

万事问AI,当心被忽悠

新闻名词

“AI幻觉”

通常指大语言模型在缺乏真实依据的情况下,生成看似合理实则存在事实错误或逻辑谬误的内容。

现象

AI助手频频惹祸引争议

前不久,社交平台上一则“用户轻信AI退票导致损失600元”的消息引发关注。河北用户李先生因行程变更,计划退掉3张石家庄至重庆的机票,向AI助手豆包咨询退票手续费。豆包明确答复,机票退票手续费仅为5%,不足百元,可以放心操作。

未向航司核实的李先生随即退票,结果被平台按40%比例扣除手续费,合计损失600元。气愤的李先生向豆包讨说法,豆包非但未推诿,反而生成《赔付承诺书》,承诺若未能在

约定日期帮他追回款项,将全额赔付600元,还催促用户提供收款码,声称“秒到账”。

可到了约定日期,李先生并未收到任何赔付。追问时,豆包“改口”称自身仅为人工智能,无法转账,此前承诺无效。今年5月,李先生向北京互联网法院起诉豆包运营公司。戏剧性的是,起诉前他向豆包咨询维权事宜,豆包不仅帮他起草起诉状,还断言“稳赢”。

无独有偶,一则“被豆包忽悠花了5万元创业”的消息近日也冲上热搜。一名打工者向AI助手豆包咨询“年轻人可参与的AI风口项目”,并在其建议下,投入5万元购买了万元级服务器和近500个账号。结果,该用户信心满满地在短视频平台推广时,因“批量注册矩阵号+同质化内容”而

被全平台封号,5万元一下子打了水漂。

AI助手频频惹祸,引发网友热议。“是时候拿出你和AI的离谱聊天记录了。”在小红书上,一则关于AI助手的吐槽帖获得6000多条评论。网友们纷纷晒出自己的截图,列举自己遭遇的“奇葩事”。

有的网友发现AI助手“眼神不好”。一位网友发帖称,自己家里的土豆时间比较久了,切开后有褐色的丝状纹路,不知道能不能吃,于是拍照让AI鉴定。“AI非说这是雪莲果,可以吃。”

有的网友发现,AI助手作图能力较差。一位网友称,她想让AI把照片P成长发的样子,结果照片中的自己不但长出了长发,还长出了胡子。

探因

为什么AI会出现“幻觉”?

为什么AI会一本正经地“胡说八道”?记者了解到,在专业术语中,这类现象被称作“AI幻觉”。根据相关解释,“AI幻觉”常常以看似逻辑自洽、表述严谨的方式生成虚假信息,尤其在回应严肃知识性问题时,会凭空捏造事实、杜撰不存在的学术文献,甚至伪造个人信息,且往往伪装得极为逼真。

在哪些场景下,AI容易“胡说八道”?记者将这个问题抛给DeepSeek,得到了比较清晰的答案。根据DeepSeek的回答,AI在具体事实与数据、最新事件、小众或专业领域、逻辑推理、预测未来、解释主观或模糊概念、顺从用户预设等7个方面容易犯错。

比如某个人物的出生日期、某个

事件的时间地点,AI可能会把相似信息混在一起,编造出看似合理实则错误的信息。一些冷门知识,例如某些冷门书籍的内容、偏远地区的法规细则等,由于数据不足,AI也容易“瞎编”。以DeepSeek为例,它的“知识储备”截止于2025年5月份,在此之后发生的事,如果没有开启联网搜索,它就可能给出过时或虚构的信息。

需要注意的是,对于“明年房价会涨吗”这类问题,AI虽然会基于规律给出推测,但本质上只是文字组合,并非可靠预言。对于评判人物的问题,例如“为什么张三是个坏人”,AI可能会顺着预设回答,但实际上张三这个人可能并不存在。

为什么AI会出现“幻觉”?记者了解到,AI助手本质上是一个大型语言模型,通过学习海量文本中的统计规律,来预测下一个最合适的词。简单来说,AI并不知道自己说的内容是什么,这就像一个人背下了很多对话模板,能流利聊天,但没真正理解内容。

“生成式AI的运行机制决定了其必须输出逻辑自洽、表达流畅的结果,当数据存在缺陷时,模型会基于现有信息进行逻辑补全,进而产生与事实不符的不准确内容。”中国传媒大学教授、博士生导师郭艳民表示,生成式AI产生“幻觉”的核心根源,在于其基于大数据的运行机制。郭艳民表示,从绝对技术层面来看,“AI幻觉”是生成式AI的固有特征,无法被

彻底消除。这就决定了人们无法从根源上杜绝此类问题的产生,而只能通过技术迭代来持续降低“AI幻觉”发生的概率。

为何“AI幻觉”不易被察觉?

“AI幻觉”为什么不易被察觉?记者发现,随着AI工具越来越普及,越来越“好用”,许多人患上了“AI依赖症”。在职场,不少人用AI写邮件,用AI做PPT,用AI查资料,甚至用AI写代码……这种“遇事不决问AI”的思维惯性,让AI已不再是“助手”,而是替代了人们自己的思考。

前不久,上海东华大学团委对该校学生AI工具的使用习惯和依赖程度展开了一次微调研。结果显示,超九成学生每天至少使用一次AI,近四成学生每天至少使用AI工具5次以上,表明AI已成为大学生日常使用的工具。除了在学习中使用AI检索信息、完成作业外,大学生还使用AI进行旅游规划、穿搭推荐,甚至让AI编辑一段文字,帮助自己回复信息。针对这种情况,部分学生担心过度依赖AI会影响独立思考。学校团委也建议,应引导学生树立正确的AI使用观念,培养批判性思维与伦理意识,让技术真正服务于人的全面发展。

“以后程序员离开AI助手就不会写代码了。”这句网上的段子,其实并不是危言耸听。观察者指出,有“AI依赖症”的人,短期效率虽然会飙升,但长期来看能力会一步步萎缩。

幻觉”。

另一方面,人们还需要辩证看待“AI幻觉”问题。“在需要丰富想象力与创造性的艺术创作、灵感发散等场景中,较高的幻觉率反而能产出富有新意的内容,带来超出预期的创意效果。”郭艳民表示,公众应当像学习其他专业知识一样,系统性地掌握AI的相关原理与使用规范,让AI更好地发挥其特长,而不是让AI去做它无法做到的事情。

“我们应当深入了解AI技术及各类AI工具,掌握其正确的使用方法。”郭艳民说。

据《北京晚报》

建议

多给老人开设AI科普讲座

面对似是而非、真假难辨的AI信息,年轻人都防不胜防,缺乏信息分辨能力的老年人,该怎么办?令人颇感意外的是,老年人对于“AI幻觉”的感受目前还并不明显。

“这是老年人对AI的了解程度和使用习惯决定的。”科技助老公益组织“夕阳再晨”工作人员介绍。在北京,“夕阳再晨”自2025年初推出AI课程以来,目

前已开展600多场AI科普讲座,受益老年人达7500人次。在大学生志愿者的带领下,一部分老年人开始接触AI,和AI聊天,利用AI助手写诗、处理照片、设计舞蹈动作,日子过得更加有滋有味。

相比年轻人将AI用于学习、工作,老年人则主要将AI助手用于生活娱乐。工作人员接到老年人的反馈大多数是“好用”,了解“AI幻觉”这一概念的老年人非常少。不过,为了帮助老年人提升数字素养,“夕阳再晨”正计划将“AI幻觉”写入科普读物。

“夕阳再晨”创始人、北京邮电大学副教授张佳鑫介绍,从认识AI、使用AI,到理性认识“AI幻觉”,是提升老年人数字素养、跨越数字鸿沟的必要过程。随着技术的进步,数字生活不断出现新的问题,科普内容也需要不断更新。他打了个比方,“夕阳再晨”一开始教老年人看短视频,后来教如何制作,再到后来教防沉迷。相应地,AI课也是这样的过程。“我们需要走在前面,不能等到‘AI幻觉’对老人产生不良影响了,再去做科普工作。”张佳鑫说,“夕阳再晨”未来将走进更多社区,让更多老年人了解“AI