

# 智能思政的伦理风险审视及应对策略

张 铤

**摘要：**智能思政在创新思想政治教育方式、提升思想政治教育效能的同时，其带来的伦理风险问题亦不容忽视。智能思政伦理风险主要表现为师生主体地位的消解与情感关系的淡化、“信息茧房”的滋生与算法伦理的失范、教育公平的式微与技术价值的迷失以及学生隐私的失守与数字侵权责任的游离等。究其深层肇因，主要在于智能算法内部运作的复杂性与不透明性、人工智能伦理与法律规范的相对滞后和师生科技伦理素养与风险处置能力的欠缺。对此，要冷静分析、积极应对，通过构建“价值—主体—规范”三位一体的策略框架，加强智能思政伦理风险治理，促进人工智能向上向善，推动新时代思想政治教育与人工智能技术发展双向赋能。

**关键词：**智能思政；伦理风险；德法共治

**中图分类号：**B82;D64   **文献标识码：**A   **文章编号：**1003-0751(2025)09-0103-06

当今世界，人工智能技术的快速发展，尤其是以ChatGPT、DeepSeek为代表的生成式人工智能的出现和应用，为思想政治教育的创新发展提供了技术条件。在此背景下，思想政治教育与大数据、云计算、人工智能等新兴技术的融合正在不断深化，思想政治教育智能化发展速度加快，并由此形成了“智能思政”这一人工智能时代“思想政治教育创新发展的新形态”<sup>[1]</sup>。智能思政的出场，在创新思想政治教育方式、提升思想政治教育效能的同时，其带来的伦理风险问题亦不容忽视。智能技术对思想政治教育的风险挑战不仅仅是直接式的“科技冲击”，而且是影响思想政治教育运作机制和实践形态的深刻变革。智能思政的伦理风险兼具人工智能伦理问题的共性与思想政治教育自身的特殊性。鉴于此，为了促进智能思政的健康发展，有必要厘清智能思政伦理风险的具体表征及其深层肇因，并据此提出可行性治理策略。

**收稿日期：**2025-07-18

**基金项目：**浙江省社科联课题“新媒体时代大学生政治价值观生成机理与培育机制研究”(2025N012)；浙江省教育科学规划课题“新媒体时代大学生政治价值观培育研究”(2023SB103)。

**作者简介：**张铤，男，同济大学马克思主义学院副教授，同济大学中国特色社会主义理论研究中心特约研究员(上海200092)。

## 一、智能思政伦理风险的多维表征

思想政治教育与人工智能技术并非天然的耦合关系。人工智能的嵌入给思想政治教育发展提供了新可能，也带来了潜在的伦理风险与隐患。具体来说，智能思政的伦理风险主要体现在伦理关系、伦理行为、伦理价值和伦理责任等四个维度，需要引起我们的高度警惕。

### 1. 伦理关系维度：师生主体地位消解与情感关系淡化

伦理关系主要指渗透道德观念的自律性社会关系，其强调人际互动时需遵循的准则。随着人工智能的快速发展和迭代更新，智能机器或将具有自己的“意志”，拥有“自主性”力量。实际上，生成式人工智能的兴起及其显现的技术特质已可见此端倪。特定环境中，人工智能的“自主性”力量会与人的主

体性产生抵牾,形成一定的张力。置于智能思政场域中,因智能思政系统无法完全理解学生复杂的情感和心理需求,如果过度依赖该系统,就可能消解教育主体的自主性与判断力,降低师生的主体地位,淡化师生间的情感交流和人文关怀。一是师生主体地位的消解。思想政治教育是教育者和受教育者双向互动的过程。教师层面,智能算法赋能下人工智能将承担更多教师的角色,如自动生成课件,开展学情诊断甚至“替代”教师批改作业、阅卷等。并且,数智技术的渗透会解构传统思想政治教育信息生产模式,甚至可能形成受教育者主导教育信息的生产和传播局面,导致教师的权威和主导地位被消解。学生层面,虽然利用智能技术有助于了解学生需求、定制学习任务,促进“‘标准化教育’向‘个性化教育’转变”<sup>[2]</sup>。但与此同时,智能助教、“AI学伴”等智能产品的过度使用也会使有些学生形成技术依赖,削减其学习主动性,消解其主体精神。此外,智能思政实践中,智能影像的过载也容易使学生形成技术依赖,陷入“泛娱乐化”困境,导致其批判精神和否定意识减弱。二是师生情感关系的淡化。思想政治教育是情感化的实践活动,主要通过情感共鸣、情感认同与价值引导使学生“亲其师、信其道”,成人达人。人工智能的嵌入使传统师生伦理中的情感联结受到挑战。数智环境中,师生之间主要借助虚拟仿真途径进行交流,然而“人—机—人”的过度虚拟交互不免会弱化师生现实中的人际交往和情感交流,从而遮蔽思想政治教育的情感属性。

## 2. 伦理行为维度:“信息茧房”滋生与算法伦理失范

伦理行为主要指个体或组织在行动中遵循伦理道德准则的规范性行为。数智传播过程中,算法推荐应遵循一定的伦理规范,在价值导向与用户偏好、内容多样性与风险防控之间维持平衡,避免算法操控与异化。算法伦理是智能思政发展的核心议题之一。人工智能在促进思想政治教育精准化的同时,“野蛮生长”的算法权力可能异化为互联网空间的一种新型权力,消解网络空间主流意识形态价值共识,侵蚀社会主义核心价值观。不仅如此,智能算法的过度推荐会使网络环境中“信息茧房”现象凸显。一是“流量至上”与教育信息的碎片化、娱乐化。算法推荐是智能思政信息推送和传播的主要手段,但智能思政中的算法推荐容易偏离算法伦理和思政本质。在“流量至上”作用下,网络信息传播碎片化、娱乐化盛行,而社会主流意识形态叙事和信息却常

常被冷落,难以被优先推荐。于是,主流意识形态话语权被削弱,偏重思想性、理论性的思想政治教育内容供给被压缩,从而影响思想政治教育功能的发挥。二是“信息茧房”与算法推荐的异化。数智时代的到来,借助语言大模型等工具,学生获取知识和信息更加快捷,但智能算法给学生持续“投喂”个人偏好的内容会滋生“信息茧房”,导致有的学生因“窄化在自我画像的信息空间中”<sup>[3]</sup>而偏听偏信,出现认知固化、偏差和价值误判等问题。智能算法推荐赋能网络传播的同时,“给网络空间带来‘劣币驱逐良币’的信息困扰”<sup>[4]</sup>,不仅会使有的学生获取信息渠道单一,与外界缺少交流,还可能滋生群体极化现象。概言之,算法推荐应把握算法的伦理边界,如果过度个性化而忽视传播的多样性,那么人工智能赋能思想政治教育的同时,就可能将受教育者锁在“算法囚笼”中,影响和支配他们的思想和行为。

## 3. 伦理价值维度:教育公平式微与技术价值迷失

伦理价值强调行为的正当性及个体与社会利益的均衡性,主要涉及对行为、选择和社会关系的道德评判。人工智能在拓展思想政治教育场景的同时,也可能因“技术鸿沟”、算法偏见而导致对不同学生群体的评价、推荐内容出现差异,从而影响教育效果,损害教育公平。另外,思想政治教育对智能技术工具的过度推崇容易陷入技术至上主义的误区。一是“数字鸿沟”与教育公平的式微。人工智能嵌入思想政治教育,目前并不能普遍赋能、惠及全体受教育者;反之,智能思政技术应用还可能扩大“数字鸿沟”。具体来说,由于存在个人技术能力、区域经济水平等的差异,仍有部分学生未能充分享受到技术发展带来的红利,未能参与到智能思政技术的应用场景中。此外,如果智能思政系统的算法设计缺少价值对齐,或其训练数据自身存在某种偏见或歧视,那么智能思政系统的输出结果也可能暗藏偏见或歧视,甚至会危及教育公平与正义。二是技术崇拜与工具理性的张扬。智能思政具有的用户精准画像、个性化信息推送和学习任务定制等功能和优势会使部分师生沉溺其中,甚至滋生“技术崇拜”,形成“技术至上主义”,而“技术至上主义”的盛行势必导致工具理性的张扬。在此境遇下,智能思政的价值理性就会被工具理性侵蚀,教育者容易被技术的表面功效所蒙蔽而迷失在“技术附魅”中,忽视教育的知识传递、价值塑造与文化传承使命,受教育者则很难体验到思想政治教育应有的情感关怀与人本温度。

#### 4. 伦理责任维度: 学生隐私失守与数字侵权责任游离

伦理责任的关键在于行为主体对自身行为后果的预见性承担,要避免因责任归属边界不清导致责任游离。智能思政场域中,人工智能技术应用过程中容易带来侵犯学生隐私和数字侵权责任界定困难等风险与后果。一是容易带来学生隐私保护的失守。人工智能赋能思想政治教育的同时,也存在教育者“控制技术”、滥用权力的责任风险。现实中,由于个别教育者责任伦理的缺失与权力约束的阙如,导致学生的数据被过度采集、不当使用以及智能监控过度等。如未征得学生同意或授权就采集学生人脸、指纹等敏感信息,并且没有妥善储存,导致数据信息泄露,学生的隐私权利因此受到侵犯。二是容易带来数字侵权责任的游离。智能思政实践中,如发生侵权事件,其责任界定是一个难题。例如,智能思政大模型可能因“机器欺骗”“模型幻觉”而误导学生,影响学生的价值判断与取向。又如,多模态大模型的出现为能自主感知外界环境、做出决策并执行行动的智能体的设计和实现创造了条件,但智能体应用于思想政治教育时如果缺少必要的监管,也可能罔顾社会道德与伦理准则,做出错误的教育决策或伤害学生的行为,从而造成严重后果。对于此类情况,按照传统的技术伦理责任规范,技术研发者(或智能体设计者)应承担主要责任。但是,数智时代技术研发者仅仅是单一的原子责任主体,完全归责技术研发者似有不妥。更何况智能思政技术研发或教育产品设计往往是团队合作行为,团队中个人的行为正当与否,已经很难简单地适用于针对个人行为的伦理准则和规范。那么,如将责任推给“智能机器”,让其作为主体承担主要责任,尚没有很强的说服力。因此,如何在智能思政实践中刺破技术“面纱”,避免侵权责任“漂移”,成为智能思政发展亟待破解的难题。

## 二、智能思政伦理风险的深层肇因

作为一种尚未成熟的前沿技术,人工智能在提升思想政治教育效能的同时,其潜藏的伦理风险也逐渐显现出来。智能思政伦理风险的生成有其内在逻辑,揭示智能思政伦理风险的深层肇因,可为采取有效治理对策提供科学依据。

### 1. 智能算法内部运作的复杂性与不透明性

“当人类不断建构复杂高科技系统时,不确定

性必然滋生,科技伦理问题必然凸显。”<sup>[5]</sup>智能算法内部运作的复杂性与不透明性是人工智能伦理风险的重要来源。人工智能算法的背后隐藏了“技术黑箱”,公众很难理解和预测人工智能算法的运作逻辑<sup>[6]</sup>。置于智能思政情境中,教育者很难保证智能算法的输出结果都符合伦理规范,也很难保证都与思想政治教育的价值诉求完全内在契合。可以说,人工智能技术的不确定性是客观存在的,此种不确定性源于人工智能数据训练、机器学习、算法运作的复杂性与不透明性及其衍生出的技术风险与负面影响。例如,虽然人工智能大模型发展方兴未艾,其在思想政治教育领域的应用前景广阔,但人工智能大模型因其深度学习模型、高创造性与自主性和多模态内容生成等技术特质以及数据质量、数据训练方式缺陷等原因,容易引发深度伪造、虚假信息、价值误导和知识产权侵犯等问题。更何况有的人工智能大模型缺乏价值对齐,“机器欺骗”“模型幻觉”等现象时有发生,影响了其技术的可靠性与可信度。坦率地讲,人类迄今无法完全还原和审查人工智能模型内部运作和决策的全过程。这种不透明性不仅导致了智能算法解释的困难,还可能引发个人信息安全问题和社会不公平现象。与此同时,智能算法的不透明性会在一定程度上形成“黑箱忧虑效应”。所谓“黑箱忧虑效应”,指的是人们对于无法完全理解或不透明的复杂系统产生的不安和焦虑感。在人工智能领域,这种效应尤为明显。需要注意的是,人工智能技术的不可控性在现实中已经显现。因此,智能思政在发挥技术正效应的同时,必须多维并举防范其技术风险尤其是伦理风险,确保智能技术可控、可靠、可信,成为思想政治教育创新发展的“利器”而不是“风险源”。需要指出的是,现实中,人工智能风险治理不可避免地要借助技术手段甚至用“智能技术”治理“智能技术风险”。然而,以“技术治理技术”难免使技术风险治理陷入自反性“悖论”中,更会加大智能思政伦理风险治理的复杂性及难度,这也是规避智能思政伦理风险需要解决的难题。

### 2. 人工智能伦理与法律规范的相对滞后

技术并非完全中立、价值无涉。智能思政实践中,人工智能会呈现技术“赋能”的一面,也会露出技术“负能”的另一面,折射出人工智能技术社会价值的两面性。需要注意的是,智能思政实践中,人工智能企业开发的智能思政教育产品并非尽善尽美。现实中,智能思政教育产品功效与思想政治教育本质属性之间不一定完全匹配,折射出技术本位与育

人本位、商业价值与思想政治教育价值之间的张力。究其深层原因,在于技术工具理性与价值理性之间的冲突。技术工具理性的泛滥会导致智能思政实践中对伦理、道德和情感等因素的忽视。智能思政要平衡工具理性与价值理性、规避技术“负能”,关键在于加强人工智能伦理治理和相关法律的制定、完善及执行,以促进智能技术应用的规范化、合法化。当前,相较于人工智能技术的迅猛发展和迭代更新,人工智能伦理规范与法律规范建设相对滞后。在人工智能技术应用和能力持续涌现的背景下,传统的技术伦理规范与法律规范如不及时调整和优化,就难以适应数智时代的新技术带来的新挑战。就智能思政而言,智能思政技术应用过程中数据训练方式缺陷或训练数据中的偏见,会影响受教育者的世界观、人生观和价值观,甚至可能使学生在性别、种族等方面遭遇歧视。这些问题迫使智能思政加强伦理规范建设,将伦理道德融入人工智能全生命周期,促进人工智能的价值对齐,使“智能机器”具备“道德感”和“道德能动性”,在思想政治教育应用中做出符合人类伦理道德的决策。“人工智能价值对齐是一个规范性问题,也是个技术性问题。”<sup>[7]</sup>道德能动性既可以自然生成,也有可能通过人工建构,“它为人工智能道德事件进行道德责任归属提供了更加完备的理论依据”<sup>[8]</sup>。

### 3. 师生科技伦理素养与风险处置能力的欠缺

“科学的精神是最强的力量,就破坏性而论,它也是最强的力量。”<sup>[9]</sup>人工智能造福社会的同时也会带来破坏性。在“人工智能+教育”浪潮席卷下,师生容易关注智能技术的正向效应而忽视其风险及破坏性效应,有的学生甚至认为“智能机器”风险仅存在于科幻电影的虚构场景中。但实际上,人工智能技术风险离我们并不遥远,防控其风险也并非危言耸听,如人工智能在对人类进行心理健康干预中就存在伦理失控的风险。面对智能思政的伦理风险及其挑战,提高师生对人工智能伦理的认知和敏感度,增强其应对技术风险的能力,不仅必要,而且紧迫。进言之,当智能技术在思想政治教育中彰显智能助学、伴学和智能督学等功能,定制式、个性化、精准化教育被广泛应用的时候,无论是教育者还是受教育者,都要有技术风险防范意识,不能对技术风险和隐患置若罔闻,或把技术善治的希望寄托于人工智能专家。当前,高校师生的科技伦理素养参差不齐,部分师生对于人工智能技术潜在风险的识别能力以及对其伦理准则(公平性、透明性、责任性等)

的理解有所欠缺,防范技术风险的能力薄弱,影响了智能思政伦理风险的有效规避和处置。智能思政伦理风险的规避和处置不容易,因为人工智能风险是由多因素交织而成的复杂过程。况且,随着人工智能技术的快速发展,即使原有的风险问题获得基本解决,但新的风险又有可能产生。客观而言,推进智能思政伦理风险治理仍然存在不少难点。例如,智能思政发展中如何弥合“数字鸿沟”?智能思政技术应用中如何避免算法偏见与算法歧视?智能思政实践中如何保护学生隐私权?如何推进智能思政伦理风险协同治理?这些都是智能思政发展中亟待解决的问题。

## 三、智能思政伦理风险的应对策略

思想政治教育与人工智能的融合正处于初始、未定的“磨合期”。智能思政的伦理风险不是人工智能伦理风险与思想政治教育的简单叠加,其风险治理颇为艰巨。对此,要冷静分析、积极应对,通过构建“价值—主体—规范”三位一体的策略框架,加强智能思政伦理风险治理,促进人工智能向上向善,推动新时代思想政治教育与人工智能技术发展双向赋能。

### 1. 价值层面:坚持以人为本,彰显智能思政育人价值导向

人工智能为思想政治教育发展提供了新的技术条件,但技术形态的重塑并不会改变思想政治教育的根本价值取向。因此,智能思政伦理风险防控要从其价值导向方面入手,强化马克思主义的指导地位,落实立德树人任务,促进思政教育与人工智能的有机结合、相互增效,避免“技术陷阱”“育人缺位”。一是坚持以马克思主义为指导。坚持以马克思主义为指导是思想政治教育生命力和活力的源泉,也是思想政治教育的本质要求。智能思政实践中,要将马克思主义主流意识形态融入智能思政的算法设计与应用中,用主流价值驱动“算法向善”,完善算法推荐服务机制,避免算法权力异化,避免技术工具理性凌驾于价值理性之上。智能思政技术价值对齐要促进伦理道德规范和设计的技术标准之间的融合,从技术和规范两个层面加强人工智能的价值对齐。二是彰显智能思政育人价值导向。思想政治教育以育人为中心,智能思政系统中的人工智能技术,究其本质,是服务于思想政治教育创新的工具和手段。如果智能思政实践被工具理性挟制,那么智能思政

伦理风险的生成和智能技术的“异化”就不可避免。换言之,人工智能技术是思想政治教育之“器”而非“道”。因此,要正确把握技术发展与思想政治教育的辩证关系,从“技术本位”走向“以人为本”。智能思政无论发展到何种程度,都要以育人为中心,而不能偏离育人价值导向,智能思政系统只是智能思政的一种技术手段而已。三是扩大优质思想政治教育资源供给。要通过高校网络文明宣讲、优秀网络文化作品展等形式,满足师生网络文化需求,构筑高校网络文明空间。要组织开发思政课示范课程库、案例库和素材库等优质资源,推动高校思政“金课”建设,加强思政课的价值引领功能。需要强调的是,应对智能思政伦理风险,要科学把握和运用思想政治教育规律。具体而言,一方面,要了解和掌握学生网络认知与网络行为规律,创新思想政治教育方式和方法,满足数智环境中不同层次的学生需求;另一方面,要增强思政引领力,教育和引导学生合理使用智能技术,避免技术依赖,尽量消减“信息茧房”、算法偏见和算法歧视等造成的不良影响。

## 2. 主体层面:普及科技伦理,增强防控智能思政风险的能力

人工智能伦理风险在技术研发和应用环节都有可能发生。推进智能思政伦理风险治理,需加强科技伦理尤其是人工智能伦理的宣传与普及,提升智能技术研发人员和师生的科技伦理素养,增强其技术风险防控能力。一是对智能思政技术研发人员开展教育培训,增强其伦理规范意识和法律意识,要求其将价值敏感设计嵌入智能思政系统,在技术路径上努力实现人工智能价值对齐。对于智能思政大模型,技术研发人员要强化算法价值对齐训练,通过基于人类反馈的强化学习(RLHF)要求“诚信优先”,并通过“生成+验证”实现动态事实核查,提高大模型的可靠性和可信度,尽量减少大模型带来的“机器欺骗”“模型幻觉”等现象,避免大模型“一本正经的胡说八道”。此外,要加强思想政治教育相关学科的知识培训,提高智能思政技术研发人员的专业能力,避免其设计开发的智能教育产品脱离实际。二是提高师生的科技伦理素养,增强其技术风险防范和化解能力。要结合伦理学、社会学、法学等学科,构建多维度人工智能伦理规范教育框架。具体而言,教师层面,可开展“人工智能伦理素养提升计划”,通过形式多样的入职培训和职业发展培训,让广大教师熟悉智能技术应用方法的同时,知晓人工智能伦理准则,“理解人工智能在知识表征与应用

方面的优势及局限”<sup>[10]</sup>,并提高教师尤其是思想政治理论课教师的技术风险应对和处置能力。学生层面,可通过专家报告、沉浸式教学、实践教学和案例教学等方式,普及人工智能技术应用、科技伦理等方面的知识。例如,将人工智能伦理纳入人工智能通识核心课程体系,通过案例教学(如自动驾驶伦理困境)深化大学生对人工智能伦理的理解。与此同时,要增强其技术风险防范意识与能力,提高学生在与算法互动过程中的自主性,避免其过度使用智能思政教育产品而导致依赖与成瘾,养成良好的网络行为习惯。

## 3. 规范层面:完善规范体系,以德法共治促进智能思政的良性发展

伦理规范与法律规范建设对于智能思政发展具有重要作用。近年,为了推动数智技术的健康发展,我国发布了《新一代人工智能治理原则——发展负责任的人工智能》《新一代人工智能伦理规范》等文件,出台了《中华人民共和国网络安全法》《中华人民共和国个人信息保护法》《中华人民共和国数据安全法》等法律法规,为人工智能的发展和正确使用提供了基本原则和制度规范,也为智能思政发展提供了基本遵循。但是,由于人工智能的规范化治理尚处于规则形成期,“法律治理框架与技术特质存在脱节”<sup>[11]</sup>。当前,我国人工智能伦理治理与相关的法律治理还有待加强,涉及智能思政等方面的伦理规范与制度规范仍比较缺乏。鉴于此,要坚持“善治”的价值目标,推进智能思政伦理风险的德法共治。一是加强智能思政伦理规范建设。智能思政伦理规范要基于技术向善与育人为本的双重锚定,坚持促进公平公正、保护隐私安全、强化责任伦理、透明可解释和不伤害等伦理原则,以价值理性优先约束技术边界,实现工具理性与价值理性的统一。需要指出的是,责任是科技伦理的题中之义,强化责任伦理对于智能思政发展颇为重要。责任伦理就是“无条件地”对自己的行为承担责任,一个成熟的人应当“遵照责任伦理采取行动”<sup>[12]</sup>。责任伦理要求智能思政实践中,在保留技术应用“回顾式责任”的基础上,更加重视“前瞻式责任”,形成一条完整的覆盖智能思政技术开发、设计和应用等环节的“责任链”。对于某些智能教育产品,在其投入使用前,要坚持“不伤害”的伦理要求,依托相关机构和伦理专家,审慎评估其潜在后果,并要求行动者对后果负责。二是加强智能思政法律规范建设。制定完善相关的法律法规,明晰智能思政伦理风险治理主体的

权利与义务,厘清智能思政伦理风险的问责机制、责任类型与形式;加强算法权力的约束与规制,把智能思政算法透明与可解释性纳入相关制度规范,化解“算法黑箱”。需要补充的是,人工智能技术发展需要伦理共同体的维系。无论是智能思政技术的研发者,还是其使用者或受益者,都要参与伦理风险治理,加强协商与合作,形成科技伦理共同体,“透过科学共同体这一伦理共同体的道德意志与道德行为担当起道德主体的责任”<sup>[13]</sup>。德治与法治两种基本治理手段,要在智能思政伦理风险治理中优势互补、融合共治。唯有如此,智能思政才能在创新发展中实现技术赋能与价值引领的有机统一。

总而言之,人工智能正以前所未有的速度和广度重塑人类社会的生产、生活和学习方式,并对思想政治教育产生深刻的影响。在此背景下,智能思政的兴起确实有其必然性和必要性。智能思政是数智时代思想政治教育技术形态演进的必然趋势,也是思想政治教育现代化发展的内在要求。与此同时,我们也要认识到,人工智能赋能思想政治教育有其独特优势,也有其功能限度。人工智能技术不是万能的,智能思政技术应用是有伦理边界的。智能思政技术应用与其伦理风险犹如一枚硬币的两面。我们要坚持智能思政发展与技术风险治理“双轮驱动”,探索智能思政发展的实践经验及其风险治理之道,推动思想政治教育与数智发展更好地相适应,使智能思政从“技术赋能”走向“育人升级”。思想政治教育是一门科学,其科学性就体现在能与时俱进地调整和优化技术形态,善于运用新技术、新工具不断推进思想政治教育内容和形式的创新,充分彰显思想政治教育的价值。

## 参考文献

- [1]崔建西,白显良.智能思政:思想政治教育创新发展的新形态[J].思想理论教育,2021(10):83-88.
- [2]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.
- [3]申小蓉,米华全.推荐算法赋能大学生价值观塑造:逻辑理路、现实忧与优化路径[J].思想教育研究,2024(7):123-130.
- [4]王天民,郑丽丽.智能思政“信息茧房”的生成机理与治理路径[J].贵州师范大学学报(哲学社会科学版),2023(5):30-42.
- [5]田旭明.数字社会的主要伦理风险及其应对[J].中州学刊,2022(2):87-93.
- [6]张铤.人工智能的伦理风险治理探析[J].中州学刊,2022(1):114-118.
- [7]闫坤如.人工智能价值对齐的价值表征及伦理路径[J].伦理学研究,2024(4):94-100.
- [8]曲蓉.人工智能能动性的伦理证成:传统观点到开放立场[J].中州学刊,2025(6):118-126.
- [9]萨顿.科学史和新人文主义[M].陈恒六,译.北京:华夏出版社,1989:45.
- [10]徐子燕,石中英.人工智能时代的真理问题与教育责任[J].教育研究,2024(10):26-37.
- [11]张凌寒.人工智能法律治理的路径拓展[J].中国社会科学,2025(1):91-110.
- [12]韦伯.学术与政治;韦伯的两篇演说[M].冯克利,译.北京:生活·读书·新知三联书店,1998:116.
- [13]薛桂波.科学共同体的伦理精神[M].北京:中国社会科学出版社,2014:44.

## Ethical Risk Assessment and Response Strategies for Intelligent Ideological And Political Education

Zhang Ting

**Abstract:** Intelligent ideological and political education not only innovates the methods of ideological and political education and enhances the effectiveness of ideological and political education, but also brings ethical risks that cannot be ignored. The ethical risks of intelligent ideological and political education mainly manifest as the dissolution of the subject status of teachers and students and the weakening of emotional relationships, the breeding of “information cocoons” and the anomie of algorithm ethics, the decline of educational equity and the loss of technological value, as well as the violation of students’ privacy and the ambiguity of digital infringement liability. The underlying causes mainly lie in the complexity and opacity of the internal operation of intelligent algorithms, the relative lags of artificial intelligence ethics and legal norms, and the lack of technological ethics literacy and risk disposal capabilities among teachers and students. To address this, we need to calmly analyze and actively respond by constructing a “value-subject-norm” trinity strategic framework, strengthening the governance of ethical risks in intelligent ideological and political education, promoting the upward and virtuous development of artificial intelligence, and promoting the two-way empowerment of ideological and political education and the development of artificial intelligence technology in the new era.

**Key words:** intelligent ideological and political education; ethical risk; governance by morality and law

责任编辑:思 齐