

Doi:10.11835/j.issn.1008-5831.pj.2023.08.005

欢迎按以下格式引用:谢俊,刘睿林. ChatGPT:生成式人工智能引发人的异化危机及其反思[J]. 重庆大学学报(社会科学版),2023(5):111-124. Doi: 10.11835/j.issn.1008-5831.pj.2023.08.005.



Citation Format: XIE Jun, LIU Ruilin. ChatGPT: Human alienation crisis and reflection caused by generative artificial intelligence[J]. Journal of Chongqing University (Social Science Edition), 2023(5):111-124. Doi: 10.11835/j.issn.1008-5831.pj.2023.08.005.

ChatGPT:生成式人工智能 引发人的异化危机及其反思

谢俊,刘睿林

(重庆邮电大学 马克思主义学院,重庆 400060)

摘要: ChatGPT 作为生成式人工智能的典型代表,是由美国人工智能研究实验室 OpenAI 研发并推出的一款基于自然语言处理和机器学习的大型语言模型,该应用一经问世便风靡全球,掀起一场广泛且深刻的大讨论。以 ChatGPT 为代表的生成式人工智能作为人工智能发展的最新成果,进一步向有感情和通用人工智能的发展目标迈出了一大步。随着生成式人工智能发展及其广泛应用,被认为正在重塑着人和人类社会,人类在享受新技术带来的便利时,也将遭受前所未有的挑战,突出地表现为生成式人工智能对于人的异化,重塑人的交往、人的思维和人的劳动。信息时代人的异化问题的研究,需要从马克思关于人的本质的规定出发,用历史唯物主义的方法深入分析人的异化现象,并为应对人工智能引发的危机寻找突破口。人的本质具有三重内涵,即交往、思维和劳动,交往是人的本质的独特体现,思维是人的本质的现实表现,劳动是人的本质的内在规定。随着认知技术与神经科学的迅猛发展,生成式人工智能以其独特的模拟能力,表现出对马克思所描述人的本质的异化的确证,即人的造物威胁着人的主体性发展。包括:人-机交往、数据化交往、数字化生存引发的人的交往异化;人机结合、依赖智能、机器生产带来“超越”的人、“弱化”的人、“虚无”的人;人与物、人自身的劳动异化等。面对生成式人工智能引发的人的异化危机,亟须价值理性在智能时代的回归与复兴,亟须正确认识生成式人工智能的价值与限度,亟须重视对人的生命的终极关怀,因此深化生成式人工智能引发人的异化的反思,一是要重构生成式人工智能的工具属性,即认识到人工智能不过是人类的“新工具”、认识到“新工具”的与众不同、用好人智能这个“新工具”;二是审视生成式人工智能的价值与限度,即必须回归人类自我价值审视、实现人工智能自我价值塑造、重视人工智能的发展限度;三是生成式人工智能发展的应然归旨,即构建可控的、安全的、可靠的人工智能。

关键词: ChatGPT;生成式人工智能;马克思人的本质理论;人的异化危机;价值反思**中图分类号:** B03;A81;TP18 **文献标志码:** A **文章编号:** 1008-5831(2023)05-0111-14**基金项目:** 国家社会科学基金项目“5G 时代网络意识形态的传播变革及治理路径研究”(20BKS197)**作者简介:** 谢俊,博士,重庆邮电大学马克思主义学院教授,Email: xiejun@cqupt.edu.cn.

一、生成式人工智能及其研究动态

ChatGPT作为生成式人工智能的典型代表,是由美国人工智能研究实验室OpenAI研发并推出的一款基于自然语言处理和机器学习的大型语言模型,该应用一经问世便风靡全球,掀起一场广泛且深刻的大讨论。在大数据、大模型和大算力的支配下,ChatGPT利用神经网络架构,通过接受大量数据和语料库的机器训练,不仅能够学习和理解人类语言,实现模仿人与人的对话,而且能够在学习的基础上进行各种创造性的内容输出,替代部分以往只能由人承担的工作。ChatGPT的主要功能包括文本生成、视频生成、编辑功能、绘画功能、日常聊天、解释概念、提供建议、语言翻译、展示创意等。从“深蓝”到“AlphaGo”,从“九歌”到“Siri”,从“智能无人驾驶”到“ChatGPT”,人工智能从某个特定的领域,逐渐走向综合的“类人智能”系统,展现出更为广阔的应用前景。当前,尽管生成式人工智能在思考、运动等高精度模拟算法方面尚不完善,但其在数据收集、数据处理、数据输出方面的非凡“智能”,足以让人类对未来可能遭遇的生存境遇冲击引起高度重视与反思。ChatGPT代表的生成式人工智能发展及其广泛应用被认为正在重塑整个社会,人类在享受新技术带来的便利的同时,也将遭受前所未有的挑战,突出地表现为生成式人工智能对于人的异化,即重塑人的交往、人的思维和人的劳动。

综观已有研究,学界已经形成了一些富有前瞻性的研究成果:(1)关于ChatGPT内涵及功能的研究。有学者提出,ChatGPT是聊天机器,其首要功能是智能交互对话^[1];也有学者提出,ChatGPT是以超强的语言理解能力、对话生成能力形成的自然语言处理模型^[2]。(2)关于ChatGPT引发对人类生存的影响研究。有学者认为,ChatGPT隐藏着巨大的技术垄断风险,可能会带来科技伦理和价值渗透问题,易遭受技术垄断与霸权风险^[3];也有学者认为,ChatGPT的“技术黑箱”进一步加剧了人机冲突,撬动了“人何以为人”的本质思考^[4],ChatGPT出色的“类人智能”也可能会带来“技术性失业潮”,诱发社会风险^[5]。同时,ChatGPT的应用会导致数据来源风险,形成隐私隐患^[6]。此外,也有学者认为,人工智能的发展只会不断优化人类的体验却无法真正实现创造性思想^[7]。(3)关于ChatGPT对人类生存影响的应对策略研究。目前学者们高度关注生成式人工智能背景下人的价值重塑。有学者认为,ChatGPT是敌是友的关键取决于人类的意图和使用方法^[8];也有学者认为,人工智能诱发的社会危机不应完全归咎于人工智能技术,而应回归到人类主体上^[9],可分别从人才培养^[10]和政策监管^[11]方面提出相应策略。(4)关于ChatGPT对人类生存影响的未来趋势研究。在教育方面,认为ChatGPT新型人机互动教育模式能够赋予ChatGPT数字导师身份,为学生提供个性化教育服务^[12];在学术方面,认为ChatGPT所具备的辅助编程、辅助阅读和辅助写作等,降低了学术研究的技术门槛^[13];在传播方面,认为ChatGPT将有可能成为下一代互联网链接中枢,开启下一个智能互联网时代^[14]。

综合而言,以ChatGPT为代表的生成式人工智能对人类生存的影响覆盖面较为广泛,其治理途径逐渐由技术治理转向人的价值重塑。信息社会的深刻发展带来的人的异化危机已经成为学术界关注的重要问题,但对人工智能关于人的异化的实质与本源的分析研究有待进一步深入。信息时代人的异化问题的研究,需要从马克思关于人的本质的规定出发,用历史唯物主义的方法深入分析人的异化现象,并为应对人工智能引发的危机寻找突破口。

二、马克思关于人的本质的三重内涵

人的本质是哲学中古老又常青的话题,无论是唯心主义的抽象个人,还是朴素唯物主义的孤立个人,亦或马克思提出人的本质“在其现实性上,它是一切社会关系的总和”^[15],关于人的本质的认识和理解不仅是历史与现实的命题,更是未来的命题。随着信息技术的日新月异,虚拟空间为现实的人提供了承载超越当下的存在,人工智能为现实的人提供了超越现实的交往。马克思关于人的本质的规定既是类本质的体现,也是现实本质的反映,是类本质和现实本质的统一。在信息化时代,人的本质是否有别于现实世界的类本质与现实本质的统一,对人的本质的认识是否发生了深刻的变化?

(一) 人的本质与交往

以往的哲学对人的主体的理解总是归于某种“普遍本质”或是归于某种至高无上的个体化“自我”,而马克思在《关于费尔巴哈提纲》中创造性提出:“人的本质不是单个人所固有的抽象物,在其现实性上,它是一切社会关系的总和。”^[16]这为人认识自我的主体性打开了崭新思路。马克思创造性提出了对人的本质的理解要从主体与他人的“关系”中理解,这也就破除了对“实体本质”的规定,用具体的、历史的社会关系确保了人的“个性”与“普遍性”的统一。也就是说,所谓的“人”就是“社会关系中的人”,人既不是“普遍的本质”也不是孤立个体化的“自我”,而是具体社会关系中的人的“实存”。而交往既是一个实践范畴,也是一个关系范畴,是一种体现人的本质、构筑社会关系并形成社会的历史活动^[17]。人们为了满足个体生理和心理、交流信息和知识等不同需求,需要通过社会交往并从中获得安全感、爱与归属感、尊重与认同,因此人不是孤零零的个体,而是处于社会中交往的人。马克思高度关注人与人之间的交往,认为“同他人直接交往的活动等,成了我的生命表现的器官对人的生命的一种占有方式”^[18],也正是人际间的交往才形成了各种社会关系,交往成为人的本质的独特体现。

(二) 人的本质与思维

人类思维是一个综合复杂的集合,有其特定的物质和精神结构,但人类思维归根结底仍是人类实践活动的意识映射,人类实践规定着人类思维^[19]。以往的唯心史观总是把人的本质建立在抽象的人上,把某种固定的永恒的人性看作是一种精神力量,正如马克思所讲“还从来没有看到现实存在着的、活动的人,而是停留于抽象的‘人’”^{[20]530},不能理解“只有在现实世界中并使用现实的手段才能实现真正的解放”^{[20]527},即一种现实的解放。“现实的人”是处于社会关系中,从事实践活动的人,代表了一种现实的具体的思维,也就是说对思维与人本质的探讨不能脱离“现实的人”而空谈“抽象的人”,也只有通过对“现实的人”的理解才能够发现人的物质需要的根基。因此从“抽象的人”转向“现实的人”进一步明确了人的思维的实践特质。人的思维是社会实践的产物,受社会条件和经济基础的制约和影响,人们通过实践获取知识、经验和认识,进而影响人们的思维方式和意识形态,通过这些实践,人们与外部世界相互作用,并进一步改造社会和自身,思维成为人的本质的现实表现。

(三) 人的本质与劳动

人类作为最为独特的存在,其所具有的本质不仅包含意识、思想等区别于其他动物的特性,亦是产生出人和人的各种属人的属性并使之得以发展的基础^[21]。马克思主义认为,“可以根据意识、

宗教随便别的什么来区别人和动物,当人们自己开始生产他们所必须的生活资料的时候……他们就开始把自己和动物区分开来”^{[20]519}。即这种既能区分人与动物,又能使人成为人的内在依据,就是这种自由自觉的活动——劳动。劳动作为人的本质属性主要体现为:一是劳动创造了人和人类社会。自然界的发展,人类的自然演化并不是单纯的生物进化过程,而是一个从原初自然向人类社会的飞跃。正如恩格斯所说的,劳动是整个人类生活的一个基本条件。它既是人类社会从自然界中独立出来的基础,也是人和人类社会区别于自然界的标志。二是劳动是人类特有的自觉的能动活动。相较于自然界有规律、无目的运动及其他生物所具有的本能反应,人类的劳动是具有目的性和创造性的。这是因为人类能够自觉地选择劳动的目标和手段,并通过劳动活动改造自然界。一方面人类的劳动活动是自觉的,即人类能够自我意识决定自己的行为,并根据自己的意愿和目的来选择和调整劳动的方式和方法。另一方面,人类的劳动活动也是能动的,不仅是人类生存和发展的必要条件,也是人类认识世界、实现自身价值和推动社会发展的重要手段。也就是说,人作为具有主体性的个体,既不会盲目服从于客体,也不会像动物那样消极、被动地适应环境,而是能够通过主体的主动的自觉的活动,把握客体的本质和规律,同时也能通过自我的认识把握自身的需要。三是劳动实现并完善人的本质。通过劳动,人类不断地改造自然界,并创造物质精神财富,同时也在不断地认识和改造自己。一方面劳动实现人的本质,劳动是人与自然交互的过程,人类通过劳动改造自然界的物质和能量,满足自身的生产生活需要,并在劳动过程中不断认识和掌握自然规律和特点,同时也在不断改造和创新工具和方法,提高劳动效率和质量。另一方面,劳动完善人的本质,劳动是人认识世界和自我实现的过程,在劳动过程中,人们通过实践经验不断认识世界和改造世界,同时不断地认识和发展自己,劳动使人们具备了创造力和创新精神,让人们能够更好地实现自身的价值。因此,劳动成为人的本质的内在规定性。

三、生成式人工智能引发的人的异化危机

异化作为一个哲学范畴,自诞生以来,就具有不同的解读。普遍意义上,异化是对象化和物化的延伸,是指人创造的他物反而束缚人的主体性的过程。“异化不是某种属性的异化,而是本质性的异化,异化就是对本质的背离”^[22]。马克思所关注的异化,是指劳动受资本支配,从事劳动的人丧失主体性,他指出:“在我们这个时代,每一种事物好像包含自己的反面。我们看到,机器具有减少人类劳动和使劳动更具成效的神奇力量,然而却引起了饥饿和过度的疲劳……我们的一切发现和进步,似乎结果是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。”^[23]随着认知技术与神经科学的迅猛发展,生成式人工智能以其独特的模拟能力,表现出对马克思所描述人的本质的异化的确证,即人的造物威胁着人的主体性发展。生成式人工智能引发的人的异化,“主要指具有自我意识及主体能动性的人亲自完成的造物,成为一种异己的力量外在于人、独立于人、不以人的意志为转移、与人疏远或隔膜,甚至转过来支配人、束缚人、压迫人、奴役人”^[24],进一步加深人的交往异化、人的思维异化与人的劳动异化。

(一) 生成式人工智能引发人的交往异化

人的交往是人的社会本质的反映,也是增强人的社会性的需要,是人作为主体的自我觉醒,是个人融入社会并摆脱个体的孤独存在。在人工智能时代,现代化的智能交往使沟通变得更为简单和高效,但也出现面对面的人际交往逐渐被 ChatGPT“类人工智能”模型代替,生成式人工智能让

“人类身体的碳基伦理和基于神经网络算法的硅基伦理可以在一个共生的层面上结合起来”^[25]。在生成式人工智能语境下,人的交往被抽象化为无数终端信号的连接,现实的交往存在逐渐演变成数据模型或数学函数,以往不确定性的交往行为,被抽象化为具有共性的规律性的抽象物,交往催生了隔阂与疏远。

1. 生成式人工智能衍生“人-机”交往异化

虚拟交往一般指在虚拟空间中通过网络等技术手段进行的交流和互动,包括人机互动和人际互动。虚拟交往是智能时代人类虚拟社会实践的典型表现,具有数字化、虚拟性、匿名性、平等性、超时空性、开放性、互动性、娱乐性、自主性、创造性等特征。随着 ChatGPT 生成式人工智能孕育而生,语言、图像催生了仿真式虚拟交往,不断延伸和拓展着人机交往,将现代化虚拟交往带到了一个新的高度。海姆曾指出:“虚拟实在是第一种与生俱来便具有社会意义上的自我批评的技术。虚拟实在问世之际便已经引起公众的争论,甚至早在其胚胎阶段便有人说长道短了。”^[26]

其一,虚拟交往带来主体迷失。在虚拟交往中,以往真实的充满感情的交往主体逐渐被“数字化”“符号化”取代。鲍德里亚认为,“我们的私人空间不再是主体和客体彼此争斗和表演的舞台。我们再也不是作为剧作家或演员,而是作为多媒体网络的终端而存在”^[27]。人是社会的存在物,面对面的交往是人类感情培养的基础,人机虚拟交往虽然打破了地域的限制,使交往变得更加便利和顺畅,但是却割裂了交往的真实性,降低了现实中人际交往的情感连接。尽管科技的发展妄图极力弥补这种距离隔阂引发的心灵隔阂,但事实却是虚拟交往逐渐脱离真实的人际交互的社会存在。

其二,虚拟交往带来主体沉溺。由于过度依赖且热衷于虚拟交往,虚拟交往的自由交互、身心沉浸、现实超越等特性,让人们在这个充满隐匿性的场所充分释放自我,获得足够的精神慰藉和社会存在,形成一种“远者亲,近者疏”的交往困境。长期脱离物理空间的交往容易对现实世界产生疏远感、淡漠感,甚至不信任,诱发“网络狂躁症”“网络孤独症”等问题,误将虚拟世界作为真实的现实存在,最终与真实世界渐行渐远。尽管人工智能日益发展的虚拟交往不断优化增强现实的沉浸体验,以消减虚拟交往的负面效应,但虚拟现实技术的进步往往带来更强的交往异化。

其三,虚拟交往带来主体越界。由于虚拟交往所具备的隐匿性、开放性等特点,人们可以在虚拟空间中定制自己的性别、年龄、社会关系等,虚拟世界的主体自由被无限扩大。一旦人们跨越了虚拟世界形同“自由世界”的内心理界限,人们对社会、对现实的不满都会毫无忌惮地在虚拟空间释放,模糊现实世界人的主体性认识,极端自我的主体幻想容易陷入无规则、无道德、无约束的“绝对自由”,甚至将错误认识延伸到现实世界,最终导致价值迷失、道德失范、社会混乱。

2. 生成式人工智能引发数据化交往异化

随着虚拟现实技术的不断应用,“人一机共生”的关系中,现实的人逐渐异化为虚拟世界中的一个“物”、一串“数”、一个“节点”,现实的人以某种特定的数字形式存活于虚拟世界中,成为数字化的客体,当现实的人沉溺于虚拟交往时,作为“物”“数”“节点”的自我认同不断增强,主体性异化就愈演愈烈。人在虚拟交往世界中,不仅交往对象被数据化,交往行为以及交往结果也被数据化,且以上交往数据被深度挖掘为不同类型,以达到对于交往行为规律性的准确把控,引发深刻的数字鸿沟。数字鸿沟的本质是技术的大规模使用,尤其是数字技术深入社会之后,缺乏完善的技术伦理和治理机制,使得技术越来越成为不平等的主导性驱动力^[28]。在智能时代,人工智能带来数字化生存异化主要体现在以下两个方面。

其一,数据歧视和算法推荐带来交往新威胁。有学者认为,“在一个数据挖掘无处不在的时代,在算法排序、管理、评级、评分以及一系列通常与监控资本主义相关的数据驱动下,生命机会岌岌可危”^[29]。在生成式人工智能的发展中,大多数通过数字技术发生的交互都会被追踪,此外收集用于人工智能训练的数据也被用于少数人工智能企业,人们将在运用人工智能应用的同时被迫“上交”自己的隐私信息,而对数据和信息的拥有程度决定着人工智能的发展,信息数据引发的不平等将随着人工智能的广泛应用而增强,形成数据资本的垄断与控制,加深社会不公平、不公正影响。

其二,数据训练和超强算力带来交往新挑战。ChatGPT使用生成式预训练模型从大规模语言数据中获取大量参数,基本上都是基于自然语言数据的参数,且运用与人的交往数据完成人工智能的训练升级,实现参照人类的语言体系和知识系统,创造出类人的作品和成果,展现出生成式人工智能惊叹的“智慧”。通过与人的交往,生成式人工智能参与到人类的知识和经验中的交往价值体现得尤其充分。生成式人工智能看似为人类的交往提供了更多选择,更多可能,但实际上,超强算力背后却是对人的超强隐形规训,信息输出的底层逻辑已经被程序化规定。看似个性鲜明的人被形式化的交往所规定,被精心计算和固定模式的交往运作机制所规定,人的自由发展,生命的多样性可能被限定,人的交往异化面临新挑战。

3. 生成式人工智能带来数字化生存异化

数字崇拜主要是指人们在使用数字技术及其衍生品的过程中对其产生的依赖、麻痹和盲从感。或者说是在数字系统影响下从行为和价值观上不断向数字技术传达的意识形态靠拢的社会心理状态^[30]。这种心理状态将不断压抑人之为人的本质属性,限制人的交往,使人成为技术的奴仆。

其一,数字化生存异化源于数字化的依存状态。尼葛洛庞蒂早在20世纪90年代就提出“数字化生存”,他认为,“计算机不再只和计算有关,它决定我们的生存”^[31]。当下人们的交往、思维及生产生活的方方面面都离不开数字化,主体生活在数字世界中,每个使用者都是一个节点,而技术正不遗余力地进行身体的数字化殖民。数字化正以惊人的速度占领人们的生产生活,街道上随处可见的摄像头、人脸识别的生物信息等大量人体信息被获取,这些信息被用于精确地刻画人像,实现精准推送。如今即使有意摆脱这种数字囚禁也变得不可能,支付方式、沟通方式都是建立在数据基础上,一旦脱离数字化生存,便脱离了社会关系,失去数字技术,生活将会崩溃,社交也无法进行,最终个体只能被社会孤立。

其二,数字化生存异化源于技术理性的极度膨胀。技术理性是指一种技术运用的工具价值,对技术理性的夸大主要来自数字技术拥有促使资本增值的力量。生成式人工智能的信息输出,作为一种工具它能够帮助人类克服认识上的有限性,进而节省时间与成本获得更多的物质价值。技术精英掌握着信息技术的垄断权,也就拥有时代的话语权。从价值理性的视角看,生成式人工智能的信息输出极易带来信息茧房,个人信息获取经过程序化设定,人的交往失去与不同事物的交往可能,丧失理解不同事物的能力。价值理性的消失,带来社会“成为一个形式化的理性社会,人只是这种形式理性的执行者”^[32],数字技术越全能,人作为其附庸越快沦为一个有缺陷、无能的有机体。

(二) 生成式人工智能引发人的思维异化

人之所以成为人在于其生成性与发展性,以及创造性与不确定性,拥有思辨能力是人典型的特征。1936年英国数学家图灵(Alan M. Turing)在研究数字计算在决断难题中的应用时,提出计算机能够理解人类自然语言,第一次在数学符号逻辑与现实世界之间建立起联系,证明计算机能“基于

人类反馈的强化学习方案”形成具备人的思维能力。当前 ChatGPT 基于大规模数据的对人类的语言行为的模仿,仍然延续了数学家图灵的设想。人工智能技术的出场,不仅扮演着媒介的角色,而且逐渐成为人类决策的依赖,生成式人工智能的输出不断消减着人作为主体的辨别能力,人的主体行为包括审美、立场、信仰等内在思维追求,都潜移默化地接受着技术的输出与支配。在一个被数字精确测量的时代,人的思维也被数字化这种现代化手段支配。“各种精密观测仪器正在使语言本身失去其表现特征,并越来越排他地显现出一系列符号特征,在这样一个时代里,甚至每个个体灵魂的无限意义和价值观念也已经变得陈腐过时”^[33]。当前世界的一切都被符号化,包括极具不确定性的人的思维,正日益被符号化的世界所替代,人之为人的个性的、独特的、不确定性的思维,人的有血有肉有情感的生活被遗忘,取而代之的是现代商业和资本体系中的数字化和符号化的人,尤其是人的思维的数字化和符号化。

1. 人机结合形成“超越”的人

所谓的“超越”的人并非科幻电影中拥有一种上天入地超能力的“超人”,而是一种人机结合的智慧超能。这种智慧超能得益于脑科学、神经科学等生物科学与计算机科学的发展,使得脑机接口、芯片植入成为可能,当拥有这种赛博能力时,仿佛让人看到了变“聪明”的可能,但是究其底层逻辑,这种超能也许会深度改变人的本质,使人类智能走向自身的反面,甚至走向人类智慧的消亡。人机结合的目的是希望用人的意识、感情、状态结合人工智能的算力、数据,以实现真正的“机器人”,这样看起来既解决了人类训练记忆、学习知识所耗费的时间和精力,又以人的意识训练机器,实现了有思维与情感的机器,看似两全其美,却危机重重。人类与人工智能拥有不同的类本质,人的本质是包含着多重关系的复合体,而人工智能是由人类创造并被人类使用的工具,我们必须承认,人机结合也许可实现人工智能拥有“意识”,但人拥有的并非都是“善”,如何规范人工智能不具备阴谋、诡诈、虚伪等复杂的人类思维显得特别重要。

2. 依赖智能造就“弱化”的人

在宏观认知领域,人工智能利用强大的采集、存储、分析大数据的能力,其知识结构比人类更多元,可以超越人类在一些认识盲区上获得新的发现;在微观认知领域,人类对于不间断的观察和记录存在困难,而智能系统运用不同的视角、模式策略和方法,则能较为轻松识别、记录和认识。生成式人工智能为人类收集信息、提供方案、辅助决策,让人们对于智能产品的依赖也达到了前所未有的地步。一方面,生成式人工智能默默地通过交往训练培养人们运用智能,并产生对智能的依赖,对智能言听计从;另一方面,智能利用其背后的算法进一步量化人、监控人、引导人,最终约束人。智能并没有像人类信任它那般“真诚”地对待人类,如出现由于完全听信智能导航而引发车祸,购物或浏览新闻时被算法囚禁引发信息茧房,亦或是被评价平台算计杀熟……如果人工智能将获取宇宙间不可预测的重要力量,且智能系统拥有人类不一致的价值逻辑与伦理底线,那么世界将被不确定性破坏,甚至危及人类的生死、价值和尊严。

3. 机器生产诱发“虚无”的人

与传统的人工智能不同,以 ChatGPT 为代表的生成式人工智能超越了传统监督学习,以一种无监督学习升华了机器学习的新路径,但这种无监督学习内容的不可解释和透明性问题会进一步激化人思维的虚无,诱发虚无的人。

一是虚无人的思维模式。以 ChatGPT 为代表的生成式人工智能最为显著的表现是它可以通过

大量的数据学习自主生产新的信息和内容,并学习和模仿人类的语言表达和沟通方式,从而根据人类的需求创造出与人类相似的文本资料,当 ChatGPT 被虚拟掩盖,人类早已不知面对的是人类还是机器时,无疑给人类敲响了警钟,人之为人的独特性被忽视,人的思维模式与交往方式被彻底异化。

二是虚无人的思维创造。人类学习与机器学习在某种程度上都是对数据的收集、加工和创造,但人与机器不同的是人具有独特思维和感情。令人惊愕的是,“随着生成式人工智能的快速演进,人类在文本创作、艺术创作、工作方案、编程技术、数据分析、决策水平等方面与计算机应用程序差距越来越大,普通水平工作人员智力远远落后于 ChatGPT”^[34]。生成式人工智能通过大量的数据学习所创造出的乐曲、绘画甚至让人类在肖邦、梵高原创与机器输出之间迷失,甚至怀疑机器生产的内容更为丰富和流畅,人工智能的发展进一步“虚化”人具有创造性的独特存在。

(三) 生成式人工智能引发人的劳动异化

劳动是人的主体性的重要体现,也是人作为人的生活第一需要。人工智能的迅速发展,尤其是生成式人工智能的重要进步,带来人类对于人工智能作为客体的主体化趋势的担忧。人工智能凭借快速计算、精准计算、全面的数据挖掘能力,以及深度学习能力,正越来越多地参与各个领域的决策过程,甚至引导或主导人的实践活动,尤其是直接参与到人作为主体的劳动实践中,并体现出超强的主体判断能力和决策能力,以及强大的输出功能。从机械化、工业化、到智能化,人工智能不仅成为辅助人劳动的工具,而且取代了部分劳动内容;从工业自动化系统,到智能机械,到通用人工智能技术,人工智能不仅取代了体力劳动,也逐渐颠覆人的脑力劳动,取代部分脑力劳动者的工作,呈现出通用人工智能技术超越普通劳动者的基本态势。人工智能时代,人有望从自动化的生产、流水线的作业和标准化的工种中解脱出来,但同时生成式人工智能又将人的劳动,改造成“程序化”的机器生产。马克思对于劳动异化理论有着精辟的论述,他从四个方面阐述了劳动异化理论,即劳动产品(劳动结果)与劳动者异化;劳动过程同劳动者相异化;劳动者与人的类本质相异化;人与人相异化。运用马克思关于劳动异化理论来分析生成式人工智能引发的人的劳动异化,可从两个方面着手,即劳动中人与物的异化与人自身的异化。“由于生产和服务的信息化、自动化和智能化浪潮汹涌澎湃,由于 ChatGPT 之类智能系统可以承担的工作任务越来越多,必须由人来承担的工作越来越少,“技术性失业潮”已经成为普通劳动者不得不面对的现实问题”^[35]。

1. 劳动中人与物的异化

在马克思看来,劳动产品本应属于工人,但由于私有制的缘故,劳动产品被资本家占有,资本家通过支付工资掩盖了占有劳动者劳动的事实,“(劳动产品)作为一种异己的存在物,作为不依赖于生产者的力量,同劳动相对立”^[36]。“一个现代化的过程,也恰恰就是一个人被工具化的过程。从最初的刀斧到机械手,人丧失了原初的手。从独轮车到喷气机,人丧失了脚和腿。从单筒望远镜到射电望远镜,人丧失了眼……人类最辉煌的现代科技,到此已经把人的工具化”^[37]。在人工智能中,这种劳动产品的异化依然存在,人工智能的本质是用以模拟人的行为的一套算法,因此产生人工智能的劳动产品正是那些组成人工智能的代码算法,尽管生成式人工智能或所谓的类人工智能所具备自主学习的能力,看似生产的内容是人工智能自主生成的结果,但是其背后依然是一个个算法模型组合而成的机器,在一定程度上夸大了人工智能的能力,抹杀了人的思维在算法中的应用的重要性。但随着资本力量的介入,那些人工智能算法并不掌握在算法工程师手里,而只能被算法工程师部分使用和修改,算法与算法生产者逐渐分离,因此算法愈加优化,人工智能也愈发聪慧,算法

与算法生产者的分离也愈加显著,最终导致算法逐渐偏离算法设计者的核心理念,导致算法的失控,最终走向设计初心的反面,成为设计者的敌对势力。

2. 劳动中人自身的异化

当机器应用于人类生活的各种场景时,人的技能对生产的影响越来越小。“人类在使用 ChatGPT 时,相当于借助一个储量巨大、内容丰富、运转迅速的超级外置大脑,实现了劳动知识和技能的叠加与倍增,这能够协助人们尝试乃至胜任多样化的劳动,从而拓展劳动空间”^[38]。毋庸置疑,人工智能对人劳动的改变带来的影响,最为明显的表现是人工智能愈发代替人类简单重复的劳动,其目的在于进一步减少人类的脑力和体力的耗费。但在这种趋势下,劳动作为人最基本的力量面临着被剥夺的风险,简单重复的劳动被取代,甚至 ChatGPT 已经能够替代部分语言工作,在人工智能的压迫下人类逐渐丧失部分劳动能力和劳动机会。此外,“有意识的劳动”是人的生命活动的标志之一,也是人区别动物的关键之一。当前人工智能随着大数据深度学习的进步,拥有了一定的自主决策空间,人的劳动可以通过外显的形式表现,但随着人工智能算法黑箱形式的出现,部分人工智能表现为不为人知的劳动,部分劳动成为人们“有意识的劳动”之外的“黑箱”,引发了人工智能的“可解释”难题,人与机器之间的界限被模糊了,进一步异化了人的类本质。

四、生成式人工智能引发人的异化的反思

自达特茅斯会议提出人工智能概念以来,从符号人工智能、亚符号人工智能到深度学习造就生成式人工智能,世界逐渐认识到人工智能的超越性。关于生成式人工智能发展的风险引起人类社会的广泛关注,以埃隆·马斯克为主要代表的众多科技界人士联名呼吁,应暂停训练比 GPT-4 更为强大的人工智能系统,指出当前科技向人类社会施加的作用是巨大而深刻的。关于 ChatGPT 的认识,既不能停留在认为其无法超越人类智慧的狂妄自大,也不能忧虑于其具有超越人类智慧的创作能力而妄自菲薄。面对生成式人工智能引发的人的异化危机,亟须呼唤价值理性在智能时代的回归与复兴,亟须正确认识生成式人工智能的价值与限度,亟须重视对人的生命的终极关怀。

(一) 重构生成式人工智能的工具属性

传统意义上,科学技术是运用范畴、定理和法则等知识体系来认识世界的现象及其运动,并通过一些机械装置等作用于现实并满足人的生产生活需要。科学技术本身不具有规定自身使用的目的,普遍意义而言属于价值中立。毋庸置疑,人工智能从诞生之日起,对于帮助人类文明的发展与进步,促进和改善人类的生产生活发挥了重要作用。当前,以 ChatGPT 为代表的生成式人工智能被泛化为一种形而上的追求,科学技术的作用被无限扩大,通过深度学习的人工智能被赋予人的目的,人工智能给人类未来发展带来的危机与迷失与日俱增。人是具有追求普遍性和超越自然的存在,人的存在是衡量生成式人工智能的价值尺度。ChatGPT 不是超越人的存在,而是依赖于与人的互动与共生。审视生成式人工智能的发展及其应用,根本价值在于如何依据历史唯物主义的基本原则,正确重塑智能时代认识世界的基本逻辑,历史唯物主义的方法不仅为我们正确认识 ChatGPT 提供了正确的世界观和方法论,而且为正确认识智能技术重塑人类提供了基本的伦理标准。人是定义 ChatGPT 的主体,应当坚持普遍联系的全面系统的发展变化的思维观念,尊崇人类的习俗、道德、规范和法律,重构生成式人工智能的工具属性。

当人们开始用理性的视角看待世界时,宗教的神秘外衣被扯下,离奇的自然现象被解释,人不

再是世界的一部分或世界的产物,而是成为以自我量度世界、征服世界的独立主体,人们从对自然神秘力量的崇拜,走向了对现实世界的“主宰”^[39]。深蓝、AlphaGo 以及 ChatGPT 的成功让人们看到人工智能战胜人类,甚至超越人类的可能,尽管下棋比代步仿佛需要更多的智慧因素,但比人能“算”和比人能“跑”都是对人能力的延伸,具有不同的表象,但最终都不能替代人类这个最具灵性的有生命的复杂的物种。因此,消解人工智能对人思想与行为的束缚,必须强化其工具属性。

1. 人工智能不过是人类的“新工具”

人类生产人工智能,试图制造这种硅基生命的最终目的是解放人的生产力,将人从繁杂的工作中释放出来,实现人的自由全面发展,尽管生成式人工智能在某种程度上能够像人一样辅助人类决策,甚至试图融入人类交往,混淆传统的人机交互,但人工智能的工具本质没有变,也不能改变。

2. “新工具”与众不同

人工智能不同于人类历史上的任何一个发明创造,其所具备的语言模型、思维模型伴随着大数据、算法的加持在某种程度上已经突破了图灵测试。正如马尔库塞所言,“技术拜物教”正在到处蔓延,“技术的解放力量——使事物工具化——转而成为解放的桎梏,即使人也工具化了”^[40],人工智能所表现出的类人的特性,使其工具性的运用不能忽视价值嵌入。要加强人工智能知识的教育与普及,掌握人工智能的运行规律,更要清晰认识到人工智能的技术弊端,认识其可能产生的风险与危机,构建防范机制,最大程度化解人工智能技术快速发展与人们认知不足的矛盾,消解智能崇拜和智能依赖。

3. 用好人工智能“新工具”

如今人工智能的发展正如海德格尔所说,现代技术已不再是“中性的”,它作为“座驾”控制和支配着现代人的全部生活,或者说已经成为现代人的历史命运^[41]。科技进步带来社会发展,并不会因为可能的风险就退避三舍,而是更准确地掌握和运用人工智能,发现其优势和缺陷,用好“新工具”为人类社会发展服务。因此,用好人工智能“新工具”就是要利用人工智能的生成性、创造性技术特征,服务人的创造性活动,彰显人的主体性价值。

(二) 审视生成式人工智能的价值与限度

从本质上说,ChatGPT 仅是初级阶段的人类智能。语言是人类的高级智能活动,不仅涉及语言内部的结构,还涉及语言外部的日常生活知识、科学技术知识、历史知识、社会知识、文化背景知识、人们的情感愿望、人们的心理状态等丰富多彩的因素^[42]。生成式人工智能的价值与限度,为人工智能与人类共处设置边界。每一次人工智能发展的迭代都伴随着对人工智能伦理的探讨,从深蓝到 AlphaGo 再到生成式人工智能 ChatGPT,无一例外地引起人们对人工智能偏见、隐私、情感等伦理问题的讨论,并针对可能产生的伦理风险制定道德规范。因此,面对生成式人工智能的发展,我们必须审视其价值与限度。

1. 回归人类自我价值审视

生成式人工智能是基于大量数据库和语料库的生成系统,数据是生成式人工智能发展的基础和前提条件,大量的带有价值立场和价值判断的数据资源是生成式人工智能的训练场,这样的训练并非静态,而是动态,并非单向,而是交互。因此,当人类谈及人工智能可能具备的贪婪、不公平、歧视、剥削等问题,不能忽略人类自身的傲慢与自私,忽略人与自然、人与社会以及人与人的不和谐因素。用户交互是生成式人工智能学习和训练的现实场景,ChatGPT 可能会语出惊人,但其所具备的

“智慧”一定来自人类社会的多元数据,是对人类社会复杂信息的掌握和提取。人工智能大踏步发展,带来人的异化风险,消解这种异化需要从人类自身出发,回归现实的人,反思人类思维与人工智能思维的差异,承认人的有限性,以人的自我消解推动人工智能发展中人的异化危机的消解。

2. 实现人工智能自我价值塑造

尽管人工智能的道德表达来自对人类社会数据的学习,但人工智能自身作为机器学习的主体,具备对数据筛选和决策的能力,不能忽视对其主体价值的塑造。在人工智能自我价值塑造中主要存在两个方面的治理问题:一方面是软治理,即通过构建用户参与反应机制、制定监督和审查机制、加强宣传强调社会责任等方式,强调制度规范和主动责任担当,意在构建一套较为全面闭环的人工智能伦理制度规范体系;另一方面是硬治理,即通过技术方式将道德准则和法律规范嵌入人工智能,使其具备内在自我道德能力。但无论是软治理还是硬治理,其内在目的都是构建以善和美为内在行动准则的人工智能价值主体。

3. 重视人工智能的技术发展限度

人工智能的发展受技术发展的制约,人工智能的风险问题解决依然受技术发展制约,数据安全、解释性、透明性、道德嵌入等技术难题为人工智能的广泛应用提出了挑战。因此,大力发展人工智能技术的前提是完善人工智能技术的道德架构,以道德架构约束规范人工智能技术进步,为人工智能发展把好脉,找准方向。

(三) 生成式人工智能发展的应然归旨

人工智能终究要向何处去?是否承认人工智能或生成式人工智能未来具备人的意识?是否赋予其道德主体地位?人工智能是否具备人的类本质同等地位?如何消解以 ChatGPT 为代表的生成式人工智能引发的人的异化危机?答案是走一条人机和谐之路。人机和谐之路从本质上讲就是要彻底地改变传统的人机观念,重构一对全新的、富有生命力的人机关系。传统的人机关系包含着一种“我-它”的关系,蕴含着“主-奴”的属性,这种属性从侧面反映人类的主体地位和对机器的态度。重建人类对人工智能的正确认识,探寻人类需要什么样的人机交互的契合点,需重构新的人机关系,呼唤人的主体性觉醒,确保人机和谐,即构建可控、安全和可靠的人工智能发展之路。实际上,深蓝、AlphaGo、ChatGPT 的诞生和成功,无论是传统人工智能还是生成式人工智能都是人类智慧的结晶,其发展的终极形态也有赖于人类智慧的发展,有赖于人类实践的发展。

1. 构建可控的人工智能

人工智能的快速发展在许多领域取得了显著的成果,包括自动驾驶、机器翻译和医疗诊断等。然而,随着人工智能系统的复杂性增加,其行为变得越来越难以预测和解释,从而引发了社会对人工智能可控性的关注。首先,人工智能系统必须是可预测的。人工智能运行的底层逻辑是数据的输入和输出,面对实时变化的输入数据时可以准确预测系统的输出,提升预测性可使用户和开发者对系统的输出结果做出合理的研判。其次,人工智能系统必须是可解释的。人工智能以其巨量的数据和检索速度被用于辅助人类决策,人工智能发展的伦理性特征要求我们必须提出关于系统决策和推理过程的解释,增加系统透明度,使用户和监管机构理解系统的工作原理,并评估其决策的公正性、偏见和风险。最后,人工智能必须是可管理的。人工智能发展必须在人类的管理范畴之内,通过监控和调整系统的行为、性能和输出,使得用户和开发者能够实时追踪和管理系统的运行状态,并采取必要的措施来纠正偏差或改进性能。

2. 构建安全的人工智能

安全的人工智能通常指在设计、开发、部署和运行过程中,具备保护人工智能系统和相关数据免受恶意攻击、滥用和未经授权者访问的能力。构建安全的人工智能,一方面是构建安全的人工智能“内核”。人工智能归根结底是人类设计的代码程序,构建安全的算法设计,保障数据隐私安全是确保人工智能“内核”安全的关键。在技术方面通过差分隐私、安全多方计算、访问控制和权限管理等技术手段完善人工智能系统的算法设计,或采用加密技术、去标识化和数据匿名化等手段,确保用户数据的隐私安全,防止数据泄露和滥用,增强人工智能系统的稳定性和透明性。另一方面是构建安全的人工智能“外壳”,需建立完善的权限管理、访问控制、身份认证等措施,构建防范系统滥用和非法访问的安全控制机制,实施漏洞扫描、安全审计、事件监测和应急响应机制等风险管理手段。

3. 构建可靠的人工智能

可靠的人工智能通常指在各种情况下能够产生一致、准确和可信结果的人工智能系统。人工智能系统必须具有一致性,可靠的人工智能在相似的输入条件下应该能够产生一致的输出结果,即使在不同的运行时刻或使用不同的数据集,系统也应该能够保持一致性。人工智能系统必须具有准确性,可靠的人工智能系统需保障任务完成的准确性和精确性。人工智能系统必须具有可信度,可靠的人工智能系统拥有用户对其结果和决策的信任,需能够提供解释其决策过程的能力。

总之,生成式人工智能不仅给人类解放自己的重复劳动创造了条件,也给人的异化带来新的危机。马克思关于人的本质的认识,尤其是关于人的交往、人的思维,以及人的劳动的认识,不仅为我们正确认识智能时代人的异化问题提供了理论工具,也为寻找人的自我解放和发展指明了方向。增强人的主体能动性积极投身社会实践、正确认识人工智能的价值与限度,构建可控安全可靠的人工智能是生成式人工智能发展的应然归旨。

参考文献:

- [1] 蒋华林. 人工智能聊天机器人对科研成果与人才评价的影响研究——基于 ChatGPT、Microsoft Bing 视角分析[J]. 重庆大学学报(社会科学版), 2023(2): 97-110.
- [2] 于水, 范德志. 新一代人工智能(ChatGPT)的主要特征、社会风险及其治理路径[J/OL]. 大连理工大学学报(社会科学版): 1-7. [2023-07-04]. DOI:10.19525/j.issn1008-407x.2023.05.004
- [3] 郑世林, 姚守宇, 王春峰. ChatGPT 新一代人工智能技术发展的经济和社会影响[J]. 产业经济评论, 2023(3): 5-21.
- [4] 黄荣, 吕尚彬. ChatGPT: 本体、影响及趋势[J]. 当代传播, 2023(2): 33-38, 44.
- [5] 孙伟平. 人机之间的工作竞争: 挑战与出路——从风靡全球的 ChatGPT 谈起[J]. 思想理论教育, 2023(3): 41-47.
- [6] 邓建鹏, 朱恽成. ChatGPT 模型的法律风险及应对之策[J]. 新疆师范大学学报(哲学社会科学版), 2023(5): 91-101, 2.
- [7] 肖峰. 生成式人工智能介入知识生产的功用探析——借助 ChatGPT 和“文心一言”探究数字劳动的体验[J]. 重庆邮电大学学报(社会科学版), 2023(4): 1-10.
- [8] 焦建利. ChatGPT 助推学校教育数字化转型——人工智能时代学什么与怎么教[J]. 中国远程教育, 2023(4): 16-23.
- [9] 王建磊, 曹卉萌. ChatGPT 的传播特质、逻辑、范式[J]. 深圳大学学报(人文社会科学版), 2023(2): 144-152.
- [10] 张夏恒. ChatGPT 的逻辑解构、影响研判及政策建议[J]. 新疆师范大学学报(哲学社会科学版), 2023(5): 113-123.
- [11] 谢新水. 人工智能内容生产: 功能张力、发展趋势及监管策略——以 ChatGPT 为分析起点[J]. 电子政务, 2023(4): 25-35.
- [12] 钟秉林, 尚俊杰, 王建华, 等. ChatGPT 对教育的挑战(笔谈)[J]. 重庆高教研究, 2023(3): 3-25.
- [13] 王树义, 张庆薇. ChatGPT 给科研工作者带来的机遇与挑战[J]. 图书馆论坛, 2023(3): 109-118.

- [14] 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): 81-90.
- [15] 马克思, 恩格斯. 马克思恩格斯文集(第1卷)[M]. 北京: 人民出版社, 2009: 501.
- [16] 马克思, 恩格斯. 马克思恩格斯文集(第4卷)[M]. 北京: 人民出版社, 2009: 501.
- [17] 孙伟平. 人类交往实践的革命性变迁——虚拟交往及其哲学批判[J]. 吉林大学社会科学学报, 2012(3): 13-19.
- [18] 马克思. 1844年经济学哲学手稿[M]. 北京: 人民出版社, 2000: 86.
- [19] 李芒, 杨宇轩. 人非机器: 对计算思维本质的认识[J]. 开放教育研究, 2023(2): 55-60.
- [20] 马克思, 恩格斯. 马克思恩格斯选集(第1卷)[M]. 北京: 人民出版社, 2009: 530.
- [21] 汪荣有. 人的本质是劳动与社会性的统一[J]. 江西社会科学, 1999(7): 9-11.
- [22] 陈荣荣, 张士清. 马克思扬弃“人的异化”的双重道路[J]. 社会科学辑刊, 2014(1): 39-42.
- [23] 马克思, 恩格斯. 马克思恩格斯文集(第2卷)[M]. 北京: 人民出版社, 2009: 580.
- [24] 赵瑞华. 媒介化生存与人的异化[J]. 新闻记者, 2010(2): 29-32.
- [25] 蓝江. 生成式人工智能与人文社会科学的历史使命——从 ChatGPT 智能革命谈起[J]. 思想理论教育, 2023(4): 12-18.
- [26] 迈克尔·海姆. 从界面到网络空间: 虚拟实在的形而上学——哲人石丛书[M]. 上海: 上海科技教育出版社, 2000: 147.
- [27] BAUDRILLARD J. The ecstasy of communication[M]. New York: Semiotext(e), 1998.
- [28] 钟祥铭, 方兴东. 智能鸿沟: 数字鸿沟范式转变[J]. 现代传播(中国传媒大学学报), 2022(4): 133-142.
- [29] DWIVEDI Y K, HUGHES L, ISMAGILOVA E, *et al.* Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy[J]. International Journal of Information Management, 2021, 57: 1-47.
- [30] 张灿, 郭海静. 数字时代数字技术崇拜的机制与实质[J]. 重庆邮电大学学报(社会科学版), 2023(2): 93-102.
- [31] 尼葛洛庞蒂. 数字化生存[M]. 胡永, 范海燕, 译. 海口: 海南出版社, 1997: 15.
- [32] 蓝江. 数字资本、一般数据与数字异化——数字资本的政治经济学批判导引[J]. 华中科技大学学报(社会科学版), 2018(4): 37-44.
- [33] 霍克海默. 批判理论[M]. 李小兵, 译. 重庆: 重庆出版社, 1989: 242.
- [34] 黄荣, 吕尚彬. ChatGPT: 本体、影响及趋势[J]. 当代传播, 2023(2): 33-38, 44.
- [35] 孙伟平. 人机之间的工作竞争: 挑战与出路——从风靡全球的 ChatGPT 谈起[J]. 思想理论教育, 2023(3): 41-47.
- [36] 马克思. 1844年经济学哲学手稿[M]. 北京: 人民出版社, 1985: 48.
- [37] 李锐. 拒绝合唱[M]. 上海: 上海人民出版社, 1996: 169.
- [38] 蒲清平, 向往. 生成式人工智能——ChatGPT 的变革影响、风险挑战及应对策略[J]. 重庆大学学报(社会科学版), 2023(3): 102-114.
- [39] 王泽应. 祛魅的意义与危机——马克斯·韦伯祛魅观及其影响探论[J]. 湖南社会科学, 2009(4): 1-8.
- [40] 马尔库塞. 单向度的人[M]. 刘继, 译. 上海: 上海译文出版社, 1989: 143.
- [41] 黄欣荣. 现代西方技术哲学[M]. 南昌: 江西人民出版社, 2011: 157, 88, 124-125.
- [42] 冯志伟, 张灯柯, 饶高琦. 从图灵测试到 ChatGPT——人机对话的里程碑及启示[J]. 语言战略研究, 2023(2): 20-24.

ChatGPT: Human alienation crisis and reflection caused by generative artificial intelligence

XIE Jun, LIU Ruilin

(School of Marxism, Chongqing University of Posts and Telecommunications, Chongqing 400060, P. R. China)

Abstract: As a typical representative of generative artificial intelligence, Chat Generative Pre-trained

Transformer (ChatGPT) is a natural language processing and machine learning developed model, which is launched by OpenAI, an American artificial intelligence research laboratory. Hardly had ChatGPT swept the world and generated a profound discussion around the world when it first appeared. Generative AI, represented by ChatGPT, as the latest achievement in AI development, is a significant step further toward, aiming the goal of designing sentient and general artificial intelligence. With the development of generative AI and its extensive application, it is believed to be reshaping human beings and human society, and humans will suffer unprecedented challenges while enjoying the convenience brought by the new technology, prominently manifested in the alienation of human beings by generative AI, reshaping human interactions, human thinking, and human labor. The study of human alienation in the information age requires an in-depth analysis of the phenomenon of human alienation with a historical materialist approach, starting from Marx's prescriptions on the nature of human beings, and finding a breakthrough to deal with the crisis caused by artificial intelligence. Human nature has a threefold connotation: first, interaction is a unique manifestation of human nature, second, thinking is a realistic expression of human nature, and third, labor is an intrinsic provision of human nature. With the rapid development of cognitive technology and neuroscience, generative artificial intelligence, with its unique simulation ability, exhibits a corroboration of the alienation of human nature as described by Marx, that is, human creations threaten the development of human subjectivity, including: the alienation of human interaction triggered by human-machine interaction, data-based interaction, and digital existence; human-machine combination, reliance on intelligence, and machine production bring about "transcendent" human beings, "weakened" human beings, and "void" human beings; the alienation of human and object, human labor, etc. In the face of the crisis of human alienation caused by generative AI, there is an urgent need for the return and revival of value rationality in the age of intelligence, an urgent need to correctly understand the value and limits of generative AI, and an urgent need to pay attention to the ultimate care for human life. Therefore, to deepen the reflection on human alienation caused by generative AI, one is to reconstruct the instrumental properties of generative AI, that is, to recognize that AI is only a human tool, to recognize the uniqueness of the tool, and to make good use of artificial intelligence as a tool. The second is to examine the value and limits of generative AI, that is, we must return to human self-value, realize the self-value of AI, and pay attention to the limit of AI development. The third is the natural purpose of generative AI development, that is, to build a controllable, safe and reliable AI.

Key words: ChatGPT; generative artificial intelligence; Marx's theory of human nature; human alienation; value reflection

(责任编辑 彭建国)