

词元

Token有了中文名,意味几何?

字里有料

日前,在中国发展高层论坛2026年年会上,国家数据局局长刘烈宏对AI领域核心术语Token的官方定译——“词元”,正式揭开面纱。这一举措看似只是一个专业术语的中文统一,实则蕴含着我国AI产业从高速发展向高质量规范迈进的深层信号,为智能时代的技术落地与商业变现划定了清晰的价值坐标。

网友热议

- “从蒸汽机的‘马力’,到信息时代的‘比特’。在智能时代,‘词元’或许正是这样一个标志性单位。”
- “术语统一只是起点,真正的考验在于,让‘词元’这一标准真正服务于产业生态的繁荣。”

Token一般指数据、信息或物理对象的符号化表示或最小处理单元,是计算机领域的数字标识符。在信息技术中,它作为随机字符串或加密数据块用于身份验证、安全访问控制,代表用户、设备或会话的授权信息。

长期以来,Token在AI领域的译法乱象丛生,“令牌”“通行证”“语元”等译法混杂出现,既造成了专业交流的壁垒,也让普通民众在接触AI知识时产生认知困惑。“令牌”源自其在计算机网络领域的传统含义,与AI自然语言处理的核心场景脱节;“通行证”则带有区块链领域的强烈标签,易误导大众对AI技术本质的理解;其他各类译法要么过于抽象,要么适用范围狭窄,难以适配大模型多场景应用的需求。

日前,在中国发展高层论坛2026年年会上,国家数据局局长刘烈宏表示,Token“词元”不仅是智能时代的价值锚点,更是连接技术供给与商业需求的结算单位,为商业模式的落地提供了可量化的可能。这也从官方角度上为Token给出了翻译:词元。有分析指出,“词元”二字精准契合了严复提出的“信、达、雅”翻译准则,“词”锚定其语言学本源,“元”凸显其作为AI处理信息的最小单元的技术本质,既准确还原了Token的专业内涵,又简洁易懂、便于传播,彻底终结了译法混乱的局面,为行业交流、学术研究和大众科普搭建了统一的语言桥梁。

“词元”的定名,更超越了单纯的术

语翻译,成为解读AI产业发展的重要钥匙。其被定义为“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”,一语道破了词元在AI产业中的核心作用。作为大模型处理信息的最小计量单位,用户输入的每一个字、AI生成的每一段话、识别的每一幅图像,本质上都是词元的消耗过程。从数据来看,我国日均词元调用量已突破140万亿,相比2024年初增长1000多倍,三个月内同比增长40%以上,这组惊人的数据背后,是AI技术从实验室走向千行百业、融入千家万户的生动实践,而词元正是衡量这份实践成效的“晴雨表”。调用量越高,意味着AI模型的应用越广泛、创造的实际价值越大。

舆论场上还有观点从商业角度对此进行分析,此前,AI产业的商业变现模式虽有探索,但缺乏可量化、可落地的核心载体,而词元的出现,让“用多少、付多少”的计费模式成为可能。不同于传统软件一次性购买的模式,如今企业使用AI服务,可根据词元消耗量精准计费,这种灵活的商业模式不仅降低了企业接入AI的门槛,更催生了围绕词元调用、分发与结算的全新价值体系,部分模型企业甚至创下20天收入超越2025年全年的佳绩,印证了词元作为“结算单位”的商业价值。同时,词元调用量的激增也对算力、电力资源提出了更高要求,倒逼“算力协同”等新基建工程加速推进,推动AI产业与绿色能源深度融合,实现可持续发展。

从政策层面,此番定名也是对AI产业治理思路的深化。从“鼓励创新、包容审慎”到“规范发展、标准先行”,治理重心正在向精细化、制度化延伸。在技术快速迭代的当下,及时对基础概念进行权威界定,既避免了概念歧义可能引发的市场混乱,也为后续监管与产业政策的精准施策提供了前提。可以说,“词元”的定译,是智能时代我国在技术话语权与产业规则层面主动作为的一个缩影。

更值得一提的是,当前,我国AI产业发展势头迅猛,不仅日均词元调用量稳居世界前列,在大模型研发领域也形成了较强竞争力,有相关数据显示,我国AI模型周调用量已连续三周超越美国,排名前五的模型中有四款来自中国厂商。在这样的背景下,统一词元的中文译名,不仅是完善AI术语体系的基础工作,更是我国AI产业从“跟随”向“引领”转变的重要体现。唯有实现术语规范、标准统一,才能更好地整合产业资源、凝聚发展合力,推动我国AI产业在全球竞争中占据主动。

当然,“词元”的定名只是起点。随着AI技术向多模态、具身智能等领域持续演进,词元的应用场景还将不断拓展,其价值内涵也将进一步丰富。未来,中国AI产业需要以词元为抓手,完善计量标准、计费体系和监管规范,让这一“价值锚点”真正发挥作用,推动数据要素与AI技术深度融合,释放更多产业活力。

莫让“侥幸”成“不幸”
安全生产,预防为主