

【新技术与教育变革】

教育人工智能系统的伦理原则与困境突破

钱小龙¹, 张奕潇¹, 宋子昀¹, 李强²

(1. 南通大学教育科学学院, 江苏南通 226019; 2. 南宁师范大学教育科学学院, 广西南宁 530001)

[摘要] 人工智能技术在促进教育变革的过程中,也给教育带来不可预知的伦理风险。作为人工智能教育应用的重要产物,教育人工智能系统备受推崇,但必须遵循相关的伦理原则。首先,从宏观层面探讨教育人工智能系统的一般伦理原则,即通过对世界各地相关政策文件的分析解读,明确人工智能伦理原则和教育人工智能伦理原则。其次,从微观层面探讨教育人工智能系统的具体伦理原则,即通过相关国际标准和规范的学习借鉴,明确教育人工智能系统开发与应用的伦理原则。最后,针对未来人工智能教育的发展和可能遇到的伦理挑战进行预测,并结合国情,提出我国教育人工智能系统未来发展的策略建议,即培养具备合格人工智能素养的社会公民、制定更加科学的教育人工智能系统伦理原则、构建教育人工智能系统应用过程的监管机制、重视人文关怀的情感教育力量。

[关键词] 人工智能; 人工智能教育应用; 伦理原则; 教育人工智能系统

[中图分类号] G434

[文献标识码] A

[文章编号] 1671-6973(2021)06-0096-09

一、背景分析

随着云计算、大数据、虚拟现实、深度学习等技术逐渐发展与成熟,人工智能技术也正以超出人们预想的速度发展。人工智能产品逐渐融入社会各领域,引起人们生产生活方式的深刻变革。从人工智能的发展态势可以发现,教育是受人工智能冲击最大的行业^[1]。“人工智能+教育”也成为世界各国的热点议题。2017年7月,国务院印发《新一代人工智能发展规划》明确指出,要促进人工智能与各学科的交叉融合,利用智能技术加快推动人才培养模式、教学方法改革,构建包含智能学习、交互式学习的新型教育体系^[2]。2018年4月,教育部关于印发《高等学校人工智能创新行动计划》的通知也提出,要进一步提升高校人工智能领域科技创新、人才培养和服务国家需求的能力^[3]。此外,美国于2019年2月发布《美国人工智能倡议》,欧洲委员会于2018年4月提交《欧洲人工智能》报告,英国于2018年出台《英国人工智能发展的计划、能力和志向》并启动以人工智能技术为核心的“现代工业战略”等。可见,人工智能教育应用在全球范围内引起了高度重视,各国纷纷将其提升至国家战略层面。

[收稿日期] 2021-09-10

[基金项目] 国家社科基金一般项目“全民终身学习视野下的国家在线教育体系发展研究”(20BSH053)。

[作者简介] 钱小龙(1976—),男,江苏海门人,博士,南通大学教育科学学院教授、硕士生导师,主要研究方向为开放教育资源、国际比较教育;张奕潇(1999—),女,江苏南通人,南通大学教育科学学院硕士研究生,主要研究方向为开放教育资源;宋子昀(2000—),男,山东乳山人,南通大学教育科学学院硕士研究生,主要研究方向为在线教育。

[通讯作者] 李强(1971—),男,湖北利川人,博士,南宁师范大学教育科学学院院长、教授、博士生导师,主要研究方向为高等教育学、教育基本理论。

人工智能以其智能化、自动化、个性化、多元化和协同化^[4],促进教育变革创新。然而,随着人工智能与教育的深度融合,人工智能教育的伦理挑战也逐步凸显。刘进等人认为,由于人工智能技术在教育中的实现基于海量数据,教育数据被泄露和滥用的风险急剧增加^[5]。此外,对技术过度的情感依赖^[6]正影响学生独立思考的能力。因此,有必要制定人工智能在教育应用中的伦理原则,以防止其造成更严重的负面影响。人工智能在教育领域的应用场景主要包括语言学习、教学辅助和自适应平台。这些应用都通过相应的教育人工智能系统来实现。教育人工智能系统就是基于智能算法和大数据技术,实现个性化学习、辅助教师工作^[7]、优化教育管理流程的与用户直接交互的软硬件系统^[8]。例如智能导师系统(ITS)可以模拟教师思维来指导和帮助学生学习、自适应学习系统能够提供个性化的学习方案、智能问答系统能够为用户提供个性化的信息服务。然而在系统的使用过程中,诸如知识产权与诚信的伦理问题也逐步凸显。此外,智能系统的使用必然会触及用户的隐私信息、智力、性格取向等信息,难免存在信息泄露的可能。因此,亟需专门针对教育人工智能系统的伦理原则,进一步促进人工智能教育的有效发展。

二、教育人工智能系统的一般伦理原则

教育人工智能系统是教育人工智能在特定教育场景的具体应用,作为人工智能在特定领域的存在形式,教育人工智能系统必须从宏观层面遵循一般性的伦理原则。按照逐层深入的逻辑关系,可以分为人工智能伦理原则和教育人工智能伦理原则。

(一)人工智能伦理原则

人工智能伦理通常被视为应用伦理学的一个范例,它关注的是人工智能的设计、开发、实施和应用所引起的规范性问题。过去十年左右人工智能的力量和潜力的爆炸性增长使人们认识到解决人工智能引发的许多伦理问题的重要性。关于人工智能可能会危及人类的工作,被恶意行为人为滥用、逃避责任,或无意中传播偏见歧视、从而破坏公平等伦理问题的担忧一直是世界各国研究的焦点。

若宾等人(Jobin et al, 2019)对全球 84 份人工智能伦理指南进行梳理和归纳,根据出现频率统计出 11 个最为重要的人工智能伦理原则,分别为透明、公平和公正、无害、责任、隐私、福祉、自由和自主、信任、尊严、可持续性和团结^[9]。2021 年 9 月 25 日,中国发布了《新一代人工智能伦理规范》,提出了增进人类福祉、促进公平公正、保护隐私安全、确保可控可信、强化责任担当和提升伦理素养 6 项基本的人工智能伦理原则^[10]。通过将上述政策文件中的原则加以进一步的整合和选择,将人工智能伦理原则总结为:透明、公正、无害、问责、隐私、福祉和自治。1)透明。透明原则指人工智能可公开、可解释、可理解、可溯源。具体而言,人工智能的决策原理、数据来源、运作模式和优势限制等都能予以描述、检查和再现。2)公正。公正原则指人工智能坚持全纳、包容、平等、普惠、公平与正义的理念,促进人工智能的益处无障碍地惠及全体,不分性别、种族、地区、健康和文化等差异。例如,人工智能背后的算法不应存在有意或无意的歧视性,确保人工智能在提供服务时能够适应各类人群的偏好和能力。3)无害。无害原则指在技术措施和治理策略上,人工智能不能对用户、资源、环境、社会造成负面影响。因此,应将技术稳健性和安全性^[11]作为首要考虑因素,既要保证人工智能正常运行,又要采取技术性措施防止人工智能被恶意利用从而造成损害。4)问责。问责原则指明确人工智能在开发、部署、应用时的责任,明确责任主体与责任分配等问题,其有赖于建立有效的问责机制。5)隐私。隐私原则指人工智能必须确保充分尊重用户隐私、提供数据保护、建立完善的数据治理机制,同时考虑数据的质量和完整性,并确保对数据的合法访问。6)福祉。福祉原则指人工智能应坚持以人为本,以促进发展、增加益处为导向,保障人类权益,促进社会可持续发展。例如,避免人工智能的不合理使用对

个体的人际关系、社会技能、身心健康带来的不利影响,以及对资源和环境带来的危害。7)自治。自治原则指人工智能要尊重用户的人格尊严、保护用户的基本权利,尤其是保障人类拥有充分的自主决策权,要寻求机器决策和人类决策之间的平衡,促进人工智能可控可信。例如,人工智能应支持用户自主选择平台或技术、与他人建立和发展关系的自由。

(二)教育人工智能伦理原则

当人工智能技术赋能教育之后,如何实现其可持续发展、以及发展中面临的挑战是什么等问题都亟待解决^[12]。王正平(2017)研究表明,教育伦理原则主要包括明辨是非善恶、保障教育活动相关者的利益、尊重学生并且与学生相互信任、履行教育责任与义务、维护教育公平和教育正义、提高学术水平与教育能力、尊重自由民主和平等、维护师生的基本权利等^[13]。如前所述,人工智能伦理原则主要包括透明、公正、无害、问责、隐私、福祉和自治。由此可见,人工智能伦理与教育伦理在很多方面有重叠之处,即福祉、无害、问责、自治、公正^[14]。

如果说人工智能伦理原则主要针对的是数据与算法产生的问题,那么教育人工智能的伦理原则更多侧重于教育层面的实践,为教育主体或其他利益相关者制定伦理规约。1)福祉。福祉原则指教育人工智能需保障与教育相关的所有主体的利益,为其带来教育福利,以及促进教育的可持续发展。例如,能否帮助教育管理者实施最佳的监管;能否使教师将更多的精力投入到学生能力的提升上;能否使所有学习者既发展个性又得到全面和谐发展;能否消除由于经济水平差异而带来的数字鸿沟^[15]等。2)无害。无害原则主要指教育人工智能需摒弃恶意,消解潜在风险,从而促进教育的向上、向善、向美。例如,需要保护学习者的隐私数据不被非法售卖获利,从而保障学习者的人身安全。3)问责。问责原则指需要明确教育人工智能的完整供应链以及教育中各行为主体的责任担当^[16],从而增强教育人工智能的可信赖性。因此,需要在人工智能系统透明性和可追溯性的基础上,建立明确的问责机制^[14]。4)自治。自治原则要求教育人工智能充分尊重教师和学生的基本人权,尤其是保障教师和学生的自主决策权。一方面,教育人工智能应尊重教师在教育活动中的主导地位,即教师有权决定利用人工智能教育产品进行教学的方式。另一方面,教育人工智能应尊重学生的主体地位,学生对人工智能教育产品及服务有自由选择的权利。5)公正。公正原则指教育人工智能需平等地使与教育相关的所有主体获益,尤其是减弱乃至消除对教师和学生这两大主要教育主体的边缘化和歧视化。首先,人工智能算法和数据本身是否对特殊学习群体带有歧视性;其次,如何提高可访问性使产品提供的内容与功能让包括残障人士在内的所有学习者都能无障碍访问;最后,教师在使用人工智能进行教学活动时,是否会因教师技能差异导致的“技能鸿沟”^[17]而影响教育公平。

三、教育人工智能系统的具体伦理原则

尽管一般性的教育人工智能伦理原则对于教育人工智能系统不可或缺,是教育人工智能系统存在的基础性条件。不过,教育人工智能系统偏向于教育人工智能的系统层面,教育人工智能系统的伦理原则应针对系统的开发与应用来制定。由于每个教育人工智能系统有其各自的特定功能,因此,除了一般性的伦理原则之外,教育人工智能系统还需要在微观层面根据自身的特殊性遵循具体的伦理原则,即教育人工智能系统开发与应用的伦理原则,从而更加细致和深入地规范人工智能教育实践。

(一)教育人工智能系统开发的伦理原则

就教育领域的人工智能系统而言,供应商是开发人工智能软件的公司或机构,客户通常是教育机构,用户则是学生。对于一个教育人工智能系统,供应商有必要理解客户并遵守对客户、用户的相关义务,以减少或避免不良的技术应用对客户、用户造成损害。因此需要系统开发者遵守一系列的基本

原则,保证系统在不同情况下进行决策时,都能做出既科学合理又符合社会规范、伦理的决定^[18]。

那么,教育人工智能系统的发展是否有伦理规范可循?1950年末,阿西莫夫(Isaac Asimov)在其著作《我,机器人》(I, Robot)中首次提及机器人三定律(Three Laws of Robotics)。第一定律:机器人不能伤害人类个体,也不能在目睹人类受到伤害时袖手旁观;第二定律:机器人必须服从人类的命令,除非这种命令与第一定律相冲突;第三定律:机器人必须尽可能保护自己的生存,只要这种保护不与第一定律或第二定律相冲突^[19]。此后,机器人三定律被广泛运用于教育人工智能系统的伦理分析并不断修改和补充。1992年,国际计算机学会(Association for Computing Machinery)为其成员制定了道德和职业行为准则,准则包括一般伦理准则和具体伦理准则。其中具体伦理准则与教育人工智能系统的开发人员有关,内容包括努力将专业工作的过程和产品达到最高质量、最高效率;获得并保持专业能力;了解并尊重与专业工作相关的现行法律;接受并提供适当的专业审查;对计算机系统及其影响进行全面而彻底的评估,包括可能的风险分析;遵守合同、协议和分配的责任;培养公众对计算相关技术及其后果的理解;只有经过授权才能访问计算机和通信资源^[20]。考虑到教育人工智能系统的核心价值,艾肯(Aiken)提出了两个元原则,即教育人工智能系统不应削弱学生在任何人类基本维度上的水平;教育人工智能系统至少应该提升学生在其中任意一个人类基本维度上的水平^[21]。基于此,可以总结出以下教育人工智能系统开发的伦理原则:1)系统应对学习者起到鼓励作用,而不是使学习者感到沮丧;2)系统应鼓励学生之间的协作以及师生之间的互动,以此帮助学生建立健康的人际关系;3)系统应促进学习者乐于助人、富有创造力等积极性格特征的发展;4)系统应能够筛选出对学习重要的信息和知识,以避免信息过载影响学生的智力和审美能力;5)系统应创造促进求知欲的环境,使学习者开发自己的兴趣和才能并能自主探索;6)系统应考虑人体工程学特征,利用技术手段避免学生在使用时造成的眼疲劳、重复性劳损等伤害;7)系统应能够帮助教师解决课堂管理问题、提升教学效果;8)系统应避免文化差异、语言差异而使学习者感受到歧视;9)系统应考虑学习者学习风格和技能水平的差异。

(二)教育人工智能系统应用的伦理原则

随着人工智能技术的广泛应用和人工智能教育的飞速发展,国内外涌现出诸多具有特色功能的教育人工智能系统,如适用于残疾学生的智能机器人、研究型自适应学习系统、以“拍照搜题、智能测评、智能错题本”为主要卖点的智能学习系统,以“精准测评、精准定位、精准提分”见长的自适应学习训练系统等。在应用过程中,教育人工智能系统中涉及伦理道德方面的隐患引发人们的担忧,其中包括教育大数据收集、处理等环节中涉及到的学习者隐私和安全问题;学习者年龄、性别、文化差异产生的数字鸿沟;教师教育技术能力差异产生的“技能鸿沟”;网络传播、网络复制^[22]导致的知识产权侵权或其他伦理道德方面的隐患^[23]。

不同的教育人工智能系统不仅需要遵循一般性的伦理原则,还需要根据自身的特殊性和实际环境制定更加具体的伦理原则。一方面,以纽约时报推广的人工智能教育系统为例,该系统能够以纽约时报书评的形式为任意类型的书创建书评,学生通过系统自动生成读书报告并直接上交。在伦理视角下,这与公然窃取信息略有不同,但实则是侵犯知识产权的行为。另一方面,对于智能监课系统,尽管大多数供应商承诺系统只是记录学生的表情和动作,不会进行全程录像。但在学生认知中有摄像头存在时,难免有一种隐私被侵犯的不适感。并且,基于系统的课堂使教师成为监控者和管理者,学生成为被监控者和被管理者,进而容易导致师生之间的不信任感与疏远感。学生可能会为了迎合教师而刻意表演、隐藏真实的自我,在一定程度上影响课堂氛围、降低教学效果。此外,系统的引入容易导致师生交互和生生交互的大幅缺失与迷茫。系统的引入使教师过分关注系统数据的光鲜亮丽,而

忽视与学生的交互,并且将学生之间在课堂上讨论交流的行为视为不规范行为,进而导致学生积极性被挫、学生思维发展受限和促进求知欲的教学环境缺失等不良后果。同时,由于系统尚未能够实现情绪情感数据的准确识别与学习干预,教师易忽视学生的情感状态,造成师生与生生情感交互的缺失,不利于学生建立情感联系,阻碍学生社会情感能力的发展。因此,需要完善的应用原则以防止教育人工智能系统带来更多类似的负面影响。根据各教育人工智能系统的特殊功能,在梳理现有人工智能教育伦理指南的基础上,提出以下教育人工智能系统应用的伦理原则:1)在教育人工智能系统投入使用前,建立针对系统决策的监督机制,保证人类始终参与决策并且能够推翻系统的不恰当决策;2)制定教育数据访问协议,协议应包括哪些人员可以访问学生个人数据,在什么情况下可以访问学生数据等;3)企业向学校提供人工智能教育产品的同时,应该告知用户该系统的工作原理和机制,以此保证系统的透明性;4)建立健全的问责制度。一方面,明确系统开发、部署、使用前后的责任主体;另一方面,建立审查机制,确保能够补救系统可预见的不良影响;5)学校引入产品前,对产品进行包容性测试,避免系统对性别、经济、文化、社会地位存在差异的学习者产生歧视;6)学校尊重教师的主导作用与学生的主体地位,明确系统始终只起辅助作用;7)营造促进学生求知欲、创造性的教学环境,鼓励学生进行探索;8)重视教育过程中对情绪和情感的识别和干预,注重情感联系的建立与连接。

综上所述,教育人工智能系统的具体伦理原则折射了教育人工智能系统面临的伦理问题,主要集中于人工智能技术自身、利益相关主体的问责、教育中情感的缺失以及个体的权利和权益四个领域。人工智能技术自身聚焦于支撑技术的数据和算法所面临的伦理问题,包括数据和(或)算法的可访问性、隐私边界、透明性和有意或无意的歧视性等。利益相关主体的问责指供应商、客户和用户三大主体间责任的界定和分配问题,即是否承担责任、承担什么责任以及承担多少责任,其依靠专门化的问责机制和审查机制来保障。教育中情感的缺失指由于数据和算法尚未能对情绪和情感进行准确识别,从而使教师和学生未能挖掘情感对教育的价值。个体的权利和权益主要指教育人工智能系统不能侵害个体的知识产权、决策权、隐私权、发展权等,尤其是要尊重和保护教师和学生这两大教育主体的权利和权益。例如,系统作为技术工具应坚持服务教育而非决定教育的取向、不能因为系统的“预测”功能而忽视教学交互以及阻碍学生的求知和主动探索等自我发展的权利。

四、教育人工智能系统发展的伦理挑战

当教育人工智能系统在开发和应用的过程中违背伦理原则,便会引发相应的伦理风险与挑战。面向未来教育人工智能系统的发展,其在人工智能技术自身、利益相关主体的问责、教育中情感的缺失以及个体的权利和权益四大伦理问题上分别面临数据与算法的缺陷、系统提供商权责不明、教育过程的情感遮蔽和教育主体角色转变四大挑战。

(一)数据应用与智能算法存在缺陷

大数据算法是人工智能的核心,人工智能教育应用与人工智能应用一样,在数据应用和智能算法上存在缺陷。在数据方面,教育人工智能系统收集大量数据以记录学习过程、推断情绪状态、提出应对策略,而这些数据的拥有者、有权访问者、隐私问题边界等问题都是亟待解决的。在智能算法方面,首先,智能算法存在难以预知的黑箱风险^[24]。由于算法过程的复杂性与不透明性,除少数研究人员之外,算法不被大多数人理解。容易导致算法责任归属与监管问题。其次,对算法的过度依赖影响学生能力发展。基于算法对个体作出的分析固然有一定的依据与可靠性,能够为学生推荐个性化的资源;但频繁使用算法得出的决策容易使学生产生依赖感,从而对探索新知与创新缺乏主动性与积极性,影响其创新思维与独立思考的能力。最后,算法本身缺乏精确性。学生是不断变化发展的个体,智能算

法的精准决策需要大量的学生数据,然而算法无法对学生的变化发展做出详尽的判断,此外还可能存在其他隐性的数据无法得到计算,从而加剧教育过程的不公平性。

(二)系统提供商权责不明

教育人工智能系统提供商的主要职责是系统的开发与完善。在系统开发和投入使用的整个过程中,其一,需要保证系统结果的可追溯性和可理解性。当客户对结果产生疑问时,能够完整再现数据的动态变化并确保用户易于理解。也要提供系统的限制与缺陷说明,使用户对系统可能产生的消极影响有充分的把握。其二,需要保证系统为受教育者带来福祉。谨防系统因种族、性别、文化差异对学生产生歧视或造成其他伤害,尊重学生的主体地位并激发学生的学习积极性与主动性。也就是说,在学生使用系统的过程中,既不会受到任何负面影响,又能够提升学习效果。其三,需要对系统测试与审查。在系统投入使用前,对系统从安全性、执行效率、执行质量等方面进行测试,保证系统决策科学合理。在系统投入使用后,定期审查系统决策是否符合规范,并根据客户意见不断维护和更新系统。

(三)教育过程的情感遮蔽

情感在教学过程中具有重要的教育价值,所有的学习都有情感上的关联,教育依赖于情感作为学习和个人发展的动力。人工智能为教育过程搭建数字化的虚拟环境,收集大量客观的教学过程数据并进行分析,然而并非所有数据都可以得到有效度量。教育者与受教育者的情感状态容易受到忽视,从而产生一系列伦理问题。这种情感遮蔽现象可以理解为对技术的过于关注而导致的师生情绪和感受的屏蔽,教育主体既体会不到自己的情感变化,也无法察觉相互间的情感变化^[25]。教师与学生、学生与学生之间的互动受到虚拟世界影响,情感交流被数字符号所取代。尽管现有的技术可以通过监测眼球活动、表情等对学习者的学习状态进行判断,但很难进行准确的情感识别。

(四)教育主体角色转变的认识问题

在现代教育观念中,教育主体包括教育者与受教育者,摒弃了以往教育者与受教育者主客体相对立的观念。由于教育人工智能系统的发展与成熟,人工智能教育应用已经能够扮演教师、学生、同伴、专家等丰富的角色,智能机器人的权利地位^[26]影响着教育主体的角色转变,从而产生一定的伦理风险。因此,需要正确认识教育主体角色转变,进而及时处理出现的各种问题,以便于将伦理风险降低到最低程度。其一,学习者的角色转变。其中最突出的是“本我”与“非我”的异位问题。人工智能分析提供的虚拟人格与学习者自身存在“非我”与“本我”的矛盾,尽管两者受控于同一精神主体,但自我的现实性和虚拟性难以达成主体性的和谐统一^[27]。如果难以实现从“非我”到“本我”的角色转变,将弱化真实情境的学习体验,降低学习质量。其二,教师的角色转变。各类教育人工智能系统在教育场景中广泛应用,尽管减轻了教师大部分的重复性工作,但存在弱化教师职能的伦理风险。现有的人工智能教育应用已经能够替代诊断、评价、反馈等部分教师职能,如果对其过度依赖,教师的权威地位将遭到消解。此外,如果教师存在伦理素养的缺失,在学生数据的使用过程中容易产生道德失范行为,对学生的隐私安全造成威胁。

五、教育人工智能系统发展的应对策略

审视教育人工智能系统发展的时代挑战并提出应对策略,对未来教育人工智能系统的发展走向起决定性作用,有利于助推人工智能成为未来教育发展的积极因素。2019年5月16日,国家主席习近平在国际人工智能与教育大会上指出,中国要积极推动人工智能与教育的深度融合,加快发展伴随每个人一生的教育、平等面向每个人的教育、适合每个人的教育、更加开放灵活的教育^[28]。由此可见,教育人工智能系统发展的伦理挑战需要坚持以人为本,并通过道德素养提升、准则完善、法律规约、人

为监管等手段构建应对路径。

(一)培养具备合格人工智能素养的社会公民

现如今比较典型的以松鼠 AI 为代表的智能教学系统、以小猿搜题为代表的拍照搜题系统、以句酷批改网为代表的作业评测系统、以好未来 GodEye 为代表的课堂辅助系统等教育人工智能系统广泛应用于教育场景中,为教育教学带来了极大便利。进而,有必要培养具备合格人工智能素养的社会公民以应对人工智能时代的机遇和挑战。数据应用引发的伦理问题归因于人们道德素养的缺失,而伦理道德修养是人工智能素养的核心因素。也就是说,在进行人工智能的学科教育时,要注重学生的道德素养培养与提升,将培养需求落实到课程与教学改革中。其一,在中小学阶段,除了教授人工智能的基础知识,还应向学生传输平等待人、尊重生命等抽象的道德观念。其二,在高等教育阶段,将伦理道德修养贯穿于人工智能教育应用的实践之中。既要求教师起正确的示范作用,提升职业道德行为的价值意识,对学生数据的使用遵循规章制度;又要加强学生的道德建设,如在思政教材中增加数据安全相关的法律法规内容。

(二)制定更加切合实际的教育人工智能系统伦理原则

科学的伦理原则是教育人工智能系统更好地服务于教育的制度保障。尽管国外现有的文件《教育中的人工智能:可持续发展的机遇和挑战》《可信赖的人工智能伦理准则》等都涉及人工智能教育的伦理框架,但仍需要结合我国国情和实际情况进一步细化与完善。从宏观来看,现有伦理原则并没有清晰的责权细分,我国伦理规范的建立需要区分所有利益相关者包括教育人工智能系统开发者、供应商、学校、教师等的责任与义务。在微观角度,可从数字鸿沟、数据隐私、可追溯性、多样性与非歧视性、人际等诸多方面做出相关规定。以数据隐私为例,系统开发者应利用技术对学习者的隐私数据进行加密以防止恶意传播与使用;供应商应提供包括系统如何采集、保护数据的措施的安全条例;学校在引入系统前进行严格的伦理审查;避免学习者对智能系统过度依赖而造成的身体与心理等方面的威胁;教师应遵守学校的相关规定,不随意访问、传播学生数据或利用数据进行科研活动等。

(三)构建教育人工智能系统应用过程的监管机制

一方面,关于数据与智能算法的风险防范,需要以技术研发做支撑、制度做保障。其一,加强中国人工智能教育应用政策的顶层设计。通过借鉴美国《教育法案》(The TEACH Act)、《家庭教育权利和隐私权法案》(FERPA),在我国《中华人民共和国计算机信息系统安全保护条例》等现有关于数据安全法案的基础上进一步优化。其二,建立健全系统开发、运用过程的监管与问责机制,供应商、机构、用户等行为主体都应明确各自的责任分摊。其三,鼓励各教育机构在教育人工智能系统决策过程中发挥主观能动性,设置专门的技术部门采取积极措施防止数据泄露,采用的措施从简单的口令技术(静态口令、动态口令、口令卡等)、传统生物鉴别技术、到较复杂的模糊身份鉴别的认证技术,乃至区块链技术。另一方面,针对教育主体角色转换引发的一系列理论问题,其一,强化教育主体的权责意识。学校加强对教师与学生的培训,使其明确在教育人工智能系统使用过程中的职责与义务。其二,规范化教育主体的权责划分。政府及有关部门制定相应的规范或制度,明晰教育主体的权责划分与伦理失范边界,并审查教育主体的失范行为,从而及时干预。此外,教育人工智能系统的责任与义务、问责方式都应有明确的界定。

(四)重视人文关怀的情感教育力量

教育环境中的人文关怀即关注学生的情感状态,是学生成长与发展必不可缺的推动因素。面对教育人工智能系统使用过程中人文关怀缺失的现象,需要対人文关怀的情感教育力量加以重视。其一,坚持育人的教育价值观导向。目前,我国人工智能技术的教育应用实际上只是浅层次的以认知为

导向的人工智能教育应用^[29],为了更好地推进人工智能与教育的深度融合,这就要求人工智能技术不能孤立地影响教育。智能技术的使用应充分考虑学生、教师、知识、教学方法与环境的整体作用,回归到教育活动育人的根本目标。其二,加强人工智能在情感识别、脑科学领域的研究。利用脑机接口技术和情感识别技术能够掌握学习者的认知水平、学习状态与情绪状态。加强情感识别与脑科学领域的研究,能够将教学过程有效量化。结合人脸识别等技术,更加全面、科学地监控与评估教学过程,提升教育主体的参与度并促进互动,提升学生的社会情感能力,优化教师的干预效率,从而实现真正意义上的个性化教学。

六、结语

人工智能为教育改革带来新的机遇,有助于促进教育个性化、智能化、终身化。本文从教育人工智能系统的发展历史与现状出发,梳理教育人工智能系统的一般伦理原则与具体伦理原则,探讨了教育人工智能系统发展的伦理挑战与应对策略。尽管现有的人工智能伦理文件对教育人工智能系统的伦理问题进行了一定程度的制度规约,但现有的伦理原则并不足以覆盖全部的伦理问题且约束力不足。围绕数据与算法、权责界定、情感遮蔽以及教育主体角色转变等伦理挑战,不仅需要完善伦理框架,还需要从道德素养提升、完善监管机制、重视情感教育等方面探寻教育人工智能系统的发展路径,以充分挖掘人工智能的巨大潜能,实现育人的最终目标。面向未来的智能时代,人工智能势必推动教育生态的进一步重塑。我们应始终秉承技术服务论和人本论,并以未来之眼审视教育人工智能系统面临的伦理问题,对可预见的伦理风险进行防范。此外,关于教育人工智能是否能够成为教育主体之一,其权利与地位问题还有待进一步思考。

[参 考 文 献]

- [1] 李德毅,马楠,秦昆. 智能时代的教育[J]. 高等工程教育研究,2018(5):5-10.
- [2] 国务院. 国务院关于印发新一代人工智能发展规划的通知[EB/OL]. (2017-07-20)[2020-12-02]. http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm.
- [3] 中华人民共和国教育部. 教育部关于印发《高等学校人工智能创新行动计划》的通知[EB/OL]. (2018-04-10)[2020-12-05]. http://www.moe.gov.cn/srcsite/A16/s7062/201804/t20180410_332722.html.
- [4] 梁迎丽,刘陈. 人工智能教育应用的现状分析、典型特征与发展趋势[J]. 中国电化教育,2018(3):24-30.
- [5] 刘进,钟小琴,李学坪. 教育人工智能:前沿进展与机遇挑战[J]. 高等工程教育研究,2020(2):113-123.
- [6] 王佑镁,宛平,赵文竹,柳晨晨. 科技向善:国际“人工智能+教育”发展新路向——解读《教育中的人工智能:可持续发展的机遇和挑战》[J]. 开放教育研究,2019(5):23-32.
- [7] 肖睿,肖海明,尚俊杰. 人工智能与教育变革:前景、困难和策略[J]. 中国电化教育,2020(4):75-86.
- [8] 曹培杰. 人工智能教育变革的三重境界[J]. 教育研究,2020(2):143-150.
- [9] Jobin A, Ienca M, Vayena E. The global landscape of AI ethics guidelines [J]. Nature Machine Intelligence, 2019(1):389-399.
- [10] 国家新一代人工智能治理专业委员会. 新一代人工智能伦理规范[EB/OL]. (2021-09-25)[2021-10-1]. https://m.thepaper.cn/baijiahao_14657892.
- [11] ACM. ACM code of ethics and professional conduct[EB/OL]. (2018-06-22)[2020-07-08]. http://tareksobh.org/online_courses/cpe300/Handouts/Handout7/ACM_Code_of_Ethics.pdf.
- [12] 任友群,万昆,冯仰存. 促进人工智能教育的可持续发展——联合国《教育中的人工智能:可持续发展的挑战和机遇》解读与启示[J]. 现代远程教育研究,2019(5):3-10.
- [13] 王正平. 教育伦理学的基础理论探究与建构[J]. 上海师范大学学报(哲学社会科学版),2017(6):24-33.
- [14] 邓国民,李梅. 教育人工智能伦理问题与伦理原则探讨[J]. 电化教育研究,2020(6):39-45.

- [15] Cruz-Jesus F, Rosalia Vicente M, Bacao F, et al. The education related digital divide: an analysis for the EU-28[J]. *Computers in Human Behavior*, 2016(56):72-82.
- [16] 沈苑, 汪琼. 人工智能在教育中应用的伦理考量——从教育视角解读欧盟《可信赖的人工智能伦理准则》[J]. *北京大学教育评论*, 2019(4):18-34, 184.
- [17] 王美, 随晓筱. 新数字鸿沟: 信息技术促进教育公平的新挑战[J]. *现代远程教育研究*, 2014(4):97-103.
- [18] 刘德建, 杜静, 姜男, 黄荣怀. 人工智能融入学校教育的发展趋势[J]. *开放教育研究*, 2018(4):33-42.
- [19] Anderson S L. Asimov's "three laws of robotics" and machine metaethics[J]. *Ai & Society*, 2008(4):477-493.
- [20] Gotterbarn D W, Brinkman B, Flick C, et al. ACM code of ethics and professional conduct[EB/OL]. (1992-10-16)[2020-12-10]. http://tareksobh.org/online_courses/cpe300/Handouts/Handout7/ACM_Code_of_Ethics.pdf.
- [21] Aiken R M, Epstein R G. Ethical guidelines for AI in education: starting a conversation[J]. *International Journal of Artificial Intelligence in Education*, 2000(11):163-176.
- [22] 李冰, 文卫华, 谷俊明. 德国数字版权法律制度的发展[J]. *现代出版*, 2014(1):78-80.
- [23] 赵慧琼, 姜强, 赵蔚. 大数据学习分析的安全与隐私保护研究[J]. *现代教育技术*, 2016(3):5-11.
- [24] 谭九生, 范晓韵. 算法“黑箱”的成因、风险及其治理[J]. *湖南科技大学学报(社会科学版)*, 2020(6):92-99.
- [25] 冯锐, 孙佳晶, 孙发勤. 人工智能在教育应用中的伦理风险与理性抉择[J]. *远程教育杂志*, 2020(3):47-54.
- [26] 陈全真. 智能机器人权利存在的由因及对策[J]. *贵州师范大学学报(社会科学版)*, 2019(3):144-151.
- [27] 孙田琳子. 虚拟现实教育应用的伦理反思——基于伯格曼技术哲学视角[J]. *电化教育研究*, 2020(9):48-54.
- [28] 习近平. 习近平向国际人工智能与教育大会致贺信[EB/OL]. (2019-05-16)[2020-12-10]. http://www.xinhuanet.com/politics/leaders/2019-05/16/c_1124502111.htm.
- [29] 郑勤华, 熊潞颖, 胡丹妮. 任重道远: 人工智能教育应用的困境与突破[J]. *开放教育研究*, 2019(4):10-17.

(责任编辑: 闫卫平, 赵磊磊)

Ethical Principles and Dilemma Breakthrough of Educational Artificial Intelligence System

QIAN Xiao-long¹, ZHANG Yi-xiao¹, SONG Zi-yun¹, LI Qiang²

(1. College of Education Science, Nantong University, Nantong Jiangsu 226019;

2. College of Education Science, Nanning Normal University, Nanning Guangxi 530001)

Abstract: In the process of promoting educational reform, artificial intelligence technology also brings unpredictable ethical risks to education. As an important product of artificial intelligence educational application, educational artificial intelligence system is highly respected, as long as it follows relevant ethical principles. Firstly, this paper discusses the general ethical principles of educational AI system from the macro level, that is, through the analysis and interpretation of relevant policy documents around the world, the ethical principles of AI and E-AI are clarified. Secondly, the specific ethical principles of educational artificial intelligence system are discussed from the micro level, that is, the ethical principles of the development and application of educational artificial intelligence system are clarified through the learning and reference of relevant international standards and norms. Finally, in view of the future development of AI education and possible ethical challenges together with the national conditions, this paper puts forward strategic suggestions for the future development of China's educational AI system which is to foster social citizens with qualified AI literacy, to formulate more scientific ethical principles of educational AI system, to construct the supervision mechanism for the application process of artificial intelligence system in education and to pay attention to the emotional education power of humanistic care.

Key words: artificial intelligence; application of artificial intelligence in education; ethical principles; educational AI system