

# 基于马克思的“机器体系”叙事逻辑建构 智能时代新型人机关系

王亚奇

(山东第二医科大学 马克思主义学院, 山东 潍坊 261053)

**[摘要]** 如何应对人工智能的快速发展给人们带来的焦虑,应回归马克思的“机器体系”叙事逻辑。马克思指出,机器体系是社会“一般智力”的物化,机器体系的应用既促进了生产力的发展,也成为资本剥削压榨工人的工具。人工智能本质上是智能技术人工物,但它的智能特性给当代社会发展带来了诸多不确定性和风险,需要在防范风险的基础上,建立人机一体化、人机共能、人机共情的新型人机关系,从而使人工智能成为推动社会进步、助力人的解放的力量。

**[关键词]** 人工智能;机器体系;人机关系;叙事逻辑

**[中图分类号]** B0-0 **[文献标识码]** A **[文章编号]** 1008-7699(2026)02-0010-10

当前,人工智能已成为各国追逐的技术热点,并深刻影响了人们的生活。人工智能的积极作用自不待言,而如何消解人工智能的负面影响,是人类主体在人工智能时代必须直面的重大理论和现实问题。人工智能能否超越并反制人类智能?人工智能的发展以何种方式、在什么程度上威胁到人类的生存?这实际上是一个从人类生存论上追问人与人工智能二者关系的问题,是人凭借自身理性的超前反思。

## 一、关于人工智能挑战与危机的叙事

人工智能的快速发展,的确给人类未来带来了诸多不确定性,由此也引发了广泛争论。由于人工智能的特殊性,人们对人工智能未来发展的后果存在分歧。有观点认为,人工智能将解决人类生存的所有问题,人与人工智能的融合而形成新的高级智能,人将拥有永生的能力。<sup>①</sup> 另一类观点则认为,人工智能作为具有颠覆性的新技术,因其快速发展带来的高风险和不可控性,有可能“取代人类”甚至“消灭人类”,从而带来“人类时代的终结”。斯蒂芬·霍金、斯图尔特·罗素和埃隆·马斯克均持此观点。然而,“那些说自己最担心人工智能的人,恰恰也是最坚决要创造人工智能并享受其所带来财富的人”,<sup>②</sup>正如马斯克既担心人工智能带来危害,又支持人工智能的发展。2024 年诺贝尔物理学奖得主杰弗里·辛顿在接受《纽约时报》采访时表示,“相较于气候变化,AI 可能对人类的威胁‘更紧迫’”。<sup>③</sup> 国内许多学者也持此类观点,如赵汀阳就认为:“人工智能的发展正是对人类智慧的一个终极测试。我有个悲观主义的预感:在人工智能成为统治者之前,人类就可能死于人工

**[收稿日期]** 2024-03-03

**[作者简介]** 王亚奇(1991—),女,山东潍坊人,山东第二医科大学马克思主义学院副教授,哲学博士。

① [美]雷·库兹韦尔:《奇点临近》,李庆诚、董振华、田源译,机械工业出版社 2011 年版,第 225-226 页。

② 西岸:《人工智能竞赛是如何开始的?》,《世界科学》2024 年第 6 期,第 57-58 页。

③ 沈淑莎:《新技术总会伴随新风险,AI 为何更棘手?》,《文汇报》2023 年 5 月 12 日,第 7 版。

智能创造的一切好事。”<sup>①</sup>孙伟平则指出：“智能科技构成了对人公开的或隐蔽的宰制，人正在沦为高速运转的智能社会系统的‘附庸’和‘奴隶’”。<sup>②</sup>还有学者把人工智能取代人类的理论归结为“能力超越式取代论”“技术失业式取代论”“物种灭绝式取代论”，并对之做出具体分析，特别是对超级人工智能完全取代人类从而导致人类灭亡的恐慌做了理论分析。<sup>③</sup>这些悲观的主张加剧了人们的焦虑。

新一代人工智能依据大模型、大数据、大算力，通过深度学习和交互学习，使其进化速度越来越快，变得越来越聪明，不断替代人类的脑力劳动。因此，许多国家已把发展人工智能提升为国家战略。但是，随着数据快速丰富、算法加速迭代、模型迅猛膨胀，人工智能快速进化，它将如何影响人类？因为与以往的技术不同，人工智能具有技术属性、社会属性和人的属性高度融合的特征，特别是在模拟人的思维属性、情感属性、意志属性等方面不断突破。以 ChatGPT 为代表的生成式人工智能利用人工神经网络、数据分析和模型训练，通过搜集创作需要的主题、情感、语言、风格等要素，对人类数千年来的文化积淀加以提炼，可以像人类一样在文学、艺术等领域完成专属的文化产品。与此同时，人工智能的语言系统日益智能化，“心智”能力不断跃升，行为能力更趋完善，并具备了一定的情感能力，与人的能力逐渐融合，成为继承和创造人类文化的“类主体”。更可怕的是，它可以在与人类交流的过程中持续学习，不断丰富自己，在开放式大模型之间可以交互学习，学习速度之快令人望尘莫及。人工智能在知识领域呈现出的这种近乎“无所不能”的状态，使人产生膜拜感。同时，人工智能的拟人化、贴心化、全天候服务，使人陷入“技术沉溺”而无法自拔。

试想，人工智能在不断的交互学习中产生自主性新知识，在执行人的指令时又不断获得自主反思的自我意识，这恐怕就会僭越人的主体意识，从而超越并控制人类。这是人工智能对人类最大的威胁。面对人工智能快速迭代所带来的巨大不确定性，杰弗里·辛顿向人类发出警告，认为随着人工智能深度学习能力的增强及人工神经网络的愈加复杂化，人工智能有可能产生自主意识，自主修改人类设定的程序，从而脱离人类控制，并最终控制人类，导致人类灭绝。那么，人工智能真的如此可怕吗？该如何把握人工智能的本质？

当前，学术界对人工智能本质的哲学把握主要有两条路径：一条是通过研究人类思维和认知的本质和规律，从而为人工智能的进一步发展提供理论支持；另一条是通过研究人工智能对人的生存带来的挑战，把握人工智能的发展对人类社会和人的生存带来的正向或负向效应。两条路径皆面临着人工智能未来发展会带来不确定性的挑战，这实际上是由技术与社会的互动决定的。人对技术的追求源于自我解放的需要，人不仅要解放自己的手脚，而且要解放自己的大脑，不仅要获得行动的自由，而且要获得心灵的自由。因此，人类在发展中不断突破想象的边界，借助技术实现想象的结果，这是主体能力对象化的显著表现。人最终不仅要制造出和人一样的机器，而且试图制造出所谓的“万能机器”。人工智能正朝这个方向发展，这为人类带来巨大惊喜的同时也带来了同等巨大的恐慌。因为，人的生活越来越程序化，确定性越来越大，透明度越来越高，人的未来生活似乎已经被计算出来了。但生命的意义却在逐渐弱化，人的生存面临巨大挑战。

面对潜存危机带来的恐慌，人类正试图从法律、道德等各个方面加以防范。例如，有研究者提

① 赵汀阳：《人工智能的神话或悲歌》，商务印书馆 2022 年版，第 155 页。

② 孙伟平：《人工智能与人的“新异化”》，《中国社会科学》2020 年第 12 期，第 119 页。

③ 陈高华、赵文钰：《“人工智能取代人类”的三类叙事及其批判性反思》，《江汉论坛》2024 年第 10 期，第 57-63 页。

出人与人工智能要实现“价值对齐”,对人工智能进行伦理约束,“将多样的人类价值观纳入考量,并确保不对机器做出违反伦理道德的行为,以确保高度自主的人工智能系统分享并接受人类好的价值观,从而做出符合人类道德和伦理标准的行为。”<sup>①</sup>有研究者则从法律角度提出人工智能应具有法律主体地位,只有明确其法律主体资格,才能解决其义务和责任等问题,从而为人工智能的发展提供法律保障。<sup>②</sup>还有研究者从哲学主体性的视角提出,人与人工智能是一种交互主体性关系,“这种主体性是参与交互实体的一种潜在行动能力,包含了交互实体的目标导向性和自主性”,<sup>③</sup>试图以赋予人工智能的主体地位,为调和人与技术之间的价值冲突提供一种关系性的哲学框架。

尽管,学术界已从哲学、伦理、法律等多重视角给出了解决人工智能潜在风险的答案,但都不足以消解公众的危机感。而且,随着人工智能的快速发展,其将全面改变人类生存方式、生产方式、生活方式的趋势亦日益凸显。在此背景下,有必要立足于历史唯物主义关于社会实践和科技发展的基本原理,揭示科技发展与社会制度的内在关联,从而科学把握人工智能的本质和发展趋势。这需要回到马克思的理论资源中寻求启迪。

## 二、回归马克思“机器体系”的叙事逻辑

每一次重大技术革新都会带来诸多不确定性,这种不确定性常常引发公众的关切和怀疑。然而,尽管对技术作用的质疑始终存在,科技仍具有推动社会进步、促进人的解放的一面。

早在19世纪,马克思在论述机器的社会作用时,就首先肯定了机器推动社会进步、促进人的解放的作用。但与此同时,马克思也指出了资本利用机器榨取工人的剩余价值。因此,在马克思看来,剥削、压榨工人的不是机器,而是资本主义制度。从一般意义上来说,机器是劳动的产物,是人的主体力量的对象化,是解放人的工具,是物化的知识力量。马克思指出,自然界并没有造出任何机器,机器是人的意志驾驭自然、进行产业劳动的产物,是人的知识力量的对象化,“固定资本的发展表明,一般社会知识,已经在多么大的程度上变成了直接的生产力,从而社会生活过程的条件本身在多么大的程度上受到一般智力的控制并按照这种智力得到改造。”<sup>④</sup>在这里,马克思提出了社会“一般智力”这一重要概念。一般智力也就是一般社会知识的积累,代表着社会一般认知水平。一般智力的形成,存在一个由个体知识向一般社会知识的转化过程,从而形成社会一般智力。同时,社会一般智力又反过来影响社会大众。在农业和手工业时代,劳动过程以劳动者为主,由于社会一般智力水平较低,因此体力劳动者只能达到最低的一般智力并应用于生产。在机器大工业时代,科技发展提升了社会一般智力,生产过程变得复杂化,然而,科技被资本垄断,并不能普及到一般劳动者。正如马克思强调的:“资本不创造科学,但是它为了生产过程的需要,利用科学,占有科学。”<sup>⑤</sup>科学作为社会一般智力提升了人们认知世界、改造世界的能力,但由于受资本控制,成为了服务于资本的工具。由此,科学便获得了成为生产财富和致富手段的使命。具体表现为,作为一般智力的科学,在资本的操纵下,成为资本家追逐剩余价值的工具,而科学在生产中的应用主要体现在机器上,一般智力被机器吸纳,形成机器生产体系,并取代工人的生产地位。这就形成了科学与劳动的对立。对此,马克思指出:“科学对于劳动来说,表现为异己的、敌对的和统治的权力”,<sup>⑥</sup>作

① 包晨婷、温波、丘亮辉:《人工智能应用的伦理隐忧及治理进路》,《自然辩证法研究》2024年第9期,第75页。

② 范进学:《人工智能法律主体论:现在与未来》,《政法论丛》2022年第3期。

③ 殷杰:《生成式人工智能的主体性问题》,《中国社会科学》2024年第8期,第125页。

④ 《马克思恩格斯全集》(第31卷),人民出版社1998年版,第102页。

⑤ 《马克思恩格斯文集》(第8卷),人民出版社2009年版,第357页。

⑥ 《马克思恩格斯文集》(第8卷),人民出版社2009年版,第358页。

为一般智力的科学成为与工人对立的力量。维尔诺把马克思的一般智力归结为人的思维能力,认为一般智力是一种“认知范式、对话演绎、语言竞争”,进而归结为“大众智能”<sup>①</sup>,这一主张实际上弱化了一般智力与资本的关系。

在马克思看来,机器是社会“一般智力”的凝结,是人在劳动中创造、使用工具的结果。机器这种在劳动需求中创造出来的复杂工具,由“一般社会知识”在生产中转化为“直接的生产力”。作为社会实践的“直接器官”,资本在组织生产时,运用众多功能各异、互通协作的机器,依据生产流程对原材料进行加工并形成制成品,达到完成生产任务的目的,这个使用不同机器进行自动化生产的整体就是“机器体系”。马克思指出,在资本的运作程序里,机器是劳动资料的最后形态,更确切地说,是“自动的机器体系(即机器体系,自动的机器体系不过是最完善、最适当的机器体系形式,只有它才使机器成为体系)”<sup>②</sup>

马克思从资本的属性上揭示机器体系的本质,认为机器体系不仅作为劳动过程中的生产资料,还成为攫取剩余价值的固定资本。将机器体系放在资本主义生产关系中加以审视,便暴露了资本通过机器体系的使用来剥削、控制工人的事实。在机器的运转下,工人应有的灵魂性技艺被机器魔力般的技能和力量取代了,工人沦为机器作用于原材料环节的“机器”,因为资本操纵下的机器“本身就是能工巧匠,它通过在自身中发生作用的力学规律而具有自己的灵魂”<sup>③</sup>,这种灵魂是资本赋予的。于是,机器体系作为固定资本在劳动中就同工人对立起来了,从而确立了能使资本得到充分发展的资本主义大机器生产方式。这一情形下,生产过程将不再从属于工人,而呈现为科学在工艺上的运用。这时,“资本才造成了与自己相适合的生产方式。”<sup>④</sup>整个过程中,资本通过机器体系确立了自我的生产方式,机器通过控制工人获得了自我的主体地位,工人则由机器的“主人”变成了机器体系的“助手”。马克思指出,在机器体系中,工人站在了生产过程的旁边,而不再作为这一过程的主要作用者。<sup>⑤</sup>可见,因为从属于机器体系,导致工人自我主体地位的丧失,机器体系作为资本的支配力量却拥有了统治权。

早在《1844年经济学哲学手稿》中,马克思就揭露了机器体系对人的排斥,“劳动用机器代替了手工劳动,但是使一部分工人回到野蛮的劳动,并使另一部分工人变成机器。劳动生产了智慧,但是给工人生产了愚钝和痴呆。”<sup>⑥</sup>表面上看,机器的应用替代了工人的劳动,节省了劳动时间,提高了劳动效率,似乎把工人从繁重的劳动中解放了出来,但实际上,却是机器的资本主义应用排斥了工人,陷工人于更加贫困的处境。工人劳动成为“有意识的机件”和“机器体系的环节”,机器体系则“在工人面前表现为一个强大的机体”<sup>⑦</sup>。这就从工人被机器体系排斥的视角揭示了现代资本主义的社会关系:在机器体系的一端,是依附于机器体系的工人,另一端则是支配机器体系的资本。

显然,机器体系不是一般物质化的生产要素,也不是客观中性的生产工具,而是社会生产关系的承载物,受资本逻辑的支配,成为获得剩余价值的资本形态。质言之,机器体系不是普通的物,而是承载着资本主义社会结构的固定资本。机器体系的本质是资本,是社会一般智力物化的死劳动,

① [意]保罗·维尔诺:《诸众的语法:当代生活方式的分析》,董必成译,商务印书馆2017年版,第82页。

② 《马克思恩格斯全集》(第31卷),人民出版社1998年版,第90页。

③ 《马克思恩格斯全集》(第31卷),人民出版社1998年版,第91页。

④ 《马克思恩格斯全集》(第31卷),人民出版社1998年版,第94页。

⑤ 《马克思恩格斯全集》(第31卷),人民出版社1998年版,第100页。

⑥ 《马克思恩格斯文集》(第1卷),人民出版社2009年版,第159页。

⑦ 《马克思恩格斯全集》(第31卷),人民出版社1998年版,第91页。



是死劳动支配了活劳动。对此,马克思指出,不能因为机器体系是资本的最适当的形式,就肯定资本主义生产关系是最适合的社会生产关系,<sup>①</sup>因为事实恰恰相反。

马克思在揭示了机器体系的资本主义本质后,亦批判了资本借助机器体系对工人的盘剥和否定。机器体系使工人劳动的独立性日益丧失,工人不得不依附于机器体系并受资本的剥削;而机器体系在提高劳动生产率的同时,也同步加剧了对工人个体的压榨程度。“机器具有减少人类劳动和使劳动更有成效的神奇力量,然而却引起了饥饿和过度的疲劳”。<sup>②</sup>这导致工人在生产中越来越被动,越来越丧失主体性而从属于机器体系。马克思在剖析机器劳动的特点时就指出,工人在操纵机器时,动作的被动性和从属性是机器劳动的特点所在,这是专业化本身的消灭。<sup>③</sup>最终,机器体系不仅导致人的主体地位与个性发展日趋衰落,更威胁到人作为“有血有肉”的生存本身。这被形容为“铁人反对有血有肉的人”<sup>④</sup>。这表明,机器体系已不单纯是生产工具,而是资本权力的有形化身,成为人的对立物。对此,涂良川认为,机器体系的叙事逻辑表明,现代资本逻辑主导下的机器体系,表现为“对象化一般智力的生产工具”,不仅拓展了人体机能、提升了人类潜能、拓宽了实践领域,还使得机器技术与社会权力的同构化进程不断深化,具体表征为资本权力呈现机器化转向、政治统治实现技术化嵌入、机器能力趋于绝对化垄断、技术占有呈现政治化特征等方面。<sup>⑤</sup>可见,机器的资本化、技术的政治化导致技术逻辑成为资本逻辑的外在表现,机器体系叙事逻辑掩盖了资本逻辑。

马克思关于“机器体系”的叙事逻辑科学揭示了工人、资本和“机器体系”之间的内在联系。如果没有资本介入,工人和“机器体系”的逻辑关系十分清晰,即工人是主体,“机器体系”是生产工具系统,工人支配“机器体系”。但在资本主义社会,资本成为社会运行的主导逻辑,资本支配着工人和“机器体系”,其逻辑关系表现为“机器体系”支配工人,实质上是资本通过“机器体系”支配工人。所以,资本逻辑是一种颠倒的逻辑关系。马克思便揭露了资本主义生产关系中这种被刻意颠倒的逻辑叙事,揭示了资本主义制度剥削、压迫的本质。马克思所处的时代虽然尚未出现人工智能,但马克思在感性对象性活动与资本生产关系这种哲学与政治经济学的双重视野中对机器体系叙事逻辑的分析,仍然是今天透视人工智能和人机关系的关键理论资源,对深刻把握人工智能的本质有重要启发意义。

当然,我们无法从马克思时代的“机器体系”直接推导出今天的人工智能的特性,毕竟二者有天壤之别。“机器体系”主要是解放人的体力劳动,是机器的综合体,虽具有自动化特征,但几乎与人的智力活动无关。而按照人类科技发展的路线图,人在模拟了自身的体力之后,一定会模拟人的智力。因此,“机器体系”和人工智能可以看作是科技发展在不同阶段上的技术人工物,前者主要是对人的体力的模拟,后者是对人的智力的模拟。与“机器体系”不同,人工智能更多的呈现出“机器的能动性”“自主性”,是能够进行“学习、推理、规划、创新”,进而自主性解决问题的智能系统,本质上是一种基于“指令—智能—反馈”进行学习决策的“多领域的综合智能”<sup>⑥</sup>。

在19世纪,马克思当然没有预测到人工智能的产生,他对“机器体系”的评价更多是从经济属

① 《马克思恩格斯全集》(第31卷),人民出版社1998年版,第94页。

② 《马克思恩格斯文集》(第2卷),人民出版社2009年版,第580页。

③ 《马克思恩格斯全集》(第47卷),人民出版社1979年版,第525页。

④ 《马克思恩格斯文集》(第8卷),人民出版社2009年版,第354页。

⑤ 涂良川:《马克思“机器论片断”的机器技术哲学叙事》,《哲学研究》2022年第9期,第32页。

⑥ 刘智慧:《民法典视域下人工智能法律规制论纲》,《学术交流》2023年第9期,第90页。

性和阶级属性上进行的,而今天更被关注的是对人工智能的法律和道德规制。因为对人工智能进化的担忧并非对于岗位替代引起的失业恐慌,而是忧虑人工智能整体进化后很可能脱离人类的控制,进而反噬人类。此种担忧主要源于人工智能在深度学习和交互学习中知识增长的不透明,尤其是其自主知识创新形成的未知性。而无论人工智能如何发展,必须遵循人的主体性这条逻辑主线去把握其本质。

### 三、建构智能时代的新型人机关系

回归马克思的“机器体系”叙事逻辑,是为了揭示人工智能的本质规定性,厘清科技进步与社会制度之间的复杂关联,为智能时代建构新型人机关系奠定必要的理论根基。

在马克思“机器体系”叙事逻辑里,机器的灵魂是资本赋予的,这是资本主义生产方式使然。在发达工业社会,科技本身已演化成为一种政治统治力量。诚如马尔库塞所言:“技术的解放力量——使事物工具化——转而成为解放的桎梏,即使人也工具化。”<sup>①</sup>从“机器体系”向人工智能的历史性演进,本质上是一场技术人工物从一般形态向智能形态的蜕变。这一过程正以前所未有的力量重塑着社会生产与生活的整体图景,不仅驱动着经济、政治、文化、教育等制度层面的显性变革,也日益渗透到精神、艺术、情感等内在经验领域,持续且深刻地影响着人的存在方式与意义世界。

特别是在资本运作的场域中,这一演进推动了平台资本主义、信息资本主义、数据资本主义等新型资本形态的兴起。在此类组织形式下,资本借助人工智能技术,将其支配逻辑从物理层面延伸至精神层面,试图实现对人从身体到心灵的全面渗透与掌控,进而使人服膺于由资本所主导的智能秩序。这也使得“资本在现实世界和网络虚拟世界大行其道”<sup>②</sup>。然而,正如当年马克思并没有否定机器带来的进步,只是否定了恶意使用机器的资本主义制度,今天面对人工智能迅猛发展而带给人类的焦虑和恐慌,也不是要阻止人工智能的发展,而是要把人工智能建立在合理的制度之上。历史地看,马克思当年的批判指向并非机器本身,而是其背后的资本主义制度。同理,对今天的人工智能也应采取同样的辩证立场:问题的症结不在技术,而在于能否为技术建构一个合理的社会制度框架。

沿循马克思对机器体系的资本主义应用的批判逻辑,可以发现,当代人工智能在本质上仍延续着与机器体系相同的资本运行逻辑,即作为资本实现价值增殖的工具。从资本主义制度的视角审视,人工智能可被纳入“数字资本主义”的范畴。所谓数字资本主义,是指通过数据分析与处理技术,将数据转化为可交换的价值形式,并以数字资本为载体,实现对劳动者剩余价值的系统性榨取。换言之,数字化技术广泛捕捉生产与生活中产生的各类数据,将其转化为利润来源,数据本身被资本化了。在数字资本主义的运作中,一切经济活动和个体行为在不停地生成着数据,这些数据被人工智能与机器学习系统实时捕获、提取,并转化为分析乃至预测人类行为的依据。所以,大众在日常实践中就无意识地扮演着数据生产者的角色,而且充当着数字资本隐蔽且日益普遍的剥削对象。

在资本的逻辑支配下,人工智能成为追求利润最大化的工具,导致其被广泛应用于替代各行业的人类劳动,由此被排挤出的工人,在资本主义条件下沦为结构性“过剩人口”。这正印证了马克思的论断:当劳动资料变成自动机,充当了劳动过程中的资本,死劳动便会支配和吮吸活劳动而同工人对立。<sup>③</sup>然而,这种对立不是机器的原罪,而是社会关系使然。当前,先进的生成式人工智能已

① [美]赫伯特·马尔库塞:《单向度的人:发达工业社会意识形态研究》,刘继译,上海译文出版社2008年版,第127页。

② 马俊峰:《数字资本主义语境下主体的消解与重建》,《山东社会科学》2023年第7期,第87页。

③ 《马克思恩格斯文集》(第5卷),人民出版社2009年版,第487页。

成为资本追捧的关键对象,人工智能技术正被少数资本垄断与控制,并在资本逻辑的支配下,通过研发自主化、类人化、智能化的技术来率先占领生产力发展的制高点,成为排斥和否定人的工具。这即马克思指出的“资本的必然趋势”<sup>①</sup>。显然,机器对人的否定性力量,并非源于技术本身,而是由隐置于其后的资本逻辑所驱动与形塑的。

可以预见,人工智能对人类劳动的大规模替代已成为一种客观趋势。大量工作岗位被机器人系统所取代,并非只是一种未来预言,而是正在发生的结构性现实。但“人有人之用”,人类具有专属的能动性 with 创造力,不会在人工智能面前束手无策。人类能够通过创造新的生存样态、生活形式以及变革生产关系,使人工智能成为促进人类解放的协同力量,而非受资本逻辑支配、用于榨取利润与加深剥削的工具。值得警惕的是,无论是马克思所揭示的机器体系对活劳动的支配,还是当下人工智能表现出的对人的排斥,其根源都不应简单归咎于技术本身,更不能因此否定科技进步,否则便可能陷入非理性的技术恐惧之中。

除了制度层面的考量,人工智能技术的内在逻辑与独特属性同样是一个关键的反思维度。只有在深入把握其技术特质的基础上,才可能为建构新型人机关系回应其发展中的伦理与社会挑战提供一种切实的路径。从技术本质上看,人工智能是建立在深度学习与交互学习基础上的系统,其运行依托于大数据、大算力与大模型的支撑。通过多层神经网络结构与反向传播等算法机制,人工智能能够实现自主特征提取与多任务并行处理,逐步形成可模拟人类认知过程的“自主化”学习与决策能力。马俊峰认为,数字技术在推动社会全面数字化进程中,塑造了一个具有高仿真、低延时、智能化、协作化运作特征的虚拟世界架构,不仅能够高度复刻物理现实,更深刻重构着人类的价值认知与存在形态。<sup>②</sup>这说明人类已经进入人工智能时代。作为具有高度自主性的智能技术人工物,此类虚拟世界装置在运行中必然会触及法律遵循和道德判断等复杂问题。其中,人工智能是否可能突破其程序预设的法律与伦理边界,是最令人担忧的。

人工智能可能带来的法律和道德问题,人们在其发展初期就有所预见。20世纪中叶,科幻作家艾萨克·阿西莫夫在其短篇小说《转圈圈》中提出了机器人学三定律:(1)机器人不得伤害人类,或因不作为而让人类受到伤害。(2)机器人必须服从人类发出的命令,除非这些命令与第一定律相冲突。(3)机器人必须保护自己的存在,只要这种保护与第一或第二定律不冲突。此后,他在《机器人与帝国》一书中进一步补充了“第零定律”,即“机器人不得伤害人类,或目睹人类将遭受危险而袖手不管”,并将其置于优先于前三定律的地位。这是以科幻叙事形式,对未来自主智能系统进行的一次前瞻性的伦理框架构建。

据报道,制药公司 Insilico Medicine 通过 AI 药物研发平台设计的治疗纤维化疾病的新药 TNIK 抑制剂已经进入临床 2 期实验,该药物是世界首个利用生物 AI 发现靶点并确定优先次序、同时利用化学 AI 生成分子的药物。此研发体现了生成式人工智能在药物发现领域的巨大潜力。<sup>③</sup>2024 年诺贝尔物理学奖和化学奖都与人工智能研究相关,这亦证实了人工智能已成为推动基础科学发展的重要工具。大量实践表明,人工智能通过促进跨学科深度融合,显著加速了科研成果的产出周期,不仅降低了科学发现的专业壁垒,推动研究走向大众化,更引发了新一轮的科技革命。作

① 《马克思恩格斯文集》(第 8 卷),人民出版社 2009 年版,第 186 页。

② 马俊峰、赵海蕴:《数字资本权力运行的三重逻辑》,《马克思主义理论教学与研究》2024 年第 1 期,第 90 页。

③ 《〈自然〉子刊:从靶点发现直至 2 期临床, AI 如何加速潜在“first-in-class”疗法开发》,腾讯网, <https://news.qq.com/rain/a/20240309A00DME00>, 2024 年 10 月 2 日访问。



为新一轮科技革命的核心驱动力,可以预见,人工智能将在更广阔的科学前沿发挥关键作用。与此同时,一个更深层的关切也随之浮现了,其“类人性”自主意识产生的可能性,已是一个不容回避的伦理与技术焦点。

需要认识到,人与人工智能的深度融合,正在促成一种超越传统主客二分的新型关系形态。面对这一转变,亟需建立一种既开放又审慎的认知框架:在积极接纳其带来的进步与创新的同时,必须清醒地认识到,作为一种发展轨迹尚未完全显现的技术,其演进方向与长期影响存在着固有的不确定性。因此,要驾驭这场变革,必须在警觉未知风险的同时,前瞻性地塑造一种与之适配的治理秩序。当前,人工智能技术的演进与滞后的监管之间的“治理鸿沟”已经显现。这迫切要求打破常规,以前瞻的视野将伦理与法律要求“前置性”地嵌入研发与部署流程,构建一个敏捷回应、动态协同的治理体系,以管控技术伴随的未知风险。

目前,全球主要经济体已相继推出一系列政策框架、专门立法及伦理指南,旨在充分释放人工智能创新潜能的同时,通过规制引导其发展方向,力求在促进发展与防范风险之间寻求平衡。如欧盟出台了《人工智能法案》,突出了人工智能发展的风险性,提出了“基于风险”的人工智能治理框架,为人工智能的发展提供了法律依据。该法案创新性地依据潜在危害程度,将人工智能应用划分为“最小风险”“有限风险”“高风险”与“不可接受风险”四个层级,并据此建立差异化的监管体系。在伦理目标上,该框架秉持“以人民为中心”的根本理念,旨在引导人工智能的发展始终服务于人类的福祉。

这一立场与学界的呼吁相呼应。如美国加州伯克利分校斯图尔特·拉塞尔教授便警告,要确保人类创造的人工智能与人类拥有相同的价值观,未对齐的人工智能系统可能会剥夺人类权力、挑战人类的主导地位,甚至将致人类灭绝。<sup>①</sup>因此,将“价值对齐”从一种抽象的伦理原则下沉为可验证、可实施的工程技术规范,是治理人工智能伦理风险的当务之急。在此趋势下,国内学界提出的“交互式价值对齐”<sup>②</sup>路径,同样为人工智能技术的良善发展提供了一种有效性和可拓展性的方法论指导。当前,我国在大力推进人工智能发展的同时,也尤为注意到了对科技风险的管控,坚持发展与安全并重,科学全面地构建了未来人工智能发展的方向与目标。

在构建新型人机关系的进程中,中国特色社会主义制度彰显了独特的结构性优势。这不仅体现为社会主义市场经济体制有效协同了市场配置效率与国家战略调控,更在于“以人民为中心”的发展思想,与马克思主义“人的全面发展”理论具有高度的价值同一性。基于此,立足中国实践的新型人机关系建构,必然遵循一条“技术创新、劳动保障、制度完善”的系统协同路径。与此同时,必须将制度优势转化为系统的治理能力。这要求对科技风险进行前瞻性的研判,通过前置法律法规、健全伦理审查、建构敏捷监管框架并凝聚社会合力,实施全过程治理。为此,我国已初步形成多层次的治理体系:在《网络安全法》《个人信息保护法》《数据安全法》等法律基础上,出台了《新一代人工智能治理原则》《新一代人工智能伦理规范》及《生成式人工智能服务管理办法》等专门规则;同时,依托《国家新一代人工智能标准体系建设指南》,推进计算平台安全框架、算法安全评估等具体技术标准的制定,并在算法可解释性、数据隐私保护、系统安全等前沿领域加强研究,旨在从技术底层提升可信性与安全性。

<sup>①</sup> Stuart J. Russell, *Human Compatible: Artificial Intelligence and the Problem of Control*, Penguin Random House, 2019; Stuart J. Russell & Peter Norvig, *Artificial Intelligence: A Modern Approach (4th edition)*, Pearson, 2020, pp. 31-34.

<sup>②</sup> 李思雯:《人工智能价值对齐的路径探析》,《伦理学研究》2024年第5期。



鉴于人工智能发展的速度之快及其伴生的不确定性,必须坚定“以人中心”的根本导向,严防技术的恶意应用,并坚决遏制具有不可接受风险的人工智能发展。为此,有必要回归马克思“机器体系”的叙事逻辑,以此为分析框架,审辨人、技术、资本与社会之间的复杂关系,重申人的主体性地位与能动作用。这意味着必须以科学而理性的态度对待科技本身,即从历史唯物主义的高度出发,在生产力与生产关系的矛盾运动中把握其方向;既充分承认科技对社会的推动作用,也须对其潜在风险保持警觉,发挥制度的约束与形塑功能,实现科技发展服务于社会整体福祉的目的。总之,唯有发展以人中心的人工智能,前瞻其社会后果,并以技术的社会逻辑驾驭其资本逻辑,才能实现对这一创造物的有效引导,规制其可能被用于邪恶目的的行为。当前出现的诸如“硅基生命将取代碳基生命”的论断,无疑过度渲染了人工智能的威胁。对此,科学的态度在于积极运用人类的理性与智慧,预见并约束技术风险。对那些触及人类生存底线的技术,如无性繁殖、基因重构、人体功能增强等试图突破生物界限的领域,施以最审慎的伦理审查与严格规制;同时,通过健全的治理,防范人工智能技术被用于犯罪活动。

在构建新型人机关系的理论探讨中,学者们基于不同视角提出了多元主张。孙伟平概括了六种可能的关系范式:从将AI视为人类的“新工具”或“新奴仆”,到对其怀有乌托邦或敌托邦式的想象——即友善的“AI神”或竞争的“敌人”,再到伙伴、继承者乃至人机融合的新型生命形态。<sup>①</sup> 马俊峰则从技术哲学的演进路径切入分析,认为传统技术载体(如文字、硬盘)作为“义肢”,主要用以弥补人类自身的生物性局限;而人工智能则实现了自身从一种外在工具转化为具身的“技术身体”。这一转变促使人类角色从技术的主导者演变为与技术互嵌、人机共存的技术互融者,二者共同进入一种“人机共存”的状态。在此状态下,人类的大量记忆、经验与社会关系以数字化方式外存,形成了斯蒂格勒理论中所谓的“第三持存”,即“数字持存”<sup>②</sup>。这一深刻变革,推动了人与技术在对象性关系中共同进化,并有望孕育出一种人机协调的社会新文明形态。

建构新型人机关系,关键在于破除根深蒂固的人机二元对立逻辑,正视人工智能已深度嵌入社会肌理并展现出“拟主体性”互动能力的事实,转而以“互构共存”的视角来理解这一关系。这意味着,新型关系的塑造必须双向考量:既遵循人工智能技术的客观发展逻辑,也须将其置于复杂的社会关系网络中审视,明确人机在协作中各自应承担的社会责任。

基于此,未来的方向应是构建一个“人机共生体”,旨在实现二者功能上的有机协同、认知上的相互增强,乃至情感层面的有限理解与共鸣。这超越了传统的主客二分与工具理性,指向一种在风险可控的前提下深度融合的伙伴关系。这一图景的实现,不仅要求人类具备有效驾驭技术的能力,亦需要完成一场接纳、适应、信任人工智能的心理转变与文化适应,在人与技术的责任共担与价值共塑中,形成一个真正意义上的命运共同体。

#### 四、结语

从“人机共生体”的视角展望,人工智能的深度应用势必会引发社会生活的结构性变革。然而,技术潜力的释放,必须被置于审慎的社会制度框架与风险管控体系之下。唯有如此,才能为人类主动性、创造性与价值主导地位的发挥提供空间,确保技术潜力转化为社会福祉。人类也方能克服对技术的无端恐惧,以更成熟、从容的姿态,迎接一个由人与人工智能共同塑造的未来。

① 孙伟平:《智能时代的新型人机关系及其构建》,《湖北大学学报(哲学社会科学版)》2023年第3期,第21-23页。

② 马俊峰、赵海蕴:《数智时代“第三持存”的教育技术困境及重构路径》,《教育文化论坛》2025年第6期,第23页。

马克思的“机器体系”理论深刻揭示了人机关系异化的社会根源,为我们今天审视智能时代的技术应用提供了一个批判性的分析框架。在中国语境下,中国特色社会主义制度的实践正在开辟出一条超越异化、实现“科技向善”的现实道路。建构新型人机关系,本质在于开辟一个以“人的全面发展”为定向的协同进化实践场域。正是在这一实践中,实现技术进步、劳动解放与社会和谐的内在统一才成为可能。

## **Constructing a New Type of Human-Machine Relationship in the Intelligent Era: A Perspective from Marx's Narrative Logic of “the System of Machinery”**

WANG Yaqi

*(School of Marxism, Shandong Second Medical University, Weifang 261053, China)*

**Abstract:** To address the anxiety induced by the rapid development of artificial intelligence, we argue for a return to Marx's narrative logic of “the system of machinery”. Marx notes that the system of machinery is the materialization of “general intellect” in the society. The system of machinery not only promotes the development of productivity, but also becomes a tool to exploit and squeeze workers. AI is essentially an artifact of intelligent technology, but its intelligent characteristics have brought many uncertainties and risks to the development of contemporary society. Therefore, we should take measures to prevent the potential risks. It is suggested that a new type of human-machine relationship, featuring human-machine integration, human-machine co-competence and human-machine shared empathy, should be established, so that AI can become a driving force to promote social progress and speed up human's liberation.

**Key words:** artificial intelligence; the system of machinery; human-machine relationship; narrative logic

(责任编辑:傅 游)