

人工智能在教育应用中的伦理风险与理性抉择^{*}

冯 锐 孙佳晶 孙发勤

(扬州大学 新闻与传媒学院, 江苏扬州 225009)

[摘 要] 教育是人类社会文明发展的重要领域,人工智能在教育领域的应用,对教育产生了革命性的影响;同时,也给教育伦理带来了一系列风险和挑战。人工智能在教育应用中导致的伦理困境和可能面临的伦理风险,主要表现在人工智能与教育主体的权利嬗变、算法推荐与学生个性的发展异化、人工情感与人机互动的情感危机、智能感知与教育数据的价值困境这四个方面。为了防范和消除这些伦理风险,“以人为本”“以德为先”“以法为界”是人工智能在教育应用中的理性选择,教育责任则是“可信赖的人工智能应该做什么”伦理准则的实践指向。

[关键词] 人工智能;智能教育;教育伦理;教育大数据;人机伦理;技术异化

[中图分类号] G420 **[文献标识码]** A **[文章编号]** 1672-0008(2020)03-0047-08

DOI:10.15881/j.cnki.cn33-1304/g4.2020.03.005

新一代人工智能技术的飞速发展,正在引领新一轮科技革命和产业变革,引发和重构新的行业结构和行业形态。教育是促进人类文明发展的重要领域,而人工智能技术在教育中的应用,正在推动互联网教育进入智能教育发展的高级阶段^[1]。利用人工智能技术提高教育服务水平,促进教育的创新发展,让人类教育变得更加智慧,这是人们对人工智能带给教育变化的期待和愿望。人工智能作为一种沉寂了几十年,然后近几年才厚积薄发的一种尖端技术,它引领和促进教育创新发展的潜力毋庸置疑。但是,人工智能在教育领域的创新性应用,也冲击和改变了整个教育系统生态和教育秩序,尤其是教育文化、教育中的人伦关系、教育结构和教育价值等诸多方面。由此,也产生了一系列伦理问题,如,削弱教师地位,侵犯学生自由,加大教育不平等,对教育正向价值的压制,对教育育人价值的僭越,教育对人工智能的依附等等。

华东师范大学唐汉卫教授认为,人工智能的深入运用将在本体意义上动摇我们原有的对教育的理解、批判和追求,从而产生前所未有的困惑和焦虑^[2]。人工智能在教育应用上的伦理问题,已经开始日益显现,例如,2019年发生的“监测头环”事件、智

能音箱劝主人“自杀”事件,2020年1月“钉钉被小学生逼疯”事件等,这些看似偶尔发生的事件其实并非偶然,它警示人们要开始注意和防范人工智能在教育应用中可能带来已知和未知的一些风险。所以,只有对人工智能在教育应用中所引发的社会和伦理问题进行研究,从而在一定的教育伦理建制内对其进行有效规范,才能发挥其对教育积极的一面,避免消极的一面;才能有效实现人工智能推进教育智能化发展的美好愿景。

一、人工智能给教育带来的“技术陷阱”

在教育领域,从来不缺乏利用新技术进行变革的热情与野心。长期以来,教育的规模化和程序化模式饱受诟病,人们一直在积极期盼和探寻着通过技术的驱动,能给这种工业化教育体系带来变革和创新。因此,人工智能从其诞生起就被赋予了更多教育技术的内涵和社会期待。从技术进化论看人工智能的创新发展,人工智能从20世纪50年代艾伦·麦席森·图灵(Alan Mathison Turing)完成的图灵测试(The Turing Test)开始,已经走过了60多年的发展历程。在这60多年的创新发展之中,人们一直在追求着一个最具雄心又很冒险的终极目标,即

^{*} 基金项目:本文系江苏省高等教育教改研究立项课题“江苏高校专业课程SPOC教学改革研究与实践”(项目编号:2017JSJG447)的系列成果之一。



创造一个能够学习和应用智能来完成各种各样任务的智能机器人^[3];从早期的教学机器、智能教学系统,到如今的自适应学习系统、智能教育机器人等等,人工智能从面对确定的环境进行有效的经验逻辑推理,发展到拥有在不确定性环境中做出准确判断、决策和行动的能力。“人工智能从简单地模拟人类的逻辑思维开始走向学习、思考;并从简单的机械行为,走向人机、机器与环境等的双向甚至是多向度的互动;从应对单一的确定性环境,走向了复杂的不确定性环境的发展路径”^[4]。

自20世纪90年代以来,随着计算机、互联网、物联网、云计算、大数据、自然语言处理等技术突飞猛进的发展,信息技术已经从2G发展到5G,从有线连接发展到无线连接,从二维空间发展到三维空间。互联网的快捷性、便利性、清晰度和诉诸不同感官的信息形态不断发生的革命,加上数字化、云计算、可穿戴计算、人工智能的整体性突破,正在推动着学校教育向泛在教育、集体化学习向个体化学习、纸笔学习向屏幕学习的转变。工业革命以来标准化教育的弊端,有可能被彻底扭转,人们梦寐以求的从“以教定学”向“以学定教”的转变,真正成为了可能^[5]。然而,人们在利用人工智能促进和推动教育创新发展的同时,人工智能的教育应用却可能引发教育主体权利嬗变、学生个性化、情感危机、教育价值丧失等“技术陷阱”。

(一)人工智能与教育主体的权利嬗变

教育主体是相对于客体而言的,它是教育过程中的教育实践的组织和实施者^[6]。对于教育主客体关系的认识,现代教育观念抛弃了以往教育者和受教育者相互对立的主客二元关系,认为教育者和受教育者是一种主体间的交往实践关系,二者之间广泛存在着相互联系与交互影响。作为主体的个人不再排斥其他主体,不再把他人看作征服、占有的客体,人类的生存已经步入共生性的存在。以课程、教材为载体的教育内容,构成他们共同作用的客体^[7]。他们“借助于人类共同的文化、语言和劳动工具”,在主体间展开交往实践活动。

什么是人工智能?这是一个看似简单却又难以给出明确界定的概念。人们对它最朴素的理解就是人造的智能,它是社会发展和技术创新的人造产物。人们对人工智能的期待不在于让它充当“工具”,而是要最终成为能够理解人,拥有与人类类似的情感

和思维方式,并且能够成为帮助人的“顾问”^[8]。“人类助手”是人类赋予人工智能最大的社会价值和期待。一场AlphaGo和世界围棋专家展开的人机大战,让人们惊喜地发现,人工智能已经拥有可能超越人类的智能,成为取代人工或者延伸人类能力的“智能帮手”;但也惊恐地意识到,人工智能的认知和智能水平将可能超过人类而威胁到人存在社会的主体性。社会学家就非常担心人工智能的大规模应用会导致失业,引发社会结构的变化。其实,社会学家对失业担忧的实质是对人的社会主体性的担忧,人工智能在教育中的应用,同样也存在着这种担忧。

从人工智能在教育应用的发展历程来看,其在不同的应用阶段,承担的教育角色和发挥的教育功能不尽相同。历史上有很多比较著名的智能教学系统,如,SOPHIE、MYCIN、GUIDON等,这些系统在培训和教学过程中扮演了辅导者、教练、评价者等多种角色,完成了基于领域知识和已有经验进行教学决策和指导的教学任务;最新集成了新一代智能技术开发的LearnSmart、Knewton平台、松鼠AI自适应学习系统、AI助教Jill Watson以及科大讯飞的阿尔法蛋机器人、美国麻省理工大学的Tega机器人、Anki的玩具Cozmo机器人、机器人教练阿尔法2号等,都能够完成智能辅导、智能答疑、智能出题、智能评价、智能学伴等教学任务,扮演着教师、学生、同伴、领域专家等教育角色。可以说,人工智能在教育“行动中”的“智能”表现并不一定逊色于人类个体,甚至超过了人的“智能”。原本在教育中由教育者和受教育者组织和实施的一些教育教学活动,如今却渐渐被人工智能和智能机器人所替代和完成。那么,这对教育会意味着什么?

近年来,人工智能在教育中的应用,实际上已经引起了人们对教育主体的担忧和思考。主要担忧人工智能的应用是否会真的取代教师:当教师遇上人工智能,究竟会发生什么?人工智能技术出现后,教师真的会被取代吗?人工智能时代,还需要教师吗?这些疑问从一个侧面,反映了人们对于教师主体性的担忧。人类的教育本质明确地告诉人们,教师肯定不会被取代,教师职业也肯定不会消亡。但是,人工智能发展的历史告诉我们,人工智能的教育应用必然会消解和重构教师的主体地位。其实,人工智能也冲击和影响了学生的主体地位,现在开发的很多学伴机器人和智能搜索引擎,已经成为学生自主探究

学习和合作学习的得力助手和学伴,有些还具有情感陪护、数字娱乐等功能。基于人工智能的教学软件能够根据个人需求定制学习,将学习者相互联系在一起,提供获取数字材料的机会,支持分散学习,并以有意义的方式让学生参与进来^[9]。

除了对教师和学生主体地位的担忧之外,还有一个令人疑惑的问题——在教育实践中,人工智能能否成为教育主体?如果能成为教育主体,那么人工智能就可以跟其他主体一样,在教育活动中有意识地认识和作用于客体,并在教育实践中表现出它的主体性。果真如此的话,我们就需要更进一步思考:人工智能在教育应用中的主体间关系是一种对称的伦理关系?还是非对称的伦理关系?其他教育主体在人工智能的作用下,其主体性是被加强还是被削弱?这直接关系到各个主体的教育责任和教育权力。如果人工智能不能被视为教育主体,那么,它在教育实践之中如何被理性地用来为教育赋能,而非赋权。这些都需要对“人工智能+教育”的本真和意义,进行思考和追溯。

(二)算法推荐与学生个性的发展异化

个性化教育是当今世界教育改革发展的重要思潮之一,也是现代教育思想的一种新境界。“教育在今天只有赢得了个性和个性发展,才能赢得社会发展的未来”^[10]。在漫长的教育改革与创新实践过程中,人们把探寻和实现学生个性化发展的更多期待赋予了技术的应用;同样,也对人工智能给予极大的热情和期待。实现学生的个体化学习是人工智能在教育实践中一个重要的应用方向,智能虚拟助手、智能导师系统、自适应学习系统,这些都是根据学习者的个人特点(语言、学习风格、偏好等)所开发的个性化学习系统^[11]。自21世纪新一轮人工智能大爆发以来,人工智能因互联网强大的计算能力和数据智能算法推荐,在个性化学习方面获得了巨大的进步。

“利用大数据分析和算法推荐,能够基于学生的学习能力或技能水平动态调整课程难度或种类,能适应每个学习者的速度和进度并实时调整内容;或在恰当的时机提供定制化的练习,实现大规模的个性化学习”^[12]。但是,利用人工智能实现个性化教学,并不一定能促进学生的个性化发展。甚至有人悲观地预测,人工智能通过自身的深度学习以及大数据的挖掘和算法,在为人类精准推荐信息的同时,人有

可能被奴役而成为“单向度”(One-Dimensionality)的人,人的自迷和自我困厄现象将难以避免^[13]。算法推荐就是利用学生的一些行为数据,通过一些数学算法,推测出学生的学习需求和学习偏好;然后,定向给学生推送学习内容、认知工具、教学活动和学习同伴。人工智能与教育大数据分析技术的结合,使人类知识被全景拆解成可计算的数字符号,学生的学习行为、过程都能够被有效地跟踪与分析,能够根据每一个学生的学习需求和行为特征量身定制学习计划,精准推送相关学习资源。

但是,这种个性化的算法推荐极易造成“信息茧房”效应。一方面,基于算法的个性化推送,呈现给学生的内容都是学生自己感兴趣的,或者是适合于自己学习偏好的学习内容。即都是被技术判断、过滤和筛选“无关的”“多余的”“杂质的”信息之后的“纯净”内容。长此以往,学生的学习兴趣和注意力就会被锁定在既定的范围内,学生获得的知识域会被狭窄化,学习内容呈现的全面性和多样性也会被破坏,部分信息被智能遮蔽和过滤,学生的思维被算法所规制;还会造成学生认知的偏差和思维的故步自封,学生在学习之中往往会变得过度自信或者盲目自信。另一方面,算法主导的智能推荐提升了教学内容的精准传播,但无形中也使学生丧失获取信息的自主权,失去了全面获取相关学习内容的学习机会和权力,而且学生过度依赖算法推荐来获取信息,容易使其行为和思维产生惰性,这样久而久之,学生自己的主体性将被技术消解。

当前,机器学习和算法推荐在教育应用中是一个不能回避的实践问题,这种看似公平合理的人工智能应用,却存在着技术导致的偏向和歧视,而且这种歧视和偏见还非常的隐蔽,完全被个性化教学的假象所遮蔽与迷惑。现代教育追求的是以学生个性为基础的全面发展,绝不是“单向度”或片面化的发展。“当技术成为一种新的极权时,人一方面顺服技术所规定的工作流程,另一方面成为工具的附属物。人的外在附属物对人的奴役,在以某种看似合理化的形式支配着人的存在。人的异化在此是不言自明的事实”^[14]。我们不禁要问,在教育中应用个性化算法推荐是教育的初衷吗?利用它要实现个性化教学,还是个性化发展?毋庸置疑,算法推荐的教育教学应用带来了很多伦理问题,它引起的“危害”是显而易见的,将会给教育带来潜在风险。

(三) 人工情感与人机互动的情感危机

教育是人类社会的一种情感实践活动。“教育没有了情爱,就相当于没有了价值倾向和生命关怀意识,也就只剩下可以独立存在的教、训、海、化这类中性的行为”^[15]。无论是教育者,还是受教育者,情感较之于理性,更能使其感受到自我的存在而主动地参与有意义的活动。人工智能为教育者和受教育者搭建了一个虚拟化、数字化的教育拟真生态环境,大量基于虚拟、远程、物联、开放、智能、聚合、协作的技术应用在教育之中,虚拟现实教学场景成了未来师生互动体验的常态化教育场景;远程人人互动和人机互动成为教育的常态交往关系;智能机器和模拟软件是学生模拟体验的常用平台;同步或异步教学让师生能随时随地开展教学活动。当津津乐道地预测着人工智能给教育环境、教育内容、教育方式方法等方面带来变化的同时,我们也要看到其对教育者和受教育者情感方面的影响。

2020年1月的“钉钉被小学生逼疯”事件,网络上相关民意沸沸扬扬,这个事件其实是人工智能在教育中应用所引发的社会伦理问题的预警,提醒人们不要单纯地关注技术给教育带来的知性变化,更要关注给教育者和受教育者带来的情感变化。一方面,教师和学生以及其他教育主体之间的关系是能够精准地建立起来的,每一个交往的教育主体都具有平等的地位,任何一个现实主体都可以根据自己的意愿塑造符号形象和身份,在网络环境中进行自由的对话与交流;但另一方面,教育主体的身份置身在虚拟与现实的二重性之中,教师和学生的具身走在“在场”与“不在场”的教学实践之中。而且在虚拟教学中,教学主体之间以及教学主体与客体之间的交互作用,仍需要借助数字化中介系统来展开和实现。因而,它是一种“主体—中介—客体”或“人—机—人”的交互模式^[16]。这对于坐在互联网终端的教师和学生而言,容易造成认知参与、情感遮蔽这种“伪参与”现象。

情感遮蔽就是教师和学生的情绪、感受被技术屏蔽,喜怒哀乐相互之间觉察不到,一是自己觉察不到自己的情感,二是觉察不到其他人的情感。教师和学生之间最为珍贵的感情交流被虚拟世界所隔离,传播和流动的只有技术眼中的“数字符号”和一些程式化的活动设置,忽视和缺失了人文关怀的情感教育力量。

微软(亚洲)互联网工程院“小冰项目”开发总监周力曾经在接受《商学院》杂志采访时表示:“好的人工智能不仅是什么都知道,而且能跟每个用户建立起感情连接”^[17]。“高智能、低情商”的人工智能是永远不可能和人和谐共处、协同工作的,也难以和人逐步建立起信任和依赖关系。喜怒哀乐的情绪变化和语言情感表达对人来说很普通,对人工智能来说却意义重大。在人工智能领域内,新一代人工智能系统的首要任务就是需要具备“感性”的情感连接能力,这样才能以更类似真实人类的方式,满足人们普遍的心理和情感需求,从而逐步建立相互信任和依赖感。有人乐观地认为,人工智能可以捕捉到学习者不同时间、不同地点的各种数据,包括面部表情、动作、情绪等。感知人类的情绪,甚至像人类一样拥有情感,这是所有人工智能的终极目标。其实,在21世纪新一轮人工智能的大爆发之中,人工智能与“互联网大脑”的结合已经赋予人工智能机器识别或表达人类情感的基本能力。

目前,对于人工智能情感识别和情感表达的应用研究,在教育教学领域中取得了一些进步。De Vicente等人研制的数字导师系统,将情感引入了教育对话领域,通过研究与学生交流得更自然有效的方式,以便提供更好的教育交互。德国人工智能研究中心在所开发的“虚拟剧场”“虚拟市场”“对话Agent”三个项目中,引入情感模型和个性特征来帮助开发儿童的想象力及创造力^[18]。微软(亚洲)互联网工程院研发的第三代人工智能机器人助理“小冰”和“小娜”,在模拟人类情感方面迈出了一大步,这对助理机器人姊妹花拥有一些类似人的感性特征,能够逐步与助理对象建立感情依赖和信任。

但是,人类在赋予人工智能“情感”的过程中,同样面临着许多问题。2012年,上映了一部电影《机器人和弗朗克》,在这部电影中,弗朗克是一个已经退休在家孤独生活的珠宝窃贼,他的儿子为弗朗克找了一个机器人来护理他的日常生活。机器人在日益努力中赢得弗朗克的信任,和弗朗克也慢慢地建立了友谊。机器人的中心职责是保持弗朗克的身心健康和活跃,于是,它为了让弗朗克的生活富有激情,在它的参与和合作下弗朗克开始重操旧业,重拾自己当年的激情。在随后的剧情发展中,弗兰克与机器人的关系使他与儿子的关系变得复杂,甚至超过了他与儿子的关系。而且,机器人慢慢有了自己的思

维,并且开始为了弗朗克学会了说谎。两者的共同生活构成了弗朗克与机器人的记忆,也有了属于彼此间的亲密感^[19]。尽管这是一个虚构的荒诞故事,但它警示人类不得不思考“赋予了情感的机器人与人相处”这一系列伦理问题。

对于教育而言,这个电影中展现的人—机伦理问题启示我们,必须直面并思考许多重要问题。有些是技术哲学层面的,有些是教育社会学层面的,我们在这里只思考教育社会学层面遇到的一些问题,主要有:富有情感的教育机器人应该如何行动,才符合教育的价值观?它跟其他教育主体(教育者、受教育者)相处到何种程度,才符合教育规范?赋予情感的人工智能在教育应用之中,是否会引起“情感泛滥”?赋予情感的人工智能会不会因为难以自控或受人控制,导致不合理的教学行为或情感表达?这些问题,需要人们从权利以及伦理道德层面去思考。

(四)智能感知与教育数据的价值困境

人工智能的发展趋势是与世界的生命系统、万物系统结合为一体,实现万物互联、人人皆知。智能感知是人工智能最主要的发展方向之一,物联网技术、工业4.0技术、区块链技术等,为人工智能的智能感知提供了良好的技术支持。在教育领域中,随着智慧校园的建设以及各种智慧平台的开通,电脑、手机、iPad、可穿戴设备,以及人脸识别、指纹识别、一卡通、监控摄像机等设备已经被广泛使用,AI人脸识别考试、一卡通智能管理、智慧校园监控、云端课堂智慧教学、VR在线互动虚拟仿真教学等系统平台,已经构成了教育领域的智能感知体系。日常大量的教育数据通过手机、iPad、电脑、可穿戴设备等,上传到云端进行处理和分析,实现了自动识别、定位、跟踪、控制、监管和信息的交换。

人工智能是以数据驱动和认知计算为核心的信息处理、分析、判断和决策系统,通过从教育大数据中发现知识,进而根据知识进行描述性分析、预测性分析、优化分析和可视化展示,并做出智能化教育决策。所以,教育大数据使人工智能在教育创新、教学变革、决策领域的应用中,有了分析和判断的“养分”,它如同生态系统里的能量一样,在不同的教育信息管理系统、不同的教育设施、不同的教育设备之间无缝地传递和交换,通过数据挖掘等方式充分发挥其应用价值^[20]。

教育大数据一部分来源于教育实践活动中学

生、教师、管理者和家长等,一部分来源于智慧校园体系中诸多的多媒体设备、传感器等。有专家将教育大数据分为教学资源、教育教学管理、教与学行为、教育教学评估四大类;从数据结构化程度来看,其包括结构化数据、半结构化数据和非结构化数据;从数据产生环节来看,有过程性数据和结果性数据^[21]。数据本身没有好坏之分,都是教师、学生、管理者和家长在智慧教育系统中的各种活动和行为所产生的。这些数据一经产生,就脱离了自己的母体,被互联网公司或学校单位控制,它们在实际运用之中存在着很多问题,有些问题是技术异化引起的,有些是人为原因造成的。目前,人们普遍关心的是数据安全与隐私保护问题,数据泄露会暴露个人身份、隐私、归属以及名誉,挑战了人的主权和尊严,这个问题已经从法律、伦理、政治、社会、技术层面受到研究者的重视。特别是随着区块链技术在教育场景中的创新性应用,数据安全和隐私保护逐渐得到有效解决。其实,在对教育大数据的创新应用中,当前最大的问题是数据权利、数据平等、避免错误决策。

目前,教育系统的信息化建设日益完善,各种智能设备应有尽有,教师和学生的一切行为基本上都暴露在“第三只眼”的监视之下,每天的数字生活痕迹都会被永久地保存起来。例如,当下教育实践中正在普及使用的AI人脸识别系统,开发者的共识和期望是利用人脸识别技术实现无感进出控制的门禁、注册登记、考勤等功能。那么,问题在于:这些设备采集的数据应该归属于谁?这些数据是否该属于我们个人?他人,包括那些记录、存储数据的公司是否有权存储和使用这些数据^[22]?因为,即使敏感性低的数据,当把它们相关联起来时,往往也能够产生一组具有高度重要意义的数据库,甚至比原始数据集更为重要,这称之为“身份重新标识”。这种重新标识可能会被人利用进行恶意的“推论攻击”(Inference Attack),即为了不正当地获得某人的材料,通过分析数据而实行的一种数据挖掘技术,类似我国所说的“人肉搜索”^[23]。特别是涉及教师 and 学生的个人数据,这些数据有可能长期持久地存储在某个服务器上,如果不明确所有权对其加以控制,将意味着数十年来个人数据被再次激活,使用或滥用的可能性持续存在。

另外,这些教育大数据有可能用于对学生行为



的预测分析,这样很容易造成“将行为模式错当实际”的结果。因为,模式是根据过去行为而确定的,不能完全决定未来的实际行为^[24]。应用这种行为模式进行推测的方式,在商业领域有一定用武之地。但在教育领域应用就会容易给学生贴上某些标签,造成刻板印象,这对于成长中的学生而言,属于一种“预测性干扰”,具有道德伦理方面的漏洞。

2019年BrainCo公司生产的“赋思头环”,在浙江金华小学课堂上的应用就引发了广泛争议,大家都在质疑学生佩戴“赋思头环”来监控专注力,是否会让其产生逆反心理?这确实是一个值得反思的问题。如果在特定场景中把它视为一个戴在头上的“监控器”;那么,学生产生人为干预后的情绪反应是必然的。这里还有一个值得深思的问题,“赋思头环”本身是一个专注于数据的量化仪器,主要通过分析采集到学生的脑电波来表征和计算专注力的大小;那么,带着一定目的和动机采集的脑电信号,是否能真实客观地表征学生的专注力?以及是否在清洗和逻辑验证时,保证了数据的真实性和多样性?如果把采集到的数据视为学生专注力的一种体现,利用这些数据给学生构建的“数字画像”客观真实吗?是否是被技术修饰合成的轮廓?

教育大数据为人类教育带来了一场从数据到知识到行动的教育变革与创新,它是教育领域中最珍贵的资源。其自身是没有价值或对错的,它的价值都是人们在教育实践中根据自己的期待和目的所赋予的。因此,教育大数据在驱动教育教学变革与创新应用的同时,我们必须注意其在应用过程中引发的技术异化、数据权利、数据平等和信息隐私等一系列伦理问题;而且,这些问题有可能会导教育不公、人格异变等不良后果。

二、人工智能在教育应用中的理性抉择

教育不同于其他领域,人工智能的应用导致的可能不仅是技术层面的伦理风险问题,更为严重的是,它可能会摧毁人类原有的教育价值和伦理体系,引发一系列潜在的、未知的伦理风险,这将对教育带来不可估量的伤害。如何突破人工智能在教育应用中所形成的伦理困境,这确实是一个需要慎重选择和认真对待的复杂问题。一些研究者试图从技术的角度来破解这种伦理困境,但是,教育中人伦关系错位、意识与责任迷失、人机信任危机、

人性全面发展异化等这些伦理风险,是难以通过技术的创新和进步来消除的。我们必须要建立具有明确道德观、价值观指引的伦理框架,必须要有相关的法规来规制。

2019年4月,欧盟委员会发布了《人工智能伦理准则》,该准则以“可信赖人工智能”为愿景,从“人的能动性和监督”“技术鲁棒性和安全性”“隐私和数据管理”“透明性”“多样性、非歧视和公平性”“社会和环境福祉”“问责”这七个方面,来构建确保人工智能足够安全可靠的伦理准则框架^[25]。其中,一个最为明确的指示和引导,就是可信赖的人工智能应该是尊重基本人权、规章制度、核心原则及价值观。这可以说是确保技术在应用上安全可靠,避免因技术不足而造成无意伤害的根本。该伦理准则对人工智能在教育中的可信赖应用,提供了良好的伦理启示,也为构建伦理框架提供了明确的进路。对于人工智能的教育应用伦理,“以人为本”是底线伦理,“以德为先”是道义伦理,“以法为界”是规约伦理,这是防范和消除伦理风险的理性选择。

(一)以人为本的价值导向

教育是人的教育,促进人的全面发展是人类一切教育的出发点和归宿。人工智能是推进教育均衡发展、促进教育公平、提高教育质量的重要手段,是实现教育现代化不可或缺的动力和支撑。在未来的教育发展之中,人工智能与教育的融合将更加深入和广泛,人工智能在教育领域的应用将逐渐走入常态化。那么,该如何减少和规避人工智能在教育应用中的伦理风险?事实上,在教育技术的发展历程中,长期受到科学主义思想观念的深度影响,教育技术的发展存在工具理性与价值理性关系的断裂问题,表现为技术思维的僭越、人性发展的圈囿和情感沟通的缺失等^[26]。人工智能在教育领域中的应用同样也深陷技术陷阱之中,从表面上看是人工智能对教育教学的赋能,而本质上却是人工智能的教育应用一味追求个性化、智能化、信息化,对工具理性的趋之若鹜和盲目崇拜,使原本“以人的全面发展为目的”的价值理性失去意义,工具理性裹挟着学生走向了人伦关系的断裂和人性的扭曲。正如叶澜教授所言:“教天地人事,育生命自觉”,教育伦理是关涉人之发展的伦理,以人的发展为中心,关注人之发展的未来性、生命性和社会性,这是构建教育伦理的价值观念和基本原则^[27]。人工智能的任何教育应用,都不

能改变这一真谛,无论人工智能多么强大,都必须服务于人的教育与人的发展这一根本宗旨。教育绝不能迷失在技术的浪潮之中,“以人为本,从学生的全面发展出发”来推进人工智能的应用,才是人工智能引领和促进教育创新的必然追求。

(二)以德为先的道义担当

在教育实践领域中,人们对人工智能变革和创新教育的社会期待越来越大,这势必使人工智能在教育中的应用与发展,不断面临着实然与应然、理性与情感、知识与德性的伦理困境。当人们从价值引领、制度约束方面来探寻破解伦理困境,防范伦理风险之道时,一个重要的、核心的、本质的实践指向是教育中“可能的、是否允许的”道德伦理,即为教育者和受教育者提供行为规范,以指导教育实践。背离伦理道德的教育实践,其危害性不言而喻。人工智能是一种“智能”技术,尽管人们普遍认为弱人工智能不具备道德行为者的主体性;但是,强人工智能肯定有了一定的主体意识、情感表达与自主性,这也是人们普遍的共识。

人工智能的发展会越来越强大,在人机共存的教育场景之中,它也和人类个体一样拥有了平等的主体地位和责任。王淑芹教授指出,人工智能技术的“是”与“应该”、“道义”与“功利”、“自由”与“责任”之间的关系,是发挥好伦理学价值规范与引领作用的理论基础^[28]。无论如何,道德在人工智能的教育应用中是不应忽视的,这是人工智能在教育应用中“归于善”本质的伦理化反映,也是以德性内化信念、以道义驱动行为的责任伦理价值的彰显。我们必须逐步建立具体的、较为全面的人工智能在教育应用中的伦理规范,确立判断行为是非善恶的道德标准,这才是将人工智能引向教育“育人”发展的意义之所在。

(三)以法为界的伦理规约

人工智能在教育应用中存在的伦理风险绝对是不可忽视的,因为它造成伤害的对象一般是我们的受教育者,而且这种伤害有可能是潜在的、长久的、难以逆转的。人类要保持对人工智能在教育实践中的有效控制,防范技术异化和价值扭曲所导致的风险,就必须要建立新的伦理制度。我们知道,人工智能在教育应用中的伦理风险,主要在于设计开发和实践应用两个环节。当开发人员在开发一个基于人工智能的教育系统时,需要遵循教育价值主导下的

技术规范和科学标准,必须遵守社会伦理的基本原则。只有在人工智能相关技术的应用和发展上加以技术伦理和社会伦理的制约和限制,形成一定的伦理制度,才能使人工智能在开发和设计环节上,真正将伦理风险减少到最低程度。

在实践应用环节,未来的教育肯定是一种人机共存的新型教育关系和教育场景。人机共存的本质是协调而非欺骗,它建立在人机认知信任和情感信任的基础之上,是一个富有挑战性的过程。这一挑战不仅仅是技术伦理层面的挑战,更为关键的是社会伦理层面的挑战。教育实践中的人机共存,需要教育制度、政策和法规,这是人工智能深度、有序、科学参与人类教育教学活动的前提和保障。伦理制度是破解和消除伦理风险的最有效举措,它是以“强制性”的形式面向人类教育实践中人机关系和社会行为的一种制度性约束和规范。

人们常言,法规不前置,现实必尴尬。因此,我们认为,人工智能在教育中的应用,必须针对教育主体建立起一系列规章公约和守则,把尽可能多的教育伦理以制度化的形式严格建立起来。伦理制度是道德规范与法律规范相衔接的一种制度,它在建立和维护人工智能在教育应用中的秩序上,发挥着非常重要的作用。

三、人工智能未来的教育责任

自20世纪90年代以来,随着传感器、移动互联网、物联网、云计算、大数据分析、虚拟现实等技术突飞猛进的发展,人工智能拥有了在不确定性环境中做出准确判断、决策和行动的能力。目前,人工智能处理和应对复杂问题的认知行为和智能水平在一定程度上与人脑相近,甚至会超越人类的智能水平,它已经在社会各个领域的应用中显示出了巨大威力和社会价值,也一再点燃和激活了人们对人工智能在教育应用中的热情和期待。正如李政涛教授所言:“技术拥有改变世界、改变人类和改变教育的伟力,它已经成为具有强大自主性的存在”^[29]。

面对人工智能有可能给人类教育带来的种种风险和危害,人们呼唤和期待通过重建教育伦理体系,重构教育伦理秩序来消除技术陷阱,化解伦理风险,除了价值前提和规范前提之外,责任前提则是“应该做什么”的伦理道德准则指向,这是对人工智能在教育应用中的应然要求,凸显其变革教育的使命感和



责任感。总之,任何一种技术,无论其如何应用,只有明确承担起自己的责任,这一责任才会成为其持续发展的驱动力和规制力。

[参考文献]

- [1]任友群,冯仰存,郑旭东.融合创新,智能引领,迎接教育信息化新时代[J].中国电化教育,2018(1):7-14+34.
- [2]唐汉卫.人工智能时代教育将如何存在[J].教育研究,2018(11):18-24.
- [3][19]Burton E,Goldsmith J,Koenig S,et al. Ethical Considerations in Artificial Intelligence Courses[J]. AI Magazine,2017(2):22-34.
- [4]孙振杰.关于人工智能发展的几点哲学思考[J].齐鲁学刊,2017(1):77-81.
- [5]袁振国.科学问题与教育学知识增长[J].教育研究,2019(4):4-14.
- [6]李富旺.教育法理要点[M].天津:天津教育出版社,2004:305-306.
- [7]冯建军.主体教育理论:从主体性到主体间性[J].华中师范大学学报(人文社会科学版),2006(1):120-126.
- [8][17]石丹.发现“机器之心”[J].商学院,2015(11):60-63.
- [9]Woollf B P,Lane H C,Chaudhri V K,et al. AI Grand Challenges for Education[J]. AI Magazine,2013(4).
- [10]戴汝潜,宛士奇.个性发展与教育改革实验[J].教育研究,1989(7):22-26.
- [11]杨现民,张昊,郭利明,林秀清,李新.教育人工智能的发展难题与突破路径[J].现代远程教育研究,2018(3):32-40.
- [12]金慧,胡盈滢,宋蕾.技术促进教育创新——新媒体联盟《地平线报告》(2017 高等教育版)解读[J].远程教育杂志,2017(2):3-8.
- [13]吴海江,武亚运.人工智能与人的发展——基于马克思人学理论的考察[J].学术界,2019(3):76-82+231.
- [14]周国文.人之解蔽:超越工具化的技术宰制之可能[J].社会科学辑刊,2007(6):44-49.
- [15]刘庆昌.教育是一种情感实践[J].河南师范大学学报(哲学社会科学版),2017(4):149-157.
- [16]罗儒国,吴青.论教学活动的虚实二重性[J].山西大学学报(哲学社会科学版),2018(1):136-143.
- [18]王国江,王志良,杨国亮,王玉洁,陈锋军.人工情感研究综述[J].计算机应用研究,2006(11):13-17.
- [20]余海燕.教育大数据价值挖掘对教学改革的启示[J].价值工程,2016(25):241-242.
- [21]苏令.大数据如何助教育更加公平优质[N].中国教育报,2018-04-17.
- [22]黄欣荣.大数据技术的伦理反思[J].新疆师范大学学报(哲学社会科学版),2015(5):47.
- [23][24]邱仁宗,黄雯,翟晓梅.大数据技术的伦理问题[J].社会与科学,2014(1):36-48.
- [25]胡凡刚,孟志远,庞茗月,刘永琪,刘敏,程玉.基于欧盟 AI 伦理准则的教育虚拟社区伦理:规范轮构建与作用机制[J].远程教育杂志,2019(6):41-49.
- [26]张宏.工具理性与价值理性的整合——教育技术发展的现实思考[J].教育研究,2016(11):28-32+53.
- [27]樊浩,田海平.教育伦理[M].南京:南京大学出版社,2000:39.
- [28]技术时代的伦理:困境与抉择[N].中国社会科学报,2019-11-12.
- [29]李政涛.现代信息技术的“教育责任”[J].开放教育研究,2020(2):13-26.

[作者简介]

冯锐,博士,扬州大学新闻与传媒学院教授,博士生导师,研究方向:网络教育应用,新媒体开发与应用,信息化教学设计、学习科学与技术设计;孙佳晶,扬州大学新闻与传媒学院在读硕士研究生,研究方向:人工智能教育应用,信息化教学设计,STEM 教育;孙发勤,博士,扬州大学新闻与传媒学院讲师,研究方向:MOOC 建设质量评估,教育大数据,教育数据挖掘与学习分析。

Ethical Risk and Rational Choice of Artificial Intelligence in the Application of Education

Feng Rui, Sun Jiajing & Sun Faqin

(School of Journalism and Communication, Yangzhou University, Yangzhou Jiangsu 225009)

[Abstract] Education is an important field in the development of human civilization. The application of artificial intelligence in the field of education has revolutionized education, and it also brings a series of risks and challenges to educational ethics. The ethical dilemma and possible ethical risks caused by the application of artificial intelligence in education are mainly in four aspects: right transformation of artificial intelligence and educational subjects, alienation of algorithm recommendation and the development of students' personality, emotional crisis of artificial emotion and human-machine interaction, value dilemma of intellisense and educational data. To prevent and eliminate ethical risks, "people first", "morality first" and "law as boundary" are rational choices, and the educational responsibility is the practical direction of the ethical code that "what reliable artificial intelligence should do".

[Keywords] Artificial Intelligence; Intelligence Education; Educational Ethics; Educational Data; Human-machine Ethics; Technology Alienation

收稿日期:2020年3月22日

责任编辑:吕东东