加速键”

当前,以生成式AI为代表的新技术革命浪潮席卷各行各业,AI搜索站在这个风口上成功“破圈”,吸引到中外厂商纷纷进场。作为全球搜索引擎巨头,Google Search拥有庞大的用户群体基础和丰富的市场运营经验,通过集成谷歌自研的AI大模型Gemini,Google Search正在进行新一轮智能化升级。今年5月,谷歌正式推出AI Overviews搜索体验,用户可通过提问、聊天等方式进行AI搜索。

从技术维度来看,谷歌拥有强大的NLP技术,这是其AI搜索引擎的核心之一。通过NLP技术,AI Overviews能够更深入理解用户的搜索意图,使其提供的搜索结果更具准确性和个性化。与此同时,谷歌在人工智能、机器学习和深度学习等算法模型方面拥有深厚的积累,这也为AI Overviews摆脱原有搜索模式的限制提供了技术支撑。

从用户维度来看,谷歌拥有来自互联网的最大索引数据库,组织和维护着超过30万个网页的索引,并与众多内容创作者、网站等建立了紧密的合作关系,具备高质量的内容资源来源,用户黏性较强。庞大的用户群体和丰富的内容资源使其能够更好地理解用户需求,并不断优化搜索算法和结果。

不过,从AI Overviews上线后的表现来看,用户体验不及预期,搜索结果的准确性有待提升。根据BrightEdge统计的数据,6月份,AI Overviews的查询量仅有7%。由于新搜索引擎可能会改变搜索引擎优化策略,谷歌惯用的传统基于链接和关键词的搜索方式需要时间来适应AI生成内容的新规则。同时,传统搜索方式可通过关键词和展示网页排名嵌入广告,而直接给出答案的AI搜索难以在模型发挥能力的基础上直接给出嵌入广告的内容。因此,谷歌以广告为主体的营收模式将受到冲击,需要探寻新的商业模式。

OpenAI也是AI搜索赛道上的种子选手,其优势在于长期深耕AI领域具备技术领先性,尤其擅长上下文理解和对话式交互,如果这一能力被进一步整合到搜索引擎中,可能会在用户体验上实现较大提升,这将助其在搜索引擎市场中快速占据一席之地。OpenAI发布了AI搜索产品“SearchGPT”,由GPT-4系列模型驱动。

根据用户测评,SearchGPT主界面保留对话框形式,但搜索内容分为答案、索引和图片三类,用户可自由选择GPT3.5、GPT4和GPT-4-Lite版本,并提供了Bing、Sydney和Fortis三个搜索引擎的选择,甚至可以选择模型提供的答案更激进或更保守。其“可视化答案”功能,尚未公布细节,推测可能会跟Sora视频生成功能做结合。从整体来看,它更像是个搜索引擎应用商店。在内容来源方面,OpenAI正在寻求与一些新闻出版商合作,使用直接的内容来源构建其搜索结果,但《纽约时报》等多家顶级新闻网站已屏蔽SearchGPT网络爬虫。OpenAI还需寻找合适的方法解决AI搜索生成内容存在的版权问题。

根据目前公开的用户测评结果来看,模型幻觉问题尚未解决,SearchGPT的搜索结果准确性有待提升。今年OpenAI的人工智能训练和推理成本可能达到70亿美元,SearchGPT在初期推出时免费使用,这将进一步加剧OpenAI的成本压力。此外,OpenAI与微软之间的复杂竞合关系也使其搜索引擎产品带来了一些不确定性。

Perplexity是最早推出AI对话式搜索引擎的公司之一,旗下Perplexity AI借助大语言模型实现了“答案引擎”的自我定位,并利用RAG(检索增强生成)技术优化大语言模型的幻觉问题,通过允许用户直接提问并从筛选后的来源总结答案,提供了一种与Google Search不同的、更准确直接的回答方式,并附带来源链接以便信息追溯,受到了市场的广泛认可。公开数据显示,目前其月活跃用户数约为1500万,并与部分浏览器厂商合作,成为默认搜索引擎选项之一。

从搜索功能来看,包括探索模块、搜索模块、学术模块、视频模块和社交媒体模块等,允许用户根据需求搜索特定领域的信息,不

仅能搜索、导出价值报告,还能分类存储,便于用户回溯。Perplexity AI最大的特点是不再简单地罗列网址,而是使用AI对搜索的网址进行总结,可以通过反复提问不断精确搜索结果,并以长文索引的形式直接给出问题答案,同时提供来源参考,且不需要注册或登录。这种将用户而非广告商放在核心位置的产品理念,得到了用户的喜爱。

不过,从收入端来看,Perplexity的主要收入来源为订阅付费,应用程序开发和广告都处于早期阶段,盈利路径尚未明晰,需要在AI搜索带来的巨大成本压力和结果准确性之间寻找平衡。而谷歌、微软等资本、人才、资源雄厚的竞争对手入场,也给Perplexity AI带来了挑战。

夸克是国内AI搜索新势力的代表之一,凭借极简的设计和一站式AI服务快速吸引到大批年轻用户。公开数据显示,其累计用户量已经超过亿,其中25岁以下的年轻用户占比超过一半。

从产品设计来看,夸克界面简洁、直观、易用。区别于其他搜索引擎,无须进行页面跳转或其他操作,即可提供集智能回答、智能创作和智能总结于一体的一站式AI服务。从功能性来看,夸克将多项AI功能整合到一起,内嵌在产品本身,用户无须进行任何额外操作,即可在PC桌面、文档、网页等不同场景中,实现无缝链接的丝滑体验。

此外,由于核心用户群是年轻用户,与现阶段使用AI搜索的主力人群高度重合,且夸克用户展现出了黏性高、活跃度高等特点,场景广泛等特点。这一方面使其拥有领先同行的深厚用户基础,构成了天然的护城河和竞争壁垒;另一方面,基于用户的实际需求出发,根据用户明确的高需求场景,持续深化AI搜索能力,实现了“由场景到技术”的发展轨迹。

对于夸克而言,如何从C端市场扩展到B端市场,打开更大商业化空间,以及如何打破年龄阶层壁垒,寻求用户新增量,有待进一步探索。

字节跳动推出的智能助手产品“豆包”,是一款集搜索、语音助手、内容生成等功能于一体的工具,其产品在手机端、PC端、浏览器插件上均有布局。豆包通过集成字节跳动自研的NLP和对话式交互技术,帮助用户更高效地完成任务。桌面版的豆包进一步扩展了其应用场景,用户可以在PC端使用豆包进行问题搜索、任务管理和文件处理等操作,提升工作与学习的便捷性。

在技术层面,豆包依托字节跳动在人工智能、NLP和生成式AI领域的积累,能够理解用户需求并提供智能化的解决方案。除了基础的搜索功能外,豆包还支持多轮对话和更复杂的任务处理,从而增强了用户的整体体验。通过字节跳动的内容资源整合,豆包能够为用户提供更精准和丰富的搜索结果。

尽管桌面版豆包的推出使其进一步延伸到PC端,但仍处于发展的初期阶段,用户体验有待进一步优化。未来,豆包有望在字节跳动的智能助手产品线中发挥更加重要的作用。

产业发展呈现五大趋势

尽管AI技术与应用集中式爆发,中外厂商纷纷涌入赛道,竞争激烈,但就现阶段而言,AI搜索技术应用、商业模式、生态建设仍处在探索阶段。

从目前来看,产业发展呈现五大趋势。一是产品力成为核心竞争力。搜索的核心在于追求信息的高效获取和精确匹配。从技术实现维度来看,满足用户的基本搜索需求并非难事,如何高效地实现新用户的引流和现有用户的留存才是真正的挑战。简约化的产品设计和一站式服务能够有效满足用户更快、更精确、更易用的搜索需求,也因此成为AI搜索发展的重要趋势之一。

从产品设计维度来看,“简约化”成为AI搜索产品的主流形态。摒弃繁复和冗杂的页面设计,AI搜索产品普遍倾向于极简主义风格,搜索界面、结果呈现界面直观明了,强调简洁之美。这样的设计不仅降低了使用门槛,也将大大提升用户的搜索效率。

不过,产品设计上的简约化并不意味着搜索功能的削弱。要让产品更“易用”,就必须先要让它更“好用”。参考苹果iPhone的案例,iPhone之所以好用,不只得得益于其极简的工业风设计,还因为其拥有以流畅稳定体验著称的iOS操作系统、计算能力优秀的自

研芯片和深度整合的软硬件生态系统等,AI能力的叠加让其产品力更突出。

一站式服务成为AI搜索的一大趋势。AI搜索不仅能够独立提供服务,还能够与各种工具和平台无缝集成,形成更加强大和综合的解决方案。例如,夸克移动端、PC端提供的一站式AI服务,通过打通移动端与PC端,为用户提供多端一体的AI服务,实现从检索、创作、总结,到编辑、存储、分享的一体化信息服务价值。这样的一站式服务能够让用户在一个更加沉浸的状态中完整地完成任务,避免了在多个产品之间频繁切换所带来的烦恼。无缝衔接的体验、高度个性化的建议和服务,极大地简化了复杂的搜索流程,在提升搜索效率和用户满意度方面将发挥明显的作用。

AI搜索呈现出与其他AI工具跨界融合的趋势,可以帮助用户把页面搜索结果自动转化为视觉美观且全面的文章或报告,还可以针对目标受众转化为相对应的文风。除了基础问答功能外,搜索产品与辅助办公工具、翻译器、健康助手等多种AI工具的深度融合,可为用户带来更多实用、个性化的功能,大幅升级用户体验。

二是从信息获取工具向信息一体化处理产品形态跃迁。随着NLP、深度学习算法、检索增强生成等核心技术不断突破,AI时代下的搜索框已经突破了传统搜索的束缚,不再仅仅局限于简单的文字搜索,而是可以提供包括文本、图片、音视频等在内的跨模态搜索体验,并逐渐演化为集搜索、存储、整合、提炼、创作于一体的全能型智能助手。比如360 AI搜索尝试“拍照即可提问”。百度、天工AI等逐渐将智能体(Agent)嵌入搜索中,将让搜索形式变得更为多元化。

与此同时,AI搜索呈现出从手机端向平板、PC,甚至智能家居、智能网联汽车等多端协同的趋势。可以预见的是,未来随着多端协同能力的不断增强,AI搜索将突破不同设备之间的边界,走进更多全新的应用场景。

此外,大模型可以通过自监督学习,从大数据中自动提取知识,实现自我进化,这也将带动AI搜索实现自动化、自运行、自升级。未来如果模型幻觉、数据隐私与安全、算力供给、模型成本等问题能够得到有效解决,搜索引擎的智能化将迎来“质”的飞跃。

三是搜索场景趋向垂直化、细分化和专业化。在信息大爆炸的今天,搜索引擎对于用户的吸引力已经不在于简单地获取信息,而是在于优质内容的滋养。传统搜索的使用路径基本是受用户自身需求驱动,通过关键词搜索找到一系列相关的网站链接。用户需要逐个点开链接,花费大量时间和精力进行浏览、梳理,才有机会从庞大的信息流中获取需要和想要的搜索结果。而在AI搜索时代,用户更多是将搜索引擎作为一个可对话的工具,希望它能更“懂”自己,用尽可能少的用户侧操作来换取尽可能优质的搜索结果。这就需要搜索引擎利用AI技术对用户搜索行为进行深入的洞察与分析,以更为精准、个性化的搜索结果取代原来的海量信息流,最大程度地减少用户的使用成本。

用户使用需求和行为习惯的变化导致搜索场景趋向垂直化、细分化和专业化。相较于泛化信息搜索,垂类搜索更加专注于特定领域的搜索结果,面对复杂问题能够提供独有的、专业化的内容,不仅优化了用户体验,也在一定程度上成为了AI搜索的“护城河”。一些垂类内容平台在AI搜索赛道上具备优势。类似小红书,知乎这样偏私域的平台,拥有大量独有的专业领域内容和用户生成内容,这些内容不仅可以作为AI搜索的语料,还可以有效缩短用户获取优质内容的路径,提升搜索效率。

值得关注的是,AI大模型的出现把搜索技术重点从传统基于关键字的简单理解转移到了对用户意图的复杂理解上。以夸克高考信息服务为例,用户不仅需要查大学、查专业,更希望能够根据自身情况搜索到更加个性化的答案,比如根据考生分数、录取分数线、就业方向等寻找合适的学校。未来,垂类搜索将逐步分流传统搜索用户,有望孵化现象级产品。

四是PC端产品将成为新的人口级应用。成长于数字化时代,泛学生人群和新生代职场人是AI“尝鲜”的核心用户群,学习和办公是其最常见的搜索场景,PC端产品更符合该类人群对AI搜索的功能需求和使用习惯。

据QuestMobile数据,到2024年1月,



TOP 10的AIGC类APP用户规模达到5376万,同比增长了3725%,用户主体是25~35岁的城市年轻男性。中青校媒面向全国高校学生发起的调查发现,84.88%受访者曾使用过AI工具,77.51%受访者认为AI工具可以在一定程度上提升工作、学习效率。另据Similarweb数据,Kimi在以年轻人为核心用户群的B站进行商业投放后,访问量呈现显著增长,增幅一度达到402.9%。

从搜索场景来看,学习和办公场景是最常见的AI搜索场景,此类场景往往对搜索结果精确性、个性化程度和知识更新频率要求较高,且具备长期可持续性特征,可以带来较高的用户黏性,这也给AI搜索带来了可成长的市场空间。根据后浪研究所的调研,近八成受访者认为AI帮助自己提高了工作效率,67.1%的受访者认为AI提供了创新思路。

从搜索方式来看,年轻人的搜索方式从搜索框延展到了快捷键、截屏、划线等方式。夸克最新PC端的“系统级全场景AI”能力,将AI能力无缝融入用户整个电脑的使用体验中。无论是在桌面、文档还是网页中,用户都能通过上述方式随时唤起夸克AI能力,实现搜索、写作和信息总结等功能。数据显示,近期用户在夸克PC端使用快捷键进行搜索的涨幅最快。可以预见的是,受益于核心用户群搜索习惯的变化,具备在PC端提供全面、系统级AI能力的产品将成为新的人口级应用。

五是AI搜索将重塑搜索行业生态。AI搜索的崛起不仅仅是对搜索算法的优化,更是对用户信息检索方式的一次根本性重塑。随着越来越多的互联网厂商、内容平台、初创公司、人工智能公司等一众新玩家先后入场,传统搜索市场的“寡头效应”正在被逐步打破。根据Gartner的预测,到2026年,传统搜索引擎的访问量可能下降25%。另据Statcounter数据,截至2024年4月,传统搜索巨头谷歌的搜索引擎市场份额似乎已降至86.99%,达到2009年以来的最低点。而以AI搜索起家的初创公司Perplexity却在成立仅一年的时间里,估值就达到了5亿美元。据Similarweb数据,截至2024年2月,其用户数已经达到5000万。

不过,AI搜索与传统搜索之间并非简单的取代与被取代的关系,而是自然发展与演进的结果。预计短期内,AI搜索将分流部分传统搜索用户,双方之间的市场份额差距将逐步缩小;但从长远来看,AI搜索与传统搜索将相互补充、相互促进、融合发展、长期并存,持续为用户提供更优的搜索体验。

未来,预计在国际市场上,谷歌、微软等老牌搜索引擎厂商将快速调整产品策略,进一步加大AI搜索投入力度。以Perplexity、OpenAI为代表的AI搜索赛道实力选手将积极探索AI搜索技术趋势,寻找商业落地模式。在国内市场,以百度、360、搜狗为代表的垂直引擎厂商将与以小红书、知乎为代表的垂直内容平台以及夸克、豆包、kimi为代表AI搜索新势力合力构建起庞大的搜索新业态,共同推动搜索行业迈向更高层次的智能化。

多管齐下推动生态建设

从整体来看,国内市场的肥沃土壤推动中国AI搜索产品在产品形态创新方面走在前沿。然而,在技术成熟度、原创性及引领性创新能力、应用广度和深度等方面,我国AI搜索产品与国外同类产品相比仍存在一定差距。以多模态搜索为例,我国AI搜索引擎在模型的多样性和综合能力上,与国际先进水平相比仍有较大提升空间。

与此同时,部分用户对AI搜索出现排斥心理。目前,大多数用户已形成了固定的搜索习惯,并倾向于使用熟悉的搜索方式。而AI搜索的出现,在某种程度上改变了这种既定的搜索习惯,让用户感到不适和不安,从而引发排斥心理。

新技术、新应用往往具有更高的技术门槛,要求用户具备一定的技术素养和操作能力。实践表明,即使是能够理解自然语言的大模型,要完成更深、更复杂的处理,也需要使用者输入提示词(prompt)来进行控制。对于普通用户而言,需要经过学习才能实现提示词的恰当应用。这种复杂性成为他们接受AI搜索的障碍。

此外,AI搜索商业化路径尚不明朗。根据Next Platform对前期训练(不含微调)的估算,GPT-3 175B模型的每次训练成本在875万~1093.75万美元之间,对应花费时

间在110.5~27.6天,每100万参数的训练价格在50~62.5美元之间。而根据Cerebras AI model studio的GPT-3模型训练服务(基于4-node CS-2cluster)的报价信息,GPT70B(700亿参数,14000亿Tokens,85天训练时间)的训练价格在250万美元/次。面对如此高昂的成本,AI搜索的发展无疑需要有持续的资金支持。然而,当前AI搜索的商业模式多为付费订阅和免费两种。若坚持免费模式,则必须积极探索其他可行的商业模式。

随着AI大模型技术的飞速跃进,对高质量、高效能算力的需求也急剧攀升,而当前算力发展却存在供需不平衡现象:一方面,整体算力结构有待优化,通用算力规模占比过高,而超算算力、智算算力总体规模较小,成为稀缺资源;另一方面,算力资源碎片化严重,数据计算与信息处理、通信网络,以及应用服务等不同环节和领域长期各自独立发展,难以实现有效的互联与协同,影响对算力的高效利用。与此同时,算力所引发的能耗增长问题也日益凸显。这些因素的交织与叠加,无疑限制了AI及AI搜索技术的进一步发展。

AI搜索技术在提升数据规模和质量方面同样面临一系列挑战。AI搜索技术的发展需要大量的数据来训练其算法,这些数据不仅要有广阔的覆盖面,还要具有高质量和高相关性的特点。然而,获取这些数据并非易事,尤其是当涉及需要处理和分析大量非公开或专有数据时。

传统搜索引擎由于长期积累的用户数据,在数据规模和质量上具有明显优势,这有助于它们巩固用户黏性和提高市场竞争力。对于AI搜索新势力来说,寻找能够替代传统搜索引擎积累的“相对私有”数据的方案,以及提供更高质量、更高价值的用户数据,是构建其核心竞争力的关键。

AI搜索产品的瓶颈并不在于大模型,而在于投喂的搜索结果。如果输入的数据存在问题,如格式不正确、内容错误或重复,都可能导致AI搜索结果的准确性受到影响,因此,数据清洗也是一个难题。数据的公平性、合规性、可解释性、时效性等要求,进一步加大了数据整合与处理的难度。

在AI技术井喷式发展的背景下,如何推进AI技术应用走深向实,实现真正的应用赋能,已经成为了行业的新热点、新课题。为建立更加健全、丰富的AI搜索行业生态,促进产业发展与进步,可从五个方面发力。

一是从各层面鼓励创新,推动互联网领域的“新质生产力”发展。当前,我国AI市场发展迅速,各类AI搜索产品百花齐放,产品形态创新层出不穷。为进一步扩大AI搜索产品优势身位,应持续鼓励以创新驱动产业的发展趋势,继续加大技术研发投入力度,以技术、产品等创新拉动用户需求,形成市场增量。

二是抓住用户需求,增强战略定力,“以人为本”推进产品优化。面临用户对新技术、新产品的心理壁垒,一方面,企业应从用户体验出发,降低使用门槛,从用户真实体验出发,实现对产品的不断优化迭代,不断发展新用户;同时,面对短期的响应空白期,也应坚定战略方向,通过线上宣传、线下体验等形式,让用户切身感受AI搜索产品的作用,最大程度保留存量用户。

三是根据产品、行业不同发展阶段制定商业化策略。当前,AI搜索市场尚在拓展阶段,应以优化产品体验、寻求客户群体增量为主要目标,从用户规模效应入手,继而打造AI搜索这一新兴市场的商业模式。从产业、企业的长远发展看,也应合理规划产品的付费点,保证维持企业运转的造血能力。

四是保证关键技术自主可控,并通过优化产品能力,避免算力滥用。算力方面,在供给侧,可通过智能算力平台实现对异构算力资源的高效管理和调度,提高资源使用效率的同时,支持AI模型的快速训练和部署,促进关键技术的自主可控。在需求侧,持续优化相关技术,实现对AI搜索需求的智能化分析,将算力优先分配给高需求的搜索。

五是强化各方面保障,促进产业生态形成。应积极推动数据共享,打破信息壁垒,促进不同企业、不同生态间的交流协作,实现数据质量的整体提升;强化版权保护,有针对性地建立健全相关法律法规,明确数据使用、隐私保护、知识产权等方面的规范,确保数据应用的合法合规,保障原创者权益;开拓更加完善的商业模式,给予企业政策、资源或资金帮扶,帮助企业维持好内容生态与商业收益间的平衡,形成更加健康、完善的产业生态循环。