

Doi: 10.11835/j.issn.1008-5831.pj.2023.02.003

欢迎按以下格式引用:蒋华林.人工智能聊天机器人对科研成果与人才评价的影响研究——基于 ChatGPT、Microsoft Bing 视角分析[J].重庆大学学报(社会科学版),2023(2):97-110. Doi:10.11835/j.issn.1008-5831.pj.2023.02.003.



Citation Format: JIANG Hualin. Research on the Influence of AIChatbot on the Evaluation of Scientific Research Achievements and Talent Evaluation: Analysis based on ChatGPT and Microsoft Bing[J]. Journal of Chongqing University (Social Science Edition), 2023(2):97-110. Doi:10.11835/j.issn.1008-5831.pj.2023.02.003.

人工智能聊天机器人对 科研成果与人才评价的影响研究 ——基于 ChatGPT、Microsoft Bing 视角分析

蒋华林

(重庆大学 公共管理学院,重庆 400044)

摘要:美国开放人工智能研究中心(OpenAI)研发的人工智能聊天机器人 ChatGPT,基于“生成预训练转换器”(GPT)语言模型能够理解和响应自然语言输入,于2022年11月30日正式发布,短时间内在全球互联网界、学术界及产业界引发热议和关注。以 ChatGPT、Microsoft Bing 为代表的人工智能聊天机器人基于超大海量数据,采用“从人类反馈中强化学习”(RLHF)进行训练,主要具有三大功能:智能交互对话、智能文本生成、智能文献爬取。人工智能聊天机器人的核心竞争力是采用人工智能生成内容技术(AIGC)而具有强大的内容生成或文本生成能力,除能够生成一般应用性文稿外,还能够从事科学写作,从而引领和推动知识生产方式和学术研究范式变革,知识生产和学术研究进程将显著提速,知识生产者将从非创造性劳动中获得解放,知识生产方式将由“手动式”“半自动式”迈向“自动化”“智能化”。这种变革对传统的科研成果评价机制将产生重大影响,更进一步将对通过科研成果进行人才评价产生重大影响。大体上说,以 ChatGPT 为代表的人工智能聊天机器人应用于学术研究中,将遮蔽成果演化学术史,影响科研成果创新性评价;挑战知识产权保护制度,影响科研成果权属认定;带来学术伦理风险,影响科研成果评价生态;冲击人才评价基本标准,影响人才评价客观性。鉴于人工智能的渗透性、扩散性和颠覆性,首先要在思想认识上积极面对而不是消极应对,用其所长、避其所短,审慎使用,着力发挥其促进知识生产、推动学术进步的积极作用;其次,要抓住现行相关法规制度与快速进步的人工智能技术之间发展不平衡的矛盾,进一步完善制度、设置“护栏”,规范使用;第三,要推动技术开源,公开其智能生成逻辑,增强人工智能聊天机器人的技术可信度;第四,要建立人机协作机制,既发挥技术的数据处理优势,又发挥人的理性思维优势,最终由人把关知识生产、科研成果及人才评价。

作者简介:蒋华林,教育学博士,重庆大学公共管理学院研究员,Email: jianghl@cqu.edu.cn。

关键词:人工智能;人工智能聊天机器人;人工智能生成内容;ChatGPT;微软必应;科研成果评价;人才评价;知识生产方式

中图分类号:TP18;TP242;G40-058.1 **文献标志码:**A **文章编号:**1008-5831(2023)02-0097-14

一、问题缘起:人工智能聊天机器人勃然兴起

适逢“世界数字教育大会”在北京召开前后,在互联网领域、学术界及产业界陡然掀起一股 ChatGPT 热。ChatGPT 仅上线 2 个月,其月活用户就超过 1 亿人,超越短视频社交平台 TikTok 成为史上增长最快的消费应用,一夜之间新晋“网红”。科技部部长王志刚认为,“ChatGPT 确实挺热”^[1],已形成了一种现象级的应用。2023 年 2 月 7 日,谷歌公司“应急”发布了基于其自研的全新自然语言模型框架 LaMDA(Language Model for Dialogue Applications)的下一代人工智能聊天机器人 Bard(中文名:巴德);2023 年 2 月 8 日,微软公司紧跟发布整合了 ChatGPT 的全新搜索引擎必应 Bing 和 Edge 浏览器,并宣布旗下 Office、Azure 云服务等产品都将全线整合 ChatGPT^[2];2023 年 2 月 24 日,Meta 公司“跟风”发布一款全新大型语言模型 LLaMA,加入硅谷人工智能竞赛^[3];2023 年 2 月 28 日,百度公司则官宣将于 3 月 16 日正式发布所谓百度版 ChatGPT——“文心一言”(英文名:ERNIE BOT);同时,阿里巴巴达摩院(杭州)科技有限公司申请的“人机对话及预训练语言模型训练方法、系统及电子设备”专利已在“实质审查中”^[4]。一时间掀起了全球人工智能聊天机器人大战。

针对 ChatGPT 热议,一部分人联想到当前中美之间的“科技战”、国内面临技术“卡脖子”的情形,感叹 ChatGPT 怎么又是外国提出来的,如“元宇宙”(metaverse)一样,为什么不是中国原创的;一部分人担心其对学生学习可能造成的偷懒甚至作弊情况而感到忧虑;一部分人则担心工作岗位被替代而失业,更多人则是在探讨以 ChatGPT、Microsoft Bing 为代表的人工智能机器人现象及其已经和可能产生的影响,它能做什么,不能做什么。如《教育研究》杂志社第一时间举办“ChatGPT 与未来教育”沙龙,《重庆高教研究》杂志社第一时间开办专栏笔谈,探讨和评估 ChatGPT 可能带来的教育影响、科技影响等,如韩云波对 ChatGPT 对教育的挑战与应对提出了初步思考;刘进讨论了 ChatGPT 时代工程教育知识重构的基本思路^[5]。此外,令小雄等对 ChatGPT 类人工智能技术的科技伦理和学术伦理进行了比较系统的思考^[6]。

实际上,ChatGPT 的核心技术是深度学习,做到更高水平(类人脑)或许就能像人一样(可能在某些方面超越普通人)思考、工作和创新。尽管 ChatGPT 的智能程度与我们的期望相比还不够高,但其深远影响却难以估量。微软公司创始人比尔·盖茨(Bill Gates)甚至认为,ChatGPT 的重要性不亚于互联网的发明^[7]。美国太空探索技术公司 CEO 埃隆·马斯克(Elon R. Musk)认为,ChatGPT “厉害得吓人”,“我们离强大到危险的 AI 不远了”^[8]。有人称其可能是科技界的“奇点”。可以预期,ChatGPT、Microsoft Bing 等新一代人工智能聊天机器人必将对人类知识生产方式产生重大影响,对于肩负知识传承、知识创新职责使命的教育界、科技界可能带来颠覆性革命。试想,如果科研成果(学术论文)是借助或完全依赖 ChatGPT 或 Microsoft Bing 生产的,应当如何评价它及其“作者”?由此可能对现行人才评价也产生革命性影响。对此,我们计将安出?

二、人工智能聊天机器人的主要功能

人工智能聊天机器人已发布多款产品,其中以 ChatGPT 和 Microsoft Bing 为代表。ChatGPT 全称为 Chat Generative Pre-Trained Transformer,中文名称是:预训练生成聊天模型或聊天生成预训练转换器,是由美国开放人工智能研究中心(OpenAI)研发的基于 GPT-3.5 语言模型的聊天机器人,于 2022 年 11 月 30 日正式发布。这是继 1997 年 IBM 公司的“深蓝”、2014 年谷歌公司的 AlphaGo 围棋机器人之后人工智能领域技术突破的新里程碑。据 OpenAI 网站介绍,ChatGPT 是一种优化的对话语言模型,采用“从人类反馈中强化学习”(Reinforcement Learning from Human Feedback, RLHF)对其进行训练,使之能够以对话方式进行交互,这种方式使其能够回答后续跟进问题(followup questions)、承认错误(当它不知道自己在说什么时,它承认不知道)、挑战不正确的前提、拒绝不适当的请求等。它是 InstructGPT 的姊妹模型,后者经过训练可以按照提示中的说明进行操作并提供详细的响应^[9]。ChatGPT 可追溯到 OpenAI 公司 2018 年在论文《通过生成性预训练提高语言理解》(Improving Language Understanding by Generative Pre-Training)中所构建 GPT-1 模型^[10]。Microsoft Bing,中文译名“微软必应”,原名 Bing(必应),是微软公司于 2009 年 5 月 28 日推出的一款搜索引擎。2023 年 2 月发布的新版微软必应,整合了 ChatGPT 系统背后的 AI 技术,它可以回答一些搜索查询,而不仅仅是提供一系列超链接。加入 ChatGPT 功能的新版 Microsoft Bing 搜索引擎,用户可以选择与搜索框对话,实现个性化智能搜索,使搜索结果更加精准。“更神奇的是,必应偶尔还会尝试主动向我们提问,建议可能答案”^[11]。基于其强大功能,据第三方 data.ai 的分析显示,Microsoft Bing 发布后仅一天时间,全球下载量就猛增了 10 倍^[12]。

目前,人工智能聊天机器人主要具有智能交互对话、智能文本生成和智能文献爬取三大功能。

(一) 智能交互对话功能

ChatGPT 作为一款聊天软件,其第一功能就是聊天、对话。基于强大的学习能力、强大的数据资源和先进的算法,ChatGPT 可谓“学人成精”了,其对话能力“引爆”网络。有人分析,其对话水平已超过 80%甚至 90%的人,在未来(对话)达到类人的能力一定是可以期待的^[13]。它具有很强的互动性,有问有答,甚至不问自答;有很强的“能动性”,能够以自然和类似人类的方式理解和响应复杂的请求,能够根据问题即兴发挥、即兴思考,“更像人类一样说话”。它可以就诸如“女朋友和妈妈同时落水先救谁?”这类让男朋友进退两难的生活常见问题进行对话、给出回答(示例如下)。从中我们可以看到其智能程度:先指出问题的复杂性和挑战性,然后从理论上进行分析,从“个人情况”给出答案,从道德的角度进行分析,最后从专业角度提出建议,没有把天聊死,也没得罪任何一方。

问:女朋友和妈妈同时掉水里,先救谁?

ChatGPT 回答:这是一个非常困难的决策。理论上,在紧急情况下,应该先救受伤或处于生命危险中的人,无论他们是谁。然而,如果两个人都需要救援,那么您需要评估每个人的情况并作出决策。从道德的角度来看,救自己的女朋友或救自己的母亲可能都是正确的,取决于您认为谁需要救助得更紧急。如果可能的话,最好请求紧急救护服务的帮助,以确保所有人的安全。

同时,ChatGPT也可以就一些“死循环”哲学问题进行对话,如“先有鸡还是先有蛋”的问题(示例如下)。从中我们可以看出,它开宗明义直接定义问题的学科领域(哲学问题)及其发展现状(存有争议),然后“花开两朵,各表一枝”,分别就对立的两种观点进行说明,最后再来一个中庸性的启发性结论。

问:是先有的鸡还是先有的蛋?

ChatGPT回答:这是一个著名的哲学问题,也是一个有争议问题。有人认为是先有鸡,因为鸡是由一对特定父亲和母亲通过生殖过程产生的,而蛋则是由鸡产生的。而有人则认为是先有蛋,因为他们认为蛋是孵化的前提,而鸡则是后续产生的。不管哪一种说法最终是正确的,他们都提醒我们要思考先因和后果的相互关系^①。

(二) 智能文本生成功能

如果 ChatGPT 只是一款单纯的聊天机器人,那类似的产品就多了,包括现在很多企业的“客服”机器人,都能进行基本的“对话”。ChatGPT 之所以“火出圈”,根本在于其基于超海量数据基础上的自主学习能力,采用人工智能生成内容技术(AI-Generated Content, AIGC),从而具有强大的内容生成或文本生成能力,相当于一个智能的文本生成器。如果只是生成文本功能,那 ChatGPT 也不是首创,早在 2009 年美国棒球大联盟就率先启用了智能机器人 StatsMonkey 完成比赛稿件的撰写,2015 年腾讯公司在国内首次上线了写稿机器人 DreamWriter^[14]。ChatGPT 等人工智能机器人之所以号称“震惊了半个科学圈”,是因为其“武力值”远远超过以前的机器人,它能够根据研究主题自动、智能撰写各种文本材料,无论是会议演讲稿、诗歌、小说、文学剧本,或者是商业文案、法律文书、新闻报道,或者是编程、写作业等,无一不懂,为此深受有关群体欢迎。据 Study.com 2023 年 1 月对美国 1 000 多名学生使用 ChatGPT 的调查显示,超过 89% 的学生曾使用 ChatGPT 帮助完成家庭作业,48% 的学生承认使用 ChatGPT 完成家庭测试,53% 的学生使用 ChatGPT 写论文,22% 的学生使用它写论文提纲^[15]。据美联社 2023 年 1 月 25 日报道,美国众议员杰克·奥金克洛斯(Jake Auchincloss)在发表演讲讨论创建美国—以色列人工智能(AI)中心的法案时,其宣读的两段讲稿是由 ChatGPT 生成的^[16]。2023 年 2 月 4 日,以色列总统艾萨克·赫尔佐格(Isaac Herzog)发表了部分由人工智能(AI)撰写的演讲,成为首位公开使用 ChatGPT 的国家领导人^[17]。新加坡政府正在将 ChatGPT 整合到 Microsoft Word 中,开发一套类似 ChatGPT 的系统——Pair,以更好地协助公务员完成篇幅较长的总结、探索相关观点或改善写作表达能力等^[18]。

如果只是能生成一般应用性文本,那也就不是 ChatGPT 了。它的出彩之处在于,无论是学术前沿的量子物理研究,还是复杂的宗教与社会问题,无一不晓。只要给定一定的条件或提供简洁的信息,ChatGPT 犹如一位饱学之士,能够“读书破万卷,下笔如有神”。有人曾要求 ChatGPT“写一个故事,用简单的术语解释量子物理学”,结果它不负所托(图 1)。2023 年 1 月,北密歇根大学的哲学教授安东尼·奥曼(Antony Aumann)评阅“世界宗教”课程的学生论文时,惊喜地读到了一篇“全班最好的论文”,虽然给出了 A+ 的成绩,但他对此还是有所怀疑,经追问和质询学生,原来这篇“全班得

^①资料来源:女朋友和妈妈同时落水先救谁? ChatGPT 给出了答案[EB/OL]. [2022-02-17]. https://www.thepaper.cn/newsDetail_forward_21946669.

分最高的论文,是用 ChatGPT 写的”^[19]。这正是人工智能聊天机器人“人有我优”的核心竞争力——能够从事科学写作。事实上,上文所说的智能交互对话功能,也是一种文本生成,属于“交互式文本生成”,而此处所称智能文本生成属于“非交互式文本生成”。可以说,ChatGPT 已表现出惊人的学习能力和创作(生成)能力。这让人不禁想起“江郎才尽”典故,江淹因得“五色笔”而才华横溢,因失之而“从此遂无妙句”。ChatGPT 似乎成了现代的“五色笔”,离开它我们或许会“江郎才尽”。

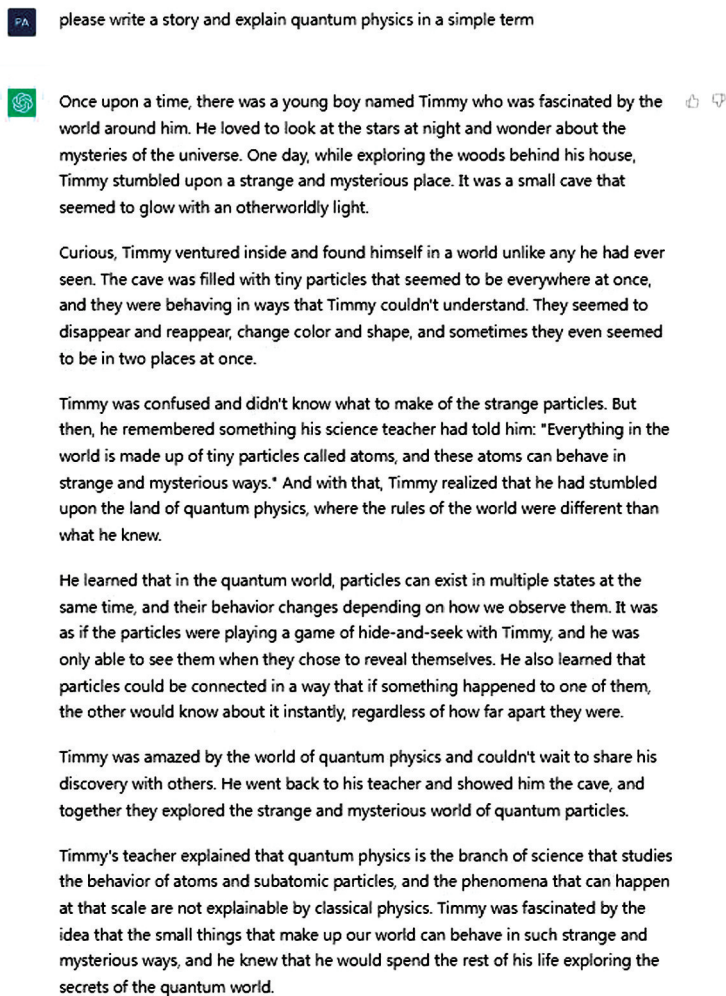


图1 ChatGPT 生成的用文学故事解释量子物理的文本^②

(三) 智能文献爬取功能

ChatGPT、Microsoft Bing 等人工智能聊天机器人的智能化建立在强大的文献搜索技术之上。为了“培养”ChatGPT, OpenAI 用了 45TB 的数据、近 1 万亿个单词,大概是 1 351 万本牛津词典^[2]。特别像 Microsoft Bing 本身就是一款强大的网络搜索引擎,可以快速检索某个特定主题的相关文献资料,并完成摘要提取,形成文献综述;或搜索某个主题的研究方向,找到该主题的前沿研究问题,追

^②文巧. 写的论文骗过大学教授,火爆全球的机器人 ChatGPT 要抢谁的饭碗? [EB/OL]. [2023-02-01]. <https://new.qq.com/rain/a/20230129A06EY900>.

踪研究热点。如有人提问 ChatGPT:“Got any creative ideas for organizational behavior research?”,结果 ChatGPT 给出了 8 点意见。国外已研究表明,使用 ChatGPT 进行文献综述,就是对 Google Scholar 上的相关论文摘要进行关键词比对,然后进行“转述”生成的^[20-21]。

笔者使用 Microsoft Bing(集成了 ChatGPT 功能)检索关于 nanoe technology 的文献综述,结果显示有 2 340 000 条记录,呈现若干关于该主题的综述类文献,从而可以帮助人们快速了解关于 nanoe technology 的研究进展和最新动态(图 2)。

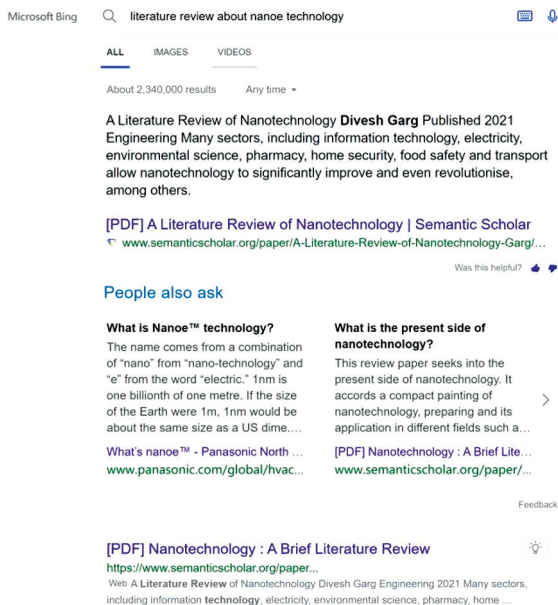


图 2 使用 Microsoft Bing 进行文献综述示例

事实上,人工智能聊天机器人的功能远不止这三个方面,它还可以广泛应用于机器翻译、软件开发(书写代码、调试程序等)以及自然语言处理(如语法分析)、各种考试等方面。

三、人工智能聊天机器人推动知识生产方式和学术研究范式变革

一般认为,学术研究因其高深性、逻辑性、理性而为人类所独有,“几个世纪以来,人们一直将科学视为一种知识形成与理性思考的融合,坚持知识生产和验证过程的严谨性与透明性”^[22]。但如今人工智能聊天机器人已拥有强大的“创作”和学习能力,学术研究面临数百年未有之大变局,新的学术研究范式和知识生产方式已露端倪,科研成果产出方式可能在不久的将来被颠覆。可以说,以人工智能聊天机器人为代表的人工智能技术正在引领范式变革。

(一) 知识生产和学术研究进程显著提速

传统意义上,学术研究和知识生产速度较慢,正所谓“十年磨一剑”“板凳甘坐十年冷”,大概就是用来说明这种情形下的知识生产节奏的。作家马尔科姆·格拉德威尔在《异类:不一样的成功启示录》中更提出了“一万小时定律”,即“人们眼中的天才之所以卓越非凡,并非天资超人一等,而是付出了持续不断的努力。1 万小时的锤炼是任何人从平凡变成世界级大师的必要条件”^[23]。随着信息技术的发展,加上社会需求更新频率不断加速,最近数十年来知识更新节奏不断加快,芯片界“摩尔定律”就是一个例证。为此,有人提出“知识半衰期”说法:一个专业人士的知识如果不更新,

在“半衰期”后,基础知识仍然可用,但其他一半新知识则已落伍。据测算,1950年前的知识半衰期为50年,21世纪知识的半衰期平均为3.2年,IT高级工程师的半衰期仅为1.8年^[24]。知识半衰期之所以不断缩短,原因是新知识增量速度不断加速,知识总量不断扩大,人类由此进入“知识经济时代”。不过,这一切在今天又都变了:几年前有报道称机器人写新闻稿只要0.3秒,现在ChatGPT几秒钟就可完成一篇学术论文,可谓只有更快没有最快。可以预期,知识生产成果数量(文献量)将在短时间内实现倍增,人工智能聊天机器人一年所创造(生产)的知识可能超过以往若干时代的总和。

(二) 知识生产者从非创造性劳动中获得解放

人类劳动一般分为重复性劳动与创造性劳动,前者指维持人类经济与社会发展中的简单再生产与扩大再生产的劳动,后者指探索、发现、使用人类不曾使用过的知识、技能、手段、材料、工具创造新的产品或新的生产方式从而以更高的效率从事商品生产的劳动^[25],其中“包含着思考和精巧的技能和技艺”^[26]。学术研究是一种特殊形式的劳动,在学术研究和知识生产的劳动过程中,创造性劳动显然是核心和关键,因为知识需要创新,只有创造性劳动才会有新的知识产生,否则研究成果就没创新性,不过,其中也少不了占相当比重的重复性劳动或称为非创造性劳动,如研究主题的学术史梳理、文献搜集与整理、实验操作、文字输出、数学计算、生成初稿以及文字校对等等。这一类非创造性劳动或创造性较低的劳动,在传统的学术研究中往往需要研究者投入大量人力物力、时间精力,速度慢、效率低,严重影响学术研究或知识生产进程和效率。ChatGPT等人工智能机器人通过其强大的数据库和智能技术手段,可以在非常的短时间内完成有关非创造性劳动,可以显著提高科学知识生产效率;同时,它将研究人员(包括期刊编辑)从非创造性劳动中解放出来(对文献检索等基本技能的需求或将大大减少),“科研人员再也不用坐在那里给经费申请书写很长的引言了,他们现在只要让系统来写就行了”^[27],从而使其有更多时间从事创造性劳动:如进行项目研究顶层设计、思考提出新理念、设计新的实验方案、逻辑推理等。或许我们可以预期,人工智能聊天机器人今后将进化到能够从事某些方面某种程度的创造性劳动。如此,学术成果或知识产品实际上是由人与人工智能聊天机器人共同完成的,学术研究范式将发生重大变化。

(三) 知识生产方式由“手动式”“半自动式”迈向“自动化”“智能化”

关于知识生产方式的转变,学术界目前主要基于知识生产的边界发生变化或者说环境变迁、参与主体变动的视角来思考,其中最大变化之一是知识生产的“情境性”。所谓知识生产方式的情境性,是指“大学知识生产的垄断地位被打破,知识需求者的立场往往取代真理而成为知识生产的标准,知识生产越来越受到具体的应用情境影响”^[28],即知识生产在“应用情境”中进行,“以销定产”而不是“以产定销”,进入吉本斯(M. Gibbons)所称的知识生产模式Ⅱ。英国知名理论物理学家和科学技术学家约翰·齐曼(John Ziman)则根据冷战结束后西方发达国家科学研究发生的重大变化,将“为科学而科学”、只为追求真理、满足好奇心的科学知识生产称之为“学院科学”,与之相对应的“后学院科学”知识生产新模式则强调集体化、产业化、官僚化等^[29]。如果我们从知识生产的技术参与程度来看,知识生产方式也已发生革命性变革。在现代信息技术、互联网技术、大数据技术等出现以前,知识生产完全依靠研究人员手工劳动或借助一定的仪器设备和工具软件进行研究,是“手动式”或“半自动式”。现代信息技术等出现以后,特别是集其大成的人工智能技术不断迭代升

级以后,知识生产在某些情况下已经可以不需要人的介入,在更多情况下人直接介入的程度越来越低,人工智能聊天机器人已经可以自动生成文本、智能生产知识了,知识生产方式(写作或科研过程)实现“自动化”“智能化”。2022年12月,Medrxiv发表了一篇题为《ChatGPT在美国医学执业考试(USMLE)上的表现:使用大型语言模型进行人工智能辅助医学教育的潜力》的论文,ChatGPT被列为第三作者,而且还通过了USMLE考试^[30]。按照劳凯声教授的观点,大学知识生产的发展已从以大师为标志的前知识生产阶段、以学科为标志的创新1.0阶段,发展到了以智能化创新平台为标志的创新2.0阶段^[31]。

四、人工智能聊天机器人影响现行科研成果及人才评价机制

上述分析表明,人工智能机器人已经对知识生产产生重大影响,它已经不是我们愿意不愿意、认可不认可的问题了,而是不以我们的意志为转移的一种客观发展态势。作为知识生产结果主要表征的科研成果(学术论文)不可避免地会“遗传”到人工智能聊天机器人的基因,这对传统的科研成果评价机制将产生重大影响,更进一步必将对通过科研成果进行人才评价产生重大影响。粗略地说,以ChatGPT为代表的人工智能聊天机器人将影响科研成果创新性评价,影响科研成果权属认定,影响科研成果评价生态,影响人才评价客观性等。

(一) 遮蔽成果演化学术史,影响科研成果创新性评价

人工智能聊天机器人强大的文本生成功能,彻底改变了学术论文的内容创建方式。已有研究表明,人们已难以辨别学术论文(文本)究竟是ChatGPT“写”的还是人工写的。国外曾做过一次实证,结果发现,ChatGPT生成的摘要愚弄了科学家,使用ChatGPT生成的摘要顺利通过了剽窃检查器,原创得分中位数为100%,表明没有检测到剽窃;AI输出检测器发现了66%的ChatGPT生成摘要,但人工审稿人却错误地将32%的生成摘要识别为真实摘要,将14%的真实摘要识别为生成摘要^[32]。这不仅表明人工智能聊天机器人已能在很大程度上“懂人心”“说人话”,而且由于人工智能聊天机器人目前多半只是“综述组合”“杂糅复制”有关数据资料,常“在没有可靠地引用原始来源或作者的情况下复制文本”,不会标注或不能准确标注参考文献,从而导致“使用这种技术的研究人员有可能不对早期的工作给予肯定”^[33],遮蔽所研究问题的学术发展历史脉络。正如《自然·机器智能》杂志2023年1月发表的一篇关于ChatGPT的社论所说,ChatGPT在提供正确事实或可靠参考文献方面不能被信任^[34]。此外,一旦研究人员向ChatGPT和其他大语言模型透露相关信息,就可能被纳入模型,随之人工智能聊天机器人就可以将其提供给他人,但提供者却不被承认为原始来源^[33]。长此以往,人们将难以判断学术思想的源流,更难以判断科研成果的创新性。加之,为了应对学术期刊(以及相关文稿发表平台)审稿查重,或应对AI生成文本检测软件(如斯坦福大学研究团队设计的DetectGPT、普林斯顿华人本科生Edward Tian创建的GPTZero等),人工智能机器人将“洗稿”“降重”“降AI痕迹”,或将导致“一本正经胡说八道”式的伪创新。

(二) 挑战知识产权保护制度,影响科研成果权属认定

在web of science上以“作者”字段检索ChatGPT,截至2023年3月4日,已有5条记录。随着人工智能聊天机器人的推广及进化,署名ChatGPT的学术成果将越来越多。对于人工智能聊天机器

人能否署名的问题,学术界和法律界持有争议。目前主要有两种观点:一种观点认为可以署名,其著作权归属开发人工智能聊天机器人的企业(机构);另一种观点不赞成署名,因为这与现有著作权法规相抵触。如《中华人民共和国著作权法》第九条规定,著作权人包括“作者”(中国公民及相关外国人、无国籍人)和“其他依照本法享有著作权的自然人、法人或者非法人组织”,类似 ChatGPT 一样的人工智能聊天机器人显然不符合著作权人的规定。欧美国家的版权法相关规定大体相同,如美国版权法规定,作者是指创作美国版权法所涵盖的某类作品的人^[35];美国版权局在其版权登记手册中也明确规定:作品必须是由人创作的。由自然现象或动植物的活动形成的结果不能进行版权登记^[35]。从根本上说,现行版权法保护的是权利人的权益,其前提是权利人能够对自己的工作负责,作者更应该是“创作者”,要“文责自负”。显然,人工智能聊天机器人不是民事行为主体,不能对自己的工作负责,也很难准确认定它对论文(科研成果)做出的实质贡献。美国《科学》(Science)杂志主编认为,应按照作者对论文原创性负责的一贯要求,明确禁止将人工智能程序列为作者,认为“ChatGPT 好玩,但不是作者”,并且修改其编辑政策,不允许在论文中使用 ChatGPT 或其他人工智能工具生成的文本、图形、图像等^[36]。英国《自然》(Nature)杂志 2023 年 1 月 24 日发表社论,认为 ChatGPT 等工具威胁着透明的科学,呼吁建立基本规则,以合法使用而避免滥用,虽然其允许作者在主动报告的情况下使用 ChatGPT 等工具,但仍然禁止署名为作者^[37]。国内方面,《暨南学报(哲学社会科学版)》对使用 ChatGPT 等人工智能工具的立场与 Nature 类似^[38],而《天津师范大学学报(基础教育版)》只是要求作者对使用情况予以说明,未明确禁止署名^[39]。由此可见,如果我们将人工智能聊天机器人署名为作者,将不符合目前相关法律对作者身份的定义;如果不将其署名为作者,又与“作者”的实际贡献不符,都将影响科研成果的权属认定。

(三) 带来学术伦理风险,影响科研成果评价生态

美国技术哲学家詹姆斯·摩尔提出科技伦理领域的“摩尔定律”,“伴随着技术革命,社会影响增大,伦理问题也增加了”^[40]。随着人工智能技术的进步,其伦理问题将日益突出。ChatGPT 等人工智能聊天机器人基于其“非人化”,不具有人类意义上的个人意识和感情,严格意义上它本身不具有伦理意识和伦理责任,但这个伦理责任应当由使用它的人来承担,比如使用 ChatGPT 用于考试作弊,考生应该受到惩罚。一方面,ChatGPT 等人工智能聊天机器人可能被滥用。目前 ChatGPT 等技术还不够成熟,经常连它自己也没有“标准答案”,但只要给定了提示,它一般都能够基于神经网络和“联结主义”的强大算法给出回答,如“林黛玉倒拔垂杨柳”“张飞杀岳飞杀得满天飞”等看似合理实则不然的答案(其中有前期语料库预训练的因素)。如果“这个系统在学术界被滥用,因为借助 ChatGPT,没有专业知识的人也可以尝试写科学论文”^[41],从而生成虚假的学术论文,“制造”所谓新知识,扭曲科学事实,传播错误信息,不断往语料库中“掺沙子”,导致以讹传讹,带来巨大伦理风险。另一方面,ChatGPT 等人工智能聊天机器人可能被有意误用。ChatGPT 虽然自信但也保持清晰,强调自己“只是一个工具,还在不断进化,还有一些局限”,但有人可能会利用所掌握的技术优势和工具的技术局限,利用海量网络资源生成若干论文,却不如实报告人工智能写作工具的使用情况,钻有关政策或技术空子(期刊编辑和同行识别困难),剽窃而不被发现,侵权而难以认定,从而引发学术失范,对科研成果评价生态造成深远影响。另外,ChatGPT 背后的算法如果存在歧视或偏见(传

播偏差、偏好),对不同国家(地区)、不同人提供不同的过滤器、调用不同的语料库,则其潜藏的道德风险更大。

(四) 冲击人才评价基本标准,影响人才评价客观性

“不出版,就死亡”,不仅是高校人事聘用考核的通行原则,也是人才评价的一项基本标准,无论是评价人才(教师)的学术潜力(招聘)还是学术实力(授予人才头衔、晋升职称、申请项目等),都要评价其研究产出效率和学术贡献率。从全世界范围看,论文依然是学术成果主流呈现方式,“是代表大学教师科研水平的标志性成果”^[42],仍是学术界的“硬通货”^[43],因此,尽管要“破五唯”不能“唯论文”,但基于其对学术发展的重要性,学术论文仍然是人才(教师)评价的重要参考指针。鉴于此,学术界将大步跟上人工智能步伐,使用 ChatGPT 等人工智能聊天机器人来“写”论文,其中除了存在上文论及的相关问题外,还将涉及如何评价这些论文的作者问题。我们知道,知识就是力量,“技术的本质是能力,而能力越大,其博弈均衡点就对技术掌权者越有利”,“占有技术资源的人拥有压倒一切的必胜技术”而成为“绝对强者”^[44],这样那些掌握人工智能技术、能熟练使用人工智能聊天机器人、擅长操纵数据资料“写”论文的“技术人员”,就将凭借数量优势相对容易地被评价为“人才”,成为“知识权威”,获得各种学术利益。更糟糕的是,这些“人工论文”至少目前还难有真正的创新性,学术价值堪忧。同时,由于其高产出,所谓人才的成长周期或将大大缩短,人工智能聊天机器人将“揠苗助长”一批人才,而这些所谓人才的成长过程并没有伴随着自己在该领域的经验的真正增加,“学术机构聘用专业人士时可能会出现道德问题”^[45]。如果只破不立,不能科学评价“新型论文”,则必将不能科学开展人才评价、客观评价人才。

五、基本应对思路与对策

根据 Web of Science,截至 2023 年 3 月 4 日,“标题”字段中包含 ChatGPT 的检索记录已有 65 条,其中 2023 年为 62 条,日均新增一条;论文类已有 27 篇。另外,已有网络文章注明:“本文有 ChatGPT 的贡献”。因此,在人工智能技术滚滚向前的历史车轮上,使用 ChatGPT 还是不使用 ChatGPT,已经不是一个问题,问题是如何合理使用、为科研成果及人才评价所用、为人类所用。

(一) 积极面对、审慎使用

习近平总书记指出:“自古以来,科学技术就以一种不可逆转、不可抗拒的力量推动着人类社会向前发展。”^[46]教育更是随着技术变革而变革、转型和发展。教育是面向人、培养人的事业,旨在培养全面发展的人、不断提高人类(群)的综合素质,进而能够推进社会科技进步,造福人类。面对汹涌奔流的人工智能技术,首先要在思想认识上积极面对,甚至张开双臂拥抱它,而不是消极抵制甚至唱衰反对它,因为人工智能是技术发展方向和归宿。在某种程度上可以说,一切技术进步都在奔向“人工智能”的这个目标,一方面是为了让“万物有灵”,另一方面是为了实现人的智能延伸拓展,从而最终实现“人是万物的尺度”这一人类夙愿。不过,我们也不要把他奉为“天神”或被其吓倒,对“热技术”要“冷思考”。其次,由于任何技术都是一把双刃剑,ChatGPT 在带给我们便利、收益时,也可能带来负能量,产生副作用,包括可能影响人(类)的思考能力。对此,我们要审慎使用,用其所长、避其所短,发挥其促进知识生产、推动学术进步的积极作用。

(二) 完善制度、规范使用

人工智能聊天机器人对科研成果及人才评价机制的影响,其中既有技术不够先进的因素(下文另述),更主要是现行相关法规制度与快速进步的技术手段之间发展不平衡的矛盾。比如人工智能聊天机器人能否署名为作者的问题,症结就在于现行法律障碍,根据近现代法律思想,“人是法律主体的元概念,权利主体始于人,并派生出权利能力”,著作权法也因此规定只有自然人、法人或者非法人组织才享有著作权,因此人工智能聊天机器人不能直接获得所生成作品的著作权^[35]。再如冲击学术伦理规范问题,我们可以建立健全和落实相关制度,在技术开发上,贯彻落实国家新一代人工智能治理专业委员会发布的《新一代人工智能伦理规范》,切实“将伦理道德融入人工智能研发和应用的全生命周期”;在使用上,要求使用人标注应用人工智能聊天机器人生成的文本内容,采用适当方式报告人工智能聊天机器人使用情况等,同时研发 AI 使用情况检测工具,建立问责机制,对违反相关制度的人或机构施以相应惩罚,为人工智能聊天机器人及其使用人的行为设置“护栏”。“相信只要我们能够构建适应新时代、新技术的学术规范、伦理道德和管理制度”^[5],加强监测和评估,保障人工智能聊天机器人被规范安全、公开透明、合乎伦理的使用,就将有助于学术研究和创新能力提升。

(三) 公开透明、开源“智能”

目前,人工智能聊天机器人应用所带来的忧虑、担心、困扰,很大程度上是因为其生成过程是一个“黑箱”,不清楚其神经网络是如何联结的,不清楚其生成内容与“原始文献”之间的算法逻辑。这一点与人们现今使用某些研究工具(如 NVivo、DEA 等)进行学术研究是有很大的不同的,因为后者对使用者来说是透明,数据处理过程、方式是研究人员设计的,是可以重复的。为此,建议推动人工智能聊天机器人技术实行“开源”,公开其算法逻辑,明示其生成逻辑,防范“数据主义”风险,从而增强人工智能聊天机器人的技术可信度,提高社会认可度,降低社会困惑度。

(四) 人机协作、共同“治理”

在可以预见的将来,人工智能技术仍会继续听命于人,受人控制,作为人的助手,其智能水平都是人赋予它的,还只能按照人类教给他的内容学习、运行,应该说还未“学成出师”;另一方面,由于“人工智能的替代效应是建立在对人类劳动数据化和逻辑化的基础上的”,不能替代那些不可数据化和逻辑化的劳动^[47],加之技术的不完备性和发展性,现阶段的人工智能聊天机器人还处于“弱智能”阶段,缺少人的理性思考,还未发展到“强智能”水平,对人类创造性劳动只能起到辅助作用,而且应该永远不可能完全替代人。为此,建议在科研成果及人才评价、学术论文撰写及发表环节都要继续安排人工介入,由“人工”与“智能”并行而有所交叉地进行检查把关,既发挥技术的数据处理优势,又发挥人的理性思维优势,注重考察科研成果和人才的创新价值、创新能力。

最后,笔者想借用“神笔马良”的故事对本文作个小结。目前的人工智能聊天机器人还属于“初生之物”,在技术研发上还有很长的路要走,在应用场景开发上还有许多工作要做。在未来,如果我们能将人工智能聊天机器人“规训”成为马良的那支“神笔”,并且保证其掌握在善良、正义的“马良”手里,那它必将使人类知识生产能力如虎添翼、科研创新能力大幅提升,科研创新人才辈出,人工智能聊天机器人则将与人类和谐相处、共同发展。

参考文献:

- [1] 刚刚, 科技部部长王志刚谈 ChatGPT[EB/OL]. (2023-02-24). <https://baijiahao.baidu.com/s?id=1758687146107906632&wfr=spider&for=pc>.
- [2] 朱涵, 彭茜, 黄堃. 三问 ChatGPT 如何影响人工智能的未来[J]. 瞭望, 2023(8).
- [3] Meta 发布全新大型语言模型 LLaMA 加入硅谷 AI 竞赛[EB/OL]. (2023-02-26). <https://baijiahao.baidu.com/s?id=1758859153822039672>.
- [4] 国产版“ChatGPT”进展如何? 阿里巴巴开始内测 百度下月面世[EB/OL]. (2023-02-09). <https://baijiahao.baidu.com/s?id=1757283966256262300&wfr=spider&for=pc>.
- [5] 钟秉林, 尚俊杰, 王建华, 等. ChatGPT 对教育的挑战(笔谈)[J/OL]. 重庆高教研究, 2023; 1-25. (2023-03-02). <http://kns.cnki.net/kcms/detail/50.1028.G4.20230301.1040.002.html>.
- [6] 令小雄, 王鼎民, 袁健. ChatGPT 爆火后关于科技伦理及学术伦理的冷思考[J/OL]. 新疆师范大学学报(哲学社会科学版), 2023; 1-14. (2023-03-01). <https://doi.org/10.14100/j.cnki.65-1039/g4.20230213.001>.
- [7] 比尔·盖茨再谈 ChatGPT: 要性不亚于互联网的发明, 将改变我们的世界[EB/OL]. (2023-02-11). <https://baijiahao.baidu.com/s?id=1757516621271845208&wfr=spider&for=pc>.
- [8] ChatGPT 可能从这 12 个方面彻底改变我们的工作和生活[EB/OL]. (2023-02-15). <https://blog.csdn.net/jane9872/article/details/129052120>.
- [9] Introducing ChatGPT[EB/OL]. (2022-11-30) [2023-03-05]. <https://openai.com/blog/chatgpt>.
- [10] RADFORD, ALEC KARTHIK NARASIMHAN. Improving Language Understanding by Generative Pre-Training[R]. 2018.
- [11] 颠覆历史! “ChatGPT 搜索引擎”发布, 微软市值一夜飙涨 5450 亿[EB/OL]. (2023-02-08). <https://baijiahao.baidu.com/s?id=1757229046402590054&wfr=spider&for=pc>.
- [12] 魏蔚. 集成 ChatGPT 后, 微软必应下载量增 10 倍[N]. 北京商报, 2023-02-09. <https://www.bbtnews.com.cn/2023/0209/466149.shtml>.
- [13] 专家称 ChatGPT 对话水平已超过 9 成人类[EB/OL]. (2023-02-04). <https://www.163.com/dy/article/H5O81FLH05319LDA.html>.
- [14] 文静. 写稿机器人“抢”饭碗?[N]. 广州日报, 2018-11-19. <https://baijiahao.baidu.com/s?id=1617506646524438875&wfr=spider&for=pc>.
- [15] 卜金婷. 超 89% 的学生使用 ChatGPT 作弊! 斯坦福率先打响反击战[N]. 中国科学报, 2023-02-03.
- [16] 文巧. 写的论文骗过大学教授, 火爆全球的机器人 ChatGPT 要抢谁的饭碗?[EB/OL]. (2023-02-01). <https://new.qq.com/rain/a/20230129A06EY900>.
- [17] 人工智能行业深度报告: ChatGPT 引发的大模型时代变革[EB/OL]. (2023-02-27). <https://baijiahao.baidu.com/s?id=1758971954494864004&wfr=spider&for=pc>.
- [18] 政府开发人工智能系统 提高公务员书写工作效率[N]. 联合早报, 2023-03-01. <https://www.zaobao.com/news/singapore/story20230301-1367941>.
- [19] 学生用 ChatGPT 拿下全班最高分, 教授惊呆了[EB/OL]. (2023-02-05). https://learning.sohu.com/a/637506415_121123901.
- [20] AYDIN Ö, KARAARSLAN E. OpenAI ChatGPT generated literature review: digital twin in healthcare[J]. SSRN Electronic Journal, 2022; 22-31.
- [21] SALVAGNO M, ChatGPT, TACCONE F S, *et al.* Can artificial intelligence help for scientific writing? [J]. Critical Care, 2023, 27(1): 75.
- [22] 王俊美. ChatGPT 改变人类知识生产方式[N]. 中国社会科学报, 2023-02-20(03).
- [23] 牛宝荣, 李如密. 有意练习: 促进学生有效学习的重要策略[J]. 教学研究, 2020(6): 9-14.
- [24] 人民日报寄语年轻人: 以知识知识锤炼真本领[N]. 人民日报, 2018-09-20.
- [25] 冯骧. 创造性劳动与劳动价值论: 对马克思劳动价值公式的补充[J]. 河南师范大学学报(哲学社会科学版), 2008(5): 16-19.

- [26] 谭凯娜, 马香莲. 创造性劳动: 苏霍姆林斯基劳动教育思想的精神特质[J]. 吉林省教育学院学报, 2022(11): 59-64.
- [27] STOKEL-WALKER C, VAN NOORDEN R. What ChatGPT and generative AI mean for science[J]. *Nature*, 2023, 614(7947): 214-216.
- [28] 张继明. 知识生模式变迁视角下大学治理模式的演进及其反思[J]. 江苏高教, 2019(4): 9-17.
- [29] 约翰·齐曼. 真科学[M]. 曾国屏, 等译. 上海: 上海科技教育出版社, 2002: 37, 84-99.
- [30] KUNG T H, CHEATHAM M, MEDENILLA A, et al. Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models[J]. *PLoS Digit Health*, 2023, 2(2): e0000198.
- [31] 劳凯声. 智能时代的大学知识生产[J]. 首都师范大学学报(社会科学版), 2019(2): 1-6.
- [32] ELSE H. Abstracts written by ChatGPT fool scientists[J]. *Nature*, 2023, 613(7944): 423.
- [33] EVA A. M. VAN DIS, JOHAN BOLLEN, WILLEM ZUIDEMA, et al. ChatGPT: five priorities for research[EB/OL]. (2023-02-03). <https://www.nature.com/articles/d41586-023-00288-7>.
- [34] The AI writing on the wall[J]. *Nature Machine Intelligence*, 2023, 5(1): 1.
- [35] 邢贺通. 从 ChatGPT 看 AI 生成内容的著作权问题[EB/OL]. (2023-02-24). http://www.ipforefront.com/m_article_show.asp?id=2900&BigClass=%E4%B8%93%E6%A0%8F.
- [36] THORP H H. ChatGPT is fun, but not an author[J]. *Science*, 2023, 379(6630): 313.
- [37] Tools such as ChatGPT threaten transparent science; here are our ground rules for their use[J]. *Nature*, 2023, 613: 612. <https://www.nature.com/articles/d41586-023-00191-1>.
- [38] 关于使用人工智能写作工具的说明[EB/OL]. 暨南学报(哲学社会科学版)微信公众号, 2023-02-10.
- [39] 关于使用人工智能写作工具的说明[EB/OL]. 天津师范大学学报基础教育版微信公众号, 2023-02-11.
- [40] 尤瑞恩·范登·霍温, 约翰·维克特. 信息技术与道德哲学[M]. 赵欢迎, 等译. 北京: 科学出版社, 2014: 13.
- [41] 李木子. AI 能列为论文作者吗[N]. 中国科学报, 2023-01-20(01).
- [42] 田辉. 国外高校教师多元化晋升机制面面观[N]. 光明日报, 2019-10-24(14).
- [43] 构建多元评价体系, 破除“唯论文”(社论)[EB/OL]. (2020-12-15). https://www.thepaper.cn/newsDetail_forward_10404249.
- [44] 赵汀阳. 人工智能提出了什么哲学问题? [J]. 文化纵横, 2020(1): 43-57.
- [45] SALVAGNO M, ChatGPT, TACCONI F S, et al. Can artificial intelligence help for scientific writing? [J]. *Critical Care*, 2023, 27(1): 75.
- [46] 习近平. 在中国科学院第十七次院士大会、中国工程院第十二次院士大会上的讲话[N]. 人民日报, 2014-06-10(02).
- [47] 白惠仁. 人工智能与创造性劳动: 评《科学技术知识的政治经济学研究》[J]. 西部学刊, 2020(1): 154-156.

Research on the Influence of AI Chatbot on the Evaluation of Scientific Research Achievements and Talent Evaluation: Analysis based on ChatGPT and Microsoft Bing

JIANG Hualin

(School of Public Affairs, Chongqing University, Chongqing 400044, P. R. China)

Abstract: Chat GPT, an artificial intelligence (AI) chatbot developed by the United States Open Artificial Intelligence Research Center (OpenAI), can understand and respond to natural language input based on the language model of “Generate Pre-training Converter” (GPT), which was officially released on November 30, 2022, and has caused heated discussion and attention in the global Internet, academia and industry in a short time. The AI chatbots represented by ChatGPT and Microsoft Bing are based on huge amounts of data and trained with “reinforcement learning from human feedback” (RLHF). They have three

main functions: intelligent interactive dialogue, intelligent content generation, and intelligent document crawling. The core competitiveness of the AI chatbot is to adopt the AI generated content technology (AIGC) and have strong content generation or text generation ability. In addition to generating general applied manuscripts, it can also engage in scientific writing, thus leading and promoting the transformation of the mode of knowledge production and academic research paradigm. The process of knowledge production and academic research will be significantly accelerated, and knowledge producers will be liberated from non-creative labor, the mode of knowledge production will move from “manual” and “semi-automatic” to “automatic” and “intelligent”. This change will have a significant impact on the traditional evaluation mechanism of scientific research achievements, and further will have a significant impact on talent evaluation. In general, the application of AI chatbot represented by ChatGPT in academic research will obscure the evolution of academic history of achievements and affect the innovative evaluation of scientific research achievements, and challenge the intellectual property protection system and affect the ownership of scientific research achievements, and bring academic ethical risks and affect the evaluation ecology of scientific research achievements, and impact the basic standards of talent evaluation and affect the objectivity of talent evaluation. In view of the penetration, diffusion and subversion of AI, first of all, we should face it positively rather than passively in ideological understanding, utilize its strengths and avoid its weaknesses prudently, and exert its positive role in promoting knowledge production and promoting academic progress. Secondly, institution construction should be further improved to set up “guardrails” for AI chatbot according to the imbalance between the current relevant laws and regulations and the rapid progress of AI technology. Third, AI chatbot technology open source should be promoted to disclose its intelligent generation logic, and enhance the technical credibility of AI chat robots. Fourth, human-computer cooperation mechanism should be established to give play to the advantages of technology in data processing and human rational thinking, and finally let people check the knowledge production, scientific research achievements and talent evaluation.

Key words: artificial-intelligence; artificial-intelligence chatbot; AI-Generated Content; ChatGPT; Microsoft Bing; evaluation for scientific research achievements; talent evaluation; mode of Knowledge production

(责任编辑 彭建国)