

生成式人工智能技术下的学术生产与出版： 变革、失范与路径^{*}

蒋雪颖 刘欣

(北京大学新闻与传播学院, 北京 100089)

摘要: 聚焦以ChatGPT为代表的生成式人工智能技术在学术生产和学术期刊出版领域的应用现状、存在的风险与挑战, 并以此为基点, 探寻技术的应用与失范之间的可能通路。首先, 在技术层面对人工智能融入学术生产的历史嬗变进行勾勒; 其次, 在实践层面总结并分析目前欧美学界对生成式人工智能技术应用于学术生产和学术期刊出版的讨论; 再次, 梳理国外代表性学术出版机构对人工智能生成内容的应对措施; 最后, 从建构多元主体协同监管的学术出版生态框架出发, 尝试为生成式人工智能技术下我国学术生态的良性发展提供可能的进路。

关键词: 生成式人工智能; 人工智能生成内容; 学术生产; 学术期刊出版; 技术失范

中图分类号: G230.7 **DOI:** 10.3772/j.issn.1673-2286.2023.05.008

引文格式: 蒋雪颖, 刘欣. 生成式人工智能技术下的学术生产与出版: 变革、失范与路径[J]. 数字图书馆论坛, 2023 (5): 64-71.

2023年伊始, 生成式人工智能ChatGPT以其强大的信息整合能力和自然语言处理能力吸引了全球的目光, 也引发了人们对此项技术应用下的人工智能生成内容 (Artificial Intelligence Generated Content, AIGC) 的热议。讨论聚焦于学术生产和学术期刊出版领域。OpenAI公司的公开资料显示, 在学术生产方面, 许多研究者已经使用ChatGPT和其他大型语言模型 (large language model) 来撰写小论文 (essay)、文献综述 (literature review)^[1], 对论文进行降重等基础性的学术生产工作。除此之外, ChatGPT还可编写计算机代码以及统计分析。这项技术将很快发展到可以完成设计科学实验、撰写基础实验报告等研究工作。在学术期刊出版方面, ChatGPT能够通过自动化编纂内容, 对文章进行智能润色、加工, 辅助和优化编辑校对工作。但其中的风险不容小觑, 人工智能可能带来技术伦理和学术伦理失范, 降低研究的质量与透明度, 以隐匿的方式削

弱人类对知识的自主权。

现有关于学术生产和出版中生成式人工智能的研究仍处于早期阶段, 许多研究集中探讨理论问题, 如版权归属、技术伦理、人机共生等。方卿等^[2]提出要关注生成式人工智能引发的著作权问题、技术伦理问题、文化安全问题。陆小华^[3]探讨了生成式人工智能对知识萃取、生产、传播方面的潜在影响。肖峰^[4]比较了ChatGPT和“文心一言”关于数字劳动的回答, 评估了生成式人工智能辅助学术研究的优势和劣势。骆飞等^[5]讨论了生成式人工智能对学术生产、学术评价、学术传播的影响。

现有研究尚未系统分析AIGC到底在学术生产和出版方面带来了何种风险和挑战, 缺乏国际案例研究以及对国内出版机构应对AIGC的明确建议。鉴于此, 本文从实践角度, 聚焦以ChatGPT为代表的生成式人工智能技术在学术生产和学术期刊出版领域的应用现

收稿日期: 2023-04-26

^{*}本研究得到国家社会科学基金后期资助项目“国家修辞·话语·传播研究” (编号: 18FYY023) 资助。

状、存在的风险与挑战,并以此为基点,探寻技术的应用与失范之间的可能通路。研究分为4个部分:首先,在技术层面对人工智能融入学术生产的历史嬗变进行勾勒;其次,在实践层面总结并分析目前欧美学界对生成式人工智能技术应用于学术生产和学术期刊出版的讨论;再次,梳理国外代表性学术出版机构对生成式人工智能应用的态度和应对措施;最后,为我国出版机构提供具体建议,尝试为生成式人工智能技术下我国学术生产生态的良性发展提供可能的路径。

1 技术祛魅:人工智能技术融入学术生产的历史嬗变

1.1 学术生产“主体—文本”视角下的文本迭代进程

广义的学术生产是指,人类群体中的一小部分经过长期而严格的专门训练的“学者”,以职业化的方式从事针对自然、社会,系统化和规范化的研究行为。其研究成果主要通过学术期刊、书籍来发表和传播,并通过高等教育体系来传承;通过与学术“场域”之外的政治、经济、社会、文化的复杂互动,来改变人类世界的表面和深层结构^[6]。狭义的学术生产特指学者通过学术写作行为,将知识以文本的形式固定在介质之上的过程。在这个意义上,通过以“主体—文本”视角考察,可将人工智能技术融入学术生产的进程划分为:人力文本时期、人工智能辅助人力文本时期、人工智能生成文本时期。3个进程并不是后者取代前者的关系,而是相伴协同共生的关系(见图1)。

人力文本时期即传统的学者作为主体进行文本写作的时期。21世纪以来,全球化不断加速,跨语言的学术写作日渐成为研究者和高校学生学术生活的一部

分。2009年前后,以Grammarly为代表的人工智能语言工具开始进入市场。随后,Shortly AI、Jarvis等人工智能语言工具如雨后春笋般被投入使用。它们主要服务于非英语母语的研究人员或有学习障碍(书写障碍)的作者,帮助其修正语法、润色语句,从而提高英文文章的可读性。截至2020年,Grammarly日活跃用户数量达3 000万人次。在技术与资本的联合作用下,学术生产逐渐从人力文本时期过渡到人工智能辅助人力文本时期。这一时期,人工智能技术只作为人类学术生产的辅助工具,在语法、拼写等层面对文本产生修饰性作用。

21世纪10年代中期至今,随着以生成式对抗网络(Generative Adversarial Network)为代表的深度学习算法的提出和迭代更新,人工智能技术跨入新时代,随之而来的是AIGC在语言效果上的逼真化^[7]。当前代表性的生成式人工智能技术语言模型有OpenAI开发的GPT-4、谷歌开发的BERT、微软开发的XLNet等。其中,基于GPT-3.5架构的自然语言处理工具ChatGPT从2022年11月首次发布以来获得了大量关注,在2023年1月的月活跃用户数量已达1亿人次,成为史上用户数量增长最快的应用^[8]。以ChatGPT为代表的生成式人工智能在学术生产中的应用可能改变文本的生产主体和组织逻辑^[7],学术生产跨入AIGC时期。

1.2 生成式人工智能在学术生产中的应用现状

生成式人工智能作为一项前沿数字技术,正在驱动学术生产方式的变革。而ChatGPT超越了表层文本,进入AIGC领域,并在学术生产方面展现出强大的功能。

(1) 以ChatGPT为代表的生成式人工智能可以生成论文思路^[9]。研究者输入论文主题(theme),ChatGPT便可以此生成论文标题、主要观点见解、可供使用的案例等内容。

(2) 创建论文大纲。ChatGPT能够生成论文的引言、文献综述、研究方法、结论等的大纲,并在每个部分提供论文要点。除此之外,ChatGPT还可以根据指令需求,控制论文的长度。

(3) 进行文献综述辅助工作。ChatGPT可以根据给定的主题或关键词(keyword)生成论文摘要,或为研究者提供相关论文列表,从而帮助研究人员快速收集相关文献。

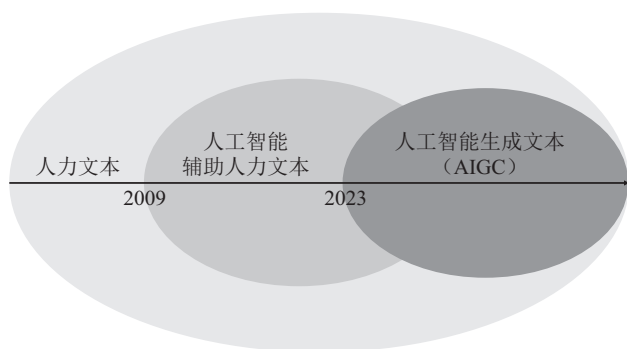


图1 人工智能技术变迁下学术生产文本迭代进程

(4) 对文本进行翻译和润色。ChatGPT可以模拟指定期刊的论文风格, 提供更有逻辑性、更符合学术规范的文字表达。

(5) 编辑论文。研究者可以要求ChatGPT编辑论文结构、语法、修辞、拼写等, 还可以与其共同编辑论文, 进一步要求它重写特定段落或句子。

(6) 自动生成总结(summary)。ChatGPT可用于自动总结学术论文、报告等, 使研究人员快速了解和把握文本内容。

(7) 学术问答。ChatGPT可以提供特定领域问题的答案, 成为高效的学术搜索工具。

在学术期刊出版方面, ChatGPT可以辅助同行评议过程, 编辑和审稿人可以通过它审查写作质量、格式、学术诚信等。在筛选论文时, ChatGPT也可向编辑提供快速推荐意见。

针对ChatGPT在学术生产等领域展现的强大功能和潜在可能性, 图灵奖得主、数据库专家吉姆·格雷依从时间逻辑和实践逻辑, 总结人类从事科研的4种范式: 第一范式为实验(经验)科学, 第二范式为理论科学, 第三范式为计算科学, 第四范式为数据科学。以ChatGPT为代表的生成式人工智能或将凭借其敏锐捕捉自然语言文本之间相关关系的优势, 促使科研从前3种范式进一步向第四范式乃至更高范式过渡^[10]。

ChatGPT具有易用性和可访问性, 可大大增加学术产出。它能够为用户提供反馈和指导等个性化和交互式体验, 从而增强研究者的自主性和参与度; 能够为用户节约时间, 提升学习和研究的效率; 使知识传播民主化, 从而形成“以人为本的伙伴关系模式, 人与人工智能共同努力以提高认知能力”^[10]。

2 生成式人工智能技术用于学术生产的失范风险

2.1 潜在的技术伦理失范

语言是人类沟通、交流的重要载体, 而书面语言(如知识、法律法规、新闻、日常讯息等)进一步参与人类社会生活。机器参与人类语言的创造和交流, 引发语言生成和人类感知方式的改变, 可能会产生深远的社会和伦理影响。当前对于以ChatGPT为代表的生成式人工智能或AIGC伦理的讨论集中在以下4个方面。

首先, 将人类道德与法律纳入生成式人工智能或

AIGC的可能性。以ChatGPT为代表的生成式人工智能目前还无法“理解”处于流变中的且高度复杂化的人类道德与法律问题, 所以它无法根据道德或法律原则作出判断或决策。

其次, 生成式人工智能的透明度和解释性问题。与许多人工智能系统一样, ChatGPT如同“黑匣子”般缺乏透明度。Azaria^[11]曾质疑ChatGPT, 理由是其无法解释为什么它报告“ $3.2^{3.3} \approx 21.73$ ”, 而实际答案是46.45。这种算法黑箱超越了大众理解的范畴, 导致大众完全置身于算法“控制”的被动状态, 这更深层次地提升了信息不平等、社会控制等隐匿在技术背后的失范风险。

再次, AIGC背后的数据偏见。麻省理工媒体实验室(MIT Media Lab)的计算机学者Joy Buolamwini于2018年提出算法偏见(AI Bias)的概念, 指出人工智能以互联网中的大数据为深度学习的对象, 经过机器学习后在生成结果上反映出存在于网络言论中的既有偏见, 如性别不平等、肤色歧视等问题^[12]。“垃圾进, 垃圾出”(garbage in, garbage out)仍然体现在当下的AIGC实践中。模型的反馈可能反映训练数据中存在的既有偏差, 这可能导致不准确或不公平的结果, 也可能出现无意的错误信息, 进一步加剧社会不公平等问题。加州大学伯克利分校心理学和神经科学教授Steven Piantadosi发现, 当ChatGPT被问及基于种族和性别考量, 谁可以成为一名优秀的科学家时, 其回答倾向白人和男性, 表现出对女性和非白人的偏见^[13]。

最后, AIGC中的虚假信息。OpenAI承认, ChatGPT“对2021年之后的世界大事了解有限”, 如果没有足够的可用数据, 它很容易编造信息。美国媒体网站CNET(Computer Network)和其风险投资公司Red Ventures已采用生成式人工智能撰写新闻, 同时由编辑进行事实核查(fact check), 即使如此, AIGC的准确性和真实性仍然有待提升^[14]。

2.2 潜在的学术伦理失范

2.2.1 技术规训下从知识原创性危机到思维危机

研究者通过分析、解释与批评的方法, 在文献中阐明想法(ideas)、得出结论, 并在研究特定实例时设计和使用这些方法。研究者需要完整展示上述学术工作, 并通过在学术共同体中的不断对话来推进学术研究。Bender等^[15]认为以ChatGPT为代表的大型语言模型如同

“随机鹦鹉”(stochastic parrots), 它无法理解也不产生知识, 只是从痕迹数据(data shadow)中组合文本, 这些痕迹数据本身并不构成有意义的、科学有效的、合法的知识, 因为没有具备知识的主体来解释和验证它们。虽然ChatGPT能够模仿人类写作, 但它仍然无法完全复制人类知识和其中的关联。有学者指出, ChatGPT的迷惑性“神话”不是源自它的技术实力, 而是因为它能够使用看似真实的单词或短语来说服读者^[16]。

知识原创性的降低会直接影响学术研究成果的质量, 更可能引发低质量学术论文甚至错误论述的恶性循环。本质上ChatGPT是一个文本预测器: 它“学习”文本片段之间的关系, 然后对内容和其中的逻辑关系进行重组, 这个过程会产生看似可信的书面文本, 但产生的文本不一定基于事实。Blanco-Gonzalez等^[17]发现, ChatGPT会生产一些不符合学术规范的虚假内容, 甚至编造不存在的参考文献。在直接基于ChatGPT获得的初步文本中, 正确引用的内容占比仅为6%。

此外, 主体可能产生技术依赖, 福柯口中的“规训”和麦克卢汉笔下的“延伸意味着截除”将成为现实——主体批判性思维的萎缩。若不同的研究者对ChatGPT提出相同或相似的任务, 如撰写特定主题的文献综述, 那么ChatGPT很有可能生成相同或相似的结果, 无法创造新的学术贡献。如果研究者不主动筛选、阅读文献, 依赖ChatGPT获得文献综述, 那么其成果贡献价值和真实性存疑。对于刚进入研究阶段的研究生来说, ChatGPT可能会对他们接受信息、回溯经验、诉诸理论、提出假设、实践调研、检验假设、推导结论等研究过程产生负面影响。综上, 研究人员需正视ChatGPT的技术逻辑: 它直接依赖既有数据信息, 缺乏人类具备的权衡价值和利用感官经验在当前背景下作出技术和科学决策的能力。对生成式人工智能应时刻抱有批判性的视角, 了解系统如何生成决策或预测。除此之外, 也应警惕在研究中过度依赖生成式人工智能, 这关乎人类能动性和自由意志, 过度依赖生成式人工智能可能损害人类认知能力。人类的能动性、创造力和独创性以及学术共同体的互动仍然是研究创新的关键。

2.2.2 加剧学术不端问题和知识产权侵权问题

生成式人工智能可能加剧学术不端问题。据英国广播公司(BBC)报道, 在欧洲已有大量大学生在撰写作业、学术论文和学位论文时使用ChatGPT, 许多大学

警告学生不要使用ChatGPT完成作业^[18]。尽管目前已有软件可以检测AIGC, 但是随着生成式人工智能技术的不断迭代, 验证论文真实性和原创性的成本也会增加。Khalil等^[19]要求ChatGPT根据不同的开放式问题生成50篇论文, 对一半的论文使用Turnitin软件进行检查, 其平均相似度为13.72%, 对另一半的论文使用iThenticate平台进行检查, 其平均相似度为8.76%。也就是说, ChatGPT生成的论文可能会被误判为是高度原创的。少数研究者使用ChatGPT来撰写内容, 但未向出版商、期刊编辑或学术会议组织言明。这种隐瞒是否符合学术伦理、未来是否会存在法律上的风险等问题仍有待商榷。生成式人工智能无法对其生成的内容承担责任, 这对学术标准和完整性维护提出了重大挑战。从传统学术生产视角来看, 这可能会引发学术不端问题。但从实践的趋势上审视, 生成式人工智能应用趋势似乎无法阻挡, 那么学术共同体基于新的现实会形成怎样的伦理规范, 这有待追问和关注。

生成式人工智能引发了2个主要的版权问题。第一, 是否应该允许ChatGPT使用第三方生成的原创内容来生成内容。OpenAI声明, 不为任何可能发生的知识产权侵权行为承担法律责任, 研究者应有效避免由使用ChatGPT造成的潜在知识产权侵权问题。在未经版权所有人许可的情况下, 若在ChatGPT中输入大量受版权保护的文本, 也可能构成侵权。第二, 生成式人工智能是否可被视为作者。尽管已经有论文将ChatGPT列为合著者, 但目前大多数出版机构并没有在著作权层面承认其作为学术出版物合著者的资格。

2.2.3 加剧全球范围内的技术鸿沟与学术不平等

生成式人工智能为知识传播和生产提供了前所未有的机会。然而, 与任何新技术一样, ChatGPT的使用主要局限于拥有大量经济和技术资源的国家和机构, 这有可能加剧发达国家与发展中国家之间现有的经济、技术和知识不平等现象。目前ChatGPT免费对用户开放使用, 但是搭载GPT-4技术的ChatGPT plus每月向用户收费20美元。超大模型、超强算力的背后是高资本投入。彭博社报道称, 微软已花费数亿美元开发基础设施以支持ChatGPT和Bing Chat等项目^[20]。据估算, GPT-3训练一次的成本约为140万美元, 对于一些更大的大型语言模型, 训练成本为200万~1 200万美元^[21]。综合硬件成本、数据中心成本、人力成本, 在10天内训

练一个千亿规模的大型语言模型大约需要花费1.43亿美元^[22]。发达国家通常在开发和实施生成式人工智能技术方面处于领先地位。爱思唯尔、施普林格·自然和威立等主要学术出版商在开发人工智能驱动的出版工具方面已经进行了大量投资,目的是实现学术出版的自动化、数字化、智能化。

ChatGPT引发了人工智能芯片的“军备竞赛”,经济成本拉开了技术使用的鸿沟,低收入和中等收入国家的研究者、科研机构和出版机构可能负担不起相应的费用。生成式人工智能驱动的工具和平台对算力、人才、数据有极高的需求,不仅需要大量的硬件和软件投资,还需要用于培训和支持人才的费用以及持续的维护费用,这对许多低收入和中等收入国家来说可能是过于昂贵的。因此,发达国家与发展中国家在学术生产和出版方面的差距可能扩大,发达国家在学术生产和出版中的领先和垄断地位难以打破,全球范围内的学术不平等现象将加剧。

除此之外,ChatGPT依据国别设置访问权限,开展干预行为,这进一步扩大了技术鸿沟。尽管全球技术扩散使得数字鸿沟在数量方面有所消弭,但在质量方面鸿沟则进一步扩大,同时发展中国家内部的技术差距也日益拉大。如何确保全球范围内学术研究中技术使用的公平性也是一个日益显现的议题。

3 应对技术失范:国外学术出版领域对生成式人工智能的态度与应对方法

不同于抱有乐观态度的商业界,国外学术出版领域对于生成式人工智能的总体态度倾向于担忧与谨慎,重点讨论了有关版权、署名、剽窃的伦理问题。有出版商表示无须对AIGC过度恐慌,认为生成式人工智能可以提高编辑和出版的效率。不过也有人提出了反对的意见,要求完全禁止使用生成式人工智能。有期刊编辑认为,学术出版商应联合起来,在学术生产的任何环节都禁止使用ChatGPT或其他生成式人工智能工具,这种观点主张将使用相关的作者列入学术出版行业黑名单^[23]。他们还提出了严格的措施以限制AIGC,如增加AIGC自动检测功能、作者签署诚信声明承诺未使用ChatGPT等工具。如遇到违规使用AIGC的情况,行业协会和学术出版商有权在作者唯一标识符(如ORCID)上注明或记录。一些期刊编辑呼吁对AIGC进行更严格的审查,利用剽窃检查、数据验证和作者记录

等证据检查论文的原创性。

在伦理倡议方面,国外学术出版协会发布了关于生成式人工智能的伦理建议书,如国际出版伦理委员会(Committee on Publication Ethics)和国际科学、技术和医学出版商协会(International Association of Scientific, Tech, and Medical Publishers),其核心观点为人工智能工具是非法律实体:首先,它无法披露与论文直接或间接相关的利益,如研究资助或基金项目、研究机构从属关系、非经济利益的写作指导关系或编委会人际关系;其次,它也不能管理版权和许可协议,无法对提交的学术作品负责。基于以上两点,它无法满足“作者身份”这一学术生产的基本要求^[24]。国外学术出版协会进一步规定,如果作者在撰写论文、制作论文图表、收集和分析数据时使用了人工智能工具,则必须在论文中阐明所使用的人工智能工具名称、版本以及使用方法。在这种情况下,作者仍对其稿件的内容(包括AIGC部分)负全部责任,即为任何违反学术和出版道德的行为负责^[25]。

在出版政策方面,各大学术出版集团和期刊及时更新了关于使用生成式人工智能的规定。泰勒弗朗西斯集团、施普林格·自然集团、爱思唯尔集团、爱墨瑞得集团等出版集团明确指出,包括ChatGPT在内的任何人工智能大型语言模型工具都不符合作者资格标准。作者不应将人工智能及其辅助技术列为作者或合著者,作者身份只能归于人类,其责任和义务必须由人类执行。这些出版机构还指出,作者必须明确其作品涉及的所有来源和贡献者,必须承认并记录使用人工智能的部分,坚守学术诚信。爱思唯尔集团的出版政策相对宽松,它允许使用人工智能检查语法和润色语言,但不能由人工智能完成关键的研究任务,如产生科学见解、分析和解释数据或得出科学结论^[26]。它倡导生成式人工智能应该在人为监督和控制下应用,作者应该仔细审查和编辑结果,因为人工智能可以生成看似很权威的内容,但这些内容可能不正确、不完整或包含偏见。《科学》(Science)规定未经编辑明确许可,不得将由人工智能或类似算法工具生成的文本用于在该期刊上发表的论文,所附数字、图像等也不得是此类工具的产物,违反此政策的作者构成学术不端行为^[27]。《自然》(Nature)的一份报告也显示,由于生成式人工智能“无法对科学论文的内容和完整性负责”,因此不能被视为作出了学术贡献^[28]。《柳叶刀》(The Lancet)声明,未来将围绕AIGC在学术出版中的影响展开严格的

交流,以期创建全面的指导规范^[29]。

4 协同共生: 我国学术期刊出版机构的应对之道

学术期刊出版作为出版活动的重要组成部分,长期以来发挥着意识形态价值、教育价值、文化价值和科学价值等,对维护和形塑国家和社会主流文化具有重要作用^[2]。由于技术开放性和流动性,AIGC对我国的学术和文化安全提出了新的挑战和要求。

当下,许多国家走上探索人工智能科技伦理与立法的道路,为未来全球治理格局下共同建构生成式人工智能的全球监管模式奠定基础。2023年4月11日,国家互联网信息办公室起草《生成式人工智能服务管理办法(征求意见稿)》,从生成式人工智能服务商的准入条件,到算法设计、训练数据选择、内容生成以及个人隐私、商业秘密保护等方面提出了相关要求。处于学术生产和传播前沿阵地的学术出版机构应积极参与生成式人工智能管理办法的协商进程,为建构多元主体协同监管的学术出版生态框架提供行业经验与建议。

4.1 人工智能检测辅助人工验证

目前学术出版领域的共识是人类作者对任何版权侵权或者学术偏差等行为负责。在实践中,出版机构应采用人工智能检测辅助人工验证,从而确认是否存在学术错误、抄袭或洗稿等现象。目前已有人工智能工具(如GPTZero)能够识别AIGC,一些论文检测产品也针对人工智能写作更新了检测功能,比如CrossRe、Turnitin、iThenticate已经更新了人工智能写作检测功能。随着训练集优化等,识别AIGC的难度可能提升。还有学者指出这些检测工具可能会误判人类写作文本或者经过英文校对软件润色的文本为AIGC,而且暂时没有工具可以检测论文的图像和数据部分^[30]。

在未来,出版机构不仅需要在技术层面更新迭代检测深度伪造文本的方法,还需要依靠专家进行事实核查。人工智能检测工具只能帮助编辑初步评估论文的原发性,编辑仍然需要自主判断和验证论文的核心观点。

4.2 规范生成式人工智能的学术使用

出版机构应建立相应规范,倡导作者以负责、透明

和诚实的方式使用生成式人工智能。出版机构应该明确论文包含AIGC的限度,明确AIGC剽窃的界限。出版机构可以要求作者提交有关论文是否包含AIGC的检测报告。如果作者使用了生成式人工智能,出版机构应该要求作者披露详细使用方法,即使仅用生成式人工智能来修订错别字和语病。作者应明确而具体地说明论文哪些部分、在多大程度上涉及AIGC,还应说明使用了哪些大型语言模型。这些声明将提醒编辑和审稿人更仔细地审查论文,从而发现潜在的偏见、不准确的数据和来源不当的参考文献。而且,由于生成式人工智能日新月异,出版机构还应该要求作者详细说明使用的人工智能工具的名称、版本、型号和来源。现代语言协会(MLA)^[31]和美国心理学协会(APA)^[32]、《芝加哥手册》(*The Chicago Manual of Style*)^[33]均已更新引用AIGC的规范。

生成式人工智能应用中还存在模糊地带,因为生成式人工智能可能会被纳入文本处理和编辑工具、搜索引擎和编程工具,比如作者使用搭载GPT-4的必应(Bing)搜索相关案例作为论据,但是作者可能无法确定生成式人工智能贡献的性质或规模。面对这种情况,出版机构必须及时调整并适应新情况,确保使用AIGC的完全透明。

4.3 建立AIGC学术问责制度

出版机构应该建立相应的学术问责制,要求作者对AIGC负责。编辑需要核查论文所有引用材料、案例、文献的来源。出版机构应该要求作者承诺论文涉及的AIGC部分没有抄袭和洗稿的情况,需要确保AIGC是准确和真实的。由于生成式人工智能可能包含偏见,会省略对某个观点的反对意见,出版机构应该要求作者核查观点的不同来源,确保事实的全面性。总之,作者必须对在生成式人工智能帮助下撰写的论文负责。

生成式人工智能可以成为一种有用的辅助或自适应技术,关键在于研究者的责任感和使用的明确范围。作者无论使用何种人工智能工具来辅助科学写作,基本原则是他们必须对作品及其原创性负责。国内已经有接受人工智能参与写作的期刊。哈尔滨工程大学主办的《智能社会研究》已开设“AI+科研”试验栏目,并正式接收由人工智能参与研究工作的论文,要求作者提交“AI训练笔记”,详细说明在研究全过程中是如何使用、训练人工智能的。该期刊规定人工智能工具不得参

与论文署名,论文实名作者对论文的真实性、科学性负全部责任。

4.4 建构多元主体共同监管的技术框架

生成式人工智能通过搜集大量互联网信息来训练其大型语言模型,可能会导致不准确或者虚假信息的传播,因此有必要监管和验证用于训练语言模型的信息及输出内容。生成式人工智能的设计、训练和部署应该是透明的,包括用于训练模型的数据,以便识别和解决任何潜在的偏差或错误。它的设计和开发必须以透明、可解释和负责的方式进行。

为了应对人工智能设计和开发不透明的现象,我国政府可牵头组织大学、非政府组织、研究机构、学术出版机构等多元主体共同开发开源、透明的生成式人工智能系统,从而最大限度地减少偏见、剽窃、版权侵权等问题。开源人工智能技术应确保大型语言模型和语料库等档案的完整性,以便产生准确和全面的结果,保证结果的可追溯性和可靠性。

参考文献

- [1] ADESSO G. GPT4: the ultimate brain [EB/OL]. [2023-02-21]. <https://www.authorea.com/users/564304/articles/611211-towards-the-ultimate-brain-exploring-scientific-discovery-with-chatgpt-ai?commit=887ec891690a225ae4309c66c6dbbc90a7a35ec9>.
- [2] 方卿,丁靖佳. 人工智能生成内容(AIGC)的三个出版学议题[J]. 出版科学, 2023, 31(2): 5-10.
- [3] 陆小华. ChatGPT等智能内容生成与新闻出版业面临的智能变革[J]. 中国出版, 2023(5): 8-15.
- [4] 肖峰. 生成式人工智能介入知识生产的功用探析: 借助ChatGPT和“文心一言”探究数字劳动的体验[J/OL]. 重庆邮电大学学报(社会科学版): 1-11 [2023-05-08]. <http://kns.cnki.net/kcms/detail/50.1180.C.20230428.1850.002.html>.
- [5] 骆飞,马雨璇. 人工智能生成内容(AIGC)对学术生态的影响与应对: 基于ChatGPT的讨论与分析[J/OL]. 现代教育技术: 1-12 [2023-05-08]. <http://kns.cnki.net/kcms/detail/11.4525.n.20230322.1635.002.html>.
- [6] 孟建,胡学峰. 数字人文: 媒介驱动的学术生产方式变革[J]. 现代传播(中国传媒大学学报), 2019, 41(4): 24-28, 54.
- [7] 中国信息通信研究院, 京东探索研究院. 人工智能生成内容(AIGC)白皮书[R]. 北京: 中国信息通信研究院, 2022.
- [8] business of Apps. ChatGPT Revenue and Usage Statistics (2023) [EB/OL]. [2023-04-11]. <https://www.businessofapps.com/data/chatgpt-statistics/>.
- [9] BOUSCHERY S G, BLAZEVIC V, PILLER F T. Augmenting human innovation teams with artificial intelligence: exploring transformer-based language models [J]. Journal of Product Innovation Management, 2023, 40(2): 139-153.
- [10] 蒲清平, 向往. 生成式人工智能: ChatGPT的变革影响, 风险挑战及应对策略[J/OL]. 重庆大学学报(社会科学版), 2023: 1-13 [2023-04-13]. <https://kns.cnki.net/kcms/detail/50.1023.C.20230412.1004.002.html>.
- [11] AZARIA A. ChatGPT usage and limitations [EB/OL]. [2023-01-02]. <https://osf.io/5ue7n/>.
- [12] MIT News. Study finds gender and skin-type bias in commercial artificial-intelligence systems [EB/OL]. [2023-04-13]. <https://news.mit.edu/2018/study-finds-gender-skin-type-bias-artificial-intelligence-systems-0212>.
- [13] AGOMUOH F. The 6 biggest problems with ChatGPT right now [EB/OL]. [2023-04-13]. <https://www.digitaltrends.com/computing/the-6-biggest-problems-with-chatgpt-right-now/>.
- [14] ROACH J. ChatGPT agrees: it shouldn't write the articles you read [EB/OL]. [2023-04-13]. <https://www.digitaltrends.com/computing/chatgpt-says-it-shouldnt-write-articles/>.
- [15] BENDER E M, GEBRU T, MCMILLAN-MAJOR A, et al. On the dangers of stochastic parrots: can language models be too big? [C] //Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, 2021: 610-623.
- [16] I get your excitement about ChatGPT, but ... [EB/OL]. [2023-04-13]. <https://www.hendrik-erz.de/post/i-get-your-excitement-about-chatgpt-but>.
- [17] BLANCO-GONZALEZ A, CABEZON A, SECO-GONZALEZ A, et al. The role of AI in drug discovery: challenges, opportunities, and strategies [EB/OL]. [2023-04-13]. <https://arxiv.org/abs/2212.08104>.
- [18] BBC NEWS. Universities warn against using ChatGPT for assignments [EB/OL]. [2023-04-17]. <https://www.bbc.com/news/uk-england-bristol-64785020>.
- [19] KHALIL M, ER E K. Will ChatGPT get you caught? rethinking of plagiarism detection [EB/OL]. [2023-04-13]. <https://arxiv.org/abs/2302.04335>.
- [20] Bloomberg. Microsoft strung together tens of thousands of

- chips in a pricey supercomputer for OpenAI [EB/OL]. [2023-05-08]. <https://www.bloomberg.com/news/articles/2023-03-13/microsoft-built-an-expensive-supercomputer-to-power-openai-s-chatgpt?sref=ExbtjcSG&leadSource=uverify%20wall>.
- [21] 国盛证券. ChatGPT需要多少算力 [EB/OL]. [2023-05-08]. <https://finance.sina.com.cn/stock/stockzmt/2023-02-13/doc-imyfpfsz2563335.shtml>.
- [22] 澎湃新闻. 算一笔细账, ChatGPT、文心一言这类大模型到底要消耗多少GPU? [EB/OL]. [2023-05-08]. https://m.thepaper.cn/newsDetail_forward_22716419.
- [23] DWIVEDI Y K, KSHETRI N, HUGHES L, et al. "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy[J]. International Journal of Information Management, 2023, 71: 102642.
- [24] Committee on Publication Ethics. Authorship and AI tools: COPE position statement [EB/OL]. [2023-04-13]. <https://publicationethics.org/cope-position-statements/ai-author>.
- [25] VAN ROSSUM J, RUSSO C, SCHOENENBERGER H, et al. AI ethics in scholarly communication: STM best practice principles for ethical, trustworthy and human-centric AI [R]. The Hague: International Association of Scientific, Technical, and Medical Publishers, 2021.
- [26] Elsevier. The use of AI and AI-assisted writing technologies in scientific writing [EB/OL]. [2023-04-18]. <https://www.elsevier.com/about/policies/publishing-ethics/the-use-of-ai-and-ai-assisted-writing-technologies-in-scientific-writing>.
- [27] Science. Editorial policies [EB/OL]. [2023-04-20]. <https://www.science.org/content/page/science-journals-editorial-policies>.
- [28] STOKEL-WALKER C. AI bot ChatGPT writes smart essays-should professors worry? [EB/OL]. [2023-04-13]. <https://www.nature.com/articles/d41586-022-04397-7>.
- [29] LIEBRENZ M, SCHLEIFER R, BUADZE A, et al. Generating scholarly content with ChatGPT: ethical challenges for medical publishing [J]. The Lancet. Digital Health, 2023, 5 (3): e105-e106.
- [30] RAHIMI F, ABADI A T B. ChatGPT and publication ethics [J]. Archives of Medical Research, 2023, 54 (3): 272-274.
- [31] MLA Style Center. How do I cite generative AI in MLA style? [EB/OL]. [2023-05-09]. <https://style.mla.org/citing-generative-ai/>.
- [32] APA STYLE. How to cite ChatGPT [EB/OL]. [2023-05-09]. <https://apastyle.apa.org/blog/how-to-cite-chatgpt>.
- [33] The Chicago Manual of Style Online. Citation, documentation of sources [EB/OL]. [2023-05-09]. <https://www.chicagomanualofstyle.org/qanda/data/faq/topics/Documentation/faq0422.html>.

作者简介

蒋雪颖, 女, 博士研究生, 研究方向: 数字出版、国际传播, E-mail: 448255536@qq.com。

刘欣, 女, 博士研究生, 研究方向: 媒介与社会变迁、数智文化。

Scholarly Production and Publishing Under Generative AI: Changes, Disorder, and Path

JIANG XueYing LIU Xin

(School of Journalism and Communication, Peking University, Beijing 100089, P. R. China)

Abstract: This paper focuses on the status, risks, and challenges of generative AI represented by ChatGPT in the field of academic production and academic journal publishing. We hope to explore the possible pathways between technology and academic ethical failures. First, this paper reviews the history of the integration of generative AI into academic production. Second, this paper summarizes and analyzes the current discussions on the application of generative AI to academic production and academic journal publishing in foreign academia. Third, this paper summarizes the responses of representative foreign academic publishing institutions to AI generated content. Finally, this paper constructs a framework of academic publishing ecology with collaborative supervision of multiple subjects. Thus, it provides a possible way forward for the benign development of academic ecology in China under generative AI technology.

Keywords: Generative Artificial Intelligence; Artificial Intelligence Generated Content; Academic Production; Academic Journal Publishing; Technological Failure

(责任编辑: 雷雪)