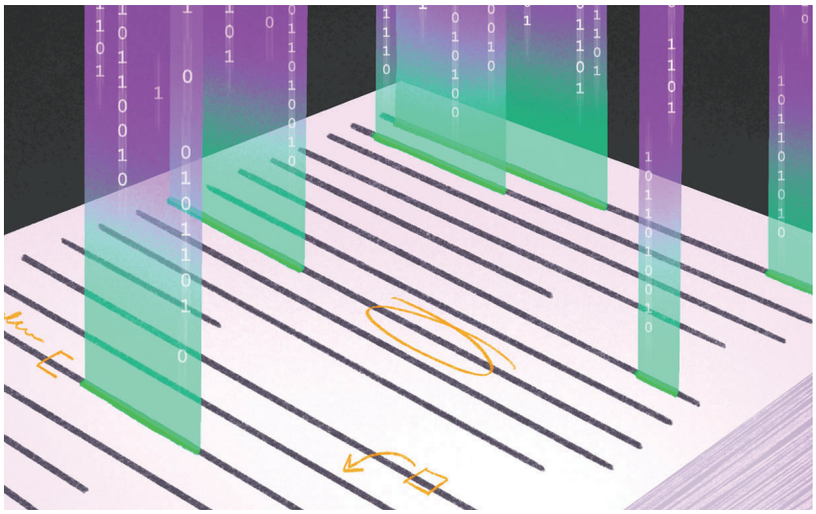
科技新闻

·深度报道·

AI生成的文本在研究论文中呈上升趋势

1/5的计算机科学论文可能包含AI撰写的句子

文/Phie Jacobs



学术论文中或掺杂大量AI撰写的文本

(图片来源:《Science》)

一项跨学科的大规模研究揭示了科学家使用人工智能(AI)撰写手稿的普遍程度,自OpenAI的文本生成聊天机器人ChatGPT横空出世以来,AI辅助论文写作呈稳定增长趋势。在某些领域,此类生成式AI的使用几乎已成为常规操作,高达22%的计算机科学论文显示出大语言模型(LLM)介入的痕迹,这些模型正是相关计算机程序的核心技术基础。

这项发表于《Nature Human Behaviour》的研究,分析了2020—2024年间发表的超百万篇科学论文和预印本,主要通过检测摘要和引言中AI生成文本高频特征词的出现频率变化来追踪趋势。美国路易斯维尔大学的研究素养与传播讲师Alex Glynn评价:“这项研究令人印象深刻”。他发现大语言模型修改的内容在计算机科学等领域更为普遍,或将为检测和规范这类工具的使用提供指引(Glynn未参与该研究)。他补充道:“或许这场讨论需要重点聚焦特定学科领域。”

当ChatGPT于2022年11月首次发布时,许多学术期刊为避免由计算机程序全部或部分撰写的论文泛滥,紧急制定政策限制生成式AI的使用。然而,研究人员和调查团队却逐渐识别出大量明显带有大语言模型辅助撰写痕迹的科学手稿,其中包括“重新生成响应”或“我的知识截止日期”等异常短语。

负责编纂记录学术论文中疑似AI使用案例数据库Academ-AI的Glynn表示:“表面上看,这颇具荒诞意味,但其潜在影响令人深感忧虑。”那些明显由AI生成的论文在经过多轮同行评审和编辑核查后仍得以发表,这暴露了期刊质量把控的漏洞,也体现出大语言模型“AI幻觉”(生成虚假或误导性信息)问题的严重性。

遗憾的是,随着技术的进步以及使用AI的作者更擅长掩盖痕迹,识别AI的“手笔”愈发困难。为此,科学家开始寻找更细微的大语言模型使用迹象。在这项新研究中,美国斯坦福大学的计算生物学家James Zou及其同事从ChatGPT问世前发表的论文中选取段落,使用大语言模型进行摘要总结,继而基于概要生成完整段落,最后用这2类文本来共同训练词频统计模型。该模型学会了根据“关键性”(pivotal)、“复杂性”(intricate)或“彰显”(showcase)等科学写作中非常用词汇的高频使用来识别AI文本特征。

研究人员将该模型应用于2020年1月至2024年9月期间发布于预印本服务器arXiv和bioRxiv及15种《Nature》合作期刊的1121912篇预印本和期刊论文的摘要及引言部分。分析显示,在ChatGPT发布后,经修改的内容出现了急剧增长。James Zou感叹:这一趋势出现得如此之早,

领域/发布平台	预估AI使用率(论文占比/%)
计算机科学	22.5
电气工程与系统科学 (arXiv)	18.0
统计学 (arXiv)	12.9
生命科学(bioRxiv)	10.3
物理学 (arXiv)	9.8
《Nature》合作期刊 (arXiv)	8.9
数学 (arXiv)	7.8

注:分析显示,截至2024年9月(即ChatGPT发布近2年后),科学论文中AI生成文本的数量已呈现激增态势(数据来源:Liang W, Zhang Y, Wu Z, et al. Quantifying large language model usage in scientific papers[J]. Nature Human Behaviour (2025). https://doi. org/10.1038/s41562-025-02273-8)

“意味着该技术在问世之初就被迅速使用了”。

不同学科领域的增长速率存在显著差异,这可能反映了对AI技术熟悉程度的不同。James Zou解释道:“我们发现增幅最大的领域恰恰上是与AI关联最密切的学科”。

截至2024年9月,22.5%的计算机科学论文摘要显示出大语言模型修改痕迹,电子系统与工程科学紧随其后,而数学论文摘要仅占7.7%。生物医学科学和物理学等学科的占比也相对较低,但James Zou指出:大语言模型的使用在所有领域都持续上升,“无论利弊,大语言模型正在成为科学研究过程本身不可或缺的组成部分。”

德国图宾根大学的数据科学家Dmitry Kobak对这项新研究印象深刻,他在《Science Advances》发表的研究表明,2024年发表的生物医学研究摘要中约有1/7可能是在AI的辅助下完成的。Kobak评价:“这项统计建模非常严谨”。

他补充说,AI在科学出版中使用的真实频率可能更高,因为作者或许已经开始从手稿中剔除“危险信号”词语以规避检测。例如,“深入探究”(delve)一词在ChatGPT发布后出现频次激增,却在被识破为AI文本特征后迅速减少。

尽管这项新研究主要针对摘要和引言,但Kobak担忧作者会日益依赖AI撰写论文的文献综述章节。这最终可能导致这些章节内容趋于同质化,并在未来形成一个“恶性循环”,即新型大语言模型在其他大语言模型生成内容上训练。

James Zou和团队正在筹划举办一场由AI完全代理撰稿人和评审者的学术会议,他们希望借此验证这些技术能否独立生成新假设、新技术和新见解。他说:“我预计会有很多相当有趣的发现,以及更多相当有趣的错误。”

(译自《Sciene》,2025,389(6760))

《Science》对微软量子计算研究的更正重新引发讨论

文/Adam Mann

围绕难以捉摸的马约拉纳粒子(曾被寄予构建稳健量子芯片厚望)的争议持续分化整个研究领域。

就像薛定谔那只著名的猫,支撑量子计算机的量子比特也能同时存在于2种状态。量子计算研究的一个特殊分支试图用神秘的马约拉纳准粒子制造量子比特,似乎也以一种双重现实的状态蹒跚前行:既被确信,又遭质疑。

经过多年争议,《Science》对微软赞助的研究人员2020年的一篇论文追加了一则更正。该论文曾宣称在微小的“纳米线”内制备出了马约拉纳粒子,这被视为构建更强健量子计算机的重要一步。但该期刊的决定(遵循一项大学调查建议,但调查未发现科学不端行为)似乎难以平息作者与批评者之间的拉锯战,后者仍以涉嫌筛选数据为由要求撤回该论文。

此项更正同样不大可能平息围绕微软赞助的马约拉纳研究的广泛争议,微软已向该领域投入逾10亿美元资金。2025年2月,