

中国版ChatGPT 跟风还是创新？

本报记者 齐旭

面对全球爆红的 ChatGPT，百度闪电应战。2月7日，百度官宣将在今年3月推出类 ChatGPT 项目——“文心一言”。一天后，百度 CEO 李彦宏将“引领搜索体验的代际变革”列为 2023 年第一季度 OKR 关键任务。

ChatGPT 成为火爆全球的国民级应用后，谁能率先推出中国版的 ChatGPT，成为业界关注的焦点。百度看似匆匆迎战，实则已经为此磨剑十年。错失移动互联十年之后，百度希望以 ChatGPT 为起点，找到新的增长曲线。而百度 AI 的“十年一舞”也将在“文心一言”正式推出之时交出答卷。

百度 AI 战略的“救命稻草”？

很多人将 ChatGPT 称为人工智能领域的里程碑和分水岭。

自 2022 年 11 月底正式推出后，ChatGPT 在全球范围内疯狂“圈粉”，上线仅 5 天用户数量便突破 100 万，上线两个月月活数超过 1 亿。作为 AI 自然语言生成领域的颠覆性产品，这个会聊天、懂代码、通翻译的工具除了好用之外，还为许多行业提供了巨大的想象空间和潜在的价值。

研发出 ChatGPT 的 OpenAI 已经给谷歌、百度这些利润丰厚的搜索帝国拉响了“红色警报”。

一位接近百度的业内人士告诉《中国电子报》记者，在百度三大业务支柱——搜索、云服务、自动驾驶中，搜索引擎业务是“铁饭碗”，其带来的广告收入为公司盈利带来了超过 50% 的贡献，是填补另外两大支柱业务巨额支出的重要经济来源。

近几年，搜索时代退潮、移动互联网红利见顶，百度的业务搜索和信息流增量等核心业务已经触碰到“天花板”。前有字节跳动层层阻击，后有京东、拼多多、美团等互联网“后浪”步步紧逼，百度增长受到双向挤压。尽管资本市场依旧会将百度、阿里巴巴、腾讯并称为 BAT，但百度逐渐掉队怕已是不争的事实。

ChatGPT 的横空出世，让多年未见战争的搜索市场又见硝烟，百度必须做出强有力的反击，捍卫自己的基本盘。这也是 ChatGPT 爆红之火，李彦宏迅速做出回应的原因所在。不过在外界看来，百度在以 ChatGPT 为标杆重新评估搜索业务的同时，或许也为其未来最为倚重的 AI 战略找到了“新闸口”。

“All in AI”，是百度过去 10 年最重要的战略。以 2013 年建立美国研究院为起点，百度在 AI 方向上的探索长达 10 年。10 年来，百度的研发投入累计超过 1000 亿元，其中核心研发的投入占比连续多个季度超过了 20%，这也让百度成为 BAT 中最早进行人工智能转型和投入最多的互联网厂商，使百度与谷歌、微软、亚马逊并列全球四大 AI 巨头。

日前，百度官宣，其开发的类似 ChatGPT 的项目“文心一言”（ERNIE Bot）将在 3 月份完成内测，面向公众开放，成为第一个跟进 ChatGPT 相关项目的中国厂商。据透露，“文心一言”正在做上线前的冲刺，但按照谷歌和微软对 ChatGPT 类应用投入紧追不舍的节奏，“文心一言”开放内测有可能提前。官

宣第二日，李彦宏曝光了 2023 年第一季度 OKR 关键任务——“引领搜索体验的代际变革”。有业内人士推测认为，“文心一言”将首先供给百度自家的搜索引擎，“搜索+生成”或将成为该项目应用落地的第一个爆点。

事实上，对于 ChatGPT 类生成式 AI 工具，百度“蓄谋已久”。去年 9 月，李彦宏公开表示，随着发展，人工智能在“技术层面和商业应用层面”，都有方向性改变。近期流出的一份百度内部讲话中，李彦宏再次强调：“技术已经到了临界点，类似 ChatGPT 这样的技术如何运用在搜索场景上，未来一年，在这方面我们非常有机会。”

接近百度的相关人士向记者透露，百度搜索早就接入了文心相关技术。2019 年 3 月，百度已经开发了文心 ERNIE 1.0 系统；2020 年，百度搜索开始应用文心大模型技术，逐步优化搜索的相关性、深度问答和内容理解等。

对于百度来说，推出“文心一言”的用意或首先在于“防御”，捍卫自己的搜索帝国。前几日微软将 ChatGPT 集成到自家 Edge 浏览器和必应搜索后，必应搜索应用程序和 Edge 浏览器在 App Store 排行榜上的排名迅速蹿升，应用程序的全球下载量在一夜之间猛增 10 倍。全球搜索市场占有率第一的谷歌公司也紧随其后宣布将推出 AI 对话式机器人 Bard，并将此功能整合到自家的搜索引擎中，以挽回已经被动摇的市场。

中国人民大学应用经济学院教授潘伟告诉《中国电子报》记者，像百度这类有技术、有产品的公司，最怕的是丢失用户和流量。如果行业中有同类项目“先发制人”且快速聚拢了流量，那么留给百度的窗口期会很短。这种情况下，它采取的往往是跟进策略，在用户的使用中不断完善自己的产品，再考虑能否超越的问题。

媲美 ChatGPT 的机会点

ChatGPT 是为数不多以技术为核心的互联网现象级产品，其成功要素有三点。一是 ChatGPT 有强大的技术底座 InstructGPT 模型。在此前三代的演进中，模型参数和数据规模越做越大，ChatGPT 又引入了强化学习机制，添加人工标注的数据，通过人类的反馈，有针对性地进行优化。二是 ChatGPT 在数据质量和多样性上非常讲究。OpenAI 雇佣了一个数十人的数据团队，专门针对数据多样性、标注体系进行精巧的设计，让数据发挥出了强大的作用。三是在回复速度上，ChatGPT 也领先一截。

记者此前就“百度要推出中国版 ChatGPT，会和 OpenAI 产生竞争关系吗？”向 ChatGPT 提问。ChatGPT 给出了“肯定”的回答：“如果百度真的推出了一个中国版的 ChatGPT，那么它很可能与 OpenAI 产生竞争关系。两家公司都试图通过提供优质的人工智能模型和技术服务来吸引客户和支持者，因此市场上可能会出现竞争情况，影响两家公司的市场份额和营收。”

想要超越 ChatGPT 并不容易。最近，谷歌刚刚推出 Bard 就遭遇“翻车”，因答错问题市值蒸发了 1000 亿美元，无非是因为用户对 ChatGPT 类竞品的期望值和关注度过高。谷歌如此，百度也不例外。

的历史积淀，生态建设也在起步期，谷歌云找到的差异化路径是，主打协同办公、数据分析、云上安全等 PaaS/SaaS 产品。

如今，AI 计算正在成为云大厂的新一轮竞争焦点。由于微软抢先一步打出“ChatGPT”这张王牌，谷歌虽在奋力追赶，却已然丧失先机。

伴随开源生态的逐步繁荣，技术已不再是 AI 大模型的难点所在，重点是要做大量的模型训练。ChatGPT 已经迭代至第三代，上线仅两个月，活跃用户突破一亿。用户越多，意味着数据量越大，模型训练量越大，迭代升级的速度越快。

郝峻晟指出，这类技术产品通常马太效应明显，强者愈强，留给后来者的机会不多了。他认为，目前在云基础设施方面，微软与谷歌已无太大区别。未来的竞争焦点主要是如何在元宇宙、人工智能等新兴领域提供更多排他性、独占性强的优势产品。微软更多的还是会聚焦在 B 端（如办公软件、云计算和人工智能相关产品）以及游戏等相关领域；谷歌则在互联网领域的优势更明显，未来可能会在 C 端市场推出更多定制化产品。

王娟认为，技术发展是一个长线行进、短线进发的过程，占了先机也不一定就是绝对胜出。当下的内容生态中，微软的表现的确是更优的。AI 大模型昂贵的训练成本促使微软与 OpenAI 建立了亲密合作关系，也为云服务寻找到了理想的生态路径。但在大规模智能化云服务场景中，例如自动驾驶、智慧医疗等领域，谷歌仍然有自己的优势。

AI 计算，云厂商的新竞赛

最新财报显示，亚马逊、微软、谷歌等全球头部云厂商的云计算业务均出现营收增速下滑的现象。这让他们急切希望找到新的突破口，打开增量市场。ChatGPT 的全球爆

当被问及谁能推出抗衡中国版 ChatGPT，一位业内专家告诉《中国电子报》记者，他首先想到的就是百度：“BAT 各有各的优势，比如百度的技术、阿里的运营、腾讯的产品。从这个角度来看，ChatGPT 作为一个以 AI 技术为核心的产品，很适合长于技术、弱于运营的百度。”

“文心一言”背后的文心大模型已经历三次迭代，训练数据超过千亿，这为百度比肩 ChatGPT 奠定了一定基础。虽然百度对文心大模型相关参数信息披露不多，但《Technology Review》两年前的一份报告提到，百度在通用语言和理解评估竞赛中击败了微软和谷歌。根据百度此前发表的论文，文心大模型 3.0Titan 包含 2600 亿个参数（ChatGPT 使用 1750 亿参数的 GPT3.5 模型），是当时中国最大的密集预训练模型，在 68 个 NLP 数据集上的性能已优于现阶段最先进的语言模型，包括 ChatGPT 所应用的 GPT3.5 模型。

百度打造的“文心大模型家族”，已经形成“模型层+工具与平台层+产品与社区层”的整体布局。2022 年 11 月底，百度发布 11 个大模型和基于文心大模型的两大产品，即 AI 作画产品“文心一格”和产业级搜索系统“文心百中”，如今已经在不同领域成功“试水”，比如“补全”《富春山居图》、数字人度晓晓 AI 画作卖出 17 万元等。

虽然一些人工智能资深人士认为，在类 ChatGPT 技术上，中美基本持平，但华为诺亚方舟实验室语音语义首席科学家刘群坦言，中国在技术上还是有差距的。其中一个基础模型本身的差距，虽然我们训练了很多万亿级模型或几千亿级模型，但训练的充分程度，是远远不够的。“我估计到现在为止，没有哪个模型能‘吃’ChatGPT 那么多数据。”刘群说。

在全球科技创新产业专家、海银资本创始合伙人王煜全看来，ChatGPT 类项目中国很多头部企业也都在做，但更多的是

放在后台做，老百姓无法参与。这样一来，“藏在深闺”的高科技到最后反而可能落伍。如今，基于 ChatGPT 的创业慢慢都已经成规模了，中国的 AI 头部企业已经错过了第一波。

或将成为新的增长曲线

“我们有这么酷的技术，那我们能不能把它变成人人需要的产品？”李彦宏此前表示，“这一步其实才是最难的，也是最伟大、最能够产生影响力的。”此前，谷歌 AlphaGo 这类 AI 人机交互技术在围棋游戏上确实达到了人类达不到的高度，但它的泛化能力有限，其商业前景则远不如 ChatGPT。

王煜全指出：“用户在这个时代需要的永远不是产品而是服务，需要的是规模化的服务。因为有人工智能，企业的规模化服务能力能够做到个性化，用‘我’的能力来解决‘你’的需求。”

目前，ChatGPT 尚不支持中国用户注册，即使谷歌或其他 AI 国际大厂下场，短时间内也不太可能为中国提供专属内容。此外，中文的多义性有天然的高门槛，ChatGPT 这类产品背后是整个互联网庞大的语料库，对于国内平台来说自然无法避免合规性的问题。因此，除了自然地理理解与回答用户的各种刁钻问题，中国版 ChatGPT 也需要比国外同类产品更严格的审核机制，甚至在早期可能还需要人工审核。

这就意味着，填补 ChatGPT 在国内市场形成的天然空缺，百度有很大机会。专家告诉记者，百度同时拥有“文心一言”技术和搜索市场优势，相当于中国的“OpenAI+谷歌”。

相关研究报告显示，截至目前，百度搜索引擎的用户数已突破 10 亿，是中国最大的搜索引擎。今年 1 月的百度 Create AI 开发者大会上，百度搜索亚太架构师辜斯缪介绍，百度搜索 2023 年最重要的工作是把整个检索系统变成“检索+生成”双模系统。这就意

国内外科技公司ChatGPT技术布局及应用场景

	芯片	深度学习框架	AI大模型	ChatGPT自应用场景	AIGC应用场景
百度	昆仑芯	飞桨	文心NLP大模型 文心-CV大模型 跨模态大模型	智能搜索智能云自动驾驶 智能地图汽车智能化 解决方案智能家居	AI作画、AI写作、A编剧 AI语音、AI视频创作、数字人
微软	/	CNTK	MT-NLG	智能搜索智能办公	AI歌词创作系统 定制语音技术
谷歌	TPU	TensorFlow	Switch Transformer	自动驾驶智能搜索智能地图	AI作画、AI生成视频、A编曲
阿里巴巴	含光800	EPL XDL	AI模型 M6	阿里云钉钉	AI海报设计
腾讯	紫霄	PocketFlow	混元大模型	内容创作、检索、推荐	AI写稿
360	KAMIN018	XLearning	/	智能搜索	AI框架安全监测
字节跳动	/	LightSeq	DA-Transformer	文本分析、Pico	AI视频创作、AI语音、数字人
科大讯飞	CSK400X	/	中文预训练模型	同声传译、内容审核、内容分发	智慧音效
京东	/	Optimus	领域性大模型 K-PLUG	智能城市、供应链管理 智能零售、智能客服	AI语音、AI写作、数字人
网易	/	/	/	在线教育	AI作文、AI口语老师、AI翻译

来源：本报记者根据公开材料整理

（上接第 1 版）

从 Windows 到 Office，再到 Azure 云服务，微软在企业服务领域构筑了显著优势。郝峻晟表示，ChatGPT 将为微软 ToB 软件增加更多样化、更丰富的辅助性功能，让其智能化水平得到进一步提升，同时也将帮助企业用户实现降本增效、提升数字化水平。

例如，未来 Outlook 或许可以自动帮助用户撰写邮件，用 Teams 开会时会有人工智能小秘书自动帮忙撰写会议纪要。料想这种智能程度的企业级软件将会极大地提升用户黏性，帮助微软进一步稳固企业服务市场的龙头地位。

北京计算机学会数字经济专委会秘书长王娟对记者说道：“微软旗下的 Office、Outlook 等系列办公软件开启智能化服务，不仅驱动了微软整体架构体系的智能化升级，也是整个软件行业升级的标志性事件。”

ChatGPT，微软云业务的流量密码

作为数字底座、新型基础设施，云计算是微软和谷歌一切业务的根基。ChatGPT 还有可能会对二者未来在云市场中的排名产生影响。

微软最新财报数据显示，尽管收入增速持续放缓，但微软智能云（主要包括 Azure、企业服务、GitHub 和服务器产品等）在微软整体业务中已占据举足轻重的地位，有望反超 Office 成为微软第一大业务。

“云业务依赖于用户和场景开发，ChatGPT 带来的庞大算力需求无疑将为 Azure 带来新增量。”王娟说道。

相较之下，谷歌云虽然仍维持着 30%~40% 的收入增速，但它与亚马逊 AWS、微软智能云之间仍有差距。业内人士分析称，由于既缺少 AWS 的先发优势，又缺少微软云

也有观点认为，中美国情不同、云市场发展阶段不同，中国云厂商有自己的路要走。“ChatGPT 仍是个不成熟的产品，目前炒作的意味更浓，未来能不能真正带来产业价值还是未知数。”信息通信业观察人士项立刚对《中国电子报》记者说道，“ChatGPT 的运行将带来大量数据，这对云业务的确是有推动作用的。但它实际上只是一个语义识别的智能化应用，其开发者 OpenAI 也曾明言，预计 2024 年营收才能达到 10 亿美元。就现阶段而言，它对微软云业务的拉动作用非常有限。”

他认为，经过了高速增长期，无论是亚马逊、微软、谷歌这些国际云大厂，还是像阿里云这样的国内云计算头部厂商，业务都面临瓶颈预期，增速减缓是必然的。未来云业务的突破还是需要更多地与场景结合，挖掘真正能够带来产业价值、社会价值的机会点。

狂欢之后，未来何去何从？

2月8日，百度官宣大模型新项目文心一言。多位业内人士认为，百度最有可能成为最先推出媲美 ChatGPT 产品的厂商。2月9日，阿里巴巴也爆出了正在研发阿里版 ChatGPT 的消息。据内部人士透露，目前该模型正处于内测阶段，未来可能与其智能办公软件平台“钉钉”进行深度合作。此前，腾讯人机对话专利获授权，也传出了腾讯版 ChatGPT 即将到来的消息。

然而，郝峻晟坦言：“中国厂商要想做出一款 ChatGPT 的产品还需要很长一段路要走。一方面 AI 大模型训练需要投入大量时间与成本，我们已经落后了；另一方面，中文语料数量远远小于英文语料的数量，而中文的模糊性也远比其他语言要高，这会造成算法在计算过程中的一些问题。”

意味着，对于百度，生成式 AI 和搜索引擎将相互叠加，而不是替代。

“‘检索+生成’双模系统最终提供的服务质量会远高于现在待机下的搜索服务质量。对比原来单元对话式的搜索模式，即搜索一个问题得到一个结果，未来的搜索模式可能变成多轮交互形式，但这跟对话式的多轮不一样，它不追求聊天，而是让用户可以更高效地向搜索引擎提出需求，同时让搜索迭代和调整适应用户需求。比如搜索一张图片，再用语言告诉搜索引擎怎么修改这张图片，最后产生一个真正定制化的答案。”辜斯缪说。

ChatGPT 类工具带来的庞大算力需求无疑需要云的支撑。在这方面，百度同样具有优势。AI 和云正好是百度两大业务支柱。根据 IDC 的报告，百度智能云在 AI 服务的维度排名中国第一。

百度云一直是百度 AI 技术的“收款码”。比如在制造业、能源、交通、政务这几个快速增长的垂直市场，百度智能云可以直接产生云收入。并且，智能云的利润率也在上升，在扩大运营规模的同时，百度在低利润率的定制项目中推动标准化解决方案，提高经营利润率。

据百度方面介绍，其在人工智能四层架构（底层芯片、深度学习框架、大模型以及上层应用），有全栈的布局。在基础架构层，百度自研了昆仑芯及太行 DPU 等 AI 算力芯片；在应用平台层，整合了百度 AI 飞桨深度学习框架和百度文心大模型后，能够实现从数据存储到 AI 模型训练、生产、部署、测试的全链路、批量化生产。

基于这样的架构，百度的 AI 应用开发的模式不再是手工作坊式的“一项目一训练”，而是更像流水线工厂——AI 底层技术被抽象出来，形成通用化、模块化的平台套件，百度只需要按照实际业务需求“搭积木”。

百度失去了移动互联网的十年，寄希望于 AI 赢得未来的十年。ChatGPT 是百度又一个新起点，若抓住这次机会，或让百度找到新的增长曲线。