

# 首批过审大模型向全社会开放 AI原生应用面临落地检验

本报记者 秦泉 北京报道

8月31日凌晨,百度宣布文心一言率先向全社会全面开放。用户可以在应用商店下载文心一言

## 首批获批

除文心一言外,抖音的云雀、百川大模型、清华的智谱华章、中科院的紫东太初等大模型也通过备案,可正式上线面向公众提供服务。

今年3月16日,文心一言开启邀请测试。实际上文心一言的基础模型文心大模型早在2019年就在国内率先发布,至今已经快速成长。截至目前,百度文心一言大模型已经迭代到3.5版本。与3.0版本相比,3.5版本训练速度提升了2倍,推理速度提升了17倍,模型效果累计提升超过50%。

“文心大模型3.5版本,不仅是技术上的升级,更是安全上的升级。”百度创始人、董事长兼首席执行官李彦宏强调,“在数据质量、生成效果和内容安全性上,都得到了明显提升。”而此次向全社会开放,将使文心一言更进一步。

李彦宏表示,当文心一言向数以亿计互联网用户大规模开放服务后,能够获得大量真实世界中的人工反馈,这将进一步改进基础模型,并以更快速度迭代文心一言,创造更好的用户体验。

天使投资人、资深人工智能专家郭涛对记者表示,文心一言向全社会开放可以增加用户基数和影响力,提升模型质量和性能、加速技术进步和创新,构建成熟的产业生态体系,打造品牌形象和建立口碑等,为企业未来的发展提供更广阔的空间。

中经传媒智库专家、互联网时评人张书乐对记者表示,任何一种

App或登录文心一言官网体验。据悉,除百度文心一言外,抖音的云雀、百川大模型、清华的智谱华章、中科院的紫东太初等8个大模型均通过审批可向全社会全面开

放。这其中既有科技巨头,也包括初创企业与科研院所。多位业内人士在接受《中国经营报》记者采访时表示,任何一种大模型,其核心都是深度学习。仅从模型本身而言,面向全社会开放必须进行压力测试,确保能够承载起来自全社会的数据调用和运算压力。全社会开放后,这些大模型可以通过实践来修补应用,查漏补缺。

大模型,其核心都是深度学习,即算法是学习方法,哪怕再先进都需要通过“学习资料”拾级而上,完成“小升初”“初升高”。目前,文心一言通过自行深度学习(百度投喂)、内测用户训练(测试期的学习方法调整,来自用户的学习资料积累、来自用户需求)的考试,进入到全社会开放,即海量学习资料的涌入和同样量级的考试(随着用户的满意度和各种调整,以调整学习方法和考试成绩),进而实现真正意义上运用正确学习方法来对海量学习资料的深度思考。

浙江大学国际联合商学院数字经济与金融创新研究中心联席主任、研究员盘和林认为,C端是大语言模型普及的基础,得C端者得天下,本轮生成式AI之所以火爆,就是因为C端应用上获得成功。一旦产品开放,C端就能够接触到产品应用,而对于文心一言,亦可以通过实践来修补应用,查漏补缺。

记者注意到,除文心一言外,抖音的云雀、百川大模型、清华的智谱华章、中科院的紫东太初等大模型也通过备案,可正式上线面向公众提供服务。其他大模型暂不在备案名单中。

不过,经过记者测试,目前,仅有“文心一言App”可以在App Store

和安卓应用商店下载或通过文心一言官网直接体验,其他通过备案的大模型暂时没有提供下载服务。

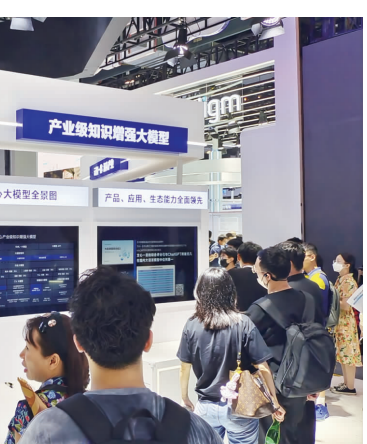
今年7月10日,《生成式人工智能服务管理暂行办法》(以下简称“《办法》”)发布,由国家网信办、国家发展和改革委员会、教育部、科技部等七部门审议通过,自2023年8月15日起施行。《办法》鼓励生成式人工智能技术在各行业、各领域的创新应用,生成积极健康、向上向善的优质内容,探索优化应用场景,构建应用生态体系。《办法》规定,提供具有舆论属性或者社会动员能力的生成式人工智能服务的,应当按照国家有关规定开展安全评估,并按照《互联网信息服务算法推荐管理规定》履行算法备案和变更、注销备案手续。

张书乐认为,大模型必须走向全社会开放,才能完成大数据意义上学习资料的累积,达成深度学习。文心一言等大模型向全社会开放,会加速其他大模型走向全社会开放的脚步,对于国内大模型在各个领域针对垂直领域进行“专研”探索,是一个催化剂。此外,在全社会开放下,强大的中文语境和针对中文语境而研发的大模型,将会和以英文语境为内核驱动的国外大模型之间,形成一个语言理解能力和创作能力上的分野,有利于国内大模型真正服务于国人,并和国外大模型之间建立护城河。

盘和林表示:“之前,很多企业因为对C端应用生成式AI有顾虑,尤其是担忧合规性,百度文心一言出来之后,其他大模型应该也会重新聚焦到C端,实际上B端是次要选择,大多数厂家嘴上说着B端,但对C端市场都心向往之。”

张书乐认为,大模型必须走向

全社会开放,才能完成大数据意义上学习资料的累积,达成深度学习。文心一言等大模型向全社会开放,会加速其他大模型走向全社会开放的脚步,对于国内大模型在各个领域针对垂直领域进行“专研”探索,是一个催化剂。此外,在全社会开放下,强大的中文语境和针对中文语境而研发的大模型,将会和以英文语境为内核驱动的国外大模型之间,形成一个语言理解能力和创作能力上的分野,有利于国内大模型真正服务于国人,并和国外大模型之间建立护城河。



目前,仅有“文心一言App”可以在App Store和安卓应用商店下载或通过文心一言官网直接体验,其他通过备案的大模型暂时没有提供下载服务。视觉中国/图

## 押注AI原生应用

AI原生应用是指以AI为核心驱动力,从头到尾完成用户任务,让用户尽可能少地付出工作来实现其目标的应用。

赛迪顾问《IT 2023》研究显示,2023年国产大模型开始爆发式增长,仅2023年1—7月,就有共计64个大模型发布。

上述研究指出,中国在大模型方面紧跟国际前沿步伐,2021年在美国引领的大模型技术快速发展的同时,中国也开启了大模型开发热潮,逐步涌现出一批具有行业影响力的大模型。2023年,受ChatGPT驱动,大模型发展迈向新阶段,国产大模型一时间呈现出爆发式增长态势。截至2023年7月,中国累计有130个大模型问世。

不过,从基础大模型到行业大模型,大模型正在落地之路上一步步快速推进,依然有“最后一公里”亟待打通,也就是AI原生应用的落地。

李彦宏曾指出:“新的国际竞争战略关键点,不是一个国家有多少个大模型,而是你的大模型上有多少原生的AI应用,这些应用在多大程度上提升了生产效率。如果我们能挤上牌桌,获得竞赛入场券,中国将拥有更壮大的数字化产业,数字经济规模将获得巨大增长。”

记者注意到,除了向全社会全面开放外,此次百度还将开放一批经过全新重构的AI原生应用,让广大用户充分体验生成式AI的理解、生成、逻辑、记忆四大核心能力。

AI原生应用是指以AI为核心驱动力,从头到尾完成用户的

任务,让用户尽可能少地付出工作来实现其目标的应用。AI原生应用能够赋能用户,使其拥有更强大的能力;同时也能够提高效率,降低成本,创造价值。

张书乐指出,AI原生应用,是让用户以一键式、“傻瓜”式方式体验各种AI垂直能力的入口,类似从Photoshop到美图那般,通过极简工具让更多人能够理解、使用AI,并真正运用到生产生活中,加速AI从极客玩具向大众工具的蜕变。

盘和林认为:“AI原生应用应该和ChatGPT的插件功能类似,我们说大语言模型,说生成式AI,其实这类AI还有一个称呼,叫通用人工智能,而通用人工智能真正的魅力并非在聊天机器人一个领域,而在于利用AI来完成大量辅助工作事宜。”

据悉,在Baidu World 2023大会上,百度发布了多款全新的AI原生应用,涵盖了文本、图像、音频、视频等多种形式和领域。这些应用都是基于百度自主研发的生成式AI技术平台——文心千帆(Wenxin Qianfan),该平台是全球首个知识增强千亿大模型平台,具有强大的自然语言理解和生成能力。

李彦宏表示:“业界有一种说法,大模型时代来了,每一个产品都值得重做一遍。但谁真正重做了一遍呢?百度要做第一个把全部产品重做一遍的公司,不是整合,不是接入,是重做、重构。”

# 大厂竞相布局 AR赛道拐点临近?

本报记者 吴清 北京报道

元宇宙热潮消退背景下,作为元宇宙基础设备的VR(虚拟现实)/AR(增强现实)发展也开始承压,但这并不妨碍国内外大厂继续加码AR赛道。

继苹果宣布于明年推出首款头显设备Vision Pro后,日前,国内AR厂商Rokid发布消费级OST(Optical See Through)个人空间计算平台Rokid AR Studio。Rokid创始人、CEO祝铭明表示:“今年和明年是行业最重要的拐点。”

如今,互联网巨头如Meta、谷歌纷纷布局AR,苹果计划明年推出AR头显,国内腾讯、字节跳动等也在抢占AR赛道,甚至罗永浩此前也宣布其新创业公司Thin Red Line选择AR方向再度创业,“我们相信AR就是下一代计算平台”。

在5G、物联网、人工智能(AI)等技术不断突破之际,AR被认为可能是继电脑、手机后消费电子领域的下一个爆点。

《中国经营报》记者注意到,目前,AR领域的三大主流场景分别为影音、办公与游戏。中国AR线上市场前三品牌为雷鸟创新、XREAL、Rokid。相比在B端推进缓慢,在人工智能浪潮加持下,中国消费级AR眼镜市场正逐步进入快速成长期,2023年因此被认为是国内消费级AR市场发展的关键之年。CINNO Research的最新数据显示,2023年第二季度国内AR消费级市场销量为5.2万台,同比增长251%,环比增长19%,AR行业加速成长。

## 竞相布局AR赛道

据悉,相较之前的AR产品,Rokid AR Studio在人机交互方面的特色在于,不依赖手柄或遥控器,只用眼睛、手势和声音,就能实现实时操控。

同时,在硬件设计、视觉体验、算力性能等方面都有相应提升。而这也是此前苹果推出的Vision Pro发力的方向。只不过国内AR厂商后发先至,抢先发布。

值得注意的是,Rokid目前与国内互联网巨头阿里巴巴、B站、腾讯、字节跳动都有合作,比如在观影

## 背后的考量

“目前,AR行业的企业有越来越多的动作和产品,整体上看是大的行业周期到了。”雷鸟创新CEO李宏伟此前表示。

对于消费级AR眼镜的再度风起,易观分析文化消费行业资深顾问廖旭华认为原因有很多,包括技术、资本和市场。因为之前有B端市场的实践,消费端目前的显示方案和产品方案逐渐成熟,各家出新品都不难。另一原因是此前元宇宙

## 机遇与挑战并存

实际上,这已不是第一次AR/VR的起势。在过去10年里,已几经起伏。

2016年前后AR/VR曾迎来一波风口,仅从2015年初到2016年7月,全球VR/AR企业数量便从200家猛增至1600余家。但风口很快消退,惨淡的销量导致AR/VR热潮无法持续,更无法吸引开发者进入,最终一部分AR/VR公司撤资转行,少部分沉淀下来的AR/VR企业则转向了B端。

2021年元宇宙概念爆发后,作为其技术支撑的VR/AR再次被推向风口浪尖。不过2022年元宇宙热潮消退后,VR/AR热潮也随之

场景上,Rokid Station国内版预装了爱奇艺、优酷、腾讯等几大主流视频平台。在办公领域,则引入阿里巴巴无影云电脑,支持轻量化办公等。此前Rokid的AR眼镜还上线了北京、天津等多家苹果专卖店。

真正引发全民对AR关注的,还是今年6月份苹果宣布将推出头显设备Vision Pro。有供应链人士透露,Vision Pro在2022年年底~2023年年初启动量产,主要放量将在2023年。

事实上,目前AR赛道的创业层出不穷,其中最具有代表性的是徐驰创

## 立AR独角兽XREAL

立的AR独角兽XREAL。徐驰曾在英伟达参与GPU计算平台的研发,而后加入AR领域先行者MagicLeap早期团队,2017年回国创立XREAL。XREAL于2019年实现商业化,但其仅用两年就成为AR眼镜领域的头部企业。据IDC发布的2022年AR/VR报告,XREAL以9.8万台出货量排名第一,在全球消费级AR眼镜市场占有率达57%。

徐驰坦言,今天苹果的Vision Pro让很多人关注到AR领域,不过AR的应用生态需要在一个更漫长的阶段中去完善。Vision Pro目前像一辆“肌肉车”,希望将武器库中所有的技术都用上,因此Vision Pro本质上是一款开发机,而非一款给普通消费者的产品。

“XREAL觉得一个更优质的AR产品,硬件的剪裁与取舍应该从用户的视角出发”,徐驰对记者表示,而衡量指标只有一个,就是来自市场销量的反馈,这是最真实的反馈。

记者梳理后发现,从2022年至今,国内已有二十余款AR新品问世。2022年,XREAL、雷鸟创新、

新增的点评,而微软HoloLens2的售价则接近5万元,国产AR眼镜在价格上已经与中端智能手机价格基本一致,正迈入推广普及的前夜。

“即便目前价格已降到2000元附近,但毕竟还是一款千元级别的产品。能否让用户真正受益并驱动购买,是推动普及的关键要素。”一位长期关注VR/AR赛道的行业分析师告诉记者。

整体而言,AR技术仍处于起步阶段,上游还处于核心零部件和技术的攻关阶段,仍存在清晰度不高、光学/场景畸变等问题。由于技术问题,面向市场的产品销量还远未达预期,内容端生态也需不断丰富,电子和光学元件有待升级,而这一切短时间内或难以改善。

同时,目前市场上AR厂商云集,竞争压力凸显,产品的同质化比较严重,AR下游还未出现杀手级应用,离规模化推广仍较远。

君盛投资从2018年开始关注AR赛道,君盛投资执行事务合伙人李昊认为,当前仍处于VR/AR行业的融资低潮,但在全球范围内,头部企业对AR技术的储备在加速,而且产业链成熟度有进入拐点的迹象。以此为节点,君盛投资先后投资了多家初创企业,包括亮风台、珑璟光电、影目科技、视讯科技等,如今均已成为各细分领域的龙头企业。

谈及AR将成为下一个计算终端的话题,李昊表示,任何一个新的计算终端其实都与产业链成熟的节奏同步。首先是硬件供应商的成熟,随后是老平台的应用厂商将应用移植到新平台上,慢慢形成用户的认知和需求,然后销售量快速起来,再促进产业链的成熟与价格的下降,从而慢慢孕育出属于新的终端交互方式的应用与内容公司。

中国正成长为全球AR/VR最重要的市场之一。IDC数据显示,2021~2026年中国AR/VR市场将以42.2%的年复合增长率高速增长,在涨幅方面超越美国和西欧等其他八个区域,位列全球首位。

而这对于中国AR产业链而言,将是风口和风险俱在,机遇和挑战并存。

度、续航等诸多问题得到了改善。“无论AR还是VR,仍然都处于一个很早期的阶段,所以需要大家有更多的耐心,等到那个爆发点的到来。”徐驰坦言,归根到底,现在行业拼的就是终端的体验是否成熟。简单来说,就是消费者能否买到足够成熟且价格合理的终端产品。

徐驰认为,生态的发展对于终端的数量有一个基本需求,终端数量不够,不可能有爆款的内容出来,也不会有厂商愿意在一个体量很小的平台上去做内容。如果用一个个具体、量化的标准,是某款AR眼镜做到了500万~1000万这样的销量。“我们有一个大胆判断,这一

次平台更迭,中国会跑在世界的最前面。”徐驰表示,中国有机会更好参与甚至引领AR行业发展,这源于中国具备的四大优势:智能硬件的供应链基础、互联网科技企业人才资源、庞大的市场规模、国内对创新企业的政策扶持等。

中国正成长为全球AR/VR最重要的市场之一。IDC数据显示,2021~2026年中国AR/VR市场将以42.2%的年复合增长率高速增长,在涨幅方面超越美国和西欧等其他八个区域,位列全球首位。

而这对于中国AR产业链而言,将是风口和风险俱在,机遇和挑战并存。

“我们有一个大胆判断,这一

度、续航等诸多问题得到了改善。

“无论AR还是VR,仍然都处于一个很早期的阶段,所以需要大家有更多的耐心,等到那个爆发点的到来。”徐驰坦言,归根到底,现在行业拼的就是终端的体验是否成熟。简单来说,就是消费者能否买到足够成熟且价格合理的终端产品。

徐驰认为,生态的发展对于终端的数量有一个基本需求,终端数量不够,不可能有爆款的内容出来,也不会有厂商愿意在一个体量很小的平台上去做内容。如果用一个个具体、量化的标准,是某款AR眼镜做到了500万~1000万这样的销量。

“我们有一个大胆判断,这一