



风险与治理：生成式人工智能与高校 思想政治教育发展

杨 忠

(福建师范大学 化学与材料学院, 福建 福州 350117)

摘 要: 生成式人工智能赋能思想政治教育发展是思想政治教育对国家发展战略需求、学科自身发展革新与思想政治教育主客体发展需求的回应。受生成式人工智能自身发展限制与思想政治教育特殊性影响,生成式人工智能赋能思想政治教育过程中可能造成意识形态、学科解构和伦理道德风险,可以通过提高科技创新能力、破除技术迷信、嵌入伦理道德标准等手段进行治理。

关键词: 生成式人工智能; 高校思想政治教育; ChatGPT; 风险与治理

中图分类号: G711

文献标识码: A

文章编号: 1671-931X (2024) 03-0068-06

DOI: 10.19899/j.cnki.42-1669/Z.2024.03.010

随着信息技术的高速发展,世界各国以科技创新为核心,竞相在数字化时代中谋求转型,力求抢占数智化时代先机。在此背景下,人工智能作为数智时代的核心要素席卷全球,新时代人工智能依托强大算力与模型改进,初步具备了类人的学习能力和逻辑能力,由弱人工智能逐渐发展为强人工智能。其中,生成式人工智能作为强人工智能时代的标志,是一种利用深度学习模型能够根据复杂多样的提示生成类似人类的内容的技术。^[1]生成式人工智能代表性产品有 ChatGPT(Chat Generative Pre- Trained Transformer)、DALL-E 2、Stable Diffusion、Gen-1、Image Video 等,通过文字、图像、视频、代码等生成功能介入人类社会的方方面面,为包括教育在内的各个领域带来深刻变革。

一、生成式人工智能赋能思想政治教育发展的逻辑起点

生成式人工智能赋能高校思想政治教育发展是大势所趋。从宏观、中观、微观不同层面分析可以发现,生成式人工智能赋能思想政治教育是思想政治教育对国家发展战略需求的回应,是思想政治教育学科自身发展革新的呼唤,也是思想政治教育主客体个性化发展的需求。

(一) 国家战略: 建设创新型国家和世界科技强国

为抢抓人工智能发展的重大战略机遇,构筑我国人工智能发展的先发优势,早在 2017 年,国务院印发《新一代人工智能发展规划》,在国家战略层面上对智能教育提出了新要求。规划强调,要利用智能技术加快推动人才培养模式、教学方法改革,构

收稿日期: 2023-11-08

基金项目: 2021 年福建师范大学思政工作及辅导员工作专项资助课题“中国共产党百年青年思想政治领导工作经验研究”(项目编号: VR-2109)。

作者简介: 杨忠(1995—),女,福建福州人,福建师范大学化学与材料学院助教,研究方向: 思想政治教育、高等教育史。

建包含智能学习、交互式学习的新型教育体系。要建立以学习者为中心的教育环境,提供精准推送的教育服务,实现日常教育和终身教育定制化。^[2]

思想政治教育作为高校进行思想引领、价值引导的主要途径,承担着立德树人的根本任务和使命,是人才培养不可逾越的重要环节,始终服务于国家战略发展。以生成式人工智能赋能高校思想政治教育是对国家发展战略的响应,也是对智能教育的深化。一方面,处在时代之中,思想政治教育不可避免地会被卷入人工智能浪潮之中,相较于被动卷入、被迫迎合,在国家战略引导下主动迎接,发扬“以我为主,为我所用”的精神,更有利于发挥思想政治教育的引领作用,从而提升思想政治教育的时代感;另一方面,通过人工智能赋能思想政治教育,推动思想政治工作传统优势同信息技术高度融合,通过变革思想政治教育的方式方法,使思想政治教育活起来,更具亲和力、针对性和吸引力,从而发挥出其人才培养效能。

(二) 学科建设: 思想政治教育学科内涵式发展

思想政治教育学科的内涵式发展,主要通过学科规划、人才队伍、人才培养、科学研究、学科基地、学科环境等要素建设,提高学科核心竞争力。^[3]生成式人工智能为与思想政治学科内涵式发展紧密相关的学科人才队伍建设、教学内容与方法设计、多学科交叉融合等内容提供助力。

1. 生成式人工智能有助于构建高素质思想政治教师队伍

受毕业院校、专业训练、自主学习能力等因素影响,高校思想政治教师水平良莠不齐,水平不佳的教师难以带出高水平的学生,通过生成式人工智能设计思想政治教师队伍训练机制有助于教师提高教学水平。如教师可以通过生成式人工智能评估自身教案撰写水平,开展教学演练评估教学效果,亦可以通过生成式人工智能检测专业功底。

2. 生成式人工智能有助于提升教学内容的时效性,增强教学方法的新颖性

在以往传统的思政教学中,对于热点的捕捉与教育时机的掌握不足,容易出现教学内容与受教育者需求脱节,受教育者的思想困惑无法及时得到澄清等问题,而生成式人工智能的及时反馈性则可极大增强教学内容的时效性。以 ChatGPT 为例,ChatGPT 实现了从一对多到一对一的教学转变,受

教育者刚开始产生思想困惑,就可以向生成式人工智能提问,通过一问一答的形式,ChatGPT 生成与提问相关的文字、图像、视频等跨模态资料,充当受教育者的智能学习助手,协助受教育者及时澄清思想困惑,实现精准思政目标。再者,相较于传统的课堂教学模式,以生成式人工智能为中介对学生进行教育的方式更加新奇,更容易将大学生的注意力吸引到思想政治教育中来。

3. 生成式人工智能有助于拓宽学科发展视野

生成式人工智能集数学、计算机科学、脑科学、哲学、心理学等多门学科于一体,通过生成式人工智能,思想政治教育与不同学科产生联结,构建“思想政治教育+计算机科学”“思想政治教育+数学”等跨学科体系,从而拓宽思想政治教育研究视野,推动思想政治教育创新发展。

(三) 个体需求: 教育主客体个性化发展

思想政治教育要求教育者不断提升思想政治教育效能,培养全面发展的社会主义建设者和接班人。在这一过程中,教育者因材施教的成本是很大的。在许多学校,受师生比例与教学难度的影响,难以推进因材施教。同时,资料整理、方案制定、活动实施等工作任务占用了思想政治教育者的工作时间,使其与受教育者深入交流互动的频率受到影响,一定程度上使得思想政治教育表面化,无法深入受教育者的内心。通过生成式人工智能,高校思想政治教育者可以通过智能分析了解所教学生的个人情况,包括学习进度、思想动态等内容,为教师因材施教提供学生个人情况分析报告。同时,在强大算法与数据的支撑下,生成式人工智能能较快地整理好文字资料或提供教师所需要的专业资料等,一定程度上减少了教师的时间成本。可以说,“生成式人工智能极大地提升了个体的宏观直觉水平”^[4],提高了教师的工作效率。

思想政治教育的受教育者期待更加精准、个性、生动有趣的思想政治教育内容。在传统思想政治教育课堂中,大学生抬头率低的很大原因是觉得教育内容枯燥乏味。而通过生成式人工智能,能够从学习规划、学习内容和学习方法三大方面激发受教育者的学习兴趣,提升受教育者的学习主动性。从学习规划来看,生成式人工智能能够为受教育者提供定制化的思想政治教育学习内容、建议、方案,帮助受教育者掌握学习节奏;从学习内容来看,生成式人工智能通过提供受教育者感兴趣的学

习内容,激发受教育者自主学习的兴趣。从学习方法来看,生成式人工智能能够帮助受教育者跨越时间、空间的限制,创设互动场景,使受教育者身临其境地感受历史、当下与未来,感受此地、他地与世界。

二、生成式人工智能赋能高校思想政治教育发展的风险挑战

生成式人工智能赋能思想政治教育应国家战略、学科建设与个体发展需求而生,为思想政治教育发展带来机遇。但受人工智能自身发展限制,其所内含的技术黑箱、技术依赖等问题为思想政治教育发展带来挑战。

(一) 意识形态风险:西方意识形态渗透削弱主流意识形态主导权

1. 资本逻辑裹挟生成式人工智能,西方意识形态话语渗透

在生成式人工智能领域,西方资本具有先发优势,迅速占领了世界生成式人工智能市场,如 OpenAI、Stability AI、微软、谷歌等。西方科技型企业掌握了生成式人工智能的话语权,并通过强强联合不断挤压我国科技企业的生存空间。如微软与 Open AI 合作将 ChatGPT 内嵌于微软现有的科技产品中,形成叠加效应。在此背景下,一方面,当前主要的生成式人工智能被“喂养”的数据天然地带有西方的意识形态特征,可能助推西方意识形态的传播。一旦错误的意识形态披上生成式人工智能这一非政治化的外衣,受教育者的警惕心就会大大下降,错误思想隐匿于看似可靠的文本之下,为受教育者带来错误的价值引导,从而生发意识形态渗透风险。另一方面,资本逻辑消磨主流意识形态权威。出于利益的考虑,生成式人工智能以“流量至上”,其制造的信息中往往出现了很多未经证实、低俗媚俗、重复低劣的内容,而严肃、客观并能够引导人们形成正确价值观念的内容则被迫“出局”^[5],主流意识形态权威消磨于无意义的内容供给之中。

2. 技术黑箱遮蔽意识形态,主流意识形态权威受到挑战

黑箱是指一种系统或装置,其内部的结构、工作原理或信息处理过程不为用户所知,用户只能观察其输入和输出,而无法了解其内部运作机制。^[6]在生成式人工智能赋能思想政治教育过程中,教育主体无法观测到生成式人工智能内容输出的机制,

输出的内容往往经过了人工智能底层模型的二次处理,无法保证输出内容的客观性和准确性。如果西方意识形态嵌入生成式人工智能模型中,其生成的内容就必然带有西方的价值偏向,加之其病毒式的传播速度,非主流意识形态将快速传播至不同受教育者之中,主流意识形态权威受到挑战。

(二) 学科解构风险:技术依赖消解思想政治教育学科主体性

正常状态下,在生成式人工智能赋能思想政治教育过程中,思想政治教育者是主体,受教育者是客体,思想政治教育是核心,生成式人工智能是辅助手段。但技术依赖或技术成瘾一旦形成,生成式人工智能将僭越思想政治教育主体,掌握思想政治教育供给侧的主动权,从而消解思想政治教育学科主体性,出现本末倒置风险。

1. 技术依赖打造“信息茧房”,思想政治教育内容供给空间受到挤压

为提升用户黏性,生成式人工智能会根据受教育者的反馈生成其感兴趣的内容,过滤其不感兴趣的信息,且不会分辨被过滤的信息是不是有价值的思想政治教育信息。也就是说,生成式人工智能实质上掌握了内容供给权。一方面,受教育者将困于生成式人工智能打造的“信息茧房”之中,表现出认知偏差、思维僵化,进而出现“圈层化”现象,每个圈层之间沟通交流少,相互割裂,最终形成圈层隔离,从而导致“群体极化”现象出现,不利于形成主流价值观共识。另一方面,在“信息茧房”中,受教育者接收到的是生成式人工智能的定制化信息供给,为抓住受教育者的“眼球”,关于世界观、人生观、价值观、道德观等思想政治教育内容的生产空间会遭到其余更有“爆点”的内容的挤压,造成思想政治教育内容供给减少。

2. 技术崇拜掩盖信息失真,思想政治教育内容供给质量下降

当思想政治教育主客体相信生成式人工智能可以解决一切问题,迷失于生成式人工智能之中时,对生成式人工智能供给信息的判断力就会随之下降。以 ChatGPT 为例,在与其对话时,输出的答案会出现“一本正经的胡说八道”情况。在技术崇拜遮蔽下,胡说八道的答案不会受到用户的批判和质疑,受教育者倾向于把其作为正确答案而接受。更有甚者,在生成式人工智能思想政治教育内容供给中的错误被思想政治教师所指出时,技术依赖和

崇拜会导致受教育者坚信人工智能供给的信息,不愿吸收不同的意见,对思想政治教育效果造成了负面影响。再者,思想政治教育不仅是一项政治性实践,也是一种情感养成。^[7]在无法交流对话的前提下,思想政治主客体的情感联系也就日渐薄弱,思政教育的育人效能随之消减。

(三) 伦理道德风险:技术滥用催生伦理道德失范

1. 信息滥用造成隐私泄露

生成式人工智能赋能思想政治教育离不开大数据的支持,通过全景式监控,生成式人工智能会对教育者、受教育者等不同用户进行精准画像,用户的个人信息、困难情况、心理状态、兴趣偏好等私密信息都为生成式人工智能所掌握。一方面,当前尚未有对生成式人工智能监管的有效措施,大量的个人隐私数据一旦为有心之人利用,可能造成用户人、财、物的大量损失,甚至导致用户的“社会性死亡”。另一方面,被精准画像后的用户与现实中的用户未必完全一致,“虚拟我”与“本我”可能存在不同的状态,思想政治教育工作者如果仅仅根据人工智能提供的用户画像对受教育者进行判定,那么“简单、抽象的信息、数字、符号替代了人的全面本质”^[8],受教育者受到误判,这与思想政治教育的育人目标背道而驰。

2. 技术滥用导致学术不端

生成式人工智能具备参加考试、完成作业、撰写论文等功能,这就给了部分技术伦理淡薄的教育者和受教育者可乘之机。或出于便利,或追求更

好成绩,或谋取更多利益,思想政治教育主客体将本体的学习自主权、创造自主权让渡给生成式人工智能,以指令引导生成式人工智能产出更标准的答案、更优质的文章。在现实中,生成式人工智能所产出的作业答案或者论文有时难以辨认,使用者也不进行标注,容易导致学术不端现象。同时,使用者成为不当获利的主体,思想政治教育向上向善的培养目标大打折扣。

三、生成式人工智能赋能高校思想政治教育发展的治理之策

生成式人工智能赋能思想政治教育面临着意识形态、学科解构、伦理道德等风险,可以通过提高科技创新能力、破除技术迷信、嵌入伦理道德标准等路径进行治理,促进生成式人工智能耦合思想政治教育,发挥育人效能。

(一) 提高科技创新能力,牢牢把握意识形态主导权

1. 以科技创新掌握智能自主权,推动生成式人工智能本土化

国内科技公司领衔进行生成式人工智能研发,加大科研力度,打造国产化生成式人工智能。从根本上说,国内生成式人工智能发展得越成熟,其赋能思想政治教育所带来的风险就越小,或者相对可控。目前,国内科技企业已进行有益尝试,如百度的“文心一言”(如图1)、阿里巴巴的“通义千问”(如图2)等。需要注意的是,如果对本土的生成式人工智能不加监管,同样可能造成资本的裹挟与生成内容的娱乐化。



图1 百度“文心一言”操作界面

2. 以教育逻辑引领商业逻辑,保证主流意识形态权威

在打造本土化生成式人工智能赋能思想政治教育路径时,由教育部、科技部等部门对生成式人工智能的数据来源和处理过程进行监管,有意识地用社会主义核心价值观、习近平新时代中国特色社会主义思想等思想政治教育内容对生成式人工智

能进行预训练,减少博人眼球内容产出的比例,强化程序设计开发人员的监管,避免设计主体的思想偏见植入生成式人工智能中,同时建立内容反馈机制,对于使用者反馈的不当内容及时进行审查和校对,多渠道规避意识形态渗透风险。要打破技术黑箱,消除可能存在的西方意识形态弥散风险,以此维护主流意识形态权威。



图2 阿里巴巴“通义千问”操作界面

(二)破除技术迷信,突出思想政治教育学科主体性

1. 破除技术迷信,是突出思想政治教育学科主体性的大前提

高校要有意识地提升思想政治教育主客体的人工智能素养,主动谋划,积极行动,打破智能崇拜与技术依赖。通过专业人员或机构对教育主客体开展培训,帮助教育主客体在学习交流过程中提升数字素养,了解生成式人工智能的基本逻辑,使教育主客体由被动适应者转变为主动掌控者,破除对生成式人工智能给定答案的偏听偏信。当生成式人工智能“祛魅”,复归为思想政治教育手段,教育主客体就能够更加明晰思想政治教育学科的主体地位,也更能合理利用生成式人工智能以满足个人的思想政治教育需要。

2. 发挥主观能动性,优化思想政治教育内容供给

对于思想政治教育者而言,要吃透马克思主义等理论知识,扩大主流意识形态的输入占比,涵养提供、审核、判断思想政治教育内容供给质量的能力,加强与受教育者的情感沟通与交流,主导生成式人工智能、受教育者与自身的三角关系,构筑良

性循环。对于思想政治教育对象而言,要主动接触、学习、吸收思想政治教育内容,提升信息判别能力,敢于质疑批判生成式人工智能供给内容,进而反哺思想政治教育内容供给,促使生成式人工智能输出更高质量的、形式更加丰富的思政教育内容。思想政治教育本质上是人的教育,是以人为本的教育,通过发挥个体的主观能动性,加强个体间的良性互动,有助于巩固思想政治教育的主体地位,优化思想政治教育内容供给。

(三)嵌入伦理道德标准,守好技术赋能的伦理道德底线

1. 建立健全生成式人工智能发展法律法规,由外而内约束道德失范行为

我国新一代人工智能治理专业委员会发布的《新一代人工智能伦理规范》就是有益的尝试,规范提出了六项基本伦理规范,分别是增进人类福祉、促进公平公正、保护隐私安全、确保可控可信、强化责任担当、提升伦理素养。^[9]为生成式人工智能赋能思想政治教育提供了基本遵循,在人工智能管理相关文件基础上,完善生成式人工智能赋能思想政治教育伦理规范实施细则有利于规避如信息泄露等因缺乏监管而导致的伦理风险。以外在规定制

度明确伦理道德风险的主体责任、规范要求、处理办法,以切实地执行落实相关制度规定,为生成式人工智能赋能思想政治教育提供强有力的法律保障。同时,顺应人工智能发展趋势,完善学术伦理道德规范相关规定,有助于约束生成式人工智能使用者的学术道德失范行为。

2. 将伦理道德嵌入生成式人工智能逻辑中,从内而外避免伦理道德风险

以 ChatGPT 为例,其训练分为三个步骤,分别是通过人工标注的方式生成微调模型、训练一个可以评价答复满意度的奖励模型及利用第二步训练好的奖励模型,通过强化学习算法来优化答复策略。^[10]在这一过程中,标注团队发挥着重要作用,将思想政治教育队伍专家纳入标注团队中,将相应的伦理道德嵌入模型之中,通过奖励模型对模型输出效果进行正向引导,对生成式人工智能进行训练,从而使其能够由内而外避免伦理风险。

综上所述,生成式人工智能赋能高校思想政治教育是时代发展的必然趋势,对立抗拒不是应对这一趋势的应然姿态,合作共赢才能真正推动思想政治教育发展。但思想政治教育改革牵一发而动全身,故而革新的步子不宜大、宜稳,在生成式人工智能尚未发展成熟时,不宜盲目引入思想政治教育之中,还须徐徐图之,可以通过试点方式先行探索。此外,就生成式人工智能与高校思想发展研究主题而言,可跨学科、多领域、持之以恒地开展研究工作,进一步丰富、完善研究内容,如此才能为可能产

生的风险做好理论准备与应对。

参考文献:

- [1] Michel-Villarreal R, Vilalta-Perdomo E, Salinas-Navarro DE, et al. Challenges and Opportunities of Generative AI for Higher Education as Explained by ChatGPT[J]. Education Sciences, 2023 (9): 856.
- [2] 中国政府网. 国务院关于印发新一代人工智能发展规划的通知[EB/OL]. http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm, 2023-11-07.
- [3] 李辽宁. 内涵式发展: 新时期思想政治教育学科建设的思考[J]. 思想政治教育研究, 2013(3): 6.
- [4] 喻国明, 苏健威. 生成式人工智能浪潮下的传播革命与媒介生态——从ChatGPT到全面智能化时代的未来[J]. 新疆师范大学学报(哲学社会科学版), 2023(5): 86.
- [5] 徐学彬, 徐登献. 算法推荐嵌入思想政治教育之“理”“机”“危”与“策”[J]. 理论导刊, 2023(9): 126.
- [6] 阮一帆, 王智博. ChatGPT之于思想政治教育的伦理风险及应对策略[J]. 社会主义核心价值观研究, 2023(3): 52-53.
- [7] 米华全. 智能思政伦理风险的生成逻辑、表现形式及防控机制[J]. 中国电化教育, 2023(2): 111-117.
- [8] 陈青松. 网络信息技术对大学生的不良影响[J]. 探索与争鸣, 2007(12): 50.
- [9] 中华人民共和国科学技术部.《新一代人工智能伦理规范》发布[EB/OL]. https://www.most.gov.cn/kjbgz/202109/t20210926_177063.html, 2023-11-07.
- [10] 丁磊. 生成式人工智能——AIGC的逻辑与应用[M]. 北京: 中信出版集团, 2023: 83-84.

[责任编辑: 石俊华]

Risk and Governance: Generative Artificial Intelligence and the Development of Ideological and Political Education

Yang Zhong

(Fujian Normal University, College of Chemistry and Materials, Fuzhou, Fujian, 350117, China)

Abstract: The development of ideological and political education empowered by generative AI is the response of ideological and political education to the needs of national development strategy, the development and innovation of the discipline itself, and the development needs of the subject and object of ideological and political education. Influenced by the limitations of the development of generative artificial intelligence and the specificity of ideological and political education, the process of generative artificial function empowering ideological and political education may result in ideological, disciplinary deconstruction and ethical and moral risks, which can be governed utilizing improving scientific and technological innovation, breaking down technological superstition, and embedding ethical and moral standards.

Key words: generative artificial intelligence; ideological and political education; ChatGPT