

别被你的 AI 忽悠了

当前,我国AI大模型的应用呈现迅猛增长态势。QuestMobile发布的最新数据显示,DeepSeek上线次月(2025年2月),活跃用户规模就突破了1.8亿。但在AI应用大众化的同时,也有不少网友惊呼:与AI的蜜月期还没有结束,就发现它在骗我!

有人用AI查资料,发现AI会编造假论文,期刊名称、论文标题、作者、网址等信息一应俱全,可全是编的;曾有一条“80后死亡率突破5.2%”的假新闻广为流传,据辟谣者分析,最初的信源可能来自AI对话……

这其实不是新问题,而是在大模型诞生之初就存在的老毛病——AI幻觉。甚至有专家认为,AI幻觉是大模型的“基因病”。因为AI大模型训练依赖海量数据,但数据的真实性却无人担保。AI大模型训练过程中,是将海量数据高度压缩抽象,输入的是内容之间关系的数学表征,而不是内容本身。AI与人类对话时,模型基于统计概率生成文本,追求流畅性而非真实性,很容易编出看似流畅

但完全不符合事实的内容。

既然AI幻觉是“基因病”,那这病能治吗?要怎么治?

首先,AI幻觉也带来了AI的“超能力”,没必要根除。

AI幻觉也有好的一面。AI自主生成内容是大语言模型的突出特点,是最具想象力的技术突破。作为大模型与生俱来的特点,AI幻觉决定了大模型的创造力,对实现通用人工智能(AGI)非常重要,并可在科研、文艺等领域发挥作用。

比如,游戏开发、动漫设计、小说创作等领域欢迎奇思妙想,AI幻觉可以突破思维定式,为创作提供灵感。再如,科学研究中往往也需要非常规思路来破局闯关,AI不走寻常路的创造力,可以推动科研工作出现“AI幻觉—实验验证—理论重构”的新范式。

其次,AI幻觉虽不可根除,但可以控制。对于AI幻觉,我们不能放任自流。一方面,AI幻觉产生大量虚假信息,这些虚假信息

息被新一代AI系统学习,就会形成“数据污染—算法吸收—再污染”的恶性循环。另一方面,在企业生产、医疗、法律等容错率很低的领域,AI幻觉可是“要命”的大毛病。因此,我们要从技术研发和管理机制上双管齐下,防范AI幻觉信息的泛滥。

技术层面,检索增强生成(RAG)技术、多模型交叉验证、动态知识更新机制等,都已被证明是减少AI幻觉的有效手段。未来,还需要继续提升训练数据质量与清洗标准,严格检测数据的来源和真实性,减少噪声干扰;引入强化学习与人类反馈机制,提升AI逻辑推理能力;增强AI大模型检测能力,用AI监管AI。

用户层面,对AI输出的答案不要直接采信,可以用多个AI模型分别生成答案再交叉验证、对照取优,对关键任务输出更有必要采取人工交叉验证,结合文献数据库、行业知识图谱等专业工具作出判断。

监管层面,“标注AI生成内容”已成为我

国监管部门推动的方向,这体现了对生成式人工智能的透明度要求,将提高AI生成内容的可追溯性和可解释性。未来,我们还应加快人工智能法立法进程,建立全周期全链条监管体系。

实现“不被AI忽悠”目标,既需要算法工程师构建底层安全防线,也需要全社会共同建设可信AI生态。期待中国AI产业解决好这道AI幻觉控制难题,在创造力与真实性之间取得平衡。



碳市场扩围加速绿色转型

日前,生态环境部发布《全国碳排放权交易市场覆盖钢铁、水泥、铝冶炼行业工作方案》,将全国碳市场覆盖范围由发电行业拓展至高耗能高排放工业行业。在当前外部环境不确定性因素增多背景下,我国推动碳排放权等要素在更大范围内畅通流动和优化配置有何意义?扩围后对全国碳市场交易规模、交易机制等产生什么影响?在扩围过程中,如何处理好发展和安全、效率和公平的关系?

田智宇



强化战略定力

全国碳市场是利用市场机制落实“双碳”目标的重大举措,是我国深化生态文明体制改革的重要制度创新。自2021年启动运行以来,全国碳市场在推动电力企业节能降碳、降低行业总体减排成本、彰显我国积极应对气候变化担当等方面发挥了重要作用。

从全球看,国际形势变乱交织,绿色低碳发展竞争博弈日趋复杂。当今世界百年未有之大变局加速演进,大国博弈、地缘政治冲突不断演进,进一步加剧全球政经格局走势的不确定性。绿色低碳投资和消费规模快速扩大,可再生能源、电动汽车、储能等对传统能源体系、产业体系的变革性影响不断显现,主要经济体之间绿色低碳竞争、气候治理博弈更加激烈。

从国内看,我国步入加快绿色化、低碳化的高质量发展阶段,统筹发展、减排和安全形势更加严峻。可再生能源、电动汽车、节能环保产业等发展迅猛,对高质量发展产生了强劲推动力。但同时,提高能效和优化调整能源结构进入攻坚阶段,传统高耗能高排放行业供大于求、低水平竞争等问题不断凸显,新能源、电动汽车等产品出口和产业发展面临部分西方国家的遏制打压,我国绿色低碳高质量发展任务更加艰巨。

在日趋复杂的国内外形势下,我国稳步扩大碳市场覆盖范围,体现了绿色低碳发展的战略定力,我国走绿色转型之路是长远之策,绝非权宜之计;体现了要素市场化配置改革的决心,建设全国统一、竞争有序的碳市场将持续推进;体现了积极应对气候变化的信心和行动,不论国际形势如何变化,我国将持之以恒积极稳妥推进碳达峰碳中和。

培育竞争优势

近年来,我国钢铁、水泥、铝冶炼行业能效水平持续提升,但企业之间绿色低碳发展水平不均衡。比如,粗钢、水泥熟料生产中仍有约10%的产能能效达不到基准水平。同时,在供大于求背景下,钢铁、水泥行业利润总额大幅下降,企业增加节能降碳投入的动力普遍不足。碳市场扩围后,有利于激励企业持续挖掘节能降碳潜力,加快低效和落后产能出清,形成以“降碳”为标尺的产业健康发展生态,推动行业发展摆脱“内卷式”恶性竞争。

引导企业优化面向长远的研发和投资决策。钢铁、水泥、铝冶炼行业既是碳排放的重点行业,也是绿色低碳技术创新的重要载体。作为我国初级产品和产业链供应链安全的重要保障,钢铁、水泥、铝冶炼行业在产能更替、技术工艺升级等方面都有巨大发展空间。尽管启动阶段对配额进行免费分配,但通过发挥碳价信号的引导作用,可以引导经营主体优化研发和投资决策,避免长期高碳锁定风险。同时,碳价信号还可以促进产学研用金、产业链供应链等协同创新,不断催生新技术、新业态,实现传统产业“焕新”和新兴产业发展互促共进。

积极掌握国际绿色低碳博弈主动权。当前,全球贸易发展与环境、气候议题不断交织,绿色壁垒增多,并且向市场准入、供应链重构等领域不断延伸。我国在稳步扩大制度型开放过程中,在绿色低碳规则、规划、管理、标准等领域,在碳定价、碳排放统计核算监测、技术服务能力等方面仍存在许多薄弱环节。在碳市场扩围的同时,加快夯实法规标准、数据统计、技术服务等基础能力,不断扩大全国碳市场国际影响力,对提升我国绿色低碳领域的制度性话语权具有重要意义。

提升配置效能

碳排放看不见、摸不着。有效发挥市场机制在控制碳排放中的作用,在扩大碳市场

覆盖范围的同时,还要以碳达峰碳中和为引领,进一步深化能源资源等价格形成机制改革,加强碳市场与电力市场、用水权交易市场、排污权交易市场等有效衔接,并一体纳入要素配置改革总盘子,实现降碳减污扩绿增长系统效益最大化。

强化碳市场与其他政策工具协同作用。我国推动“两高”行业绿色低碳转型的政策工具众多,包括产能调控、节能降碳改造升级、严格污染物排放标准要求、绿证强制消费等,一些大气污染防治重点地区还提出了煤炭消费减量要求等。在碳市场扩围的同时,要注重增强宏观政策取向一致性,加强不同政策工具的高效协同和优化组合,把握好改革的时度效,在避免增加经营主体不合理负担的前提下,有效发挥各种政策工具的长处,形成推动绿色低碳发展的部门合力。

促进区域、行业绿色低碳协调发展。我国不同地区资源禀赋、发展阶段和水平等存在差别,不同行业在国民经济中的地位和作用存在差异,并且区域格局、产业结构、能源流向等仍在调整变化。在碳市场扩围的过程中,要处理好发展和安全、效率和公平的关系,既引导各类要素向先进生产力、经济增长极等集聚,也要保障好初级产品供给和产业链供应链安全,形成优势互补、协调推进的绿色低碳高质量发展格局。

(作者系中国宏观经济研究院能源研究所研究员)

求真

职场困扰、升学压力、人际关系……当下,不少人在面对心理压力时,倾向于向交互式人工智能(AI)倾诉。AI因即时响应、无评判态度等特点,成为部分人的“情绪出口”,甚至在不少青少年人群中,AI作为倾诉对象的受欢迎程度已超过家人。把AI当作倾诉对象,靠谱吗?

北京大学第六医院主任医师黄薛冰表示,青少年是数字世界的原住民,对交互式的AI软件接受快,这和青少年的心理特点有关,他们的交友需求较为强烈。另外,青少年在成长过程中,其自我意识以及自信心尚在建立中,与他人社交时容易缺乏自信,有时感到羞怯,担心别人评价自己、不接纳自己。青少年和AI交流时没有这样外在的评价,会更加放松

“法务协商”骗局高发且难立案。消费者增强风险防范意识的同时,建议短视频平台对“法务协商”等关键词强制添加风险提示,并进行严格筛查,尽可能压缩黑灰产引流空间。

本报记者 王宝会

近期,多地金融监管部门密集发布风险提示,提醒公众警惕非法代理维权中介的诈骗行为,揭示贷款协商背后的风险,以保护消费者合法权益。

刚大学毕业不久的张明(化名)因创业失败,背负了50余万元的信用卡债务,频繁的催收电话让他不堪其扰。日前,他在短视频平台上刷到一条自称“专业法务”的广告,宣称可帮助网贷逾期用户通过法务协商实现延期息。于是,他上网搜索相关帖子,看到多个“成功案例”,遂与短视频博主签订合同并支付定金。然而,不久后银行致电张明,称有人以他的名义提供虚假材料进行投诉,导致其延期和减免息费诉求被拒。

张明的遭遇并非个例。当前,非法代理维权中介经常利用网络论坛、自媒体平台等渠道发布信息,冒充各种咨询公司或法律顾问骗取金融消费者信任,发布代理维权成功的虚假案例,诱导金融消费者委托其进行代理。

消费者一旦签订协议,往往需要先支付高额费用。这些不法分子通常诱导消费者借新还旧或申请过桥垫资,最终收取垫资费,导致消费者债务重组后的综合资金成本远高于正常银行贷款,甚至陷入连环债务陷阱。此外,消费者在提供身份证件、金融账户、家庭成员等重要信息后,个人信息还可能被非法利用或出售。

北京京臻律师事务所律师杨高范指出,“法务协商”骗局高发的原因主要有两方面:一是庞大的群体性需求,债务人在催收压力下容易病急乱投医;二是债务人普遍缺乏对正规流程的认知,易被“成功案例”“专业法务”话术和伪造案例迷惑。此类骗局立案的根源在于不法分子多采用“线上引流+虚拟身份”模式,资金流向复杂、隐蔽性强、取证成本高。

国家金融监督管理总局、中国人民银行、中国证监会近日联合发布的《关于推进金融纠纷调解工作高质量发展的意见》明确提出,金融管理部门要会同有关部门加强金融领域“代理维权”乱象的联合整治,强化金融纠纷源头治理。金融机构也在加强风险提示,比如,在今年的“3·15”金融教育宣传活动中,浙江农商联合银行辖内嘉善农商银行聚焦维护权益主题,通过情境化消保视频和组织金融知识讲座等形式,提醒用户警惕非法代理维权的危害及风险。

有专家建议,除此之外,短视频平台等也应“对‘法务协商’等关键词强制添加风险提示,并进行严格筛查,尽可能压缩黑灰产引流空间。

中国社会科学院金融研究所银行研究室主任李广子表示,打击非法代理维权金融黑灰产业需要从多方面入手。一是加大执法力度。围绕消费金融等重点领域开展专项整治行动,对违法行为保持高压态势,提高违法成本。二是强化技术防控与大数据监测。执法部门可与各类平台进行合作,利用人工智能等技术筛查违规内容,打击非法金融信息传播。三是持续加强消费者教育。通过案例警示与科普宣传等方式,提高金融消费者风险识别和防范能力。

本版编辑 孟飞 李苑美 编倪梦婷 来稿邮箱 jrbgzb@163.com

□ 本报记者 吴佳佳

AI 互动不能取代真实人际关系

地吐露心里话,这是孩子们喜欢AI的原因。

尽管AI能提供即时情绪安抚,但青少年的健康成长离不开真实的社会互动。“社会化过程中,个体需要通过他人的反馈调整行为认知,学习社会规范。”黄薛冰说,真人交流中的评价、表情、肢体语言等,能帮助青少年跳出自我视角,学会共情、理解社交边界。而长期依赖AI,可能导致情感识别与表达能力退化,甚至出现“机器人化”倾向——情感淡漠、社交疏离。

“AI互动可作为心理压力调节的短期辅助手段,但不能取代真实人际关系。”黄薛冰建议,家长和学校应关注青少年的社交需求,通过团体活动、心理健康教育等方式,引导其建立健康的社交模式。