

人工智能的伦理困境及其规避进路

——基于马克思主义科技观分析

杨双彪

(郑州大学 马克思主义学院, 郑州 450000)

摘 要:人工智能作为以算法为核心、模拟人类活动的智能机器,它的迅速发展在推动社会进步、促进人类解放的同时,也引发了一系列的伦理风险挑战,即人的主体性消解、人的意志遭受侵犯和责任归属认定困境。对此,规避人工智能所带来的伦理风险,必须遵循马克思主义科技观引导、提高社会科技伦理水平、构建人工智能监督体系、健全人工智能法律规范,从而确保人工智能为人类服务、最大限度地促进人的自由全面发展。

关键词:人工智能;伦理风险;马克思主义科技观

中图分类号:B82-057

文献标识码:A

文章编号:1009 — 2234(2023)06 — 0052 — 05

近年来,人工智能得到迅速发展,特别是自动驾驶、AI 绘画等产品的广泛应用,无不说明人工智能时代已经到来。但人工智能在解构传统伦理关系的过程中,也带来了新的伦理矛盾,在这样的情况下,如何使人工智能达到技术进步与伦理治理的平衡,成为当今社会所不可回避的重大课题。基于此,本文主要聚焦人工智能当前面临的伦理困境,在描绘人工智能时代镜像的基础上,探讨人工智能伦理困境的多维表现、进而提出解决伦理困境的规避进路。

一、人工智能伦理困境的时代镜像

以算法为核心的人工智能通过对人类思维过程的模拟,在推理、判断等方面表现的越来越类人化,甚至在某些领域已经超越人类。人工智能的发展促进了生产力的提高,给予人类更多自由时间,有效促进了人类的自由全面发展。但在实践活动中越来越相对独立的人工智能,对人类主体性地位也提出了严峻挑战。

(一)智能机器的类人摹效

人工智能是以互联网信息技术为基础,以基于具有强大数据采集分析能力的算法为核心,以对人类思维过程的模仿和超越为目标的现代科学技术。

通过对其智能化水平的划分,主要分为“弱人工智能”和“强人工智能”。目前我们普遍使用的人工智能都是“弱人工智能”,即仅让机器拥有某种智能行为。而“强人工智能”,则是指能够通过深度学习,产生脱离程序具有独立意志的智能机器,此时的人工智能已经超出机器的界限,其行为不再是人类意志的延伸,而是能够自主做出决策并实施的“类人”。需要指出的是,无论是“弱人工智能”还是“强人工智能”,都能够在一定意义上发起具体的社会实践活动^[1],并在实践过程中根据需要做出调整。从这个意义上说,人工智能与人类社会中传统的机器已经截然不同。

特别是随着大数据算力的提升和深度学习算法的改进,人工智能展现出了与人类高度相近的认知能力、逻辑能力甚至情感能力。具体来说,人工智能不仅可以分析数据、协助人类进行决策,还能够进行艺术创作这种人类独有的实践活动,事实上,人工智能显然对传统的机器概念进行了消解。马克思认为,生产形式在机器出现后的最大改变就是人类不再参与特定的生产过程,而只是“站在机器旁边”监督机器的运行,把人类对商品的生产过程转化为对机器的操作过程。而人工智能通过复杂的数据网络模拟人

收稿日期:2023 — 06 — 08

作者简介:杨双彪(1997—),男,河南民权人,硕士研究生,研究方向:马克思主义发展史。

的行为模式和思维方式,使机器也拥有“类人”的智能,在生产过程中不仅增加了劳动者的剩余劳动时间,还在一定程度上取代了人类劳动,将人从重复、繁重的体力劳动和脑力劳动中解放出来,促进人的自由全面发展。但需要指出的是,新技术的发展大多具有复杂性和不确定性,人工智能在推动社会进步的同时,也不可避免地引发了各种不良后果,如何确保人工智能向善发展,成为当前摆在人类面前的重要课题。

(二)科技发展视域下人的解放

“人”是马克思一切思想批判的核心问题,实现人的自由全面发展也是马克思主义的根本追求。科学技术作为一种革命性的力量,每一次的技术进步都会促进人的解放和社会的发展,而人工智能在生产生活中的广泛应用,使人的自由全面发展有了现实可能性。首先,人工智能使人的劳动由异化转向自由自觉。劳动是人的存在方式,也是人的基本权利,但在很多时候劳动也是一种沉重的负担,甚至劳动者的艰辛劳动不仅得不到财富,反而会加深劳动者自身的贫穷,这就是劳动的异化。马克思认为“‘劳动’,按其本质来说,是非自由的、非人的、非社会的。”^[2]因此“只要肉体的强制或其他强制已停止,人们就会像逃避瘟疫一样逃避劳动。”^[3]但是随着人工智能代替人类从事体力和脑力劳动,人类逐渐摆脱劳动的束缚,有更多的时间自由自觉地从事更能体现人类价值的其它活动。

其次,人工智能推动社会生产力的提高。正如马克思所言,“手推磨产生的是封建主的社会,蒸汽磨产生的是工业资本家的社会”^[4]¹⁴²。人工智能自诞生以来就通过变革生产方式、降低生产成本和提高劳动生产率等方式,促使生产力落后的传统企业改造升级,促进社会生产力的普遍提高。人工智能在发展生产力的过程中与生产资料相结合,提高社会生产效率从而显著增加了劳动者的自由支配时间。人工智能与劳动对象相结合,变革了以往严格繁重的生产模式,可以面对生产过程中各类变化随时调整,有效提高劳动生产率水平特别是劳动者的安全水平。最后,人工智能创造出更多的自由时间。劳动实践活动在人工智能时代变得更加自动化和智能化,以往重复且繁重的工作被大量的智能机器所取代,使社会能够在更短的时间内生产出更多的物质财富,有效缩短了人们生产必要物质资料的时间,在此基础上,人类所能够支配的自由时间大大增加,从而使人们拥有更多的时间和机会去实现人的全面发展。^[5]

(三)实践向度的主体性思考

人工智能以互联网大数据为载体并通过复杂

算法的运算,获得了人所特有的一些基本特征,在推理、思考等方面表现得越来越像人,传统的人机界限不断被模糊化,特别是人形机器人的不断突破,对“人”的概念提出了严峻挑战。在这个意义上,主体还是不是人类的专属,或者说人工智能是否可以在哲学层面被认为与人类等同的“主体”,成为当前对人类主体性哲学思考的当务之急。

马克思认为“同活动对象的客体相反,主体是活动的承担者和执行者”。^[12]在马克思看来,现实的实践活动是确立主体性的根本原则,^[1]人类正是通过实践活动奠定了全部社会生活和人类历史的基础。在实践中,人将自己的活动与意识相分离,使意识不仅成为自身的一部分,还可以审视自身的活动。也就是说,意识作为人主体性的重要特征,不仅能够直观反映人的实践活动,还具有相对独立性,可以审视甚至规划人的实践活动。从而引出马克思对主体的另一规定性,即主体能够进行主观能动的创造性活动,主体所从事的实践活动并不只是依靠本能,而是拥有明确的目的性和计划性,并且主体能够根据实践的需要制作和使用工具,这也是人之为人的本质特征。

随着人工智能技术的突飞猛进和广泛应用,越来越多的智能系统从事着各种劳动任务,我们甚至可以说,这些人工智能系统就是为了代替人类劳动而产生的。特别是随着近些年计算机算法和深度学习技术的突破,人工智能不仅能够根据人类编制的算法进行生产劳动,还可以根据需要“制造和使用生产工具”,尽管当前的人工智能制造和使用生产工具还比较初级,主要还是通过受人驱使的方式进行数据的采集和程序的编写。在这样的情况下,劳动实践还是不是人的专属活动,就值得我们进一步反思和追问了。但是明确的是,人的唯一主体地位正被人工智能所挑战,为此必须重新思考人与人工智能之间的人机关系,确保人的本质的复归。

二、人工智能伦理困境的多维透视

马克思曾指出:“在我们这个时代,每一种事物好像都包含有自己的反面,……技术的胜利,似乎是以道德的败坏为代价换来的。”^[6]人工智能的发展不仅是科技产业的重大革命,还引起了社会伦理的深刻变革,带来了一系列人工智能伦理困境,如人的主体性消解、决策自主权的丧失、伦理责任归属等各种问题。

(一)人的主体性消解困境

人的主体性主要体现在两个方面,一方面是在自然世界中的主体性,人类能动地改造自然的实践过程是其主体性的关键因素。另一方面就是社会生活中的主体性,人类社会是人们在长期的生活实践

中演变来的,人无疑是社会生活的主体。然而人的这种主体地位受到了人工智能的严峻挑战,特别是随着人工智能逐渐成为人类社会不可或缺的一部分,人类主体性也遭到了不断地异化。马克思在《1844年经济学哲学手稿》中完整提出了异化命题,即异化主要为原本由人所创造的、对象化的客体,本应该服务作为主体的人,但这种客体反而控制了作为主体的人,这种主客体关系的颠倒就是异化。^[1]恩格斯就曾以英国纺织工人为例,仔细描绘了机器如何将原本乡村中精神健全的农民,转化为纺织工厂里愚钝和痴呆的工人。遵循同样的逻辑,原本为人类所创造用以服务自身的人工智能,现在却反过来成为控制人类的手段和工具。

首先,人工智能会导致人类主体能动性的削弱。人类主体能动性就是人类区别于其它物种的独特属性和专有能,如认知能力、审美能力。但原本独属人类的领域不断被人工智能所侵占,人工智能可以写诗、绘画、甚至给予他人情感慰藉,导致人类主体性的能力被极大地消解。其次,人工智能导致人类社会交往的虚拟化。人工智能与互联网的深度融合,使得人与人之间现实的社会交往逐渐被基于网络的虚拟社交所取代,但这种虚拟的社会交往实质就是人类交往的异化,同时也是对人社会属性的背离。最后,人工智能使人类异化为“智能机器的附庸”。马尔库塞在《单向度的人》中指出,现代社会的畸形发展,使科技人员创造的工具和装置不仅没有使人们更好地解放和自由,反而成为奴役人们的工具和手段。^[7]在高度智能的机器面前,普通人只会丧失作为主体的个性化特征,逐渐地沦为庞大且复杂的智能机器中不起眼的零部件。

(二) 决策自主权丧失困境

人工智能技术的发展也会引发决策自主权的丧失。媒体为了增加浏览量,往往利用大数据算法来分析用户的兴趣偏好,并根据所刻画的用户画像进行精准化的信息推送。个性化的算法推荐导致用户的信息域被智能遮蔽和过滤,使人的思维只能在算法规制的“信息茧房”中进行活动,这不可避免地影响个体的价值取向,严重削弱了人类对信息的甄别能力。在此意义上,人工智能算法虽然提高了民众获取信息的准确性,但是却剥夺了民众选择数据的自主权。最典型的例子就是2016年美国大选的“剑桥分析事件”,有证据表明在美国总统大选期间,特朗普团队利用人工智能针对性地向选民推送片面甚至虚假信息。人工智能对人的决策施加影响不仅存在于政治领域,在日常的网络浏览中亦会如此,算法会根据人们平常浏览或搜索的数据推荐相应的信息。这些固然会高效快捷地将感兴趣的信息

展现在人们面前,但也让人们丧失对其它信息的敏感度和感知力,甚至影响其对现实世界的认知。

另一方面,人工智能所具有的信息搜集和数据分析能力是人类难以匹敌的,基于此,人们往往利用人工智能对数据进行处理分析,并在此基础上做出“最优解”的决策,如医生通过医疗系统给病人的治疗方案救助病人,金融公司通过智能系统确定投资方向和比重等等。这种对人工智能产生的过度依赖,不仅严重削弱了人类对问题的主动探索能力,还消解了人类在解决问题中对自身能力的培养。人的自由全面发展受到严重阻碍,如医生在对医疗系统形成依赖后,就会逐步丧失某些依靠长期实践经验积累的医疗技能。在这样的背景下,人类对人工智能的过度依赖的后果,只会“使物质力量具有理智生命,而人的生命则化为愚钝的物质力量”。^[6]这无非是说,随着人工智能越来越多地介入人类的选择决策过程,人类的行为选择也逐渐失去自我思考,在此意义上,人工智能将可能成为“永远的独裁者”。

(三) 伦理责任归属困境

通常情况下,一项新技术的发展往往具有复杂性和不确定性,在应用过程中会出现难以避免的消极影响,正因如此,正确处理因技术发展引发的责任问题就显得尤为重要。而人工智能的发展使传统的科技伦理规范失效,正如威尔伯斯所说:“在日益科技化的现代社会中,人工智能技术的广泛应用使得道德责任的归属复杂化。”^[8]无人驾驶是当前人工智能最典型的应用领域之一。无人驾驶相比人类更加专注且不易疲劳,大幅减少了因司机驾驶过错导致的交通事故,特别是对于没有驾驶能力的特殊人群来说,无人驾驶无异于是对生活轨迹的重建。但需要指出的是,无人驾驶背后隐藏着一个巨大的责任伦理困境,即颠覆了传统驾驶那种以司机过错为基础的责任规范体系,因为无人驾驶导致的交通事故中,既不存在司机主观上的故意,也不存在司机因疲劳驾驶等造成的无意过错,道德和法律的责任分配此时变得更加复杂和难以确定。

事实上,相关法律的滞后也是人工智能责任判定困难的重要原因。马克思曾指出,责任是建立在自由意志基础之上的,只有人才是承担责任的主体,因此在传统的法律判定主要基于人的行为进行责任划分,如果发生机器伤害人的事件,则主要由机器的所有者或制造商承担相应责任。但随着人工智能技术的发展,人工智能也在一定程度上拥有了自己的判断决策能力,人类很难决定智能机器的具体行为。在这样的背景下,出现人工智能导致的故事该如何分配责任,迄今为止都没有一个统一的法律规定。通常我们对责任归属的判定,主要依据法律

条文中的相关规定。但人工智能的发展速度远远超过相关法律条文的制定速度,致使当前人工智能领域面临责任分配困难的法律难题,只能将人工智能失控归咎于产品缺陷,让人工智能及其产品的设计者承担相应责任。而长期的责任分配失衡,势必会挫伤人工智能研发人员的积极性,最终限制人工智能的发展。

三、人工智能伦理困境的规避进路

人工智能所引发的伦理困境是技术发展中的必然产物,对此既不能因噎废食放弃人工智能发展,也不能对所产生的伦理问题视若无睹。而是应该探究当前所产生伦理困境的真实动因,在马克思主义科技观的指导下,构建多方协同治理体系,推动人工智能朝着为人类服务的方向稳步前进。

(一)遵循马克思主义科技观

马克思主义科技观是马克思主义关于科学技术的系统性的观点和方法,我国发展人工智能,必须坚持马克思主义科技观的基本立场和基本原则。马克思主义认为,“科学技术是人的创造物,是人类认识世界、改造世界的手段和工具,”^[9]科学技术还是推动社会变革和时代进步的关键因素。但是需要指出的是,科学技术具有二重性,一味强调科技对社会发展的推动作用并不客观,科技在给人类带来巨大收益的同时也会带来难以预料的风险。正是基于这样的认识,发展人工智能必须要始终坚持人类主体性地位,任何科学技术,都只是基于人类需求创造出来的客体,如果将主客体位置颠倒,不仅会阻碍人类对科技的正确认识和运用,还会误导人类对科技负面影响根源的把握。

马克思主义强调认识来源于实践,但人工智能所拥有的知识并不是实践的产物,它的“智能”来源于人类自身。人类通过发挥自身的能动性,将人工智能的能力转化为人类自己的能力,从而促进人类自身实践能力的提升。科技的发展就是为了给人类争取更多的自由时间,最终实现人类的解放,是以对于科技的研发永远不会停滞下来,但对于发展科技引发的一系列伦理问题,必须始终坚持马克思主义科技观的指导,坚持以人为本的理念。人是审视人工智能发展的价值尺度,实现人的解放也是人工智能发展的根本追求,为此人工智能技术的研发要始终强调人在社会发展中的主体地位。

(二)提高社会科技伦理水平

科技的发展固然能够方便人类的生活,但也有可能给人类带来灾难性的后果。爱因斯坦就曾认为“怎样用它,究竟是给人类带来幸福还是灾难,全取决于人自己,而不是取决于工具”。^[10]将人工智能发展所带来的伦理问题一味归咎于科技本身,是有失

偏颇且不客观的,事实上,当今人工智能各种伦理问题的产生,很大程度就是因为科研工作者道德责任感缺失。事实确实如此,如果科研工作者的道德理念淡漠,那么此前发生的基因编辑这种伦理悲剧很容易在人工智能领域重演。通过对科研工作者进行正确的道德约束,可以有效避免技术异化对人类产生的不利影响,确保人工智能研究始终朝着为“人”服务的方向前进。

提高公众的知识素养也是解决人工智能伦理问题的重要途径。库兹韦尔将人们对待新技术的态度分为三个阶段,首先,人们会惊叹新技术为人类所带来的便利;其次,人们了解到隐藏在新技术背后的弊端,并对此产生恐惧;最后,研究出一条科学的发展道路,实现科技为人类服务的初衷。近年来人工智能威胁论甚嚣尘上,特别是部分媒体夸大其词地报道,更是使民众对其产生惧怕心理。基于此,提高民众相关的知识素养显然已成一项必要工作,这需要政府运用多元化手段,培养民众对人工智能技术的科学认知。让民众了解到人工智能并不会威胁人类安全,反而会对人类的发展做出重大贡献,从而有效缓解人类所产生的恐惧心理,推动人工智能得到健康发展。

(三)构建人工智能监督体系

当前人工智能发展过程中所产生的一些伦理问题,其主要原因就是缺乏有效监管。基于此,根据当前人工智能发展实际,建立起一整套切实可行的监管体系,成为当今时代的首要任务,而制定一个多方位、全流程的监督系统,首先需要确保政府在人工智能监管体系的主导地位。政府需要设立各级各层的人工智能监管机构,加大对人工智能技术的监管,从算法逻辑、专业机构的检测报告等方面保障人工智能技术的安全可控性,确保即使有意外发生也有切实可行的应急举措。还要定期对人工智能相关企业进行伦理审查,防止人工智能技术被不法分子因追求私利危害人类,对于有可能危害人类尊严权利的相关问题,一旦发现要立即整改并严惩不法分子。

此外,通过民众监督更有利于有效发现人工智能的风险。民众是人工智能最广泛的使用者,在这种情况下,赋予民众对人工智能监督的权力,更有利于让民众在使用人工智能过程中,及时发现风险并将风险扼杀在初始阶段,确保人工智能及其产品符合伦理法规要求。最后,来自人工智能研发者的监督也十分必要。科技本身无法与其研发者分裂开来,科研工作者在开发人工智能及其产品时,有责任也有义务在设计研发阶段,确保“人工智能算法的可解释性、可验证性和可预测性”。^[11]不仅如此,科

研人员还要在人工智能产品使用阶段全程监管,一旦发现问题及时干预,确保研究始终坚持为人类服务。

(四)健全人工智能法律规范

人工智能发展过程中出现的各种伦理困境,仅靠道德感化是远远不够的,因为伦理道德只是借助说服教育等软性手段,并没有相应的强制力,在这样的情况下,法律无疑是化解当前人工智能伦理问题的有效措施。法律通过国家强制性手段,最大可能地限制人工智能发展过程中出现的无序性,有效降低因人工智能伦理问题给社会带来的危害。基于这样的认识,首先要着眼于人工智能的法律主体关系,充分论证法律上的“主体”概念,并在确定人工智能法律主体地位的基础上,厘清人工智能在无人驾驶、医疗机器等应用过程中的权利、义务和责任归属,对于有可能出现的伦理事故,划分好开发者、使用者及智能机器的法律责任。

此外,还要加强法律对数据信息和算法的治理。严格保护用户的个人信息,防止不法分子因利益对个人信息的收集、传播与滥用,同时还要建立健全关于人工智能的法律问责机制,对于在应用过程中出现的侵害个人隐私、制造算法偏见的法律主体,必须根据实际情况和法律标准进行严厉打击。最后,要规范人工智能的生产使用原则,采用法律手段对智能机器的生产制造加以限制。在研发生产阶段,将设计规范写入法律条款之中,明确开发人员的法律责任,在智能产品的使用过程中也要进行实时监控,避免使用者超越法律的界限。总之,通过对人工智能进行全方位的法律规范,能够减少科技滥用对人类造成的伤害,最大限度地解决当前伦理困境。

[参 考 文 献]

- [1] 姜延强.人工智能的伦理困境与正解[J].道德与文明,2022(01):131-139.
- [2] 中共中央编译局.马克思恩格斯全集:第42卷[M].北京:人民出版社,1979:254-255.
- [3] 马克思.1844年经济学哲学手稿[M].北京:人民出版社,2000:55.
- [4] 中共中央编译局.马克思恩格斯选集:第1卷[M].北京:人民出版社,1995.
- [5] 孙闪闪.人工智能时代人的发展的新机遇与新挑战[J].沈阳工业大学学报(社会科学版),2018,11(06):571-576.
- [6] 中共中央编译局.马克思恩格斯全集:第2卷[M].北京:人民出版社,1985:79.
- [7] [美]马尔库塞.单向度的人:发达工业社会意识形态研究[M].刘继,译.上海:上海译文出版社,1989:112.
- [8] Katinka Waelbers. Technological Delegation: Responsibility for the Unintended, Science and Engineering Ethics, 2009, 15(01).
- [9] 李桂花.科技的人化[M].长春:吉林人民出版社,2004.
- [10] [德]爱因斯坦.爱因斯坦文集(第3卷)[M].许良英,等,译.北京:商务印书馆,2010:78.
- [11] 郭锐.人工智能的伦理和治理[M].北京:法律出版社,2020.
- [12] [德]阿·科辛.马克思列宁主义哲学词典[M].郭官义,等,译.北京:东方出版社,1911:483.

[责任编辑:侯庆海]

The ethical dilemma of artificial intelligence and the way to avoid it ——Analysis based on the Marxist view of science and technology

Yang Shuangbiao

(Zhengzhou University, Zhengzhou, Henan, 450000)

Abstract: As an intelligent machine with algorithms as the core and simulating human activities, the rapid development of artificial intelligence not only promotes social progress and promotes human liberation, but also causes a series of ethical risk dilemmas, that is, the dissolution of human subjectivity and the will of human beings. Dilemma of Violation and Attribution of Responsibility. In this regard, to avoid the ethical risks brought by artificial intelligence, we must follow the guidance of the Marxist concept of science and technology, improve the level of social science and technology ethics, build an artificial intelligence supervision system, and improve artificial intelligence legal norms, so as to ensure that artificial intelligence serves human beings and maximizes. Promote the free and comprehensive development of human beings.

Key Words: artificial intelligence; ethical risk; Marxist view of science and technology