

# 教育人工智能伦理问题与伦理原则探讨

邓国民<sup>1</sup>, 李梅<sup>2</sup>

(1. 贵阳学院 教育科学学院, 贵州 贵阳 550005; 2. 北京开放大学 教育学院, 北京 100081)

**[摘要]** 文章旨在揭示智能化时代的教育将面临哪些新的伦理问题,并在人工智能伦理和教育伦理的基础上,探讨教育人工智能的伦理原则。运用文献研究和逻辑推理方法,系统梳理了教育人工智能应用的特点和优势,及其可能引发的新的伦理问题:(1)角色层面,系统创建者、教师、学生和监测员等利益相关者将面临更加复杂的伦理困境;(2)技术层面,涉及自动化决策伦理问责、算法伦理和数据伦理等方面的问题;(3)社会层面,可能转变就业市场、改变人际关系和引发新的社会不公平。因此,需要对教育人工智能伦理问题的复杂性有充分的认识,同时需要遵循一般的人工智能伦理原则和教育伦理原则,其中最核心的内容为二者之间的交集,具体包括:(1)福祉;(2)是非善恶;(3)公平正义;(4)人权和尊严;(5)自由自治;(6)责任和问责。

**[关键词]** 教育人工智能; 人工智能伦理; 教育伦理; 数据伦理; 算法伦理; 伦理原则

**[中图分类号]** G434 **[文献标志码]** A

**[作者简介]** 邓国民(1981—),男,湖南醴陵人。教授,博士,主要从事信息化环境下的教师专业发展和教育人工智能研究。E-mail: dam1981@163.com。

## 一、引言

当前,人工智能、大数据和物联网等新技术的发展正在推动人类社会从信息时代进入智能化时代。信息时代的教育仍然带有明显的工业化特征:借助现代信息技术,提高教育资源创建、传递和共享的效率,打破时空的限制,实现人才培养的批量化和效率的最大化。然而,信息时代的教师资源仍然短缺,而且教师深陷于大量琐碎的日常工作中,使得真正意义上的因材施教难以实现。

教育人工智能是在教育领域应用的人工智能技术及系统,能够凭借其高度智能化、自动化和高精准性的数据分析处理能力与主动学习能力,承担一部分过去只有教师才能够胜任的工作,从而有效地解决信息时代的教育所面临的一系列问题:(1)利用基于自然语言处理和深度学习的自动问答技术,能够自动回答学生

提出的问题<sup>[1]</sup>;(2)自动评价学生的学习与作业,并提供符合其认知状态的反馈和支持<sup>[2]</sup>;(3)利用情感检测技术自动识别学生的身体动作和面部表情,识别学生学习过程中的情感需求,提供个性化情感支持<sup>[3]</sup>;(4)模拟教师、管理者、学习伙伴、合作者和竞争者等不同的角色,促进交流与合作<sup>[4]</sup>;(5)构建自适应学习环境,分析和诊断学习者的学习风格、学习需要,为他们提供自适应学习支持服务<sup>[5]</sup>;(6)利用决策管理技术辅助教师的教学决策和学习者的学习决策<sup>[6]</sup>。因此,教育人工智能能够更有效地支持个性化教学,甚至在某些学习支持服务方面的能力已经超越了人类教师。

然而,教育人工智能强大的数据整合与分析能力,同时也会引发一系列新的伦理问题,例如:可能会给用户的私人数据和隐私保护带来潜在的风险;模糊教师、学生和人工智能代理等角色之间的界限。因此,教育人工智能的设计、开发和教学实践应用都需要对

基金项目:2019年度贵阳市科学技术局—贵阳学院科技专项资金项目“教育人工智能伦理研究”(项目编号:GYU-KYZ(2019-2020)JY-13);2018年度北京市教育委员会社科计划重点项目“基于学习分析的在线协作学习诊断与干预研究”(项目编号:SZ201851160033)

其伦理问题的复杂性有充分的认识,并认真思考应该遵循何种伦理原则。

## 二、教育人工智能面临的伦理问题

### (一)利益相关者伦理问题

探讨教育人工智能的伦理问题,首先需要确定其涉及哪些利益相关者。汤姆塞特(Tomsett)等人从人工智能的可解释性或透明度的角度,定义了人工智能生态系统中的六种不同的角色,分别是系统创建者、系统操作员、决策执行者、决策主体、数据主体以及系统监测员<sup>①</sup>。教育人工智能是在教育领域应用的人工智能技术及系统,其利益相关者同样应该包括这些角色。首先,创建者包括开发教育人工智能的软件公司、员工及与其合作的相关专业人员和研究人员。其次,教师和学生均承担了操作员、执行者、决策主体和数据主体等角色,因为在教学过程中,教师起主导作用,学生具有主体地位,他们都会直接与教育人工智能进行交互,为系统提供输入并接收系统输出的信息。师生均会在教育人工智能输出的信息的基础上作出决策,而且会受到对方决策的影响。他们的个人信息均可能用于训练教育人工智能,丰富和完善其智能导师代理、学习者模型和教学法模型。最后,监测员是指对教育人工智能进行审查的代理机构或人员,由他们对教育人工智能的应用效果和可能产生的潜在风险进行审查。因此,在构建教育人工智能的过程中,需要同时考虑系统创建者、教师、学生和监测员等利益相关者的伦理问题。

#### 1. 系统创建者伦理问题

教育人工智能的创建者包括系统设计师、软件架构师、程序员等设计开发人员。他们在设计、开发教育人工智能的过程中,需要在综合考虑其教学应用效果的同时,兼顾教师和学生的利益。他们需要在保证教育人工智能的稳健性、安全性、可解释性、非歧视性和公平性的同时,支持教师和学生的主观能动性,重视师生的身心健康和学生综合素质的全面发展,而且需要能够支持问责追因。为了满足这些方面的要求,教育人工智能的创建者应该清楚各自承担了什么样的伦理责任。首先,设计师需要了解不同用户之间存在的差异,真正基于善良的动机设计产品,避免任何形式的伤害与歧视。软件架构师需要构建安全、可靠和透明的系统架构体系,在保护用户个人数据安全、保障系统可靠运行的同时,更好地支持问责追踪。而对程序员来说,需要设计无偏见的算法,使其能够公平、公正地对待每一位

用户。

#### 2. 教师伦理问题

人工智能技术将会实现很多职业的自动化。教师的日常工作,如教学评价、教学诊断、教学反馈以及教学决策中的很大一部分都将由教育人工智能分担,教师角色将会不断深化,最终发生转变。一方面,教师将从一些重复的、琐碎的日常事务中解放出来,能够有更多的时间和精力投入到更具创造性、人性化和关键性的教学工作中去,例如:创造和部署更加复杂的智能化学习环境,利用和扩充自身的专业知识,在恰当的时间为学习者提供更有针对性的支持。另一方面,对技术重要性的加持,将对教师工作及角色产生直接影响,例如:可能会掩盖教师的个人能力、知识及人格等方面的重要性,削弱教师的主观能动性<sup>②</sup>。因此,在智能化教学环境下,教师将面临一系列新的伦理困境:(1)如何定义和应对自己新的角色地位及其与智能导师之间的关系;(2)利用教育人工智能采集和使用学生的学习行为数据为其提供教学支持服务的过程中,该如何保护他们的隐私及人身安全;(3)在新的教学实践环境下,教师应该具备什么样的伦理知识,需要遵循什么样的伦理准则以作出伦理决策。因此,教育人工智能将对教师的能力素质与知识结构提出更多新的要求,他们不仅需要具备一定的关于教育人工智能设计、开发和应用方面的能力与知识,而且还需要具备一定的伦理知识,使其能够以符合伦理原则的方式将教育人工智能融入教育实践<sup>③</sup>。

#### 3. 学生伦理问题

智能化时代对学生的学习、能力、素养提出了更多新的要求,也使学生面临的伦理问题更加复杂化。第一,人工智能可能改变未来的就业市场,就业导向的教育很可能过时,因此,学习者对于自己所学习的内容和目标都需要进行重新思考<sup>④</sup>。第二,人的神经系统是可塑的,教育人工智能能够给学习者提供个性化的学习支持服务,但同时也可能会剥夺他们的很多思维训练和学习体验的机会,改变他们的大脑结构和认知能力,尤其是在学习者大脑发育的关键期过度使用认知技术,将可能带来不可逆转的严重后果,这对他们发展更高级的认知技能和创新能力是非常不利的<sup>⑤</sup>。第三,教育人工智能的发展,使知识产权和诚信等伦理问题变得更加微妙和复杂。如果学生利用教育人工智能自动生成论文、作业或作品,其中涉及的知识产权的归属问题,目前仍缺乏明确的伦理规范<sup>⑥</sup>。第四,教育人工智能的使用会不会给学习者带来身体的重复性劳损、视力问题和肥胖问题等,会不会影响他们的视觉、触觉和

听觉等感官的敏感性,会不会剥夺他们体验简单快乐生活的机会<sup>[11]</sup>。第五,教育人工智能可能会剥夺大量的学习者与教师和同伴交流的机会,这会不会影响他们参与有意义的人类社会活动的机会。第六,教育人工智能的认知能力、计算能力和信息处理能力在某些方面远远超越了人类学习者,会不会让他们变得沮丧,降低他们的学习动机,使其不再愿意去主动解决问题和进行理性思考。

#### 4. 监测员伦理问题

在教育人工智能场景下,监测员包括制定教育人工智能标准的机构或相关人员、数据专员等,他们主要负责对系统的一致性、安全性和公平性等方面进行测试、调查和审查<sup>[9]</sup>。安全测试一般在教育人工智能部署之前进行,而审查需要检查系统过去的输出,因此,需要一个备份系统存储教育人工智能的数据、决策及解释。监测员正是通过这些数据、解释与教育人工智能进行交互,分析和审查教育人工智能作出教学支持服务决策的一致性、安全性和公平性,并提出相应的整改措施。因此,保证教育人工智能的可解释性,需要对其作出的决策、判断和学习支持服务建立相应的存储保留机制和解释方案,供监测员审查,通过第三方机构来保证教育人工智能的设计、开发和应用是符合教育教学伦理规范的。审查之后,监测员还需要向创建者提供如何改进教育人工智能的反馈,如重新训练模型、提高模型的公平性和保护学生隐私数据的安全策略。

### (二) 技术伦理问题

#### 1. 自动化决策伦理问责问题

随着教育人工智能的自主控制和自主决策能力不断增强,它们也需要意识到自身行为的伦理含义,而不是依赖于设计人员预置的规则。因此,需要给教育人工智能引入伦理原则和规则,构建智能道德代理,使其具有一定的伦理决策能力,并能够对自己的决策、行为和后果进行判断,协调各方利益,但这样也会引发出一些新的伦理问题<sup>[12]</sup>。首先,规则与算法是由设计者、程序员定制的,而不是由教育人工智能自主建构的,权威、权力和影响力会越来越多地体现在算法的开发和实施上。在各方利益冲突的情况下,教育人工智能应该优先服务于谁的利益,设计者、程序员、教师还是学生?其次,教育人工智能的应用可能会导致学习的失败,其中潜在的责任主体包括设计开发人员、教师、学生和智能代理等,出现这种情况时应该向谁问责,软件、数据、教师和学生各自应该承担什么样的责任,目前还缺乏相应的问责机制。最后,教育人工智能本身的自动化和复

杂性,意味着我们很难确定系统的决策机制是如何运行的,也很难对其进行追溯和解释。

#### 2. 算法伦理问题

目前,人工智能主要通过机器学习和深度学习算法进行学习和训练,基于过去记录的历史数据作出决策和判断。如果利用教育人工智能对学习及其学习进行评价,那么只能根据过去的标准和方法,如特定历史文化背景下的成功衡量标准来进行,这将会使学习算法不可避免地产生历史或文化的偏差与偏见,局限于依据历史案例的平均化数据进行处理,而在处理具有创造性和创新性的学习过程方面面临极大的挑战<sup>[9]</sup>。从基本的伦理观点来看,教师和学生能动性的表达需要的不仅仅是重复过去的历史选择,而是需要强调人的创造性和创新性,不能只是停留在过去。教育人工智能可能提供更多个性化和多样化的学习路径和学习资源的选择,但目前的做法一般是根据用户过去的选择和决定将其确定为不同的个人画像,然后过滤掉不符合其个人画像的学习内容和路径。这种方式可能会限制学习者对学习内容多样化的需求,忽视人的学习和发展的灵活性、多样性和不确定性。

#### 3. 数据伦理问题

教育人工智能的正常运行需要大量数据的支持,因而需要记录大量关于学习者的能力、情绪、策略和误解等方面的数据,其中涉及大量新的数据伦理问题。首先,在数据的分析和处理过程中是否会涉及数据意外泄露的问题,教师和学生应该在何种程度上控制自己的数据,保护自己不受到伤害和歧视。其次,谁拥有这些数据,其中涉及什么样的隐私问题,如果发生个人信息泄露,谁应该对此负责。最后,对于这些数据,应该如何分析、解释和共享,如何纠正可能对个别学生的合法权利产生负面影响的偏见(有意识和无意识的),尤其是在考虑到种族、社会地位和收入不平等的情况下。因此,教育人工智能的伦理原则不能被简化为关于数据或计算方法的问题,还需要考虑数据伦理和教育伦理。教育人工智能的伦理问题,实际上是数据伦理、计算伦理和教育伦理的综合应用<sup>[13]</sup>。

### (三) 社会伦理问题

#### 1. 转变就业市场和人才培养的方向

现行教育系统的关键作用之一是培养学习者参与社会经济生活的能力。教育制度的历史与工业社会的发展密切相关,雇佣劳动仍然是当今社会及其日常生活的核心组织原则。但是,人工智能将可能代替人类的很多工作,这会不会导致失业率的上升?纵观历

次工业革命,每次都会出现类似的担忧,但最后对人才的需求量反而越来越大,因为随着人口的增长和社会的发展,对产品和服务都会提出新的需求<sup>[9]</sup>。一些简单的、重复性的劳动将由人工智能承担,人类则将从事一些富有创造性、挑战性和不可替代性的工作,例如:人工智能系统的设计与开发、人工智能哲学问题和社会问题的研究等。他们将有更多的时间和精力投入自己感兴趣的领域,真正实现人性的解放。因此,未来的就业市场格局将面临根本性的转变,社会对人才的能力和素养也会提出更多新的要求,这将对人才培养的质量标准和规格带来革命性的影响。

### 2. 改变教学过程中的人际关系

教育人工智能的技术化过程可能会忽略师生之间以及学生之间的正确关系(“我一你”而不是“我一他”),弱化我们对教育的理解,进而只是将其理解为学习技能,而不是人的培养,从而容易忽略教师和教学在教育过程中的重要性<sup>[14]</sup>。教育人工智能将对师生关系产生重要的影响,它可能会削弱师生之间和学生之间的联系水平,以及他们之间建立联系的能力,同时也可能削弱教学过程和学习过程,弱化教师作为专业人员的地位。师生关系质量对于学习者的个人自尊、学习动机和挑战新问题的信心等影响学业成绩的因素都具有重要的影响。教师对学生的友善态度和及时反馈有助于他们获得自信心,促进他们健康情感的发展。而且,我们也无法保证所有的教育人工智能使用者都是基于一种良好的意图,教育人工智能的滥用和不良使用,例如:获取他人的隐私数据,模仿他人的笔迹、语音和会话风格去获取非法利益,以及将教育人工智能生成的知识、报告和艺术作品据为己有,都可能降低教学过程中人与人之间的信任感。

### 3. 引发新的社会不公平

教育人工智能的初衷是要增进教育系统内各利益相关者的福祉,让更多的教师和学生受益,需要尽量避免因为数字鸿沟而导致的新的社会不公平。教育人工智能领域的数字鸿沟,不是简单地由社会经济地位差异而导致的对技术设备拥有和访问权的悬殊引起的,而是由一系列复杂的影响因素构成,如信息素养、人工智能素养、对待教育人工智能的态度、应用能力和使用意向等,它们会影响教师和学生应用教育人工智能的目的、方法和策略,进而影响最终的教与学效果,带来新的教育不公平<sup>[15]</sup>。历史经验也表明,新技术的应用往往并不会缩小教育的不平等,反而可能扩大教育的不平等,这其中涉及技术应用的能力、效果和效率等方面的问题。

## 三、教育人工智能伦理原则

人工智能本身的自动化、自主性和能动性等方面的特征,使其同时扮演道德行动者和容受者的角色,并具有一定的伦理责任和道德决策能力。教育人工智能的应用可能会使各利益相关者面临众多新的伦理两难问题,并带来一系列新的技术伦理和社会伦理困境。因此,教育人工智能的应用绝不仅仅是一个技术层面的问题,还涉及技术与人类的关系、科技哲学及社会学等层面的问题,如果对其伦理困境的复杂性没有深刻的认识,将可能带来不可估量的负面影响。一方面,教育人工智能是人工智能技术及系统在教育领域的具体应用,需要遵循一般的人工智能伦理原则;另一方面,它又是一种具体的教育教学实践活动,需要遵循教育伦理原则。

根据对国际电气和电子工程师协会(IEEE)2019年发布的《合伦理设计:利用人工智能和自主系统最大化人类福祉的愿景(第一版)》<sup>[16]</sup>、欧盟委员会2019年发布的《值得依赖的人工智能伦理准则》<sup>[17]</sup>以及其他人工智能伦理指南文件的分析<sup>[18]</sup>,可以发现,尽管各国存在文化传统方面的差异,但在通用人工智能伦理方面表现出一定的趋同性,主要包括可解释性、公平正义、责任、隐私和数据治理、善待、非恶意、自由自治、信任、尊严、稳健性、福祉、有效性、慎用、胜任和人权等方面。教育伦理原则主要包括:(1)能否给教育活动的全体利益相关者带来福祉;(2)能否明辨是非善恶标准;(3)是否尊重教育公平和教育正义;(4)是否保护人的教育权、自由权和隐私权等基本人权;(5)是否尊重人,是否维护人的道德尊严;(6)是否履行教育的道德责任与义务<sup>[17]</sup>。可见,人工智能伦理和教育伦理在福祉、是非善恶、公平正义、人权和尊严、自由自治和责任等方面存在交叉重叠之处,如图1所示。

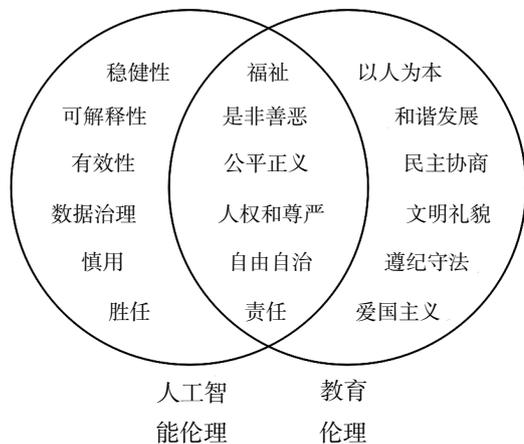


图1 人工智能伦理与教育伦理之间的关系

因此,教育人工智能伦理原则应该是二者之间的交集,代表在教育情境下,人工智能系统的技术开发与实践应用应该遵循的具体伦理原则。

### (一)福祉

福祉主要是指人工智能的设计和应用应该改善人类生活质量和外部环境,包括人类赖以生存的各种个人、社会和环境因素。它需要评估人工智能是否会让人类社会和环境变得更加美好,能否促进社会的进步,促进社会和环境的可持续发展,为任何地方的人所有并使其受益<sup>[15]</sup>。从教育领域来看,福祉主要是指教育人工智能的设计、开发和应用应该能够为教育活动的利益相关者带来福利,实现所有教育教学活动相关人员的教育利益的最大化。例如:它能否改善现有的教学环境,更好地支持学习者的学习,尤其是存在学习障碍的学习者的学习;能否提供更多、更好的学习机会,更好地做到因材施教,使所有学习者均能够得到更充分的发展;能否将教师从一些低水平的重复工作中解放出来,使其更好地投入到更有价值、更有意义、更具创造性的教育活动中去,让他们的个人价值得到更充分的实现;能否改善师生关系,使师生能够以一种更幸福、更健康的方式投入到教与学活动中去,实现个人全面和谐发展的同时,更好地保障他们的身心健康。

### (二)是非善恶

人工智能的非恶意原则主要是指其设计和应用应该是出于仁爱而不是恶意的愿望,即绝对不应该造成可预见或无意识的伤害、偏见和不公平,需要避免特定的风险或潜在的危害<sup>[16]</sup>。因此,人工智能产品的设计、开发和部署需要基于一种良好的愿望,同时还需要有一套预防风险和潜在危害的干预机制。教育伦理则强调教育行为需要基于善良的动机,遵从具有普遍价值的道德准则和崇高的道德意义,履行社会和个人的道德义务。因此,教育人工智能的设计和开发应该基于一种善良的动机,是真正为了实现教育更好的发展,为了教师能够更好地进行教学工作和履行育人的社会责任,学生能够更好地进行学习,促进他们全面和谐的发展,而不是出于其他的功利性甚至带有个人私欲的某种目的。

### (三)公平正义

公平正义原则要求人工智能对所有人都是公平公正的,并且要避免保留现有的社会偏见和引入新的偏见。因此,人工智能的决策算法和训练数据不应该沿袭设计者固有的偏见和歧视,而是应该通过评估确保这些数据对不同的社会文化均具有适当的代表性,让所

有人都有平等访问的机会<sup>[19]</sup>。教育教学活动需要维护教育的公平和正义,受教育权是每个社会公民的基本权利,不管其社会地位的高低和财富的多寡。而且,教育本身也是实现社会公平的一种方式,它为阶层之间的流动提供了一个良好的通道<sup>[20]</sup>。在部署教育人工智能系统之前,需要仔细思考会不会带来更多的教育机会,是促进教育公平,还是妨碍教育公平。首先,教育人工智能的决策算法和依托的数据是否存在偏见与歧视,能否公平、公正地对待每一位学生。其次,教师利用教育人工智能进行教学的能力水平的差异,会不会导致教学效果的差异,从而引发新的教育不公平。再次,学习者的信息素养和人工智能素养水平的不同,对于他们使用教育人工智能进行学习会不会带来学习效果的差异,从而扩大学生学业成绩之间的差距。最后,家庭社会地位和经济实力的不同,会不会造成访问教育人工智能机会的不均等,从而带来新的教育不公平。

### (四)人权和尊严

人工智能应该优先考虑并明确尊重人的基本人权和尊严。这需要在各利益相关方之间合作协商的基础上制定指导人工智能开发的人权框架,指导和规划人工智能的未来发展方向,以确定、预防和纠正人工智能可能给基本人权和尊严带来的潜在风险和危害,以维护人类的尊严和固有价值<sup>[21]</sup>。教育人工智能在教育中的应用,将可能替代教师的很多日常工作,但不能因此而动摇教师作为能够自主确立教学目标的主体地位,以及忽略教师工作本身的价值和创造性,降低教师的地位和尊严。教育人工智能给学习者提供认知支架和智能化教学支持服务,应该让学习者能够根据自己的需要和个性特征做出选择,而不是代替学生做出选择,即应该充分地尊重学生作为一个人的基本人权和尊严,而不是将其作为某种目标对象来对待。因此,教育人工智能的创建者、教师和其他工作人员都应该充分尊重师生在智能学习环境下的这些人权和尊严。

### (五)自由自治

人工智能的发展需要在以人为主导的决策和以机器为主导的决策之间保持平衡,应该促进人类的自主权和控制权,而不是伤害和破坏人类的自由。人类应该选择以何种方式赋予人工智能何种程度的决策权,以实现人类自主选择的目标<sup>[22]</sup>。因此,人工智能应该将使用者视为独立自主的人,而不是被操控的对象,不能剥夺和影响人类自由自治的权力。部署教育人工智能的学校和机构对于预置什么样的社会文化价值观念具有自治权,教育人工智能的创建者不能把自己的社会文化价值观念嵌入教育人工智能中,影响

用户对文化和价值体系的自治权。教师职业有其创造性和自主性的一面,对于自己选择什么样的教学方法,如何安排教学活动和进行教学评价,教师拥有自主权。学生对于自己在教育人工智能中的个人形象、数据和身份,拥有自主控制的权利。他们对于教育人工智能提供的教学支持服务和资源策略建议,拥有主观能动性,而且不应该由于没有遵循教育人工智能的建议而受到任何形式的惩罚。

### (六)责任和问责

当人工智能具备了一定的行为和决策自主权后,其责任和问责机制将会变得更加复杂,所以应该建立一套完善的责任和问责机制,以确保能够对人工智能及其行为决策后果进行问责,明确应该如何如何在制造商、设计师、运营商和最终用户之间分摊责任。明确的问责机制需要建立在系统透明的基础上,即应该建立完善的数据跟踪记录方案,保持系统算法和决策推理过程的透明,以便在任何情况下均可以快速找出责任主体<sup>[5]</sup>。教育人工智能,可能会代替教师的部分教学决策和教学行为,或者向他们提供教学策略建议,为他们的决策与行为提供支持。这种决策与行为后果的

责任主体和问责机制是比较复杂的,这要求教育人工智能对于自身对教育教学工作的支持和干预具有一定的预见性。对于算法的决策过程、教师使用教育人工智能进行教学的过程以及学生学习过程中的行为数据,教育人工智能要有完整的记录保存方案,从而为后续的问责过程提供依据。

## 四、结 语

毫无疑问,人工智能将对教育产生深远的影响,甚至可能带来教育教学实践领域的一场根本性的变革,转变教师的工作性质和社会角色,改变学生的学习体验。教育人工智能的出现给教育技术学带来了新的挑战,它在支持教育教学实践的过程中,体现出越来越强的自动化和主观能动性,将同时扮演道德行动者和容受者的角色。它将使教育领域的伦理问题变得更加复杂化,并引发出一系列新的伦理困境。因此,在教育人工智能的大规模设计、开发、部署和投入应用之前,需要认真考虑教育人工智能伦理原则的构建,为在教育教学实践过程中安全有效地利用教育人工智能奠定合适的伦理道德基础。

### [参考文献]

- [1] 郭天翼,彭敏,伊穆兰,等.自然语言处理领域中的自动问答研究进展[J].武汉大学学报(理学版),2019,65(5):417-426.
- [2] PERIKOS I,GRIVOKOSTOPOULOU F,HATZILYGEROUDIS I,et al. Assistance and feedback mechanism in an intelligent tutoring system for teaching conversion of natural language into logic [J]. International journal of artificial intelligence in education,2017,27(3):475-514.
- [3] MAR S,SANTOS O C,SERGIO S M,et al. Towards emotion detection in educational scenarios from facial expressions and body movements through multimodal approaches[J].Scientific world journal,2014,484873:1-14.
- [4] VIZCAINO A. A simulated student can improve collaborative learning[J]. Artificial intelligence in education,2005,15(1):3-40.
- [5] LUCKIN R,HOLMES W,GRIFFITHS M,et al. Intelligence unleashed: an argument for AI in education[M].London:Pearson,2016.
- [6] TOMSETT R, BRAINES D, HARBORNE D, et al. Interpretable to whom? a role-based model for analyzing interpretable machine learning systems [C]//Proceedings of the 2018 ICML Workshop on Human Interpretability in Machine Learning. Stockholm:CoRR, 2018:8-14.
- [7] 邓国民.教师整合技术与伦理的学科教学知识框架[J].黑龙江高教研究,2018(2):99-103.
- [8] SIAU K. Education in the age of artificial intelligence: how will technology shape learning[J]. The global analyst,2018,7(3):22-24.
- [9] TUOMI I.The impact of artificial intelligence on learning, teaching, and education: policies for the future [M].Luxembourg: Publications Office of the European Union,2018.
- [10] DAVIES C R. An evolutionary step in intellectual property rights-artificial intelligence and intellectual property [J]. Computer law & security review,2011,27(6):601-619.
- [11] 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.
- [12] SIVILL T. Ethical and statistical considerations in models of moral judgments[J].Frontiers in robotics and AI,2019,6:39.
- [13] HOLMES W,INIESTO F,SHARPLES M,et al. Ethics in AIED: who cares? [C]//20th International Conference on Artificial Intelligence in Education. Chicago:Open Research Online,2019.

- [14] GUILHERME A. AI and education: the importance of teacher and student relations[J]. *AI & Society*, 2019, 34(1): 47-54.
- [15] The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. Ethically aligned design: a vision for prioritizing human well-being with autonomous and intelligent systems[EB/OL]. (2019-03-19) [2020-03-24]. <https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead1e.pdf>.
- [16] FLORIDI L. Establishing the rules for building trustworthy AI[J]. *Nature machine intelligence*, 2019, 1(6): 261.
- [17] 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, 46(6): 24-33.
- [18] JOBIN A, IENCA M, VAYENA E, et al. The global landscape of AI ethics guidelines[J]. *Nature machine intelligence*, 2019, 1(9): 389-399.
- [19] SHAW G. *The future computed: AI and manufacturing*[M]. Washington: Microsoft Corporation, 2019.
- [20] 王本陆.教育公正:教育制度伦理的核心原则[J].*华南师范大学学报(社会科学版)*, 2005(4): 98-104, 160.
- [21] LATONERO M. Governing artificial intelligence: upholding human rights & dignity [EB/OL]. (2018-10-10)[2020-03-24]. [https://datasociety.net/wp-content/uploads/2018/10/DataSociety\\_Governing\\_Artificial\\_Intelligence\\_Upholding\\_Human\\_Rights.pdf](https://datasociety.net/wp-content/uploads/2018/10/DataSociety_Governing_Artificial_Intelligence_Upholding_Human_Rights.pdf).
- [22] FLORIDI L, COWLS J, BELTRAMETTI M, et al. AI4People—an ethical framework for a good AI society: opportunities, risks, principles, and recommendations[J]. *Minds and machines*, 2018, 28(4): 689-707.

## Research on Ethical Issues and Ethical Principles of Educational Artificial Intelligence

DENG Guomin<sup>1</sup>, LI Mei<sup>2</sup>

(1.School of Education Science, Guiyang University, Guiyang Guizhou 550005;

2.College of Education, Beijing Open University, Beijing 100081)

**[Abstract]** This paper aims to reveal what new ethical issues education will face in the intelligent age, and explore the ethical principles of educational artificial intelligence on the basis of artificial intelligence ethics and educational ethics. By means of literature research and logical reasoning, this paper systematically combs the characteristics and advantages of applications of educational artificial intelligence, as well as the new ethical problems it may cause:(1)at the role level, system creators, teachers, students, monitors and other stakeholders will face more complex ethical dilemmas;(2)at the technical level, it involves issues such as automated decision-making ethics accountability, algorithm ethics, and data ethics;(3)at the social level, the employment market may be transformed, interpersonal relationships may be changed, and a new society unfair may be triggered. Therefore, it is necessary to have a full understanding of the complexity of the ethical issues of educational artificial intelligence. It also needs to follow both the general ethical principles of artificial intelligence and educational ethics. Among them, the most core content is the intersection between the two, including:(1) well-being;(2) right and wrong, good and evil;(3) fairness and justice;(4)human rights and dignity;(5)freedom and autonomy;(6)responsibility and accountability.

**[Keywords]** Educational Artificial Intelligence; Ethics of Artificial Intelligence; Educational Ethics; Data Ethics; Algorithm Ethics; Ethical Principles