

人工智能与人类关系伦理问题

——以ChatGPT为例

王玉涵

中共江苏省委党校, 江苏 南京

收稿日期: 2023年12月18日; 录用日期: 2024年1月18日; 发布日期: 2024年1月26日

摘要

党的十八大以来, 发展数字经济已经被纳入我国重点战略规划范畴中。党的二十大报告提出以中国式现代化推进中华民族伟大复兴的任务, 同时坚持将发展经济领域事业的着力点放在实体经济方面, 加快建设网络强国和数字中国的节奏。本课题旨在以ChatGPT为例, 引出当代人工智能与人类关系的伦理问题思考, 首先对人工智能伦理学的起源与发展做出了详细概述, 接着对ChatGPT做出了简要介绍, 并探讨了ChatGPT发展路径, 分析了ChatGPT存在的伦理问题与解决方法, 最后立足于ChatGPT的视角对人机协调——未来人工智能的新型发展之路做出了深度研析, 希望能够为大家提供一定的借鉴与帮助。

关键词

ChatGPT, 马克思主义, 人工智能, 技术革命, 伦理学

Ethical Issues in the Relationship between Artificial Intelligence and Human Beings

—Taking ChatGPT as an Example

Yuhan Wang

Jiangsu Provincial Party School of the Communist Party of China, Nanjing Jiangsu

Received: Dec. 18th, 2023; accepted: Jan. 18th, 2024; published: Jan. 26th, 2024

Abstract

Since the 18th National Congress of the Communist Party of China, the development of digital economy has been included in the scope of China's key strategic planning. The 20th National Congress of the Communist Party of China report put forward the task of promoting the great rejuvenation of the Chinese nation. This paper aims to take ChatGPT as an example to explore the ethical issues in the relationship between artificial intelligence and human beings. It first provides a detailed overview of the origin and development of artificial intelligence ethics, then introduces ChatGPT briefly, and discusses the development path of ChatGPT, analyzes the ethical issues and solutions of ChatGPT, and finally, from the perspective of ChatGPT, it makes a deep analysis of the new development path of artificial intelligence coordination in the future, hoping to provide certain reference and help for everyone.

文章引用: 王玉涵. 人工智能与人类关系伦理问题[J]. 哲学进展, 2024, 13(1): 237-243.

DOI: 10.12677/acpp.2024.131036

nation of the Chinese nation through Chinese-style modernization, while insisting on focusing on the substantial economy in the development of the economic field, and accelerating the pace of building a cyber powerful and digital China. Taking ChatGPT as an example, this topic aims to introduce the ethical issues of the contemporary relationship between artificial intelligence and human beings. Firstly, it gives a detailed overview of the origin and development of artificial intelligence ethics, then makes a brief introduction to ChatGPT, discusses the development path of ChatGPT, and analyzes the existing ethical problems and solutions of ChatGPT. Finally, based on the perspective of ChatGPT, make an in-depth analysis of human-machine coordination—the new development road of artificial intelligence in the future, hoping to provide some reference and help for everyone.

Keywords

ChatGPT, Marxism, Artificial Intelligence, Technological Revolution, Ethics

Copyright © 2024 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

近年来，在创新科技的强力支持与现代化建设的大力推动下，我国社会不断迈向发展的新高度，与此同时，科技领域事业的创新发展也为人工智能技术的不断改进提供了积极助力。实际上人工智能与人类二者存在显著差异，但人工智能与人类之间的关系却又充满了复杂性与奥妙性，二者不仅相互依存，而且还彼此作用。人工智能可为人类提供更为高效、精准的服务，帮助人类解决更为复杂且困难的问题，同时在人工智能的强势助力下，人类的生产力与创造力还会更为卓越，这对于推动人类社会的长效稳定发展具有极其重要的现实意义，为了更好地展现人工智能存在的价值，为社会主义现代化建设保驾护航，以 ChatGPT 为例进而探讨了加强人工智能与人类关系的伦理问题探究势在必行。

2. 人工智能伦理学的起源与发展

人工智能(Artificial Intelligence, AI)伦理学是涉及伦理、技术和社会观念的复杂性学问。其旨在探讨和解决与人工智能相关联的深层次伦理问题，包括算法偏见、隐私、公平性、安全性以及 AI 系统对社会和个体所能造成的影响。AI 伦理学的起源与发展是一个深受科技进步和社会变革影响的历史过程。本节将追溯 AI 伦理学的历史，从早期的 AI 探索直至现如今的全球关注。

(一) 人工智能伦理学的起源

AI 技术的起源可以追溯到 1950 年代初期，当时 Alan Turing、Warren McCulloch、Walter Pitts 以及 Marvin Minsky 等人探讨了人工智能的思想，并设计出了第一台机器学习系统。1966 年，Arthur Samuel 首先提出了“机器学习”这一概念，其开发了 Alpha-Beta 算法，可以根据对游戏的实际模拟来决定接下来要采取的分支。

1970 年，Geoffrey Hinton 提出了“反向传播算法”，其乃当下机器学习中最常用的算法之一。可自动从训练数据中学习出模型，用以智能判断。1980 年，Roger Schank 和 Christopher K. Riesbeck 开发了一种名为“学习精灵”的机器学习系统，其可根据自然语言来学习和理解用户的问题、语境以及概念。

1970 年代到 1980 年代这一期间，随着 AI 技术的进一步发展，人工智能与人类关系的伦理问题也开

始逐渐引起社会各界的广泛关注。人们开始认识到，随着 AI 系统的愈发强大以及 AI 系统的普及化发展，未来必将面临诸多源自伦理和道德方面的挑战。在这一时期，AI 伦理的概念逐渐形成。尤其 Isaac Asimov 的机器人三定律，更是引发了人们对于机器伦理的深思，人们开始考虑机器行为的伦理原则。同时，计算机伦理学、信息伦理学等相关领域也开始崭露头角，研究人员开始探讨数字化时代人工智能与人类关系的伦理问题。

1990 年，Geoffrey Hinton、Yoshua Bengion 和 Yoshua LeCun 就发展了深度学习技术，使其可以模拟人类大脑的神经网络来实施信息处理。2000 年后，随着算法的进一步改进，机器学习技术及其应用也得到了迅猛发展，并此项技术已经应用到机器人、自动驾驶、自然语言处理等领域中，用于提升其智能化水平。

从 1990 年到 2000 年这一期间，计算机和互联网技术在我国得到了广泛普及，在此势态下，AI 伦理问题也变得更加复杂，在这样的时代背景下，伦理问题不再局限于机器人行为这一方面，而是扩展到了数据隐私、信息安全、社交媒体伦理等多个领域。数字时代的到来使人们开始将关注的焦点放在个人数据的收集和隐私问题这一方向。社交媒体平台的普及引发了关于虚假信息、信息操纵和个人数据滥用的伦理问题。随着人工智能在广告、金融、医疗保健等领域的广泛应用，伦理问题所涉及的范围也越来越大。

(二) 人工智能伦理学的发展

进入到二十一世纪之后，AI 技术不断升级，在深度学习、自然语言处理以及机器学习等诸多领域也实现了突破性进展，虽然实现了 AI 系统的强化，但是与此同时也引发了一系列的伦理方面问题。AI 伦理学作为一个独立领域开始崭露头角，吸引了伦理学家、计算机科学家和社会科学家的高度关注。在全球范围内，学术界、政府和行业组织开始重视 AI 伦理方面的问题。许多国家和地区制定了 AI 伦理指导原则，以指导 AI 技术的发展和應用。例如，欧洲联盟发布了《欧洲 AI 伦理准则》，旨在强调人工智能技术的发展需遵循伦理原则。美国也发布了 AI 伦理原则，以使 AI 研究和开发更具道德性和负责性[1]。

国际社会曾举办各种研讨会，以深化推动 AI 伦理研究和探索。之所以在此方面如此下功夫，主要目的就是积极推动 AI 伦理的进一步发展，并寻找解决方案，从而更好地应对 AI 技术可能带来的所有伦理挑战。

通过人工智能伦理学的起源与发展可以充分反映出技术和社会观念二者之间的相互作用。从早期的 AI 探索到如今的全局关注，AI 伦理学已经成为一个重要的领域，其不仅涵盖技术的伦理问题，而且还涉及到人工智能如何影响社会和个体的伦理挑战。

近年来，随着社会的不断发展，时代的日益更迭，AI 技术的应用范围也愈加广阔，将 AI 技术应用于诸多方面，包括医疗诊断、金融投资、安全管理、图像处理、语音识别、股票分析等领域，与此同时，AI 技术也被广泛应用于机器人领域，机器人可以使用 AI 技术来实施环境感知、自动驾驶以及机器人协作等。此外，AI 技术还被应用于客户服务方面，AI 技术可根据广大用户的行为与需求，为广大用户提供最佳匹配的客户服务，从而节约时间与资源，强化服务质量，未来 AI 技术还会发挥出更多更为显著的价值效用，更好地解决人类社会发 展进程中遇到的一系列复杂性 问题，改善社会发展面貌，为人类生活提供更多便利条件。

若想推动 AI 伦理学走向更为理想的发展前程还需要进行跨学科合作，包括伦理学、法律、计算机科学和社会科学，唯有如此，才能保障 AI 技术未来能够对人类社会的健康稳定发展带来更为积极的影响，随着技术的不断升级，AI 伦理学未来必然会不断演化和发展，从而适应新的伦理挑战和机遇。

3. ChatGPT 简介与发展路径

ChatGPT 全称是“Chat Generative Pre-transformer”，中文为生成型预训练变换模型，其乃美国公司

OpenAI 在 2022 年 11 月 30 日发布研发的聊天机器人程序，可以应用于问答、文本摘要生成、机器翻译、分类、代码生成以及对话 AI，是一款人工智能技术驱动的自然语言处理工具。ChatGPT 具有同类产品具备的一些特性，如对话能力，可在相同会话期间回答与上下文有关的后续问题，然而，其之所以能够在短期内引爆全球，根本原因就在于在网民们晒出的截图中，ChatGPT 不但可流畅地同用户展开对话，而且还能做诗、撰文、编码等。

ChatGPT 还施行了注重道德水平的训练方式，参照提前设计好的道德准则，对恶意提问或者请求都能“说不”，一旦感知到用户给出的文字提示中存在恶意倾向，都会拒绝提供有效答案。

2022 年 11 月底，人工智能对话聊天机器人 ChatGPT 推出，并迅速红遍各大社交媒体，短短几天的时间，注册用户数量就已经超过百万。

2023 年 1 月末，ChatGPT 的月活用户数量已经破亿，成为史上增长速度最快的消费者应用。

2023 年 2 月 2 日，美国人工智能(AI)公司 OpenAI 发布 ChatGPT 试点订阅计划——ChatGPT Plus。ChatGPT Plus 将以每月 20 美元的价格提供，订阅者可以获得比免费版本更稳定、更快捷的服务，并掌握尝试新功能以及优化的优先权。

2023 年 2 月 2 日，微软官方公告表示其旗下所有产品将全线整合 ChatGPT，已经被黑客们用于策划网络攻击时使用。

4. ChatGPT 存在的伦理问题与解决方法

(一) ChatGPT 存在的伦理问题

1) 隐私和数据安全

AI 系统 ChatGPT 需要依赖大量数据展开训练并加以改进，这些数据通常包括用户生成的对话和文本。用户提供的个人信息和对话内容可能会被应用到训练模型中，如此一来，便会引发人们生活中的一些隐私问题。综合来讲大致包括两个方面隐私安全问题：首先是隐私侵犯，用户的对话和信息可能被用于训练 AI 模型，但用户可能并不知情或者没有明确同意。其次存在隐私泄露问题，存储和传输大量用户数据极易引发数据泄露、黑客入侵或使用不当等风险。

2) 社交影响与健康问题

如今，ChatGPT 在社交媒体、客户支持以及在线社区中得到了广泛应用，如此一来，便容易引发一系列伦理问题。首先是社交隔离方面的问题，人们会更多地与 AI 系统进行互动，而并非真实的人类。进而就容易导致社交隔离状况的发生，使人们产生孤独感，尤其长时间与 AI 系统对话，更是如此。其次就是虚假信息、恶搞与滥用，ChatGPT 等 AI 系统可用于生成虚假信息、恶搞和滥用，容易对社交媒体平台以及信息生态系统带来潜在性的风险。再往深处研究会发现 ChatGPT 也极易引发一些心理健康问题，马克思曾说过：“人是一切社会关系的总和。”^[2]当人们痴迷于虚拟的 AI 交流时，AI 助手与人类社交互动就容易对个体心理健康带来负面影响，如焦虑、孤独和社交障碍等。

3) 算法偏见和虚假信息

通过 AI 系统的训练数据可充分反映出社会的偏见与不平等，如此一来，就容易导致模型生成具有性别、种族、宗教或其他偏见。首先就是歧视性生成，AI 模型可能会生成歧视性内容，例如性别歧视、仇恨言论或不当建议等。其次是责任问题，当 AI 系统生成有害或不当内容时，伦理责任应由谁来承担，是开发者、使用者还是 AI 系统本身？受 ChatGPT 本身设计原理如算法透明性和可解释性的影响，AI 模型通常是黑盒，难以解释其决策过程，如此一来，也就很难检测和纠正偏见或错误。

此外，ChatGPT 还存在欺诈和虚假信息方面的问题，AI 系统常会被滥用，进而便会生成虚假信息，对社会舆论、金融市场或政治决策等都会造成一定的影响。有些居心叵测的人 would 恶意使用 AI 系统来生成

虚假信息，以便谋取私利。通过 AI 系统而生成的虚假信息可能会迅速传播，扰乱社会正常发展秩序。如何验证在线信息的真实性以及如何应对虚假信息传播都是现今 ChatGPT 存在的伦理问题。面对这些伦理问题，加强教育引导以及思想认识方面的培育显得尤为重要。用户和开发者需对 AI 系统潜在风险和伦理原则进行深层次了解，而加强伦理教育的施行则能够帮助人们更好地掌握 AI 系统的应用，以防 AI 系统遭到滥用。

(二) 应对 ChatGPT 与人类关系伦理问题的解决办法

1) 内部控制

首先，从伦理指导原则的视角进行分析，大家应对伦理原则拥有充分了解和把握，准确地说，应该是 AI 开发者和公司应当明确伦理原则，明确规定 AI 系统在生成内容时应遵守的道德标准，如尊重隐私、反歧视和防止虚假信息等。这些原则应作为开发和部署 AI 系统的基础。与此同时，应对开发者和数据科学家实施相应的伦理培训，使开发者和数据科学家都能对伦理问题和可行解决方案拥有深刻认识。培训应涵盖伦理决策的复杂性、伦理审查的重要性以及如何处理伦理冲突。

其次是透明度和可解释性，大致包括两方面，首先是模型可解释性，开发者应选择可解释性较强的 AI 模型架构，并使用可解释性工具来可视化模型的决策过程。这样有助于用户理解模型生成内容机制原理，其次是元数据和解释性标记，AI 生成的内容可以附带元数据和解释性标记，提供生成内容的来源、依据和生成过程的信息，这样有助于用户更好地了解内容的产生背后逻辑[3]。

最后是伦理审查，公司应设立独立的伦理审查团队，由不同领域的专家组成，负责评估 AI 系统可能引发的伦理风险和问题。此团队应在系统开发各阶段参与并提出改进建议，此外，伦理审查团队还应实施系统的道德风险评估，分析训练数据、算法和生成内容的潜在风险，这样有助于发现伦理问题，并在早期阶段进行干预。

2) 加强外部监管

首先，政府部门应制定明确的 AI 法律法规，对 AI 技术的发展和应用加以规范，以使其符合伦理和法律标准。同时政府还应设立 AI 监督机构，负责监督 AI 系统的合规性以及对其伦理标准的遵守状况。这些机构应制定规则、开展调查和处罚不当行为。

其次是国际合作方面，国际社会应积极展开合作，制定全球性的 AI 伦理标准，确保 AI 技术能够在全球范围内遵守伦理原则。国际合作机构应鼓励 AI 开发者和研究者积极分享关于 AI 伦理问题的信息和最佳实践，这样即可推进全球社区共同学习和进步。

最后是公众参与用户权益保护，政府和 AI 公司应积极鼓励公众参与，组织开展伦理研讨会以及相关方面的调查，以便公众能够表达对 AI 伦理问题的看法，这样也有助于建立民主的决策过程。政府应制定相应的法律法规，保障用户权益，包括隐私权、信息自主权和伦理权益等。用户有权知道 AI 系统如何使用其数据，并有权对其数据的使用提出异议。

通过内部控制和外部监管二者之间的有机结合，可全面应对 ChatGPT 与人类关系的伦理问题。内部控制强调开发者和 AI 系统的责任，确保 AI 系统在设计和使用中遵守伦理原则。外部监管则提供了监督和法规框架，确保 AI 技术的发展受到道德和法律的约束，同时促进公众参与和国际合作，以构建一个健康、公平和可持续的 AI 技术生态系统。这种综合方法有助于确保 AI 技术的发展符合人类的伦理和道德期望。

5. 从 ChatGPT 看世界：人机协调——未来人工智能的新型发展之路

人机协调指 AI 系统与人类之间的协同工作。未来 AI 系统将不再仅仅作为工具，而是伴随人类一同工作，以提高人们的生活质量、解决一些复杂性的问题。人机协调乃新型人机互动模式，强调 AI 系统与

人类之间的合作与共同努力，从而出色完成任务和解决存在的问题。这一概念在医疗保健、教育、工业自动化、客户服务等领域的应用前景极为可观，有助于提高效率、创造价值、改善人类生活[4]。

人机协调将需要教育和培训，使人们能够与 AI 系统合作。这不仅包括技术培训，还包括教育人们如何理解和信任 AI 系统，以及如何与 AI 系统展开有效交互。目前，AI 系统已在医疗保健领域取得一定的突破，未来将继续发展。AI 系统可辅助医生进行诊断、处理大规模医疗数据，并提供个性化医疗建议，从而改善医疗保健服务。通过人机协调能够改变人类工作方式，某些工作可能会被自动化，但与此同时也将出现新的就业机会，例如 AI 系统的维护和管理，在此情况下，伦理问题将变得更加复杂。人们需要考虑的是 AI 系统的决策如何影响人类，以及如何保护隐私和数据安全。人机协调将鼓励创新，促进不同领域合作。科学家、工程师、医生、艺术家和其他领域专业人士将会加强与 AI 系统的合作，从而制定新的解决方案。政府和监管机构需要制定新的法律法规来监管人机协调方面的事宜，包括责任分配、数据隐私和安全标准[5]。人机协调将对人类社会带来深远影响，包括改变人们日常生活方式、教育体系、医疗保健和经济结构。人机协调必须以可持续的方式发展，这样才能更好的保障 AI 系统的长期稳定性及其社会益处。

未来人工智能(AI)发展蓝图充满了无限可能性，特别是在 ChatGPT 这样先进 AI 技术的助力下。大家可在实际人工智能条件的基础上充分发挥想象，展望未来 AI 发展蓝图，未来 AI 助手不仅是回答问题的工具，还会充当真正的个性化伙伴，对每位用户需求、偏好和情感状态加以了解，以提供个性化建议、支持和娱乐。将 AI 应用于医疗保健以及超级医疗助手领域能够快速分析患者的医疗历史、症状和基因数据，为医生提供精确诊断和治疗建议，如此一来，就会有效强化医疗保健的效率和质量。

其次，工业自动化将迈向新高度。智能机器人和 AI 系统将在工厂中协同工作，不仅能够完成重复性工作，而且还能适应变化、灵活生产，从而提升生产质量与生产效率。AI 将激发不同领域之间的创新。科学家、工程师、医生、艺术家和社会科学家将与 AI 系统协同工作，共同解决复杂问题，推动科学和技术的进步。AI 系统还可在教育领域发挥作用，个性化学习助手将根据每一位学生的学习风格和需求提供定制的教育内容。虚拟现实和增强现实技术将创造沉浸式学习体验。ChatGPT 等 AI 系统将在翻译、跨文化交流和文化遗产方面发挥关键作用，帮助人们跨越语言障碍，更好地了解不同文化。AI 将成为解决全球挑战的强大工具，包括气候变化、医疗危机和自然灾害，其可协助科学家、政策制定者和国际组织制定有效的解决方案。

随着 AI 的崛起，伦理和社会责任也变得更加重要。AI 系统必须能够达到公平、透明和负责任的运行境界，同时 AI 系统还需尊重隐私和道德。总之，人机协调代表未来人工智能的新型发展之路，将面临前所未有的机遇和挑战。人们应当提高思想认识，保障 AI 系统能够与人类协同工作，从而改善人类生活质量，为社会可持续发展做出贡献。此外，人们还应积极应对伦理和社会影响，确保 AI 技术能够最大程度地造福人类，对此，需加强全球范围内合作和持续探索。此蓝图描绘了一个充满希望和潜力的未来，其中 AI 系统不仅是工具，还是人类的合作伙伴，通过二者合力可解决各种全球性问题，提高人类生活质量，并推动科学技术创新发展，而若想实现这一愿景则需依赖严格的伦理指导、全球合作及技术创新。

6. 结语

综上所述，人工智能的发展为社会的进步提供了积极助力，使人们见识到了更为先进的科技成果。然而，人工智能的发展也为人类带来了极大的隐患与挑战，作为人工智能领域的代表性产物，ChatGPT 在改善人类社会面貌的同时也需要人们提高自身思想意识，正确看待人工智能与人类关系的伦理问题，因为未来人类需要与人工智能实施深化沟通与密切协作。总而言之，人工智能与人类之间的关系妙不可言，需要大家时刻保持警惕与开放态度，以 ChatGPT 为例，深入研究人工智能与人类关系的伦理问

题，与此同时，还需要积极维护人类的自主性与尊严，避免人工智能对人类造成过度依赖与控制，进而对人类产生负面影响。

基金项目

江苏省研究生科研创新计划项目“以 ChatGPT 为例，探讨人工智能与人类关系的伦理问题”（项目编号：KYCX23_3919）。Postgraduate Research & Practice Innovation Program of Jiangsu Province《Ethical issues in the relationship between artificial intelligence and human beings—Take ChatGPT as an example》（项目编号：KYCX23_3919）。

参考文献

- [1] 王楠, 张林, 李维中. 人工智能伦理学研究综述[J]. 计算机应用, 2019, 39(2): 1-7.
- [2] 卡尔·马克思. 关于费尔巴哈的提纲[M]. 北京: 人民出版社, 1995.
- [3] 郭宁, 马超. AI 伦理问题与数据隐私保护[J]. 计算机科学与探索, 2018, 12(6): 893-902.
- [4] 朱大雁, 王晓骏, 杨兵. 基于数据挖掘的 AI 伦理问题与挑战[J]. 计算机研究与发展, 2019, 56(4): 773-787.
- [5] 谢霆婧, 蔡斌. AI 与社会伦理的平衡[J]. 科技与法律, 2020, 17(8): 32-38.