

• 科学技术文化研究 •

## 从科学文化看人工智能发展的“可能”与“应当”

### The “Possibility” and “Oughtness” of AI Development from the Perspective of Science Culture

王前 /WANG Qian 张卫 /ZHANG Wei

(大连理工大学人文学院哲学系, 辽宁大连, 116024)  
(Department of Philosophy, School of Humanities, Dalian University of Technology, Dalian, Liaoning, 116024)

**摘要:**当代人工智能的迅猛发展呈现出多种可能路径,但科学技术上可能的事情并不都是应当做的事情。从科学文化视角看,人工智能的社会影响作为一种文化现象,需要保证人有人的用处、保护人的自然本能、避免由于“技术鸿沟”影响社会公正、防止因传承方式简单化削弱思想文化的稳定发展。有必要在人工智能发展的“可能”与“应当”之间建立理性平衡机制,开展前瞻性决策评估、动态技术治理和阶段性反思调整。中国传统文化“以道驭术”“以人为本”“顺应自然”和“德福一致”等思想原则与现代科学文化相结合,有助于协调人工智能发展的“可能”与“应当”的关系,防范可能出现的各种风险,促进人类社会的和谐发展。

**关键词:**人工智能 可能 应当 科学文化

**Abstract:** The rapid development of contemporary artificial intelligence (AI) presents various possible paths, yet scientifically and technically feasible endeavors do not necessarily equate to morally or socially appropriate actions. From the perspective of science culture, the social impacts of AI as a cultural phenomenon necessitate ensuring humanity's unique purpose, safeguarding people's natural instincts, mitigating the impact of the digital divide on social justice, and preventing the undermining of the stable intellectual and cultural development due to the simplification of the inheritance of ideas. It is necessary to establish a rational balance mechanism between the “possibility” and “oughtness” of AI development that allows for forward-looking policy decision evaluation, dynamic technological governance, and periodical reflections and adjustments. Integrating traditional Chinese philosophical principles such as “guiding Shu with Dao” “putting people first” “harmony with nature” and “alignment of virtue and happiness” with modern scientific culture can help facilitate the coordinated balance between the “possibility” and “oughtness” in AI development, prevent various potential risks related to AI development, and promote the harmonious development of human society.

**Key Words:** Artificial Intelligence; Possibility; Oughtness; Science Culture

中图分类号: G02; TP18 DOI: 10.15994/j.1000-0763.2026.08.013 CSTR: 32281.14.jdn.2026.08.013

基金项目: 教育部人文社会科学研究规划基金项目“技术设计的伦理嵌入机制研究”(项目编号: 24YJA720012)。

收稿日期: 2024年9月5日; 返修日期: 2026年6月3日

作者简介: 王前(1950-)男, 辽宁沈阳人, 大连理工大学人文学院哲学系教授, 研究方向为技术哲学、科技伦理、机体哲学。Email: qianwang@dlut.edu.cn

张卫(1983-)男, 河南开封人, 大连理工大学人文学院哲学系教授, 研究方向为技术哲学、科技伦理。Email: zhwei@dlut.edu.cn

在生成式人工智能引起社会普遍关注的今天,人工智能的迅猛发展引发了人们各种复杂心态。人工智能的普遍应用提高了工作效率,改善了人们的生活质量,“人工智能+”成为新质生产力的重要标志。<sup>[1]</sup>很多人以为人工智能的应用都具有积极意义,凡是人工智能技术可能做到的事情都应努力使之变成现实。然而,人工智能的应用带来了就业结构的明显变化,现在很多脑力劳动者的岗位也正在或已经被人工智能取代,他们的再就业面临很多新挑战。人工智能的应用还带来了某些人特别是青少年对人工智能的过度依赖,带来了人际关系和日常生活的很多新变化,新的“技术鸿沟”给跟不上人工智能技术更新换代的人们带来了很多障碍。由此看来,并非人工智能可能做到的事情都是应当做的事情。人工智能的应用是一种科学文化现象,需要从科学文化如何良性发展的视角看待人工智能发展的“可能”与“应当”。下面就此做一些具体讨论。

## 一、人工智能时代如何保证“人有人用处”?

人工智能技术的兴起与控制论的创立有直接关系,而控制论的创立者维纳早在20世纪50年代初就提出了“人有人用处”的命题,<sup>[2]</sup>这是非常有远见的。工业革命以后,采用机械化、自动化技术的大机器生产和应用替代了人们大部分体力劳动,而当代人工智能则越来越多地替代人们的脑力劳动。这种趋势发展到极致,似乎可以取代所有劳动。人们无须做任何工作,完全由社会保障衣食无忧。这种发展前景对人类来说是最佳选择吗?

马克思主义唯物史观一直主张劳动不仅创造了人,也是人之为人的本质特征。<sup>[3]</sup>劳动不仅仅是个人的生理和心理活动过程,也联系着社会价值的创造与分配、文化的传承与发展、人类本性的存续与不断提升。人类需要通过科技与生产手段摆脱危害人类健康的繁重劳动方式,创造更适合人类社会可持续发展的劳动方式,更有创意,更有价值,但不应该追求完全

脱离劳动,因为这样会导致人的本质异化,人失去了应有的用处,从而失去了协调人类社会生活、人与自然关系、人的自身生存发展的基本能力。当人类将自身置于极度智能化、自动化、机械化的环境中,而这种环境又存在破坏社会公正、生态环境、文化遗产的危险时,人类或许没有能力扭转这种局面,这显然不是人类发展人工智能技术的初衷。

如果说上面的讨论还比较抽象,下面以《庄子》中“庖丁解牛”寓言为例做一些具体讨论。“庖丁解牛”是体现人有人用处的一个典型事例。庖丁高超技艺的价值不仅在于高效、省力、有美感,还在于这种劳动使得身心愉悦,有益养生。( [4], pp.15-20 ) 假设现在能够发明一把“智能刀”,可以透视牛的所有筋骨皮肉,特别是一些极细微的骨缝和关节,然后精准地自动下刀切割,使刀的磨损降到最低限度,不仅又快又好,而且比庖丁亲自动手做得更干净利落,这时庖丁解牛的高超技艺还有存在的价值吗?如果把着眼点仅仅盯在解牛效果上,发明这样一种“智能刀”在技术上是可能的,完全可以替代庖丁的技艺。问题在于,要设计、制造和使用这样的“智能刀”,价格一定非常昂贵,人类为什么要搞出这样一种发明来代替庖丁的技艺?要保证这种“智能刀”使用的高效率,需要把很多头牛从各地源源不断地送到某一定点屠宰工厂,然后又把宰好的牛肉不断运送出去,这样做在经济上划算吗?这种“智能刀”如果出现故障,还需要人来维修,维修的成本当然也会很高。用人可以做得很好的事情一定要用智能机器,还要加上高昂成本和能耗,这种人工智能运作方式合理吗?于是,“智能刀”与社会、环境、人的身心活动等要素的关系就会逐渐浮现出来。而追求这些要素的和谐,需要关注庖丁讲的关键的一句话:“臣之所好者道也,进乎技矣”,就是说庖丁追求的是“技”之上的“道”。

中国传统技术文化强调“以道驭术”。( [4], p.135 ) “道”的引导是保证技术应用不至于引起人的本质异化,而人的用处不会“缺位”的决定性因素。从技术文化视角看,“道”的词

源本义是“道路”。人们走路需要选择合理的、优化的途径,引申为做任何事情都需要采用合理的、优化的步骤和方法。当这种要求达到最合理、最优化的程度,比如当庖丁的技艺出神入化的时候,就进入他所说的“道”的境界。中国古代讲的“道”有“小道”“大道”之分,“小道”是庖丁解牛中涉及的只考虑操作者、工具、对象三者关系的“道”,而“大道”还要考虑技术活动与社会生活、自然环境、伦理道德以及人的身心健康等要素的关系,追求整体上协调这些要素的最合理、最优化的程序和方法。中国古代很多学者都在抨击“奇技淫巧”,就是要用“道”限制这些对人类社会生活有害的技术应用。( [4], pp.7-14) 庖丁解牛的合理性在于这种劳动不仅有益于身心健康,而且因地制宜、经济实用、保护环境,完全符合“以道驭术”的要求。推而广之,假设现在一个农民将所有耕作环节全都人工智能化,他的任务只是按开关甚至连这种事情也不用管,他还是真正意义上的“农民”吗?这种生产方式对他而言就是应该追求的“幸福”吗?他如何保证自己的身体健康?他需要去健身房吗?如果他从事耕作,是否更有益于健康?我们可以找出适当耕作对人的身心健康的多种好处,如空气新鲜、新陈代谢旺盛、亲近自然、心情愉悦、有成就感,等等,而所有这些好处都是人工智能技术带不来的。

在人工智能高度发达的时代,人类要保证人有人的用处,实现对人工智能合理发展的有效控制,需要运用人类特有的,即人工智能无法取代的智能,预见和评估人工智能产品的利弊得失,引导其设计、制造、使用和调适,这一切都需要基于直观体验、想象、灵感的创造性思维能力,以及通过这种能力获得的隐性知识,即“可意会而不可言传”的技艺性知识。<sup>[5]</sup>现在一些高技术企业里的高级技师就是“庖丁”一样的人物。比如,当十分复杂的计算机系统出现故障时,有些专家凭经验能很快查明原因;有些非常精密的航天器、数控机床、机器人突然失灵时,一些专家凭直觉就能判断毛病出在哪里。医生通过人工智能设备判断患者

的病症和制定治疗方案、发明家在设计人工智能设备时发挥价值敏感性的作用、高级技工调适人工智能设备中各精密部件的匹配关系等,都需要这种人类特有的智能。人工智能技术主要以人类自身的生理和心理为作用对象。传统技术应用的后果基本上可以明确预见,而人工智能技术应用后果带有很多不确定性,需要更关注技术风险。传统技术基本上可以通过直观体验加以把握,而人工智能技术的作用过程很多时候相当隐蔽甚至处于“黑箱”状态。<sup>[6]</sup>在人工智能产品的设计、制造和使用过程中,如何保证其中的价值因素符合技术活动与社会生活、自然环境、伦理道德以及人的身心健康等要素之间的和谐关系,也需要发挥人类的特殊作用,及时发现问题和解决问题。人工智能技术无论发展到何种程度,都不可能也不应该脱离人的参与。

人工智能将来是否会自动产生人类所有的智能,具有自主意识、创造性思维能力和自我复制功能,甚至将来可能“统治”人类?这种担忧是不了解人工智能基本运行的逻辑,将人工智能“神化”所致。<sup>[7]</sup>人工智能依赖的各种算法、大数据、生成式大模型等都是对大脑左半球部分功能的模拟,相当于丹尼尔·卡尼曼(Daniel Kahneman)所说的“系统2”的慢思考,而大脑右半球和边缘系统具有的基于直观体验、想象、灵感的创造性思维能力是人类长期进化和文化积淀的产物,相当于“系统1”的快思考,<sup>[8]</sup>具有整体性、有机性和历史传承性,这是人工智能从原理上看就无法取代的。因此有学者主张“人工智能是最好的左脑,但从未发展出右脑的能力”。<sup>[9]</sup>人类特有的,人工智能无法取代的智能,对应于中国传统文化中称之为“心”的认知能力。这里所说的“心”不是指现代解剖学意义上的心脏,而是指人的从事认知和心理活动的思维器官(中国古代哲学家王阳明说过“心不是一块血肉,凡知觉处便是心”)。<sup>[10]</sup>从历史和现实的语境看,“心”的认知特征正是强调从整体的、有机联系的、体验性的角度理解事物。这种认知活动在负担过重时心脏明显感到压力,所以人们才会认同“心

之官则思”。<sup>[11]</sup>西方文化一直强调“动脑”的作用,注重逻辑推理和定量分析,工具理性比较发达,而整体的、有机的、体验性的思考往往被置于次要位置上。与“心”的认知理解相对应的有现代西方哲学中现象学、解释学、隐喻研究和具身认知的研究成果,这些成果对理解人工智能与人类智能关系也有启发意义。然而“心”不仅是认知主体、道德主体、审美主体,也是行为主体、决策主体,具有社会调节和文化传播功能,因而在体现人有人的用处方面更具代表性和实效性。当然,在人工智能时代发挥“心”的作用,不能停留在传统的思维形态层次上,而是要与现代认知科学和认知哲学研究成果相结合,实现创造性转化,将整体思维与分析思维有机地结合起来,将直觉与逻辑有机地结合起来,在一个更高层次上发挥“心”的功能。<sup>[12]</sup>

在现代意义上发挥“心”对科技和经济发展的引导作用,不同于专业化、局部化、单一化的逻辑分析和科技应用思路,它有助于避免技术发明和应用成为德国哲学家海德格尔所说的摆置人、控制人的“座架”。<sup>[13]</sup>这是一种能够操控人工智能技术而不至于为人工智能技术摆布的智慧。当感官为外物所累而不再听从内心的召唤时,当局部的、片面的外界刺激导致精神世界的碎片化时,当人工智能技术应用存在隐私泄露、非法侵害和不当控制的风险时,“心”会从整体上敏锐意识到这些“异化”现象,主动消除其负面影响。用“心”去感悟人工智能的应用,就会觉察哪些应用对人类是有益的,哪些会导致有害结果,使人们在人工智能技术应用中始终保持清醒、审慎的态度,这是人有人的用处的根本所在。

## 二、人工智能时代如何保护人的自然本性?

所谓人的“自然本性”,并非指人类的原始状态,而是人类在存续发展中必须保留的一些自然属性和本能,包括感觉、知觉、情感、行动、人际交往等等。这些自然本性的价值在

于保证了人之为本真的存在和社会生活的稳定发展。科学技术的发展已经通过各种器具强化了人类的一些自然本性,但器具的不当使用也可能无形之中弱化了人类的一些自然本性,而人工智能的应用在改变人类自然本性方面的作用尤为突出。有些人工智能产品在设计时并没有削弱人的自然本性的意图,但在应用中出现了异化现象。有些则在设计之初就将功利目标置于合理需求之上,而且不计后果,这就更需要警惕。

人工智能技术逐渐改变或替代人的某些自然本性,在绝大多数情况下是好事,如传感器增强人的感知能力,自动化减轻人的繁重体力和脑力劳动负担,网络开拓人际交往新渠道,让人们获得更多发展机会。可是这种改变或替代一旦超过合理限度,导致人的自然本性逐渐出现不可逆的衰退,就会产生一些难以控制的负面影响。人工智能技术如果过度替代人的天然感知能力,会造成人的视力和听力不可逆的减退。不少青少年从小就长时间看手机和电视屏幕,很小就戴上近视眼镜,是一个值得重视的信号。

青少年自制能力较差、有些家长对此不以为然、手机和游戏机内容日益具有诱惑力,这些因素叠加在一起,影响的不只是青少年的视力和听力,还包括对学业的专注和对现实生活的正确理解。人工智能技术的不当应用可能造成信息传播碎片化,进而影响人们思维的完整性和系统性,人所具有的直观体验的、综合的思维能力逐渐被数字化、工具化的思维能力挤压,变得狭隘、刻板、僵化。人的思维过度依赖计算机的逻辑,会逐渐丧失主观意志,越来越听命于人工智能给出的算法决策。

如果人工智能技术无限制地取代人的交往能力,这种交往将变得愈发数字化和网络化,造成人自身的知情意相统一的交往能力减弱,人与人之间直接接触的深情交往逐渐减少,甚至使一些人遇到特殊困难时变得孤立无助。一个比较典型的例子是,医生运用人工智能技术辅助诊断是必要的,但同时需要具备对病情的整体的、综合的、直觉的判断能力,具备与患

者的情感体验方面的沟通能力。医生如果缺少这种沟通能力,患者的病情很难得到准确诊断和及时治疗。在战略评估、人才评价、素质教育、创意产业、文化传播等领域,更需要充分运用人的整体性的、直观体验的、综合的思维能力,使人工智能成为人的思维能力的重要补充而不是完全彻底的替代工具。人工智能技术的不合理应用会影响人的全面发展,出现片面强化某种智能而整体失衡的现象。

现在满足人类工作和生活需要的各种智能机器人正在逐渐扩大使用范围,人类如何在保护自然本性的前提下与智能机器相处?在设计和使用针对儿童开发的教师机器人和陪伴机器人时,不应减少儿童与成人的直接接触机会,不能使儿童对机器人形成依赖,影响儿童的情感和心理健康发展;<sup>[14]</sup>在使用针对老年人开发的护理型或陪伴型机器人时,不应减少老年人与其他人直接交往的机会,要避免其子女和护理人员推卸其应尽义务和责任;<sup>[15]</sup>在使用性爱机器人时,要警惕出现让女性身体物品化和商品化倾向。<sup>[16]</sup>现在很多学者关注人类与智能机器人之间的“共情”问题。智能机器人确实也能够提供满足人类特定需要的情感服务,情感计算技术会使得智能机器人表达的情感日益“逼真”。然而与智能机器人之间的情感交流能够完全取代人际真实的情感交流吗?哪些人与人之间的真情实感是智能机器人不能够取代或者不应该取代的?这些问题都需要进一步深入研究。

在保护人的自然本性方面,中国传统技术文化强调技术活动要“以人为本”和“顺应自然”,二者有着内在的联系。“以人为本”指的是从人类根本利益出发,保障人的身心和谐、人格尊严、全面发展。“顺应自然”指的是技术活动要顺应自然规律和人的“天性”,保证技术活动与社会、自然的充分和谐。如果一项技术发明限制和伤害了人的“天性”,侵犯了人格尊严,给社会和自然带来负面影响,这样的发明就应受到禁止。道家学说主要从人与自然关系角度强调技术活动要顺应自然,([17], p.31)儒家学说主要从社会关系角度强调“人

道”要符合“天道”。([17], p.24)人们在利用工具、机械这些“物”的手段时,不能为“物”所累,不能成为“物”的奴隶。近现代很多学者也有类似看法,比如近代西方出现过人本主义伦理学的主张,康德关于“人是目的而不是手段”的观念有着深远影响,<sup>[18]</sup>现代西方学者韦伯、哈贝马斯等人也对工具理性的社会影响有过深入分析。<sup>[19]</sup>现在很多人热衷于人工智能技术应用带来的个性化的、局部的、短期的好处,对人工智能技术应用可能带来的整体的、长期的负面影响和风险很茫然,因而更需要“以人为本”和“顺应自然”这种整体的、宏观的、注重有机联系的视角。

在人工智能时代保护人的自然本性,需要提倡一种理性的反思和警觉意识。因为人工智能技术对人的自然本性的负面影响往往是在人们不知不觉中发生的,这是一种渐进积累的过程,达到一定程度后才会有明显的表现,但此时想要扭转往往为时已晚。要发现和鉴别在初期或过程中人工智能技术对人的自然本性的负面影响,必须“用心”体察人的自然本性在人工智能技术作用下的变化趋势,注意发现能够预示危险趋势的征兆,及时加以调整。人工智能技术对人的自然本性的影响有些至今还处在不明晰、不显著的状态,其积极价值往往使人们忽略其负面影响的可能存在,因而需要仔细维护和发展人的自然本能,以利于人类社会的长久发展。

### 三、人工智能时代如何维护社会公正?

人工智能技术的设计和应用本身很少有违背社会公正的初衷,但人工智能技术作为一种工具总要服从于各种具体目的,满足不同功利需要,运用不当就可能造成违背社会公正的后果,而人工智能技术高度的专业性、不透明性、不确定性可能会加剧这种局面。大数据技术的滥用可能带来侵犯个人隐私的风险,一些人可能被迫或在不知情的情况下上传自己的隐私信息,失去其自主选择的权利。<sup>[20]</sup>人工智能算法决策可能带来某种“算法歧视”,使一些社会

上的弱势群体利益受到损害。<sup>[21]</sup>人工智能技术在设计、制作、使用阶段都可能被嵌入特定价值因素,成为某些个人或企业牟取不正当利益的手段,这些手段隐藏在人工智能技术高度专业化、程序化、标准化的外在形象后面,而受到侵害的用户却浑然不知。<sup>[22]</sup>人工智能技术发展应该同基于人的全面发展的价值体系和伦理原则相适应,尊重人类的个性化差异及其享有的权利与义务,使人工智能技术发展在人类可控范围内最大限度地造福人类,这就是现在学术界热议的“价值对齐”要求。<sup>[23]</sup>然而人类并不是通过设置一个统一机构对是否“价值对齐”进行裁决和引导的,不同国家、地区、企业的人工智能技术设计者、使用者、管理者对如何“价值对齐”都有各自不同理解,因而维护社会公正面临很多隐蔽的困难。<sup>[24]</sup>现在学术界和社会上不断提出人工智能技术发展的各种伦理准则和监管措施。这些准则和措施有助于从制度设计层面维护社会公正,然而从协调人工智能发展的“可能”与“应当”关系的角度看,还有些深层问题值得思考。

其一,目前制定的各种伦理准则和监管措施基本上针对人工智能技术应用中已经出现的负面影响,这是从“后果论”角度出发的,如何在预见有可能出现负面影响时就尽早做好防范?中国传统技术文化的“以道驭术”的理念在这方面可以提供启发性的思路。人工智能技术的不当应用可能导致的社会不公正现象,实质上是技术应用与社会生活的不和谐,从“道”的宏观视角考察可以发现一些端倪,比如某种新的人工智能技术应用在给相关企业或机构带来显著利益的同时,并没有使用户群体产生更便捷、更安全、更有实效的感觉,反而要用户不断承诺提供个人信息、意向和不追责的保证,这种缺乏诚意的态度就是很可疑的。如果某种新的人工智能技术应用明显出现有意摆布、误导、强迫用户消费行为的倾向,技术供应商却不愿承担相应责任和义务,反而将问题根源归因于技术的不确定性,这就已经表现出这种人工智能技术应用的价值取向,出现社会不公正的后果很难避免。人工智能技术的价值敏感因

素不大容易识别,需要细心体察,见微知著,尽早发现苗头,及时加以限制。

其二,有些社会不公正现象是由于“技术鸿沟”造成的,而消除“技术鸿沟”同样需要通过技术创新,引导弱势群体顺利跨越“技术鸿沟”。如果将人工智能技术应用带来的“技术鸿沟”视为不可避免,所有使用者都必须接受,由此造成的不便就是必要的代价,就会任由“技术鸿沟”进一步加深加宽,使社会不公正现象更为严重。按照“以道驭术”和“以人为本”的理念,人工智能技术的开发利用应当惠及所有利益相关者,特别是那些跟不上人工智能技术更新换代的人们。在人工智能技术开发“可能”的范围内,为不熟悉新技术使用而影响日常工作和生活的人们提供适合他们具体情况的“绿色通道”,在操作层面给出更贴切的语音和文字提示,尽可能降低“技术鸿沟”造成的不公正现象的影响。在疫情监测、银行、交通、医疗等领域,都会遇到类似情况。

其三,在制定了人工智能技术发展的各种伦理准则和监管措施之后,应当由谁来具体落实?除了各级伦理审查和监管机构外,公民特别是利益相关者的参与也是十分必要的,但公众参与监管往往止步于人工智能技术应用的高度专业化和不透明状态。这里的关键是建立有效自律和他律机制。<sup>[25]</sup>他律需要有具备价值敏感性、有能力识别伦理和法律问题、能够有效履行监管职能的利益相关者代表,通过特定机构具体落实监管措施;而自律则需要培养专业技术人员的高度社会责任感和伦理意识,培养人工智能技术使用者理性的自我意识和自主能力。如果专业技术人员不能自觉根除利用人工智能技术牟取不当利益的私欲,由此产生的社会不公正现象总有可能绕过层层监管不断出现。如果使用者缺乏自控能力,总有可能在不自觉状态下被操控成为受害对象。现在有些接触人工智能技术的青少年出现过度依赖难以自制的问题,学校和家长往往采取断网和强制隔离的手段,很难从根本上消除过度依赖现象。必须从自律角度培养青少年对待人工智能技术的理性态度,真正将人工智能视为必要的工具

而不是摆布自己的依赖对象,才符合“以道驭术”的理念。在设计制造面向老年人、幼儿和缺乏使用技能的人群的人工智能产品时,应该仔细体察这些人在使用过程中可能出现的困难、风险和意外情况,尽可能增强宜人性和安全性。

在人工智能时代维护社会公正,从本质上讲是如何协调社会公德与实际效益的关系问题。按照中国传统技术文化的说法,是如何贯彻“德福一致”理念的问题。<sup>[26]</sup>对“德”的要求并不排斥对“福”的向往,但对“福”的追求不能逾越“德”的底线,否则就会失去“道”的约束,出于急功近利的考虑而带来消极后果。如果某些人工智能技术只是给少数人带来各种福祉,同时给其他利益相关者造成损害,即对“福”的追求建立在不公正基础上,这就违背了“德福一致”的理念。或许少数从人工智能技术某种应用中得到格外好处的人,并未意识到这样可能造成对他人的不公正后果,因而需要培养人工智能技术的设计者、制造者和使用者普遍具备这方面的伦理道德意识,注意发挥人工智能技术造福全人类而不是少数群体的作用,共同维护人工智能时代的社会公正。

#### 四、人工智能时代如何加强文化传承?

人工智能技术应用的一个突出特征是用技术人工物的操作取代人的切身体验,而这种体验是人工智能无法模拟、记录和保存的。文化传承中的“非物质文化遗产”部分就需要文化继承人的直接参与和亲身传授。美国著名技术哲学家米切姆教授在考察了中国书法写作后就曾指出,一个儿童在用毛笔练习书法,他使用时能感受到与传统相遇,仿佛牵着祖辈的手。如果送给他一只新的人工智能毛笔,可以按照需要自动写出非常漂亮的毛笔字,他会感觉到在牵谁的手呢?这种人工智能毛笔会很快更新版本,卖得很快,这对儿童练习书法又有什么帮助?尽管人们用起来省力,效率又高,但中断了历史文脉,发挥不了通过练习书法涵养身心的作用,这种发明对培养儿童又有什么意

义?毛笔据说是战国时代的发明,一直传承到现代,传承的一个重要载体就是人类世代之间通过口手相传积累延续的操作体验,其中有很多隐性知识。如果发明了一种人工智能的自动毛笔,这种传承也就中断了。这是偏离人类文化传承之“道”的。

在对文化传承的理解方面,存在着一种将思想文化成果简单地技术化、数字化、工具化的倾向,而人工智能技术在这方面做得可能比人类更好。不仅历史上的典籍文献通过数字化可以长久完整保存,人们日常的工作文件、交往记录、图片资料、音频视频等资料都可以及时存储在“云空间”里,随时调取。人工智能技术可以将物质文化、制度文化和精神文化的各种细节特征完整地记录下来,这方面的文化传承功能远远超过了历史上的各种相关努力。然而文化传承仅仅靠人工智能技术就足够了吗?完整的文化包含着人工智能技术无法替代的切身体验,包含类似庖丁具有的那种“可意会而难以言传”的技艺、对重大社会事件和文化形态意义的理解、在长期文化氛围中养成的精神气质、蕴含普遍哲理的实践智慧,这些都不是可以简单数字化存储的。哲学界经常提到的一个思想实验是普特南的“缸中之脑”。<sup>[27]</sup>假设这个“大脑”可以接受文化传承所需要的所有信息,并且在与外界的互动中感受到自身正在从事文化传承的相关活动,但意识不到自身的真实处境,是否可以认为文化传承通过“缸中之脑”也可以进行?人们从直觉上很难接受这种判断,因为文化传承有着真实的物质文化、制度文化和精神文化的载体,需要通过人们的亲身感受将文化各种要素联结成有机的系统,才可能将文化形态完整、准确、生动地复制和传承下去。而人工智能在这一过程中提供的技术手段尽管是强有力的,但仍然是一种工具。

## 结 论

通过上述讨论可以看出,有必要在人工智能发展的“可能”与“应当”之间建立理性平衡机制,这就需要开展前瞻性的决策评估、动

态的技术治理和阶段性的反思调整。

就开展前瞻性的决策评估而言,人工智能发展需要在各种“可能”方向上预先进行是否“应当”的评估,根据人类能力和人工智能技术应用相互适应的情况,决定哪些探索方向可以继续前行,哪些应该及时“叫停”。在对人工智能技术进行决策评估时,专业技术人员、企业管理者和政府部门的负责人可能更偏重“福”的方面,而专业领域之外的人文学者、社会管理部门和普通民众可能更偏重“德”的方面,这很可能造成对“福”的偏重掩盖“德”的问题,而对“德”的偏重并没有抓住隐含在人工智能技术应用中的核心问题,起不到有效作用。所以,人工智能发展的决策评估特别需要专业技术人员与技术伦理学者相互理解、共同参与、密切配合,开展卓有成效的对话。当前在很多国家得到广泛推广的“负责任创新”模式,<sup>[28]</sup>为在人工智能发展的决策评估中贯彻实施“德福一致”要求提供了重要启示。人工智能发展中的技术创新活动要在“负责任”的前提下展开,通过前瞻性决策评估使各种风险和隐患及时得到化解。

对人工智能发展进行动态的技术治理,可以通过制度化的途径发挥作用,对一些可能存在风险和争议的开发项目进行伦理审查和监管,合理调整人工智能技术发展策略和使用方式,形成循序渐进取代人类各种能力的良性局面。也可以通过舆论监督和科学普及的途径发挥作用,转变人们的观念,建立人工智能技术应用的理性秩序。对人工智能发展的技术治理也包括通过科技伦理教育的途径影响理工科学生以及在职科技工作者的职业操守,使人们自觉地约束自己不去做违背科技伦理和法律的事情。人工智能技术应用中的道德自律需要从行动做起,使行动成为习惯,以此涵养性情,逐渐形成道德良知。人工智能产品的技术设计要贯彻“道德物化”的理念,<sup>[29]</sup>需要充分考虑保护人的自然本性,在可能造成负面影响时能够及时发出警示或限制,比如青少年阅读电子屏幕时设置定时休息功能。从教育方面,需要培养人们理性使用人工智能产品的技术素养,提

高保护自身的生理和心理健康的自觉性。

在人工智能发展的“可能”与“应当”之间建立理性平衡机制,不可能是一蹴而就的事情,需要不断进行阶段性的反思调整。人们应该与人工智能技术的应用保持一定的“心理距离”,即人们在心里意识到人工智能技术毕竟是一种工具,不能无限制地贴近,必要时应寻找替代的生活和工作方式,以保持这种距离感。这种“心理距离”能够给反思留出足够的空间。在人工智能发展的“可能”与“应当”之间的理性平衡,应当逐渐形成一种文化氛围,不只影响科技工作者、产品制造者和企业管理者,也影响所有人工智能技术产品的使用者,影响他们的思维方式和行为习惯,这样才能从根本上保证人工智能技术发展中各种相关要素之间的充分和谐。

#### [参考文献]

- [1] 黄汉权. 深刻领悟发展新质生产力的核心要义和实践要求 [EB/OL], 求是网, [http://www.qstheory.cn/dukan/qs/2024-06/01/c\\_1130154169.htm](http://www.qstheory.cn/dukan/qs/2024-06/01/c_1130154169.htm). 2024-06-01.
- [2] Wiener, N. *The Human Use of Human Beings: Cybernetics and Society* [M]. Boston: Da Capo Press, 1988, 7.
- [3] Banaji, J. 'Modes of Production in a Materialist Conception of History' [J]. *Capital & Class*, 1977, 1(3): 20-44.
- [4] 王前. “道”“技”之间——中国文化背景的技术哲学 [M]. 北京: 人民出版社, 2009.
- [5] 王前、于雪. 从体验到领悟: 科技伦理教育的价值目标与实现路径 [J]. *自然辩证法研究*, 2024, 40 (3): 125-131.
- [6] Bélisle-Pipon, J. C., Monteferrante, E., Roy, M. C., et al. 'Artificial Intelligence Ethics Has a Black Box Problem' [J]. *AI & Society*, 2023, 38(4): 1507-1522.
- [7] Larson, E. J. *The Myth of Artificial Intelligence: Why Computers Can't Think the Way We Do* [M]. Cambridge: Harvard University Press, 2021, 3-19.
- [8] Kahneman, D. *Thinking, Fast and Slow* [M]. New York: Farrar Straus Giroux, 2012, 19-39.
- [9] 洪小文. 计算机是最好的左脑 [N]. 中国青年报, 2016-11-15 (11).
- [10] 王阳明. 传习录 [M]. 陆永胜 译著, 北京: 中华书局, 2021, 505.
- [11] 孟子. 孟子 [M]. 杨伯峻、杨逢彬 译注, 长沙: 岳麓书社, 2016, 224.

- [12] 王前. 谈“心”: 从传统到现代 [M]. 北京: 人民出版社, 2024, 240–248.
- [13] Botin, L. 'Why Heidegger Makes Sense in Contemporary Philosophy of Technology' [J]. *Foundations of Science*, 2022, 27(2): 345–350.
- [14] Selwyn, N. *Should Robots Replace Teachers?: AI and the Future of Education* [M]. Cambridge: John Wiley & Sons, 2019, 12–19.
- [15] Jenkins, S., Draper, H. 'Care, Monitoring, and Companionship: Views on Care Robots from Older People and Their Carers' [J]. *International Journal of Social Robotics*, 2015, 7: 673–683.
- [16] Peeters, A., Haselager, P. 'Designing Virtuous Sex Robots' [J]. *International Journal of Social Robotics*, 2021, 13(1): 55–66.
- [17] 王前. 中国科技伦理史纲 [M]. 北京: 人民出版社, 2006.
- [18] 康德. 实践理性批判 [M]. 韩水法 译, 北京: 商务印书馆, 1999, 95.
- [19] Carson, C. 'Science as Instrumental Reason: Heidegger, Habermas, Heisenberg' [J]. *Continental Philosophy Review*, 2010, 42: 483–509.
- [20] Mai, J. E. 'Big Data Privacy: The Datafication of Personal Information' [J]. *The Information Society*, 2016, 32(3): 192–199.
- [21] Heinrichs, B. 'Discrimination in the Age of Artificial Intelligence' [J]. *AI & Society*, 2022, 37(1): 143–154.
- [22] Aizenberg, E., Van Den Hoven, J. 'Designing for Human Rights in AI' [J]. *Big Data & Society*, 2020, 7(2): 2053951720949566.
- [23] Gabriel, I. 'Artificial Intelligence, Values, and Alignment' [J]. *Minds and Machines*, 2020, 30(3): 411–437.
- [24] Firt, E. 'Calibrating Machine Behavior: A Challenge for AI Alignment' [J]. *Ethics and Information Technology*, 2023, 25(3): 42.
- [25] 王前. 技术伦理通论 [M]. 北京: 中国人民大学出版社, 2011, 63–68.
- [26] 安乐哲、郝大维、彭国翔. 《中庸》新论: 哲学与宗教性的诠释 [J]. 中国哲学史, 2002, (3): 5–17.
- [27] Wray, K. B. *Knowledge and Inquiry: Reading in Epistemology* [M]. Toronto: Broadview Press, 1985, 186.
- [28] 王前、菲利普·布瑞. 负责任创新的理论与实践 [M]. 北京: 科学出版社, 2019, 14–29.
- [29] 张卫. 内在主义技术伦理学研究 [M]. 北京: 人民出版社, 2023, 113–121.

[责任编辑 孟建伟]