

人工智能教育应用中的伦理风险及其规避

罗生全 谭爱丽 钟奕军

摘要：随着物联网、大数据、量子计算等科技的快速发展，人类进入人工智能时代。智能时代，智能技术与教育的不断融合，在为教育带来便捷、培养时代新人的同时，也衍生出复杂的伦理问题。人工智能教育应用中的伦理风险主要表现为：“我—他”师生的传统伦理关系错乱、数据壁垒和鸿沟带来的合法权利丧失以及确定性算法和不确定性发展之间的伦理背离。究其原因，与“一切皆技术”的算法崇拜主导着当前的人工智能教育应用，人工智能教育应用与固有制度观念的时空分离以及人工智能教育应用中的多元主体责任壁垒密切相关。为合理规避这些伦理风险，促进其可持续健康发展，应坚持“以人为本”的教育理念，建立系统完善的学生隐私保护规制，构建各责任主体共同参与的人工智能教育应用制度体系。

关键词：人工智能；人工智能教育；伦理风险；规避

中图分类号：G434 **文献标志码：**A **文章编号：**2096-6024(2023)02-0079-10

教育是与时俱进的事业，具有可塑性、延展性，且是持续发展的。进入21世纪以来，随着教育与智能技术深度融合，人工智能技术引发了第四次教育革命。2017年全国两会期间，“人工智能”一词首次写入国务院政府工作报告之中。同年7月，国务院颁布的《新一代人工智能发展规划》提出要致力于构建新一代人工智能发展格局，并立足从“法律法规、伦理规范、政策”^①等方面提出相关保障措施。2019年6月，国家新一代人工智能治理专业委员会发布《新一代人工智能治理原则》，强调发展“负责任”的人工智能，突出了“和谐友好、公平公正、包容共享、尊重隐私、安全可控、共担责任、开放协作、敏捷治理”^②八条原则。2021年9月，国家新一代人工智能治理专业委员会发布《新一代人工智能伦理规范》，旨在将伦理道德与人工智能相融合，为以人工智能为活动主题的自然人、法人和其他相关机构等提供伦理指引。2022年，国务院印发《关于加强科技伦理治理的意见》，致力于完善科技伦理体系，提升科技治理能力，实现科技自立自强。由此可见，国家日益注重

作者简介：罗生全，西南大学西南民族教育与心理研究中心教授、博士生导师，主要从事智慧教育、课程与教学论研究；谭爱丽，西南大学教育学部硕士研究生，主要从事课程与教学论研究；钟奕军，成都高新区新光小学校长，主要从事中小学教育研究。

^① 新一代人工智能发展规划 [EB/OL]. (2017-07-08) [2022-05-10]. http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.html.

^② 发展负责任的人工智能：新一代人工智能治理原则发布 [EB/OL]. (2019-06-17) [2022-05-10]. https://www.most.gov.cn/kjbgz/201906/t20190617_147107.html.

构建与人工智能相关的伦理规范，以此规范相关从业者的日常行为。当前，教育不断与人工智能相融合，携手致力实现公平、全纳的教育，这也使得教育中人工智能的发展与伦理的联系越来越紧密。人工智能技术在给教育带来便捷的同时，也衍生出了复杂的伦理、法律和安全问题。譬如：侵犯学生知情同意权，削弱教师地位；拉大教育间的差距，导致教育发展不平衡、教育不平等。同时，由于目前的人工智能教育应用没有完善的技术应用规范和配套的法律法规来监管，容易产生大量的伦理问题，而这些都需要人们进行理性反思和审慎应对。因此，本文将着眼于当前我国教育的现实状况与未来的发展需求，聚焦人工智能教育应用中伦理风险的类型、风险诱导原因和风险消解策略，以期为人工智能教育事业的健康发展提出治理之策。

一、人工智能教育应用中伦理风险的界定

为了探清人工智能教育应用中伦理风险的内涵，我们必须了解伦理与伦理风险的内涵。“伦理”一词由来已久，主要是指人伦关系及其规则、规范。伦理与社会上各成分不断融合，形成了诸如医学伦理、网络社会伦理、家庭伦理等伦理分支。伦理学可分为规范伦理学、元伦理学和应用伦理学。规范伦理学试图回答关于正确行为规范的问题。规范伦理理论主要包括义务论、结果论和美德伦理。元伦理学是哲学的一个分支，它立足于解决关于是非本质的问题。应用伦理学涉及将伦理理论或原则应用于具体的现实问题。伦理风险作为现代社会衍生的一种新型风险类型，有学者认为，它是在人与人、人与社会、人与自然、人与自身的伦理关系方面由于正面或负面影响可能产生的不确定事件或条件，尤指其产生的不确定的伦理负效应，诸如伦理关系失调、社会失序、机制失控、人们行为失范、心理失衡等。^①这种负面影响一旦产生，将会给人与社会带来灾难性的后果。人工智能教育应用伦理属于教育应用伦理学的研究范畴，它在属于自己特定的教育生态中生长。其研究重点是从人自身出发，而非仅关注技术本身，关注人们对人与人、人与“类人”的关系的看法，探讨人、“类人”的行为的合理性、合法性。

综上所述，随着现今的人工智能教育中智能技术的不断应用与升级，随之产生的伦理问题也不断显露。而消解人工智能教育中的伦理问题也在人们不断追问与反思的过程中形成自在的伦理旨趣，表现为在一定的社会条件下，通过制定相关准则、标准和制度规避人工智能技术带来的风险，以此促进人、教育、智能技术的和谐共处。相较于研究旨趣及其价值的确认，梳理人工智能教育应用中的伦理风险表现、风险形成原因和风险消解策略等基础性问题更为重要，而这也是本研究的重心所在。

二、人工智能教育应用中可能的伦理风险

教育在长期发展的过程中会逐渐形成模式固化及规模固化，这种状态不利于教育持续向上发展。当前，数字技术的迅速发展使得教育领域正处于新技术变革的活跃期，因而，人们期待创造出一个能完成各种任务的职能系统，从而解决当前的教育困境。辩证唯物主义强调，事物是存在矛盾的，在看待事物的过程中要坚持用一分为二的全面的观点。^②人工智能技术的发展内部同样也存在着矛盾，我们既要看到人工智能技术为教育领域带来的种种便利，如定制个性化的学习资源、方案，实现精确教学，也要看到在人工智能技术应用过程中逐渐显露的人工智能相关监督政策不健全、技术伦理规范边界难以界定、严重依赖技术产生的算法偏差、学生某些基本技能的丧失以及智能技术自

^① 陈爱华. 高技术的伦理风险及其应对 [J]. 伦理学研究, 2006 (4): 95-99.

^② 王晓琳. 浅析矛盾的同一性及其在事物发展中的作用 [J]. 探索, 1999 (6): 69-71.

身的“类人性”等问题。虽然这些问题暂未对教育系统产生颠覆性的影响，但小范围内的个案屡见不鲜。因此，我们需要以清醒的头脑，认识到人工智能教育运用中存在的问题，如学生所佩戴的智能“防走神”头环侵犯学生隐私^①，智能机器人人脸表情识别技术不利于学生情感认知发展^②，智能教育应用造成的“智能鸿沟”造成教育的不公平^③，以及技术可能阻止学生获得第一手经验和参与学习活动等，这些都将引起人们对人工智能教育应用伦理的担忧，引发新的教育问题乃至社会问题。为聚焦人工智能教育的良好应用，我们结合教育现状从以下三个方面分析人工智能教育应用中存在的伦理风险。

（一）“我一他”师生的传统伦理关系错乱

一直以来，教师作为一种独立的社会职业，需要一定的知识基础才能进入课堂开展教学，学生想要学习知识就必须进入课堂。教师充当“权威者”的角色，拥有知识、经验，学生在教师的权威下处于被动服从的劣势地位，学生是被教师支配的客体，必须服从教师指令。因此，传统师生之间由于教师具有知识权威逐渐形成了注重师道尊严的“我一他”传统伦理关系。

伴随着人工智能技术在教育领域中的深入应用，一方面，智能技术可以帮助教师解决教学中的复杂问题，将教师从单调的日常任务中解放出来，并协助教师培养学生更高层次的思维能力。另一方面，它给教育和教师带来多方面挑战。有研究表明，在数字驱动的知识经济中，传统的正规教育系统（包括教育范式）正在经历剧烈变化。现今的教学不只局限在教师与学生二者之间，而是形成了由教师、人工智能技术、学生所组成的“三元”新型教学关系。在这种新型教学关系下，教师这种职业正在发生根本性改变。一方面，教师最主要的职责不再是充当教学关系中传授者的角色，而是成为学生在整个学习过程中辅导者的角色。另一方面，智能机器人能够将学生学习情况的诊断分析和教师个性化教学方案的制定建立在大数据和机器学习的基础上，并且超过个人有限的知识背景和认知偏好，进行客观的学习诊断和教学方案设计，这就使得教师极易面临失去“学科地位”和“教学地位”的双重困境。^④目前已经存在智能教师助理以及智能导师系统，在教学过程中主要有自动出题与批阅、课程辅导与答疑、智能教研等应用^⑤，甚至出现了“教学代理”^⑥的新概念。智能机器反应速度比人类快，因此它的教学效率、深度学习的能力以及自我改进能力等均远超人类。有学者提出，互联网技术的发展正在改变在固定的学习时间、地点进行学习的方式，这意味着教师不用亲自传授知识内容，学习也无须传授学生生活中必须的全部知识。^⑦在人机共存的智能时代，尽管智能技术能够帮助教师承担一些基础性工作，但教师的身份同样存在被智能机器逐步取代的危险。这非但没有释放教师的压力，提高他们的工作积极性，反而会因此引发教学主体价值被遮蔽的隐忧，即教师是否会被人工智能所取代，人工智能时代的教师是否还有存在的必要，等等。显然，智能技

① 小学生戴“头环”防走神？教育部门：责令暂停使用 [EB/OL]. (2019-11-01) [2022-05-10]. <http://edu.people.com.cn/n1/2019/1101/c1053-31432543.html>.

② 课堂人脸识别让孩子变“戏精”？“AI+教育”伦理难题待解 [EB/OL]. (2019-05-02) [2022-05-12]. https://edu.youth.cn/jyzx/jyxw/201905/t20190502_11942636.html.

③ 人工智能会对教育造成多大的颠覆？ [EB/OL]. (2017-12-04) [2022-05-10]. https://www.edu.cn/xxh/ji_shu_ju_le_bu/rgzn/201712/t20171204_1570992.shtml.

④ 塞尔登·阿比多耶. 第四次教育革命：人工智能如何改变教育 [M]. 吕晓志，译. 北京：机械工业出版社，2019：146.

⑤ 卢宇，马安瑶，陈鹏鹤. 人工智能+教育：关键技术及典型应用场景 [J]. 中小学数字化教学，2021（10）：5-9.

⑥ 何克抗. 教学代理与自适应学习技术的新发展：对美国《教育传播与技术研究手册》（第四版）的学习与思考之六 [J]. 开放教育研究，2017（5）：11-20.

⑦ 杨现民，张昊，郭利明，等. 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）：30-38.

术并未完全讨得人们的欢心,反而让人们在内心构筑了一条不信任的心理防线,或者说是一种对智能技术无所不能言说的理性反思。

(二) 数据壁垒和鸿沟带来的合法权利丧失

在信息爆炸和人工智能革命的时代,数据是智能的基础。学生的数据均被人工智能教育系统所收集、存储并作为教学的依据使用,数据边界越发模糊,数据冗杂、数据两端对象不清等现象均会引发数据伦理风险问题。

在数据安全方面,各教育权利主体通过智能设备自动化收集、处理,控制着教育领域内海量的数据信息,涉及众多教育管理人员和师生的隐私。^①在数据收集方面,商业教育系统提供商正在不断地利用法律漏洞和隐蔽的数据获取渠道收集数据。在数据开放方面,越来越多的行业参与者正在不断开发面向数据的商业模式。目前智能机器会在“沉静”中收集学习者的全部数据,数据的过度收集易加剧学生隐私泄露的风险。^②在智能时代的教育中,数据主体对于智能机器收集数据的类型与方法不清晰,隐私边界越发模糊,人在社会中变得越来越“透明”。比如,2019年初,在贵州试行的智能校服事件^③,就在网络上引发了人们对此类智能产品是否侵犯学生隐私的热议。在现实生活中,人们只需授权给智能教育主体,接下来的学生数据收集、存储、使用等全由智能教育主体自主进行,这就增加了数据隐私泄露的风险,从而引起外界对“数据如何使用”“超越边界”等问题的质疑。

由于目前智能机器拥有数据的绝对控制权,教师、家长、平台开发者、管理者都可以看到智能技术收集到的所有数据。在数据存储方面,防火墙漏洞、软件病毒都可能引发数据的意外泄漏,要想做到毫无漏洞的数据保护墙可谓难上加难。在人工智能教育中,如何保证所收集到的数据不外泄,谁应该拥有所收集到的数据,拥有数据的这个人是否合理合法的,这些问题都需要解决。只有这些问题得到解决,数据主体的合法权利才能得到保障。

就目前的智能教育平台而言,各大互联网教育平台在提供服务的同时也在收集教育用户的大量信息。需要警惕的是,这些互联网平台由于技术的垄断有可能实现信息的垄断。信息垄断不仅会造成数据壁垒与鸿沟,形成“信息茧房”,甚至还会削弱学生对个人数据的控制权、支配权,从而使得学生的合法权利受到损害。“人工智能的数据分析能力已经远超人力,人在人工智能面前成为无任何遮掩的透明体。”^④智能机器依靠数据对学习者的分析,但教育自身的复杂性以及智能时代学习的零碎性加大了人工智能获取完整且高质量数据的难度^⑤,数据采集不全以及学习者自身具备的发展性使得数据不准确,均会导致教育决策结果的不准确。此外,教师依据数据作出教学决策,这样很容易侵犯学生的合法权利,从而引发诸如教育不公平等一系列教育教学问题。因此,必须对相关数据以及数据使用者和经手者的使用权限范围进行明确规定,并采用安全度更高的验证方式对使用人员进行验证。

(三) 确定性算法与不确定性发展之间的伦理背离

人工智能的关键所在是算法,教育中人工智能算法的运行主要是利用学生在学习过程中产生的

① 陈磊,刘夏,高雪春.人工智能视域下教育治理的现实挑战与路径选择[J].中国教育科学,2020(6):24-30.

② 李世瑾,胡艺龄,顾小清.如何走出人工智能教育风险的困局:现象、成因及应对[J].电化教育研究,2021(7):19-25.

③ 智能校服进校园,实时定位监控等功能引争议[EB/OL].(2019-02-20)[2022-05-10].<https://www.jiemian.com/article/2877551.html>.

④ 杨建武.人工智能技术的伦理风险及其协同治理[J].中国行政管理,2019(10):44-50.

⑤ 张坤颖,张家年.人工智能教育应用与研究中的新区、误区、盲区与禁区[J].远程教育杂志,2017(5):54-63.

数据进行公式化的分析与计算，从而推测出学生学习的当前状况、需求和学习偏好，然后精确地将新的学习内容推给学生。智能数据分析技术使得学生的个体行为被分解成一个个易于理解、评判的数字符号。但由于算法背后隐藏的“技术黑箱”，人们很难理解具体算法分析教育数据的逻辑，算法也很难向人类解释它分析教育数据的具体方式。^①“算法黑箱”使得使用者对于个性化学习内容、个性化教学内容的算法生成方式一无所知，教师无法判断他依据算法作出的决定是否正确，是否符合学生要求，会不会对学生成长造成负面影响。在现实中，不论教师是否理解算法，只要智能技术被应用于教育实践，教师都要被动接受算法输出的结果。可以说，算法成为隐藏在学校教育后的控制者的概率很大，若不对算法加以规范，教师很容易失去教育自由。

算法决策的基本原理是用过去的的数据预测未来的趋势，输入数据的主要来源是智能技术对学生各类数据的记录。有研究表明，用于训练人工智能系统的输入数据中存在的任何偏差都会持续存在，甚至还有可能被放大。智能技术虽然能够根据历史性数据预测未来，这在概率统计学意义上也具有合理性，但在无形中打破了教师自行制订教学方案和学习者自行制订学习方案的自主意识，仅根据往届学生的数据或者学生之前的数据来执行教学，会使教师忽视当前学生的实际情况，不利于教育教学的自由发展。同时，由于教育活动中学生的未来发展具有不确定性、嬗变性和灵动性，依照固定公式的算法存在对学生成长的认知偏差与偏见，这与教学现实的复杂性、创造性和可持续发展相矛盾。

“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。”^②基于算法的推送，人工智能教育会在无形中使学生丧失获取信息的自主选择权以及学生的知情权，学生可能会过于依赖算法来获取信息，个性化辅导、个性化学习计划均容易使学生的行为和思维产生惰性，使其陷入“只获取不思考”的窘境，反而不利于全人教育的实现。

三、人工智能教育应用中伦理风险产生的缘由

人工智能在教育领域的广泛应用是一把“双刃剑”。一方面，人工智能技术进入教育，加速了知识生成，人们利用智能技术开发出新的教育形式。另一方面，对于智能技术应用于教育中为何会产生这样的伦理风险，我们目前所秉持的教育理念是否恰当，这些问题都需要探讨。对这些问题的回答有助于我们明确人工智能教育应用的范围和边界，厘清人工智能教育应用中伦理风险产生的导火索，是我们规避伦理风险的“基石”。

（一）“一切皆技术”的算法崇拜主导着当前的人工智能教育应用

当今世界教育发展的趋势之一是科技成为促进教育发展的关键引擎。只要出现新技术，人们都会尝试将其应用于教育，希望教育能够因此进步。人工智能教育的核心是利用算法和技术模型对教育现象进行算法推演，将教育问题抽象为数据问题，对其进行公式化处理，从而解决教育问题。这种依据数据做“决策”的运行模式会造成“一切皆技术”的算法崇拜在教育领域蔓延，教育不断被算法化。本来技术是由人所创造的，应为人类服务，但现今技术却反过来支配人类、控制人类。算法时代，由于算法的功能性、神秘性，人们将其视为无所不能的技术，人们在心理上逐渐形成类似拜物教的算法崇拜。但这些新技术的应用也一直处于不确定状态，现实生活中可能会出现无论必要与否，教育教学中都一律运用智能技术的情况。

^① 谭维智. 人工智能教育应用的算法风险 [J]. 开放教育研究, 2019 (6): 20-30.

^② 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): 42-47.

算法时代，每个人都被抽象成一组数据。具体来说，社交媒体、在线学习环境（如大规模开放在线课程）、智能辅导系统、学习管理系统、传感器和移动设备正在生成越来越多的动态和复杂数据，这些数据包含学生的个人记录、生理数据、学习日志和活动，以及他们的学习成绩和结果。智能技术根据学生的学习习惯、学习轨迹、学习偏好、知识基础等对其进行分类，为每个学生标记上不同的“标签”，赋予学生“算法身份”，学生由一个不可分割的个体被拆解成一个个便于分析的成分。教育逐渐以数据作为分析材料，进行公式化的个性推送，算法精确逐渐成了教育权威的代名词。如此一来，教师容易过于依赖智能技术的判断而丧失对学生进行教学与评价的主观能动性。此外，因为学生自身的特征是很难被数据完全涵盖的，所以在现实教学过程中，教育界不能强行运用智能技术去收集数据、分析数据、得出策略，甚至为了智能教育的发展强行让学生穿智能校服、戴智能手环，监控学生的一言一行。在教育情境中，以技术为主导或者强行全覆盖地使用智能技术，只会导致数据采集的谬误，这样的做法颠倒了技术使用与教育教学成效的主次，增加了教育风险与伦理问题，不利于教育的长期向上发展。

（二）人工智能教育应用与固有制度观念的时空分离

快速的技术进步和相对缓慢的教育采用，导致技术准备和教育应用之间的差距不断扩大。技术进步与理论制度制定的步伐同样也不一致。当前，人工智能教育应用的理论制度研究一直滞后于实践创新，造成人工智能教育实践运用与理论探索、制度制定的步伐不一，主要表现在两个方面。

第一，法律制度制定的进度与人工智能技术发展进度不匹配。人工智能不断与教育相融合，学术界和产业界在人工智能教育应用方面都投入了较大的人力与财力，以教育为中心对人工智能技术应用领域进行深入开发，使得人工智能在教育领域具有广阔且明朗的发展前景。但由于人工智能发展速度快，制度与法律完善的速度跟不上人工智能技术发展的速度，这样可能会导致大量企业进入市场出现无序竞争以及不端行为。因而，社会需要加快构建完备的法律法规体系，规范人工智能技术相关主体的权利与义务，规范各主体行为，从源头上解决人工智能教育应用的伦理规制问题。

第二，人工智能教育应用缺乏监督管理。人工智能应用的监管主体主要包括国家强制力量以及社会舆论力量。从目前的人工智能法律上看，国内更多的是从产业政策促进、扶持和发展的角度对人工智能提供行政指导，如国家新一代人工智能治理专业委员会发布的《新一代人工智能治理原则》以及《新一代人工智能伦理规范》、国务院印发的《关于加强科技伦理治理的意见》均反映了这一点。但目前国家并未明确制定人工智能的相关监督条例，现有的法律法规很难监管人工智能相关应用的各个方面，法律监督体系不够完善且威慑力不足。而且，社会大众媒体报道的人工智能教育应用中，内容大多是智能技术的新阶段、新成果，很少有关于伦理问题的专题报道。这在一定程度上暴露出社会媒体对人工智能教育应用的风险管控的监管力度较弱，对人工智能教育应用的持续健康发展不利。

（三）人工智能教育应用中的多元主体责任壁垒

长久以来，政府在教育中处于主力军的位置，集投资者、办学者和管理者于一身，政府包办和统管学校，出台相应规范、法律。之前，教育数据管理的责任主体主要包括中央政府、教育部、地方教育局、学校校长、教师。但目前，随着大数据时代的到来、计算机和互联网的发展，数据管理的责任主体增多了，智能技术的开发者和智能机器的监测者均被看作责任主体。在社会中，这些责任主体均有知晓数据的权利。人工智能教育稳步运行的关键在于多元主体责任的划分、分配与平衡，但这种责任划分与平衡的状态并不会自动出现，各主体因身份不同存在着利益分歧与责任壁垒。现今，无论是社会公众还是政府，都对智能技术主体有期待，认为智能技术主体除商业交易属性之外，

都应该有服务受众的使命，都应该承担更多教育管理的职能。然而，智能技术主体由于自身具有对数据的拥有权和支配权，更多地注重盈利而轻视服务大众的使命。如何改变现今智能技术主体秉持的这一观点，是我们现今亟须考虑的问题，若是真等到智能技术大规模充斥整个教育教学之中时再去考虑规范智能技术教育，可能为时已晚。

在智能时代，数据该由谁合法地去管理，如何去管理，按着什么规则去管理，已经成为现今教育管理方面关注的问题。但现今在数据管理方面，政府尚未出台详细的规章制度规范人工智能教育中各主体的行为。智能技术的应用使得教育决策、数据管理不再只是政府和专业人士的事情，学校、科研机构、智能技术主体等都有责任。除此之外，我们还要衡量的是智能技术这种“类人类”是否也具有相应责任，它在享受权力的同时，是否也需要承担一定的责任。另外，智能技术这一“类人类”是跟教师、学生、技术开发者这些“准人类”联系紧密的，如若智能技术收集到虚假数据或者算法出错导致“伤害”教师与学生的情况发生，这时候谁应当担负责任？人工智能教育的责任主体多元，但现今国内没有统一且明确的标准。我们亟须做的就是统一责任认知，从而规避伦理风险。

四、人工智能教育应用中的伦理风险规避

人工智能技术的广泛运用具有两面性，如何规避人工智能教育应用中的伦理风险，是一个需要特别重视且迫切需要解决的问题。因此，我们要充分认识人工智能技术的“善”和“恶”，做到存“善”去“恶”，真正发挥人工智能技术“助教”的作用。

（一）“以人为本”教育理念主导人工智能教育应用

2019年欧盟委员会发布的《人工智能伦理准则》中列出了构建“可信赖人工智能”的七个关键条件——人的能动性和监督能力、安全性、隐私数据管理、透明度、包容性、社会福祉、问责机制。该准则认为“可信赖的人工智能”的必备要素就是应尊重基本人权、规章制度、核心原则及价值观，同时应在技术上安全可靠，避免因技术不足而造成的无意外伤害。^① 欧盟这份《人工智能伦理准则》强调尊重人权，以人为本，对我们规避伦理风险具有启示意义。在我国今后的人工智能教育应用中，坚持以“人”为中心是教育持续健康发展的关键引擎，也是人工智能参与课堂教学所要遵循的基本准则。

教育是面向人的教育，它的根本目标是促进人的全面和谐发展。人工智能技术是实现教育公平，培养新时代高质量人才、师资队伍的重要手段之一，是建设教育强国的重要推动力。未来人工智能与教育会融合得更加紧密，教育中的人工智能会变得常态化。目前，人工智能在教育中的应用，强调对学生个性化的赋值、对智能资源的推送、对学情状况的诊断，这些使得智能技术在学生发展中占据主导地位，影响着人伦关系。我们不能把智能技术摆放在人类“主体”位置的上位，必须坚决地将其视为机器，将其视为只是为人类教育服务的工具。智能技术仅仅对教育的效率、效益有所帮助。

教育伦理关注人的长期优质发展，任何智能技术都不得违背这一初衷。学生是学习的主体，无论人工智能发展得多么强大，人工智能教育都必须遵从“人工智能为教育服务”的根本宗旨，坚持以学生为中心的原则。在教学中，教师不能将智能技术视为万能的，教师只能利用智能技术获取学生相关数据，并且只能将数据视为开展教学的参考。另外，教师要以一种动态的眼光看待数据，要从数据中探究学生为何会呈现这样的数据，要结合数据与课堂真实情境看到学生的需求，并分析数

^① 欧盟发布人工智能伦理准则 [EB/OL]. (2019-04-11) [2022-04-11]. <http://scitech.people.com.cn/n1/2019/0411/c1007-31024475.html>.

据呈现更深层次的原因,更好地为促进学生素养提升服务。我们应主张“冷”技术的支持者和“热”人性的支持者进行深入对话,支持教师与学生的终身学习,提供教师教育计划,加深教师与学生技术的理解,更好地利用教学实践规避风险。另外,技术的应用要遵循教育教学的基本规律与原则。一方面,要根据学生的发展特点与需求,在合适的时机恰当地使用智能技术,同时控制人工智能技术的使用频率,不能为了使用而使用,忽视学生真实需求和能力进步。另一方面,由于人工智能自身的“类人性”,不能在学生大脑的发育期过度使用认知技术,这样可能会对学生的认知、情感和创新能力的发展产生不可逆转的严重后果。总之,我们要坚持“以人为本,立足学生的全面发展”的理念,在教育中深入地引进人工智能,这才是推动教育发展的题中应有之义。

(二) 建立系统完善的学生隐私保护规制

教育教学作用的直接主体是学生,人工智能技术在教育教学应用中影响最大的主体应该也是学生,不论其影响是积极还是消极,这些影响均是长久的、不可逆的。同时,智能时代,隐私包含的内容逐渐增多,如信息隐私、空间隐私、自决隐私等,其中自决隐私指的是机器人在未被人授权的情况下替人做决策,现今智能辅导系统、智能教学系统便有侵犯学生自决隐私的倾向。^①又因为“人工智能没有像人那样的社会属性和社会关系,机器的自主性是技术性的,没有道德自主性”^②。完善的制度供给是智能技术应用于教育的基本保障,人类要有效控制现实中人工智能的“行为”,就需要建立合乎规范的制度机制,规范教育实践中各主体的行为,同时加强对各主体的伦理教育,建立恰当的伦理观,从而规避伦理风险。

第一,人工智能教育主体虽具有数据收集权,但仍要秉承“尊重”的原则实施数据收集权、数据决定权。尊重原则是教育管理者应用人工智能时要遵循的首要原则,是对其权利进行规约的道德力量,也是以人为本管理理念的体现。^③人工智能教育不是单纯的技术教育,也是立德教育。在现今所提倡的人工智能素养中,相关人员要尊重生命、尊重他人的隐私,只有这样才能构建持续发展的教育生态圈。人工智能系统收集的情感数据应考虑文化差异、背景因素、教师观察和学生意见,数据需要在信息实时生成速度上和 Information 质量上保持平衡,以避免形成基于此类数据训练的算法中隐含的偏差。

第二,政府必须坚持以“善”为技术活动的出发点,制定支持终身学习的政策,提供教师教育计划,提高教师素质,贯彻“学生这一数据主体有权利知晓数据采集的类型、方法”的原则,引导数据主体积极参与智能隐私技术的研发和数据的日常监督,增强数据主体的数据自我保护意识,加大数据保护力度。此外,还要探究人工智能背景下学生隐私保护的适应性规制与规范要素,出台相应的法律法规来规范智能技术的行为,保护个人数据免遭恶意收集、未经授权的披露、商业利用和其他滥用,并制定“机器人违法”的处罚条例,增强法律监管力度,确立好智能技术相关人员(创建者、使用者、执行者、监测者等)的权利与责任,尽可能把教育伦理全部以制度化的形式严格建立起来,防止在人工智能教育实践中出现权责不清的现象。

第三,对于智能技术的行为应加强监察、审计等专门监督以及社会监督和舆论监督,使得技术主体能全方位接受监督,形成科学有效的教育行政权力运行制约和监督体系。对于在智能教育应用中违反行为条例的人,国家可以对此创设一个结合监控、大数据和人工智能的社会信用系统,让失信的人很难迈出一步。对人工智能在教育领域中的应用,政府要谨防一家独大的局面出现。因为,

^① 张娜,侯静怡,贺威,等. 中小学人工智能应用的伦理教育研究 [J]. 中国教育信息化, 2021 (23): 49-52.

^② 宋灵青,许林. 人工智能教育应用的逻辑起点与边界:以知识学习为例 [J]. 中国电化教育, 2019 (6): 14-20.

^③ 李晓岩,张家年,王丹. 人工智能教育应用伦理研究论纲 [J]. 开放教育研究, 2021 (3): 29-36.

这种垄断会扼杀创新，限制人们的教育选择，损害教育工具市场。可以采取的措施是将智能教育应用放置于更广泛的商业环境中，让教育工作者和学者可以有更广泛的选择权，探索新的途径和更有效的教育形式来改进教育与研究。

第四，社会既要鼓励研究者研发智能新成果，又要制定相关法律防止技术的不健康发展。要谨防“算法黑箱”的出现。由于教育管理者有决定何时收集数据、如何利用数据的权力，所以有必要对教育管理者进行伦理课程的培训并对他们的数据使用权设置边界，使他们知晓自己的活动范围，在界限范围内活动，防止数据滥用。对于涉及伦理规范的技术，社会要不断实践，用文本的形式明确开发人员的法律责任和义务，要将外在的制度规范内化于人心，让技术主体、教育主体知晓并明晰人工智能教育应用伦理的重要性，从而规范智能技术应用。

（三）构建各责任主体共同参与的人工智能教育应用制度体系

教育作为一项复杂事务对本国的发展有着举足轻重的作用，它从来都不只是政府和学校的事情，整个社会都应该参与到教育管理中来。对于教育引进智能技术所产生的伦理风险，政府、非政府组织以及个体都应参与到解决过程中来，开展协作配合，共同治理，通过商谈达成道德共识，制定相应的研发与使用规章，确立促进行业健康发展的伦理原则。同时，政府应敦促出台相关切实可行的政策法律，通过道德塑造与法律引导“双轨”并进的方式促进智能技术的健康发展。

首先，要构建人工智能教育应用的管理体系。政府要在智能技术标准规范的前提下，制定一套专门关于人工智能教育应用的管理标准和制度。并在社会中加大对人工智能教育应用伦理制度的普及力度，让社会活动主体养成伦理素养，形成伦理共识，让智能技术开发者和数据管理者接受社会伦理道德准则的约束、舆论的引导和监督。此外，还应增强管理体系的风险预测能力，做到安全、可靠、可控。

其次，要规范建构明晰的、文本的人工智能教育应用的法律体系。人工智能教育应用过程中出现的数据泄露、算法失误等问题，在现今已有的法律制度上均没有明确的处理措施，智能技术设计者、开发者、使用者和用户对于自身的责任认知并不清晰，因而要完善法律体系建设，建立明确的问责机制。这个问责机制是长期的、透明的，应该致力于构建透明的数据收集系统，规定安全的数据跟踪方案，以便在出现伦理风险的情况下快速准确地找出责任主体承担相应的责任。

再次，在教育行业层面，学术界和产业界应建立互惠互利的关系，开展富有成效的合作，确保智能技术以相关的理论框架为指导，多举措并举规避伦理风险。学术界应积极探索适用于智能技术产业界的理论框架，并为解决可能出现的伦理风险问题贡献学术界的解决方案，积极借鉴国内外关于智能技术的治理经验，促使各行业、各主体参与到智能技术风险规避中来。习近平总书记在中共中央政治局第九次集体学习时强调，“要整合多学科力量，加强人工智能相关法律、伦理、社会问题研究，建立健全保障人工智能健康发展的法律法规、制度体系、伦理道德”^①。因而，教育行业不仅要与外界协作，其自身内部也要加强交流研究，必须长期反思兼预测，形成教育生态圈。在产品开发阶段经过实证测试，利益相关者尤其是人工智能制造商要预测人工智能可能存在的偏见类型，并在产品开发过程中尽早尝试减少算法偏见。随着对可能出现的偏见类型的认识，可以鼓励学生加入人工智能开发团队，并为未来健康人工智能的设计作出贡献。

最后，要构建多方协同教育治理的新格局。《中国教育现代化 2035》提出“形成全社会共同参与的教育治理新格局”的战略任务，即构建社会、政府、学校多方共同参与的教育治理格局，实现教

^① 金德楠. 习近平关于道德建设的重要论述探析 [J]. 湖北社会科学, 2019 (11): 20-24.

育治理现代化。在智能时代，需要保证参与教育伦理风险规避的各治理主体地位平等，共享数据信息，使得教育信息资源能够被全方位掌握，共同服务伦理风险的规避问题。

人工智能技术在近些年发展迅速甚至有时赶超人脑，在社会各个领域显示出巨大的社会价值。“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。”^① 但无论技术多么先进，其都应以“善”为起点，承担自己应有的责任。如今，我们认为人工智能的责任就是以人为本，为教育服务。只有让技术成为推动道德进步的驱动力，达到技术手段与教育理念的内在统一，才能规避教育应用中的伦理风险。国家必须认清现状，抓紧顶层设计，重建教育伦理规则与体系，多举措并举，在运用智能技术提升教育现代化水平的同时，也要注意规避其在运用过程中衍生的诸如伦理风险等问题。这就需要我们不断加强对人工智能技术发展的支撑力、推动力和引导力，为智能化教育奠定基础，为加快推进教育现代化保驾护航。

（责任编辑 曹周天）

Ethical Risks and Avoidance in Educational Application of Artificial Intelligence

Luo Shengquan Tan Aili Zhong Yijun

Abstract: With the rapid development of Internet of Things, big data and quantum computing, human beings have entered the era of artificial intelligence with continuous integration of intelligent technology and education, bringing convenience to education and cultivating new talents, but also deriving complex ethical issues. Ethical risks are manifested in the confusion of traditional ethical relations between teachers and students, loss of legal rights caused by data barriers and gaps, and ethical deviation between deterministic algorithm and uncertain development. The reason is because of the dominance of “everything is technology” algorithm in current AI educational application, time separation between AI educational application and existing institutional concepts, and responsibility barriers of multiple subjects. In order to avoid these ethical risks and promote sustainable development, we should adhere to “people-oriented” educational concept, establish systematic and complete student privacy protection regulation, and build AI educational application system with the participation of responsible subjects.

Key words: artificial intelligence; AI education; ethical risk; avoidance

About the authors: Luo Shengquan is a professor from Center for Studies of Education and Psychology of Ethnic Minorities in Southwest China of Southwest University. Tan Aili is a postgraduate from Faculty of Education of Southwest University. Zhong Yijun is the principal of Xinguang Primary School.

^① 宋灵青，许林. 人工智能教育应用的逻辑起点与边界：以知识学习为例 [J]. 中国电化教育，2019（6）：14-20.