

# 小词元如何撬动AI大产业

## 新华视点

近期,随着人工智能(AI)技术的快速迭代与产业化加速,词元(Token)作为大模型处理信息的基本单位,成为公众日益熟知、高度关注的热词。我们每使用一次AI大模型,都要消耗一定数量的词元,一次复杂的智能任务可能要消耗数万个词元。小小的词元,如何撬动AI大产业?当前还存在哪些挑战、短板?“新华视点”记者进行了调查。

### 调用量两年间增长超千倍

“今天北京的天气如何?”“能帮我生成一个AI短视频吗?”“我要去演讲,帮我做一份PPT”……如今,在工作生活中遇到问题,找大模型助手“AI一下”越来越常见。

这些AI应用的背后,都离不开词元。深圳行胜数字技术有限公司CEO郝放表示,词元是信息时代和AI时代的水电,信息时代的度量单位是字节,AI时代的度量单位是词元。

深圳云天励飞技术股份有限公司董事长兼CEO陈宁表示,词元是数字化的价值载体,可计量、定价、交易,是智能时代的价值锚点和连接技术供给与商业需求的结算单位。一个词元约等于0.75个英文单词,背后对应着真实的算力、电力和基础设施成本。

词元正串联起新的经济链条。如上游GPU(图形处理器)芯片、AI服务器等为词元提供算力支持,中游云服务商、大模型厂商负责词元的工业化生产、分销与商业化运营,下游AI智能体、企业级应用等是词元消耗的最终场景。

在广东韶关,一个个数据中心日夜不停处理着大湾区的算力请求,有了这些算力,词元才能源源不断产生。而在下游的深圳华强北,玩具、眼镜、乐器等普通硬件产品,加载AI智能体、有了“新大脑”后,便成为词元的消耗载体。

来自国家数据局的数据表示,我国日均词元调用量从2024年初的1000亿跃升至2026年3月的140万亿以上,两年间增长超千倍。第三方机构预测,到

2030年中国AI推理词元消耗量将增长超300倍。

中国(深圳)综合开发研究院数字经济研究中心主任李恩汉说,工厂、港口、跨境电商、金融、物流等高频场景有海量的词元调用需求。当把AI嵌入生产、运营和服务流程,把问答工具升级为真正能“跑业务”的智能体,词元就能变成实实在在的产值和税收。

今年初,以“龙虾”为代表的AI智能体爆发,推动词元调用量激增。不少海外用户调用中国大模型API(应用程序编程接口),在中国完成推理运算,结果回传境外,“电力未出境,但价值跨境交付”。部分中国大模型收入爆发式增长,20天的累计收入已超过2025年全年总收入。

### 仍难标准计量 治理还有难点

“词元是新的大宗商品。”在英伟达2026年度开发者大会上,英伟达创始人兼首席执行官黄仁勋说。

不少业界人士表示,虽然风口已至,发展空间广阔,但当前有关词元的计量、交易与管理,还有不少问题待解。

——仍难标准计量,“热多冷、闭环少”。腾讯研究院相关负责人表示,不同的词元价值不一样。百万个词元,如果是用来闲聊的,可能仅值0.01美元;如果是用来写代码的,可能值200美元;如果是用来做法律文档审阅的,可能值1000美元。这也导致一个可能的情况:不到5%的词元消耗,创造了超过80%的可测量价值。

在词元生产端,也存在类似情况。同时,在编程、视频生成、聊天等之外,AI应

用尚未大范围普及,不少企业的收入增长建立在资本开支扩张和短期订单上,还未形成稳定、可复购的商业模式。

——规则尚不完善,不利于长期发展。陈宁认为,当前我国大模型正在服务全球用户,词元可直接跨境交付,不需要通过集装箱和海关;但现有的贸易统计框架、分类还无法准确描述这种新型贸易。同时,大量AI服务创造的价值,游离于国内生产总值(GDP)统计之外,也可能带来一定信息遗漏和盲区。

——安全隐患凸显,谨防新型骗局。国家安全部微信公众号安全提醒,词元在使用中存在一定的安全风险,包括泄露劫持风险、伪造篡改风险、诈骗陷阱风险等。不法分子可通过跨站脚本攻击(XSS)、公共Wi-Fi嗅探等方式,窃取、

截获未加密的令牌。同时,不法分子还有可能制造“虚假令牌”,诱导用户泄露身份证号、手机号等隐私信息。

“如何进一步提升词元流通的安全,这是当前亟待解决的一个问题。”郝放说。在一些社交平台 and 私密社群里,词元被包装成新一代“暴富密码”,“现在囤词元,就像早年囤比特币、囤域名,是通往财富自由的新风口”“Token就是数字石油!现在不囤,以后稀缺了哭都来不及”“我有内部渠道,Token价格打三折!囤货转手就赚”等话术层出不穷。

深圳市人工智能行业协会执行会长郎丽艳表示,词元只是一个度量单位,真正有价值的是算力和高质量数据;要防止有人浑水摸鱼,以词元为噱头制造新型骗局。

### 完善治理机制 提高安全防范意识

在不少业界人士看来,要持续健全有关词元的标准、市场、公共治理等机制,促进相关产业高质量发展。

陈宁建议,加强实时算力供给与应用需求匹配能力,实现算力资源的精准匹配、动态调度与高效利用,构建与之相适应的统计制度,助力词元成为数字经济中的通用计价单位,贯穿AI服务、数据交易与算力租赁全场景。

合肥零次方机器人有限公司联合创始人李宜哲表示,具身智能是AI智能体的物理载体,可将数字世界的词元价值延伸至物理场景,实现“智能决策+物理执行”一体化。要尽快确定标准,打通词元在人、机、物三方融合中的通用价值。

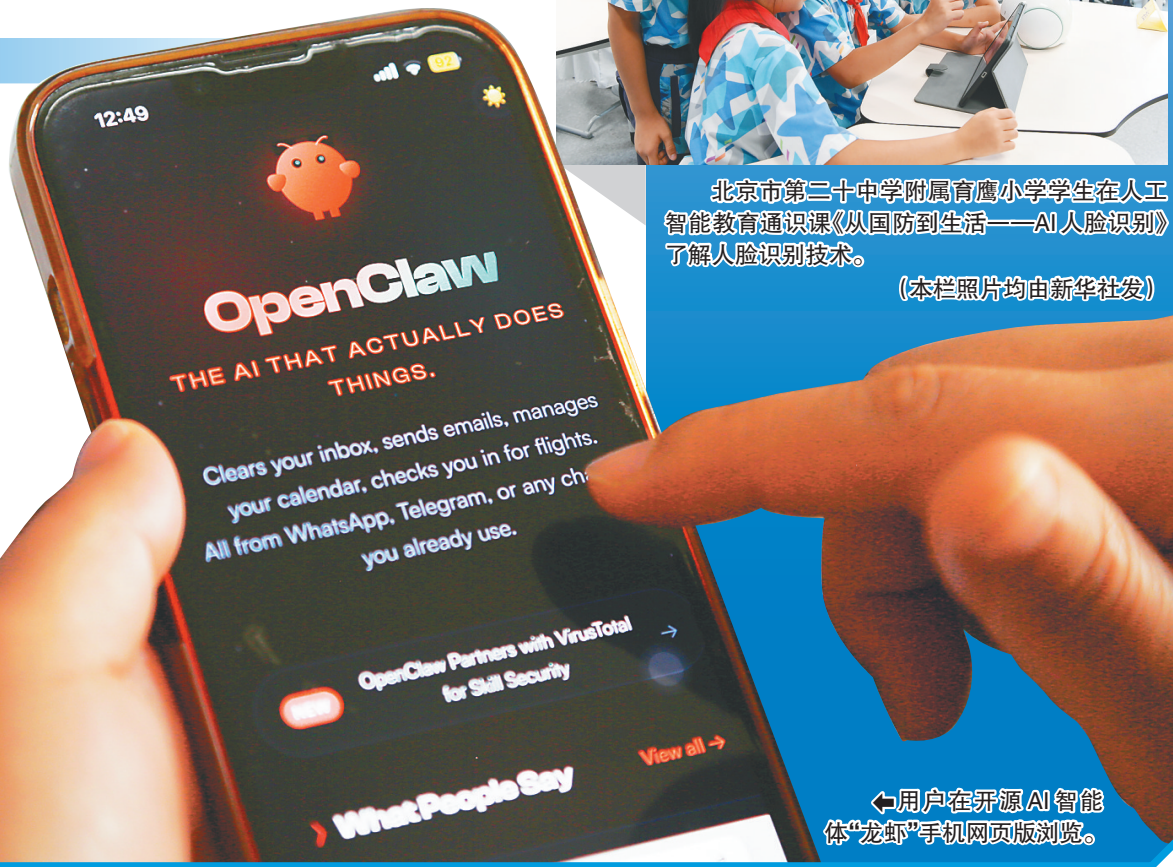
要持续推动技术突破,提升经济性、安全性。上海交通大学副教授、无问芯穹联合创始人戴国浩建议,研究打造敏感数据端侧脱敏、高智能任务上云执行的端云一体架构,通过推

理优化有望降低词元成本,支撑安全、低成本、规模化的词元应用。

不少业界人士认为,要研究词元应用的能耗与碳治理,推广绿色算力认证,重视环境成本,并持续推动构建数字经济全球治理新秩序。郎丽艳表示,要建立可信、高效、公平的价值度量与分配体系,才能让数据、算力、算法、智能体充分释放价值。随着词元机制不断完善,AI将全面融入产业流程,重塑生产、分配、交换、消费全链条。

有关部门提醒,面对词元热潮,我们既要理性看待其价值,又要注意信息安全、隐私安全,提高安全防范意识,做到了解、善用。如遇诈骗、信息泄露或发现非法活动,应及时向有关部门反映。

(新华社深圳5月13日电 记者孙飞 陈宇轩)



人们在浙江省杭州市文三数字生活街区的AI黑科技市集上体验DeepSeek的人工智能大模型。



北京市第二十中学附属育鹰小学学生在人工智能教育通识课《从国防到生活——AI人脸识别》了解人脸识别技术。

(本栏照片均由新华社发)

用户在开源AI智能体“龙虾”手机网页版浏览。

## 中国空间站首次人类“人工胚胎”实验进展顺利

新华社北京5月13日电(记者李国利 刘艺)记者13日从中国科学院空间应用工程与技术中心了解到,随天舟十号上行太空的人类“人工胚胎”实验样本,已装置于中国空间站实验模块,目前实验进展顺利。

这是世界首次在太空开展的人类“人工胚胎”实验。

5月11日,包括“人工胚胎”在内的41项空间科学实验项目,随天舟十号货运飞船抵达空间站。当晚约10时,“人工胚胎”实验样本被在轨航天员装入空间站实验模块。

“目前实验进展非常顺利,预设好的自动化系统每天都会为它们更换新鲜的培养液。”“人工胚胎”空间科学实验项目负责人于乐谦介绍,他们将通过这项实验对关乎人类未来在太空长期驻留、生存、繁衍等问题展开前期研究。

人工胚胎,是用干细胞构建的与真正胚胎非常相似的一种结构。“那么,人类‘人工胚胎’,就是以人类干细胞为原材料制备的。”于乐谦强调,“这不是真正的人类胚胎,不具备发育成为个体的能力,但可作为模型用于人类早期发育研究。”

据了解,“人工胚胎”实验样本包括两款模型,一种是放在子宫细胞上,一种是置于微流控芯片里,旨在了解太空微重力环境对人类胚胎早期发育的影响。与之完全相同的实验样本,也同步在地面实验室开展。

根据计划,人类“人工胚胎”在太空完成5天的实验周期后,实验样本将在轨冻存并择机下行,后续回到地面实验室进行天地对比分析。

“期待通过天地实验样本的发育对比,探索空间环境对人类胚胎早期发育的影响因子,解决人类在太空长期生存面临的风险和挑战。”于乐谦说。

## 湖南全力救治 浏阳华盛烟花厂爆炸事故伤员 已有29名伤员出院

新华社长沙5月13日电(记者帅才)记者从湖南省卫生健康委、长沙市卫生健康委获悉,截至5月13日上午9时,湖南浏阳华盛烟花厂爆炸事故累计出院人数达29人。另有31名伤员仍在院接受治疗,其中,5名重症患者经过精心救治后已有4人转为轻症。

5月4日,湖南长沙浏阳市华盛烟花制造有限公司车间发生爆炸事故后,湖南省卫生健康委、长沙市卫生健康委启动应急响应,调集中南大学湘雅医院、湘雅二医院等多家医疗机构的重症医学科、烧伤科、骨科、急诊科、呼吸科多学科专家,赴浏阳进行伤员救治工作。

湖南卫健、应急、交通等部门统筹资源,由省级急救专家牵头,确保伤员得到高效、科学救治。

“我们第一时间赶到浏阳市中医院,与当地医生一起抢救16名伤员。”参与救治的湘雅医院烧伤科副主任医师张丕红介绍,他们为伤员开展伤口清创、烧伤处置等工作,部分轻症伤员经及时对症治疗已出院回家。

浏阳市人民医院救治了多名重症伤员。浏阳市人民医院医务部主任丁胜强说,对伤势较重的人员,医院医务人员联合湘雅二医院医疗团队围绕抗休克、脏器功能支持等开展专业诊疗。

“经过全力救治,大部分伤员处于康复阶段。”长沙市卫生健康委党组书记刘佳勇介绍,专家组实行“一人一策一团队”精准救治模式,按照危重症、重症、轻症分级分类治疗。随着急救关键窗口期平稳度过,目前医疗救治进入常态化康复诊疗阶段。后续医务人员将对出院伤员建立常态化随访机制,开展康复服务,全力守护患者的生命健康。

## 铁路部门将推出老年旅客淡季周中购票优惠服务

新华社北京5月13日电 记者13日从中国国家铁路集团有限公司获悉,为更好地服务广大老年旅客美好出行需求,铁路部门近期将推出老年旅客淡季周中购票优惠服务,相关优惠车票将于5月15日起陆续发售。

国铁集团客运中心相关负责人介绍,此次优惠服务面向年满60周岁及以上使用中华人民共和国居民身份证、港澳居民居住证、台湾居民居住证、港澳居民来往内地通行证、台湾居民来往大陆通行证、外国人永久居留身份证、临时身份证明及居民户口簿购票的老年旅客。

相关旅客在线上或线下购买开车时间在5月29日至6月30日期间(周一12:00至周五12:00,6月18日至22日端午节假期运输期除外)的周中时段的部分动车组列车车票时,可享受执行票价9折优惠。铁路12306在相关优惠车次后标注“敬”字,享受优惠的车票在订单详情和本人车票页面标识“长者优惠”字样。

年满60周岁及以上的持有残疾人证、伤残人民警察证、国家综合性消防救援队伍残疾人员证的中国公民,在享受原有优待票价的基础上,可再享受9折优惠。成功购票出行的老年会员旅客,同时享受普通常旅客会员的3倍积分优惠,积分可用于兑换火车票或办理席位升席业务。



新华社发

# 中国量子计算再破纪录 独家揭秘“九章四号”



“九章四号”量子计算原型机局部。 新华社发

新华社合肥5月13日电(记者陈诺 戴威 何曦悦)目前世界上最快的超级计算机,求解特定数学问题约需10的42次方年。由中国科学家最新研制成功的“九章四号”量子计算原型机只需要25毫秒,比超算快超过亿

亿亿亿亿倍。

中国科学技术大学潘建伟、陆朝阳、张强、刘乃乐等组成的研究团队联合国内多家科研单位,成功研制出1024个量子压缩态输入、8176模式的可编程量子计算原型机“九章四

号”,首次操纵和探测高达3050个光子的量子态,再度刷新光子信息技术世界纪录。国际知名学术期刊《自然》13日发表了该成果。

什么是量子计算机?它和我们日常用的电脑有什么区别?“九章四号”为何那么强?这次成果意味着啥?新华社记者独家采访科研团队,带来权威解读。

问:啥叫量子计算机?

答:量子计算机是一类遵循量子力学规律进行高速数学和逻辑运算、存储及处理量子信息的物理装置。与经典计算的普通比特不同,量子计算机用的量子比特不是只有“0”或“1”,而是可以同时处于“0”和“1”的叠加态。借助叠加态与纠缠态,量子计算能够在同一时间并行处理大量计算路径,在特定问题上实现指数级加速。

目前主流量子比特技术路线包括超导、离子阱、光量子和中性原子等。“九章四号”是光量子计算原型机,使用光子来编码量子比特,通过对光子的量子操控及测量来实现量子计算,是“九章”(2020年)、“九章二号”(2021年)、“九章三号”(2023年)量

子计算原型机的再升级版。问:1024个量子压缩态、8176模式、3050个光子是啥意思?

答:“1024个量子压缩态”可以通过通俗理解为这台量子计算机运行时的“高能燃料”。在量子计算中,压缩态光子具有极其特殊的量子特性,是构建复杂量子纠缠的基础。

“8176模式”指的是光子在计算网络中可以穿梭的“路径”或“维度”的总数。你可以把这想象成一个有着8176个出口的超级三维立体迷宫,光子在里面进行极复杂的干涉游走。

“3050个光子”就是3000多个量子比特。此前,“九章三号”探测和操纵255个光子,“九章四号”提升了10倍多,意味着系统能够代表和处理的计算状态空间呈指数级增长,标志着人类操控微观量子世界的规模有了数量级的跨越。

问:“九章四号”最大的技术突破在哪?

答:光量子计算走向更大规模一直被一个叫“光子损耗”的“拦路虎”阻挡;随着光学网络越来越大、越来越复杂,光子在里面极易容易跑丢,从而大幅削弱计算能力。

此次最大的技术突破在于团队首创的“可编程时空混合编码”架构。过去如果要扩大规模,只能通过增加光学器件,导致设备庞大且损耗极高。“九章四号”让光子在时间和空间两个维度上同时发生干涉,极大提升了整个网络的连通性,又兼顾控制了物理器件的规模,进而获得了对高达3050个光子的操纵和探测能力,带来算力的指数级提升。

问:这次成果意味着啥?

答:目前“九章”系列是极为强大的专用量子模拟机,它只擅长解决“高斯玻色取样”这样特定的数学问题。这个数学问题短期内可用于图像识别、图论计算等,长远还能生成玻色纠缠态,是未来打造高稳定通用量子计算机的关键。

“九章四号”在规模与低损耗的双重领先优势下构建“万亿量子模式的三维簇态”和未来的“容错光量子计算硬件”提供了可能,为未来的通用量子计算机奠定了坚实基础。

实现通用量子计算机,需要操纵上百万个量子比特,同时也要具备纠错能力。这些都需要在现有量子计算原型机基础上不断迭代,慢慢实现。