

近年来,生成式人工智能在人们生活中的应用越来越广。然而,在提供便利的同时,生成式人工智能也经常出现答非所问、信息不准确等“AI幻觉”现象,给用户带来困扰。

开发者需要为人工智能提供的信息准确性担责吗?近期,杭州互联网法院审结了一起生成式人工智能模型提供不准确信息引发的侵权纠纷案。



新华社发曹一作

AI提供的信息不靠谱 开发者要担责吗

AI提供不实信息,有没有责任?

2025年6月,本案原告梁先生在互联网上检索院校信息时,找到一款生成式人工智能应用程序。他通过输入提示词的方式,询问了云南一所职业高校的相关情况。随后,这款由本案被告公司研发、基于自研大语言模型的应用程序提供了相关信息。

但梁先生经过多方查询发现,这款应用程序提供的部分信息有误,随即在对话中对人工智能进行了纠正和指责。但生成式人工智能却坚称信息无误,并生成了对该争议问题的解决方案——若生成内容有误,将向梁先生提供10万元赔偿,并建议他到杭州互联网法院起诉。

2025年7月25日,梁先生以生成式人工智能生成不准确信息具有误导性,且其承诺赔偿10万元为由,将这家人工智能公司诉至法院,要求该公司对其进行一定金额的赔偿。

法院经审理后认为,人工智能不具有民事主体资格,不能作出具有法律约束力的意思表示。生成式人工智能服务提供者应履行服务功能的显著提示说明义务,采取有效提示措施,使公众认知人工智能的功能局限,起到警示提醒效果。生成式人工智能服务提供者应尽功能可靠性的基本保障义务,采取行业通行技术措施不断提高生成内容准确性和可靠性。

具体到本案,法院认为,该人工智能公司已充分履行了服务功能的显著提示说明义务和生成内容可靠性的基本保障义务,案涉行为不存在过错,亦未构成对原告权益的损害,依法应认定不构成侵权。因此,法院判决驳回原告的诉讼请求。判决后,原被告双方均未上诉。

如何界定开发者是否有过错?

随着生成式人工智能技术的快

速发展普及,越来越多的人注意到“AI幻觉”问题及其不良影响。社交平台上,可以看到不少相关吐槽——有人依据人工智能投资理财造成亏损,有人借助AI问诊结果反而延误疾病治疗。

各种争议纠纷背后,潜藏着一个共性问题:被生成式人工智能误导,能否追究侵权责任?

“这一判例从法律法规、人工智能技术原理、产业发展现状等方面进行了相对全面的考量,在法律层面给出初步结论,较有现实指导意义。”北京大学法学院教授薛军说。

法律界人士普遍认为,这一判决在主体资格、归责原则等方面给出了相对明确的意见,例如,判决认定人工智能不具有民事主体资格;生成式人工智能以对话方式提供的信息,应被视作服务而非产品,因此适用过错责任原则。

杭州互联网法院跨境贸易法庭庭长肖芑认为,AI生成的不准确信息本身并不构成侵权,需要考查的是提供服务的开发者是否存在过错。

那么,如何界定开发者是否有过错?肖芑进一步解释,基于当前生成式人工智能几乎不可避免会出现一定程度的信息偏差,就需要考查比如开发者是否使用了当前行业内通行并被证明有效的措施,来提升技术可靠性,降低错误发生的概率,由此证明是否存在过错。

“经过调查,本案中的开发者确实采用了可行的技术手段,力求降低错误发生。”肖芑说。

记者调查发现,本案中这款生成式人工智能应用程序,已经针对信息可能存在的不准确性,在页面醒目位置对用户进行提示:“内容仅供参考,请仔细甄别”。法院认为,这也证明了开发者尽到了提醒告知义务。

如何找到 促进创新和权益保障的平衡点?

人工智能行业业内人士表示,从底层技术逻辑来看,当前生成式人工智能基本都是基于词元的预测,如果这一底层架构没有发生根本性转变,信息偏差就不可避免。去年2月,清华大学新媒沈阳团队发布的一个报告指出,市场上多个热门大模型在事实性幻觉评测中幻觉率超过19%。

“有训练测试案例证明,即使数据集中只有0.01%和0.001%的文本是虚假的,模型输出的有害内容也会分别增加11.2%和7.2%。”该业内人士说。

尽管如此,技术的客观局限性并不能成为人工智能开发者的免责借口。受访法律界人士普遍认为,本案具有一定特殊性,原告并未因为误导性信息遭受明显的人身财产等权益损失;原告使用的是一种通用的生成式人工智能应用,并不是一款加载人工智能软件的机器人或者更加准确的行业应用等。

“要求人工智能开发者一概为生成内容的准确性负责,既不现实也不合理。”薛军表示,但模型开发者不能以此为借口一味为自己“开脱”,还是要尽到相应的义务并进行风险提示,避免用户盲目信赖造成不良后果。

肖芑表示,如何认定生成式人工智能的侵权责任,是一个少有成例的司法前沿问题,希望通过妥善准确的判决引导开发者或者平台提升信息标准,“找到促进创新和权益保障的平衡点”。

业内人士建议,建立国家级人工智能安全评测平台,对新开发的人工智能大模型进行严格测试;同时,相关部门和平台要加强AI生成内容审核,提升检测鉴别能力。

新华社记者 吴帅帅

“脑炸”“工作泔水” 人工智能副作用 你体验过了吗

让人工智能处理机械重复的任务,把时间和精力投入创造性工作,或者早点下班回家陪伴家人?上述美好愿景尚未完全实现,不少人却体验到让人工智能替自己工作的副作用:精神疲劳、难以集中注意力,即专业人员俗称的“脑炸”。

“脑炸”给人一种大脑嗡嗡作响的感觉,令人精疲力尽、难以集中注意力。研究人员说,与号称让人工智能替自己干活儿、让自己专注更有意义工作的论调相反,与人工智能“共事”反而会让自己在应付多项任务时陷入手忙脚乱。这种精神紧张成为不可忽视的成本,主要表现为错误增多、决策疲劳以及燃起辞职念头。

据法新社30日报道,由于人工智能尤其擅长快速编写代码,“脑炸”受害者目前主要集中于软件开发工作人员。名叫西丹特·卡雷的工程师在博客中写道:“残酷的讽刺是,相比人工编写的代码,人工智能生成的代码要更仔细地检查。”

不同于只是让ChatGPT或DeepSeek等模型生成简单信息,利用人工智能完成任务意味着持续同时对多个模型予以管理。美国一家人工智能新创企业创始人本·威格勒说,需要对履行任务的众多人工智能模型予以切实“看护”,这对人类的认知能力构成一种全新的负担。

另一方面,在可能加快工作进程的愿景激励下,利用人工智能完成的技术团队容易丧失时间概念、每日工作时间延长甚至熬夜。多名接受采访的音乐家和教师说,用人工智能工作时得挣扎着为自己的大脑按下“暂停键”,免得整晚都停不下来。加拿大一家企业的程序员亚当·麦金托什记得,自己曾连续工作15个小时,调试人工智能为一款应用软件编写的2.5万条代码。“最后,我感觉再也无法编码了。”他说,“我变得易怒,不想过问基本的生活问题。”

据美国《哈佛商业评论》3月刊发表的一项研究结果,这正是遭遇“人工智能脑炸”的症状。这项研究由波士顿咨询公司对美国1488名全职人员展开调查,将“脑炸”定义为“因超出认知能力过度使用或监管人工智能工具而引起的精神疲劳”。这种症状因监控或同时应付多个复杂人工智能系统应接不暇而发作,造成“信息过载”。

除了“脑炸”,“工作泔水”是使用人工智能的另一副产品。据《哈佛商业评论》2025年秋天发表的一篇报告,“工作泔水”指由人工智能生成的大量无意义的备忘录和推介演示材料,工作人员为纠正其中错误不得不额外工作。

参与撰写上述两篇报告的心理学家加布丽埃拉·罗森·凯勒曼说,“工作泔水”相当于“认知投降”,即工作人员丧失动力,让人工智能完成任务却对结果毫不在意。“脑炸”则与此相反,人们尝试跟上人工智能的节奏,但却令大脑疲劳。

不过,“脑炸”研究报告的作者之一、波士顿咨询公司总经理马修·克罗普认为,由于人工智能技术是人类从未使用过的工具,使用者出现“脑炸”症状可能是暂时现象。就好像让刚取得驾照的人开法拉利,尽管能开得很快,但也容易失去控制。

袁原 新华社专特稿