

小词元如何撬动 AI 大产业

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，一次复杂的智能体任务可能要消耗数万词元。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。



3月12日，在江苏省连云港市连云区，用户在使用开源AI智能体“龙虾”电脑网页版。新华社发(王春摄)

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短剧吗？”“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时代的分水岭，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)综合开发研究院通证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，把问答工具升级成真正能“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，“热闹多、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发文提醒，词元在使用过程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。

在一些社交平台和私域社群里，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敏感、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云端成本，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。如遭遇诈骗、信息泄露或发现非法活动，应及时向有关部门反映。

新华社“新华视点”记者孙飞、陈宇轩
新华社深圳5月13日电