编者按 在我国经济由高速增长转向高质量发展的历史进程中，数字化、智能化的作用不断凸显。数字化是构筑国家竞争新优势的基础，智能化则是新一轮科技革命和产业变革的重要驱动力量，二者不断融合，共同构成中国式现代化的强引擎，推动国民经济高质量发展。为顺应这一趋势，从今日起，本报推出数智版，展现人工智能和数字领域的新技术、新业态、新趋势。

把大模型装进手机

让“私人助理”可随身携带

AI世界

◎实习记者 吴叶凡

随着“百模大战”进入拼落地、拼产业应用的“下半场”，不少手机厂商纷纷入局，表示将推出自研手机大模型。2023年底，荣耀对外公布了自研的70亿参数大模型——“魔法大模型”的规模数据。市场调研机构Canalys预测，2024年，全球约5%（约6000万部）的手机会具有端侧AI运算能力。据业内人士分析，智能手机用户规模巨大、便携性较强、应用生态完善，有望成为大模型最理想的落地平台。

为何手机厂商要花费大量精力自研手机大模型？手机大模型落地还需经历哪些考验？带着这些问题，记者采访了相关专家。

图为一位女士正在使用手机查看数据。当大模型被装进手机后，人类与手机的交互方式或将发生改变。

手机厂商纷纷自研大模型

记者梳理发现，2023年8月以来，国内外的主流手机厂商纷纷表示将把大模型装进手机，华为、小米、vivo、OPPO、荣耀、三星等企业无一例外。北京市社会科学院副研究员王鹏认为，这一趋势是技术发展和市场需求共同推动的结果。“大模型在自然语言处理、图片识别等领域取得了不少成果，消费者对于手机智能化、个性化服务的需求也在不断增长。这促使手机厂商通过在手机中引入大模型来提升用户体验和自身竞争力。”王鹏说。

但是市场上已有许多比较成熟的大模型，各手机厂商为何不将这些大模型直接装进手机，反而要花费大量时间和精力去自研大模型呢？

记者了解到，目前市面上大部分大模型都是云端大模型。它们不仅参数量巨大，同时也需要极强的算力来支持运算。这导致了高昂的成本。vivo副总裁、vivo AI全球研究院院长周围曾透露，目前和大模型进行一次对话的成本为0.012元到0.015元。如果某大模型有两亿用户，这些用户一天要用大模型进行10次对话，那么一年中，光是对话的成本就会达到上百亿元，更遑论运行、维护服务器等成本。

为解决云端大模型成本高昂的弊端，手机行业开始着手研发手机大模型以降低成本。与部署在云端的云端大模型不同，大多数手机大模型从端侧发力，部署在手机等终端设备上，力图实现纯端侧运行或端云侧协同。手机大模型的优势不仅在于成本低。从数据角度看，云端大模型的数据大部分都来自互联网，这不利于大模型适应用户个性化的使用习惯；手机大模型的数据源于用户，这使得大模型能更了解用户。因此，如果说云端大模型是个“万事通”，那么手机大模型则更像是为用户量身定制的“私人助理”。

提升手机智能化水平

OPPO高级副总裁兼首席产品官刘作虎曾说：“未来，AI可能让手机呈现新面貌。”王鹏认为，当手机大模型成为用户的“私人助理”，会带来四点变化。

首先是交互方式的改变。王鹏认为，现在人们更多的还是通过按键或触屏的方式与手机交互。随着语音识别技术的不断发展，大模型能够更加精准地识别和理解用户的语音指令。这使得语音有望成为未来用户与手机的主要交互方式。此外，在手机大模型的加持下，人们搜索信息的方式有望变得更加便利。

其次，手机的个性化服务水平有望得到加强。“手机大模型可以更好地了解用户的需求和偏好，为用户提供更加个性化的服务，比如智能推荐、智能提醒等。”王鹏认为，虽然目前一些手机已具备这一功能，但精准度仍有待提升。手机大模型的应用可大大提升手机个性化服务的精准度。

再次，手机应用程序的智能化水平有望提升。比如，对照片和音频、视频的编辑会更加智能。以国内某手机为例，其具有照片“消除路人”功能。在大模型技术加持下，手机助手能够在深度学习的基础上理解用户需求，进行智能重绘，产出照片的效果相比于其他照片处理软件也更加自然。

最后，手机在未来可能会成为智能家居的核心。“现在手机主要是作为智能家居的操控器。手机大模型落地应用后，手机就可根据用户的不同需求、不同场景和不同的天气、季节，智能化地指挥智能家居。”王鹏说。

落地仍需过四关

手机大模型固然有许多优点，但要想让它真正落地，还有很长一段路要走。

首先要解决的是数据安全和隐私的问题。由于手机大模型的数据来源是手机用户在使用过程中产生的数据，因此，如何在提升大模型精准程度的基础上保障用户数据安全，是手机厂商要解决的重要问题。王鹏认为：“相关部门应该制定手机大模型训练的规则。同时，手机厂商也要在政府和行业的指导之下，建立完善的数据保护制度，在保证用户数据不被泄露的前提下对模型进行训练。”

其次要解决的是手机计算能力和存储空间的限制问题。在端侧运行大模型，对于手机的算力、能耗、内存都提出了更高要求。如果算力过高，那么能耗可能也会变大，导致手机发热；如果占用内存过多，也会影响其他应用程序的运行。目前，即便是市面上的高端手机，其硬件配置也依旧有限，难以满足千百亿级参数大模型的端侧运行需求。对此，王鹏提出了两种解决思路：“第一是提升手机算力和存储空间；第二是通过使用更高效的算法或模型压缩技术，来适应有限的手机硬件条件。”

再次，手机大模型落地还面临着“商业关”。手机大模型需要提供更加自然、便捷的交互方式，优化用户体验。这就要求手机厂商在开发过程中，充分考虑到用户的使用习惯。“即便算力、安全性都满足了，但用户觉得不好用，手机大模型也不算是成功落地。这就需要开发者根据用户的反馈来不断对大模型进行改进。”王鹏说。

最后，手机厂商花费大量精力研发手机大模型的根本目的，在于通过市场推广盈利，实现企业可持续发展。王鹏认为，只有企业做到成本可控，生产出大家都用得起的产品，才能实现手机大模型的真正落地。

领军企业推动数字技术快速发展

◎本报记者 叶青

以人工智能为代表的数字技术正开启全球新一轮科技浪潮。新机遇将会带来哪些新挑战？数字技术将呈现哪些发展趋势？围绕这些问题，来自各个领域的专家学者在近日举办的“预见未来·前沿科创论坛”上进行了交流和展望。

数字技术的进步正在给千行百业带来一场变革。由腾讯研究院联合百余位内外部科学家和技术专家编写的《2024数字科技前沿应用趋势》报告，预测了数字技术的未来发展趋势和应用前景：越来越多的数字技术将走向应用，以人工智能为代表的新一代数字技术将引领新一轮技术发展和产业重塑。

2023年初，ChatGPT横空出世，掀起了数字技术热潮。但与此同时，也有不少人认为，数字技术的发展已经进入瓶颈期，其发展速度未来将放缓。

“数字技术进步的速度绝不会放缓。”中国（深圳）综合开发研究院院长樊纲认为，无论对于科技界、经济界还是整个社会来说，新一代数字技术都是重大机遇。其研究速度不会放缓，进步速度也不会放缓，科技工作者当前要做的就是抓紧研究。

在樊纲看来，有些应用可以被限制，但科学必须要发展。在人工智能领域，我国只有走在研究前沿，在制定标准和规范方面才有可能走向前沿。

任何技术的发展都离不开大量投入。对于数字技术来说，成本高、风险大、周期长的特点在其研发过程中体现得尤为明显。

全国工商联数据显示，2022年，中国民营企业中研发投入最多的三家公司分别为华为、腾讯、阿里巴巴。这三家公司的研发投入总计约占全国总研发经费支出的近7%，且这三家公司的研发投入均排在全球前二十，在前沿科技方面也都有各自的布局。如腾讯在2018年就开始研发与大模型有关的技术，并设有专注研究量子计算、下一代机器人等技术的实验室。

“我们要发展大企业，要有大企业来领军大研究。”樊纲认为，创新具有很大风险，大企业持续投入、持续跟进创新研究的能力更强，并能不断创造。此外，大企业的研究还可以带动小企业的研究。

数字技术的发展带来了产业变革，但也有人质疑，对大模型等数字技术投入过多资源是一种浪费。

“在新一轮科技革命与产业变革背景下，前沿科技创新成为新焦点，各国政府和企业都在前瞻部署新领域、新赛道，增加对前沿科技创新的投入是必然选择。”中国科学院科技战略咨询研究院研究员万劲波说，经过适度市场竞争，只有少数优质产品、服务和企业才能存活下来。

图说智能

巡检机器人“问诊”电力设备

近日，在安徽省黄山市徽州区国网黄山供电公司500千伏徽州变电站，电力巡检机器人全面“问诊”电力设备健康情况，确保电网安全稳定运行。图为电力巡检机器人在变电站内巡检。

Monkey：实现更准确的“看图说话”

科技日报讯（记者**吴纯新** 通讯员**汪伟颋 高翔**）1月5日，记者从华中科技大学获悉，该校软件学院白翔教授领衔的VLRLab团队正式发布多模态大模型——Monkey。该模型可精确描述图片内容，并和人类就图片内容进行深入交流。

多模态大模型是一类可以同时处理和整合多种感知数据（如文本、图片、音频等）的AI架构。近年来，它在众多场景中展现出较大潜力。据介绍，Monkey在18个数据集上的实验中表现出色，在图片描述、视觉问答任务以及文本密集的问答任务中具有优势。

据介绍，目前，几乎所有多模态大模型都需要运用网上爬取的图文对数据集。这些数据集只能让大模型完成简单的图文描述任务，难以充分挖掘图片分辨率日益增加的优势。

为解决上述问题，Monkey研发团队利用现有工具构建了一种多层级的描述生成方法。通过依次对图片进行整体简述、空间定位、模块化识别、描述赋分选取和最终总结，该方法可大幅提升图片描述的准确性和丰富程度。

“一个个工具就好比不同的零件，合理排列组合才能使其发挥最大作用。”白翔说，他所在的团队从2003年就开始从事图片识别研究。他们一起反复讨论，尝试了10余种方案后才确定Monkey的最终方案。

白翔介绍，Monkey的另一亮点是能处理分辨率高达1344×896像素的图片，这是目前其他多模态大模型所能处理的最大尺寸的6倍。这意味着Monkey能对更大尺寸的图片进行更准确、丰富、细致的描述甚至推理。

据悉，目前业内能处理的图片最大分辨率为448×448像素。若想进一步提升多模态大模型的图片处理能力，需投入高昂的算力成本。该团队成员刘禹良介绍，为解决上述问题，团队采用创新性的“裁剪”方法。他们将原始输入图片分割成多个图片块，每个图片块的尺寸小于448×448像素。他们还为每个图片块配备了一个“放大镜”，将“放大镜”放到图片块合适的位置即可“看”清更多细节。多个“放大镜”同时工作，分别“放大”不同的图片块，就能提取更多图片局部特征。

40亿数据灌注国内首个古籍处理与研究开源智能工具

“荀子”大语言模型：化繁为简 通读古今

◎本报记者 金凤

“秦淮佳丽地，城阙望中迷。柳暗青丝发，花香碧玉衣。歌楼留夜色，画阁敛春晖。细雨轻舟去，双鱼梦泽飞。”这是近日上线的“荀子”古籍大语言模型（以下简称“荀子”）以“金陵”为题，生成的一首古诗。

记者了解到，“荀子”是国内首个专门应用于古籍处理与研究的开源智能工具，由南京农业大学王东波教授研究团队联合古联（北京）数字传媒科技有限公司发布。它依托国家社科基金重大项目“中国古代典籍跨语言知识库构建及应用研究”，基于40亿字的大型混合语料数据生成。

“数据是大模型的基础。”王东波介绍，在“荀子”的研发过程中，研究团队在人工智能通用模型的基础上，灌注了繁简体《四库全书》等20亿字的古代汉语语料和文化领域的20亿字的现代汉语语料，使“荀子”具有古籍智能标引、古籍信息抽取、诗歌生成、古籍高质量翻译、阅读理解等功能。

“对于汉语言研究者来说，他们还可以利用‘荀子’完成古籍词法分析、实体识别、关系抽取、文本分类与匹配、文本摘要等工作。”王东波举例，如果要研究《史记·陈涉世家》的人物关系，就可以用“荀子”识别这篇文章中的人物名称和关系名词，再用知识图谱的方式呈现人物关系图谱，从而提高检索、查询、研究的效率。

王东波介绍，此次发布的“荀子”大模型中的基座模型，还可以让用户根据自己的需求对“荀子”进行微调，帮助用户开展更有针对性的研究。

“荀子”是怎么做到化繁为简、通读古今的？“核心是‘算力充足’并且‘饱读诗书’。”王东波介绍，“荀子”的顺利问世离不开南京农业大学提供的高性能算力基础设施支持，以及研究团队长期积累的精加工语料库。

“模型的构建受算力、场景应用等多方影响，但精准度较高的优质数据是最为关键的。”王东波表示，研究团队自2013年起，一直专注于人工精标注数据的工作。

“比如要训练大模型自动标注《岳阳楼记》中的形容词，首先需要人工标注这篇文章中的形容词。在积累了大量的人工标注后，再让机器进行学习。”王东波说，这项“坐冷板凳”的基础标注工作，他们一做就是10年。

“我们期待能将古籍的智能化研究与跨学科的人才培养相结合，让学生既有前瞻的科研视野，又能积累较为深厚的人文底蕴。”王东波表示，研究团队希望能让更多人接触古籍、品读古籍、传播古籍，让“故纸堆”重新焕发活力，推动中华优秀传统文化创造性转化、创新性发展，赓续中华文脉。

王东波介绍，“荀子”除了能让人们更顺畅地阅读古籍内容，推动古籍整理、古籍数字化、古籍活化利用与传播之外，未来还可应用于人工智能写作、人工智能教学、数字文娱等领域。

电子狗提升巡防智能化程度

2024年伊始，江苏省连云港市公安局巡特警引进了具有人脸在线比对功能的电子狗。这种电子狗将参与巡逻防控，提升巡防管制水平和智能化程度。图为民警操纵电子狗和一位小朋友握手互动。

本版图片由视觉中国提供