

未来人工智能伦理问题探究

——基于机器人、元宇宙视角

何静凡, 李蓓蓓

南京信息工程大学法政学院, 科技史与气象文明研究院, 江苏 南京

收稿日期: 2023年1月10日; 录用日期: 2023年1月30日; 发布日期: 2023年2月13日

摘要

人工智能的发展, 赋予了机器人新的智能空间, 而由元宇宙实现的虚拟的3D世界则给人工智能和机器人带来了更多可能。机器人和元宇宙在给人类带来便捷的同时, 也产生了机器与人在元宇宙中新的伦理关系。在人与机器人机交互的同一性关系中, 人类依然需要承担主体责任。

关键词

机器人, 元宇宙, 机器人伦理, 人工智能

Research on Ethical Issues of Artificial Intelligence in the Future

—Based on the Perspective of Robot and Metaverse

Jingfan He, Beibei Li

Institute of History of Science and Technology and Meteorological Civilization, School of Law and Politics,
Nanjing University of Information Science and Technology, Nanjing Jiangsu

Received: Jan. 10th, 2023; accepted: Jan. 30th, 2023; published: Feb. 13th, 2023

Abstract

The development of artificial intelligence has given robots a new intelligent space, and the virtual 3D world realized by the metaverse has brought more possibilities to artificial intelligence and robots. While robots and the metaverse bring convenience to human beings, they also create a new ethical relationship between machines and humans in the metaverse. In the identity relationship between humans and robots, humans still need to assume the main responsibility.

Keywords

Robot, Metaverse, Robot Ethics, Artificial Intelligence

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

科学技术的发展在社会发展中起着重要的作用, 现代计算机技术的发展改变了时代的面貌和人们的日常生活。人工智能已经成为当代计算机技术的变革前沿, 深刻融入到各行各业赋能增值, 从国家部委的实验室改革到江苏省科技十四五规划中都提出了人工智能相关领域。元宇宙可以追溯至 20 世纪 90 年代, 在疫情期间, 虚拟化的人类社会到达了阈值, 因此 2021 年也自然被定义为元宇宙初年。人工智能的技术对元宇宙集成来说具有重要意义, 并且在元宇宙中机器人将以虚拟化的形式存在, 这必然带来关于人与机器的关系、人的本质、机器及其算法情感道德等新问题的新讨论。

2. 机器人视角下的科学技术伦理问题

一般意义来讲, 科学是重要的生产力, 技术是实现生产力的重要途经。有关科学的概念, 伟大的思想家马克思表示: 科学属于人类经验的范畴, 自然的科学一般都包含人类的科学[1]。归纳起来, 科学是特殊的人文。关于技术, 马克思、恩格斯认为在人与自然的具体实践中, 技术发挥着重要的作用, 技术是从人类需求出发的, 具体而言是在实践中总结具体的实践经验, 总结其实践中的科学原理, 根据这些创造发明、手段、方式、方法等相加而成。科学与技术是具有相互作用的[1], 并且基础研究在科学体系与技术发展中发挥着重要作用。

由于科学研究的发展, 科技研究在结构与功能上已发生了很大的变化, 不再是传统意义上探索宇宙的奥秘作为唯一目的, 其中也夹杂很多其它的需求, 也包括科学研究组织者自身。研究过程中研究者的崭新的思辨理念可能会与当时的法律及道德观念相冲突, 并且科技活动并非站在价值这个杠杆的中间点, 其中也需要有普遍的道德准则与规范等的约束[2]。如此所产生的一个有争议的话题, 便是科技伦理。具体而言, 一个重要的例子是机器人自身的科技伦理, 在机器人伦理世界中, 需要进行关注的是机器人在无自主性的情况下, 与之相关人类的责任, 在具体的伦理道德规则面前, 机器人需要遵守什么? 随着机器人自动化程度的不断加深, 其独立性也不断的增强, 在这些因素的前提下, 又延伸出来一个问题, 作为机器人而言能否像人类实践活动一样呢? 在行为过程中能否具有某种道德能动性, 或者说机器人能否成为道德能动者或伦理能动者[3]。

机器人伦理随着机器人类型多样化, 在日常生活、医疗保健、军事战争等产生多方面影响, 可能性伦理风险可主要概括为两类, 一类是物理性风险, 另一类是非物理性风险[4], 因此对机器人的权利也需要进行限制[5]。近年来人与机器人进行交互时所产生的伦理问题得到了国内外学者的广泛关注, 国外学者英国计算机专家诺埃尔·夏基教授就曾呼吁人类应该尽快制定机器(人)相关方面的道德伦理准则[6]。尤其是在具体应用中对算法推荐伦理失范动因以及人机共处方式进行追问与探讨, 以及人机的相处模式与关系走向[7]。近年来服务型机器人发展迅速, 这种机器人的人机交互在法律、道德、情感方面的风险也十分典型, 例如儿童看护机器人、助老机器人的情感依赖分享, 助理机器人、演艺机器人等社会机器人

法律风险, AI 绘画机器人的艺术版权问题等[8] [9] [10] [11] [12]。

3. 人工智能下的机器人

机器的产生与发展与人类认识和改造世界是相互促进的, 恩格斯指出, “数学: 辩证的辅助手段和表达方式” [13]。机器的出现逐渐可以取代的是一些更为复杂的工作、计算等, 同时人工智能带来的复杂性和学习性思维能够在大数据和云计算处理时更具优越性, 为人类节约成本并提供更高效工作。但是机器人本身是并不具备承担责任的能力, 作为机器人伦理的主体, 在人工智能的迅猛发展中, 机器人演变的更加智能化、拟“人”化, 机器是否需要承担与人一样的责任? 机器对人类使用者甚至设计者, 以及其他机器产生和带来的一系列行为后果, 是否与开发和人类无关? 这一解答需要关注人与机器人的区别。

透过世界的发展, 人与世界存在着什么关系? 对于这个问题的回答, 一般而言, 人与世界的关系其中具有自由。对自由的人本身的定义, 更多成为一个聆听的, 欣赏这个世界的, 并非常见的世界的创造者、利用者以及消费者[14]。机器二字在古时候的解释是分开而行的。18 世纪以来的机械论哲学试图以“普遍数学”原理在理论上阐述“人是机器” [15]。

科学技术革命反映的是人类认识和改造客观世界的飞跃。18 世纪的第一次科技革命为我们带来了大机器生产, 19 世纪第二次科技革命则带我们走进了电气时代, 20 世纪中期至今我们享受着信息时代的发展, 当下智能化时代逐步占据了生活的很多方面, 出现了个人(家庭)服务机器人及商用服务机器人。随着机器人的发展, 关于机器定义和选择范围更加广阔, 机器人所做的工作也愈来愈复杂化, 在这中间机器人发展的很突出的一个特点是机器人更像人了, 也就是仿人机器人。从外观上看, 机器人的发展逐步“类人化”更接近于人类的身体, 例如关于放任机器人中的一些复杂设计中它具有与仿人机器人实体相同的自由度数目和关节结构, 骨骼长度也和仿人机器人实体对应部位几何长度相等[16]。从其使用的机能上, 机器人与人的身体有很大的不同, 这种不同主要表现在意识形态上, 人如何产生人的意识, 而机器人的意识是什么? 人脑在意识中发挥者重要作用, 是意识的机能, 机器人是否也一样? 人类具有一定的欲望和不同的价值观念, 这二者也成为了发生冲突的矛盾点, 而这些存在的基础便是自我意识。

那么机器人真的可以有意识吗? 他的意识的来源和人类一样吗? “机器人三原则”是由美国作家艾萨克·阿西莫夫提出的, 在发展中关于机器人学(Robotics)的概念也首次出现。其中第二定律是机器人必须服从人类命令, 这一命令的例外是与第一法则发生冲突, 这一定律已经表明从一开始机器人是不能有独立的意志的。美国哲学家普特南通过一系列的例子也未能表明机器人都具有意识, 而让机器人有意识是基于“合成有机身体部分的柔软性或硬度”的歧视[17]。此外还有机器人自身的道德准则。机器人到人类、从思维到交互、从智能思维到社会情感存在、从现实到表象、从正确到善、从外部标准要好, 内在要实践, 要从理论到经验和想象[18], 实现如此的一个复杂美好的存在。在机器人与人的交互中, 不仅是物理上的一致性, 心灵主要体现在情感交互, 而这一交互的过程也不妨碍机器人是否具有意识, 其中所主要衡量的还是法则的牵制, 牵制的主体依旧是人。

人工智能(Artificial Intelligence, AI)这一概念是在 1956 年达特茅斯会议首次提出的。人工智能是多种学科交叉的产物[19]。人工智能是通过对数据的不断收集, 对信息的提取的基础上, 通过技术的发展不断对自己改进的一种系统或者机器[20]。随着人工智能的发展, 也产生了很多问题。人工智能与哲学的关系中, 人工智能在根基的意义上具有认识论的性质。心智哲学家丹尼特讲述“人工智能在很大程度上就是哲学” [21]。其发展过程是机器不仅能模仿人的思维, 也在人工智能的发展中起着“查漏补缺”的功能, 人工智能在发展中也不可能完全等同于人类智能, 而是对于人类智能中所需要的某些功能和属性的增强以及补充延伸功能。人工智能的潜在风险问题有多个方面, 除了重要的安全问题外, 还存在着风险的时

代性和风险的全球性[10]。人工智能的哲学问题已不再是人工智能的本质问题,而是关于人的意向性、概念框架等问题,人工智能只不过是人类智能的特殊的实现方式[22]。同时人工智能也面临着不确定性推理等,不仅是推理的本身不确定性,也有对于计算机所不能应付的环境的变化,不能承受例如内部元件损坏所产生的影响,但人脑在这方面却显示着优越性[19],在智能计算的理论体系中,人类的智慧是智能的源泉,计算机是人类智能的赋能。这些都从侧面表明人的主体地位,即人工智能本身并没有创造,只是对工具的更新使用。

4. 人工智能下的元宇宙

元宇宙技术是由多个技术组合等并发而成的交叉应用组合系列。关于元宇宙的基本元素主要是“协作关系、生产资料、技术体系、经济体系”[20]。是一种融合多种新技术的新型互联网应用和社交形态。它具有多技术性、社会性和超时空性等特点[23]。2021年10月29日脸书(Facebook)创办人扎克伯格(Zuckerberg)宣布将 Facebook 改名为 Meta,这一更名一时成为全球热议的一个话题。元宇宙离我们很近,2022年6月北京大学新媒体研究院举办了元宇宙毕业典礼,2022年9月南京信息工程大学成立了元宇宙工程系。

元宇宙,是一个3D的虚拟的世界,在元宇宙中的并不是我们现实的人,而是一个虚拟的世界,在这个世界可以工作、生活、娱乐,拥有不同身份,选择另一种生活,享受平行宇宙的不同体验。这个世界将一直存在,来去自由。现阶段,人工智能在元宇宙的应用体现在数字孪生、数字虚拟人、个性化数字化身和无障碍沟通这四个方面之中[20]。在元宇宙中人工智能可以代替人的本体进行上述活动,例如我们可以在现实世界里面睡觉,而人工智能分身可以帮我们工作,交朋友,甚至约会。如此虚拟和现实之间的碰撞,在元宇宙中也会产生很多风险,因此元宇宙将是对人类“人性”的巨大考验。

在人工智能的时代,人在元宇宙的发展中占据核心和主导作用。再具体一点,元宇宙的机器人相对目前现实世界的机器人是怎样?可能更具“人类”感。不仅体现在物理操作上面,其中也体现在一些具体的意识操作层面。在元宇宙的世界中,人们或许会在寻求满足自我需要的过程中,发生冲突,进而可能导致社会处于混乱状态,并且可能会出现侵权行为等[24],这些都是需要去考虑的。当然每一次的科技革命都会引发社会变迁、权利变动和伦理革新,元宇宙的出现也不例外,其中主要有数据伦理冲击,在感知的数字化世界中的生命与神经伦理挑战,不仅如此还会产生虚拟现实之间的难以平衡的界限,会产生虚假伪造等社会认知伦理和精神伦理层面的颠覆性影响等[25]。归纳起来其主要面临的挑战问题有如何明确基本框架、元宇宙高度垄断、虚实世界的关系、隐私和数据安全等[26]。元宇宙由于底层架构技术——区块链技术的应用和发展,其中会产生一大问题便是会与用户数据隐私保护之间存在一定的冲突,加之用户虽然在元宇宙中可以自由选择性别、年龄、种族、社会阶层等等,但这实际上可能进一步加剧元宇宙中的弱势群体的处境。更重要的一点,元宇宙可能导致用户个体意识被操纵,进而影响人类自主性。关于元宇宙行为承担者其追溯的根源是自然人本身的行为,所产生的道德甚至法律后果也都是自然人所承担的[24]。

5. 结语

科学技术是社会发展的巨大动力,也是人类文明的重要标志,当前人工智能的进步发展也带来很多困难和挑战,主要体现在政治、社会和经济等方面。人工智能背景下尤其是机器人的发展让人在享受生活中处处可见智能便利,实时感受科技的进步,但同时也在人机交互中感受到了机器人逐渐“人”化的威胁,其中便存在着科技伦理这一问题。元宇宙让人感受到了现实与虚拟相互融合的世界,元宇宙也存在着隐私、数据安全等问题,但在元宇宙的主体中制定相关规则和切实承担责任的还是应该是自然人。

理性还是能否与人的心灵和谐共存, 机器人是否具有人的心灵? 至少在阿西莫夫的机器人三定律看来, 机器人在虚拟的世界中依旧是人类所主导, 人类必须承担起主体的责任。

参考文献

- [1] 郭贵春, 等. 自然辩证法概论[M]. 北京: 高等教育出版社, 2018: 43-95.
- [2] 甘绍平. 科技伦理: 一个有争议的课题[J]. 哲学动态, 2000(10): 5-8.
- [3] 段伟文. 机器人伦理的进路及其内涵[J]. 科学与社会, 2015, 5(2): 35-45+54.
- [4] 雷瑞鹏, 张毅. 机器人学科技伦理治理问题探讨[J]. 自然辩证法研究, 2022, 38(4): 108-114.
- [5] 杜严勇. 论机器人权利[J]. 哲学动态, 2015(8): 83-89.
- [6] 刘伟, 赵路. 对人工智能若干伦理问题的思考[J]. 科学与社会, 2018, 8(1): 40-48.
- [7] 周颖玉, 柯平, 刘海鸥. 面向算法推荐伦理失范的人机和谐生态建构研究[J]. 情报理论与实践, 2022, 45(10): 54-61.
- [8] 黎常, 金杨华. 科技伦理视角下的人工智能研究[J]. 科研管理, 2021, 42(8): 9-16.
- [9] 阎国华, 闫晨. 家庭泛用型智能机器人的安全风险研究[J]. 东北大学学报(社会科学版), 2018, 20(3): 227-232.
- [10] 吴汉东. 人工智能时代的制度安排与法律规制[J]. 法律科学(西北政法大学学报), 2017, 35(5): 128-136.
- [11] 杨通进. 论机器人的道德承受体地位及其规范意涵[J]. 哲学分析, 2019, 10(6): 14-33+191.
- [12] 杜严勇. 现代军用机器人的伦理困境[J]. 伦理学研究, 2014(5): 98-102.
- [13] 韦建桦. 马克思恩格斯文集(第九卷) [M]. 北京: 人民出版社, 2009: 401.
- [14] 吴国盛. 科学与人文[J]. 中国社会科学, 2001(4): 4-15+203.
- [15] 尚智丛, 闫奎铭. “人与机器”的哲学认识及面向大数据技术的思考[J]. 自然辩证法研究, 2016, 32(2): 24-28.
- [16] 张利格, 毕树生, 高金磊. 仿人机器人复杂动作设计中人体运动数据提取及分析方法[J]. 自动化学报, 2010, 36(1): 107-112.
- [17] Putman, H. and Putnam, H. (1964) Robots: Machines or Artificially Created Life? *The Journal of Philosophy*, **61**, 668-691. <https://doi.org/10.2307/2023045>
- [18] Coeckelbergh, M. (2009) Personal Robots, Appearance, and Human Good: A Methodological Reflection on Roboethics. *International Journal of Social Robotics*, **1**, 217-221. <https://doi.org/10.1007/s12369-009-0026-2>
- [19] 胡显章等. 科学技术概论[M]. 第二版. 北京: 高等教育出版社, 2006: 162-167.
- [20] 武强, 季雪庭, 吕琳媛. 元宇宙中的人工智能技术与应用[J]. 智能科学与技术学报, 2022, 4(3): 324-334.
- [21] 肖峰. 人工智能就是认识论[J]. 云南社会科学, 2021(5): 12-20.
- [22] 郑祥福. 人工智能的四大哲学问题[J]. 科学技术与辩证法, 2005(5): 34-37.
- [23] Ning, H., Wang, H., Lin, Y., et al. (2021) A Survey on Metaverse: The State-of-the-Art, Technologies, Applications, and Challenges.
- [24] 鲁照旺. 元宇宙的秩序和规则[J]. 学术界, 2022(2): 65-79.
- [25] 段伟文. 探寻元宇宙治理的价值锚点——基于技术与伦理关系视角的考察[J]. 国家治理, 2022(2): 33-39.
- [26] 方巍, 伏宇翔. 元宇宙: 概念、技术及应用研究综述[J/OL]. 南京信息工程大学学报(自然科学版), 1-25. <http://kns.cnki.net/kcms/detail/32.1801.N.20221207.1946.001.html>, 2022-12-27.