



南京邮电大学  
Nanjing University of Posts and Telecommunications

( 2023 - 2024 学年 第 2 学期)

## 《工程与社会》课程论文

题 目 人工智能引发的社会问题

所在学院 自动化学院、人工智能学院

专 业 人工智能

年级班级 B210416

学 号 B21080526

姓 名 单家俊

授课教师 任洁

社会与人口学院

2024 年 4 月 27 日



# 人工智能引发的社会问题

## 1. 摘要

AlphaGo 在 2016 年的围棋比赛中取得胜利，标志着人工智能技术的一个重要里程碑。本研究聚焦于人工智能技术的社会影响，特别是它如何引发人权、安全和责任分配方面的问题。通过深入分析这些问题的根本原因，本文旨在提出对策和建议，以促进人工智能技术的健康发展并预防潜在的社会危机。本研究的发现强调了社会各界在应对人工智能挑战时采取行动的重要性。

## 2. 文献综述

在技术进步的同时，围绕人工智能引发的社会问题的研究也持续进行。李伦在《人工智能与大数据伦理》一书中对人工智能的道德决策和主体地位进行了探讨，指出人工智能道德算法主要是那些在道德上可接受或符合伦理的算法，这些算法能够极大地提高自主系统的决策可靠性和安全性<sup>[1]</sup>。此外，《人工智能会取代人类吗？》一书中的作者讨论了人工智能将如何以多种方式影响我们的社会、经济、就业和战争等方面，强调它们会为人类所做的每一种活动带来翻天覆地的变化<sup>[2]</sup>。

关于人工智能的道德责任，学者们进行了深入的探讨。有学者认为人工智能体具备道德主体地位，但这个“道德主体”是“非完全道德主体”，非完全道德主体不具备意识和意志，也不具备情感和欲望，是“二阶道德能动者”，因此尚不足以被问责<sup>[3]</sup>。郭桑的《人工智能的自主性与责任》指出，人工智能的责任问题必然要摆脱传统的单一问责机制，走向社会—技术—认识系统中的问责机制<sup>[4]</sup>。

法律领域的研究，袁曾的《人工智能有限法律人格审视》和郑志峰的《人工智能时代的隐私保护》，探讨了人工智能的短期影响取决于谁控制人工智能，而长期影响则取决于人工智能到底能否受到任何控制，人工智能的运用与发展需要法律人格的明确<sup>[5]</sup>。当前，我国隐私保护法律体系还不够健全，特别是对于信息隐私的保护还处于起步阶段，缺少一部专门的《个人信息保护法》，难以应对人工智能时代的隐私风险<sup>[6]</sup>。

在伦理方面，赵云和张煌在《智能化战争的伦理审视》中指出智能化武器为人类带来的伦理冲击，甚至可能导致人类文明史的终结<sup>[7]</sup>。王银春在《人工智能的道德判断及其伦理建议》一文中剖析了现阶段我们对待 AI 的基本立场仍是“人类中心主义”原则，主张用一种审慎态度来对待 AI<sup>[8]</sup>。

此外，蔡自兴在《中国人工智能 40 年》中赞同国家战略，强调人工智能将日益成为新一轮产业革命的引擎，必将深刻影响国际产业竞争格局和国家的国际

竞争力<sup>[9]</sup>。韩晔彤在《人工智能技术发展及应用研究综述》中也指出人工智能应用面比较广泛,可代替人类进行各个方面的工作,可以说大大提高了人类在日常生活工作中的效率<sup>[10]</sup>。在国际层面, Kazim Emre 在《A high-level overview of AI ethics》中认为人工智能要做到以人为本强调 “In all three ethical approaches, the central subject of concern is the human being, i.e., “persons.””<sup>[11]</sup>

学者们从多个角度探讨了人工智能带来的影响和挑战,例如人工智能系统的道德决策能力、法律地位、隐私保护、武器使用的伦理问题等。总的来说,大家呼吁在发展人工智能的同时,要高度重视其伦理、法律和社会影响,追求公正、透明、可理解和可追责的人工智能系统,遵守相关国际公约,维护人类价值观。

在审视了人工智能技术所引发的众多伦理、社会及法律问题之后,本文接下来将深入探讨人工智能对人权的影响、技术安全性的重要性,以及责任界定的复杂性。这三个主题不仅核心地体现了当前对人工智能的关注焦点,也是评估和引导人工智能健康发展的关键。

### 3. 人工智能技术背后的社会问题

#### 3.1 人权

人权是人与生俱来的基本权利,体现了人的尊严和价值。马斯洛需求层次理论将自我实现定位为人的最高层次需求,强调人应成为全面发展、具有价值的完善人格。然而,人工智能技术的兴起给人权带来了前所未有的严峻挑战和冲击。

首先,人工智能技术可能会极大削弱人的自主决策权和主体能动性。人工智能系统依托大数据和算力优势,可以为人类提供看似“最优解”的决策建议。但当人们过度依赖这些“最优解”做出重大决策时,实际上已经将自主判断权割让给了机器。无论是借助人工智能做出决策,还是完全交由人工智能系统替代决策,人的主观能动性和独立意志都将受到严重影响。在医疗、法律、金融投资等关系重大的领域,人们将被动接受人工智能系统的判断,主体性和尊严遭到极大削弱。

其次,人工智能可能会导致大多数人失去就业机会,产生巨大的“无用的阶层”。随着人工智能技术的飞速发展和广泛应用,越来越多传统工种将被人工智能系统完全取代。那些难以持续自我学习和提升的人,将被边缘化,沦为“无用者”。与之形成鲜明对比的是,少数掌握人工智能技能的精英阶层将独享发展机遇。社会分化加剧,绝大多数普通人将失去赖以维系生活尊严的工作权利,成为被剥夺基本权利的“下等公民”。

此外,人工智能技术在一些领域的不当使用,也可能严重侵犯人的生存权、健康权等基本权利。比如在医疗诊断领域,如果人工智能系统存在偏差或出现失误,

可能会对患者的生命健康造成威胁。而在一些高风险场景,自主作战的人工智能武器如果失控,也可能导致大规模平民伤亡,违反人道主义原则。

人工智能技术给人权保护带来了全方位的巨大挑战。我们亟需高度重视这一问题并采取切实有效的应对措施,否则人类文明的基石将遭受空前的动摇和颠覆。人权保护无疑是人工智能发展必须直面的首要课题,需要我们在技术和伦理层面付出长期的努力。

### 3.2 安全性

人工智能系统的"不可解释性"和"黑箱"效应,给其安全性带来了巨大的不确定性风险。由于外界无法完全看透人工智能内部的运算过程和决策机理,我们难以预测和验证其行为轨迹,就无法及时发现和防范其中可能存在的安全隐患。一旦出现失控或被滥用的情况,后果将是不堪设想的。

以人工智能自动驾驶系统为例,尽管它通过持续学习可以提升驾驶水平,但如果遇到特殊情况,其如何作出决策并采取行动,往往是一个"黑箱操作"。如果决策出现失误,可能会导致严重的交通事故和人员伤亡。而要追究系统内在的具体原因,例如是算法设计缺陷还是遭到外部数据干扰,困难重重。

个人隐私和数据安全同样面临着前所未有的挑战。人工智能发展离不开海量个人数据的获取和利用,但如果这些敏感数据遭到滥用或窃取,将严重侵犯公众的隐私权和数据安全。事实上,越来越多应用程序过度索取用户权限,人们的个人信息极有可能被人工智能系统"看光"并加以利用。这种隐患已经引发广泛关注。最近,人工智能还被滥用于犯罪活动,如"AI 换脸"诈骗等,引发公众对个人信息安全的重大担忧。如何管控这种新型犯罪行为,避免人工智能沦为犯罪工具,已经成为当务之急。

此外,人工智能的应用已经拓展到战场上,人工智能武器系统一旦失控,也可能造成难以估量的危害。这些自主杀伤武器是否会违背人道主义原则对平民造成伤害,是一个令人忧虑的问题。同时,人工智能技术的军事应用还可能加剧国家间的军备竞赛,引发新一轮军事对抗,动摇地区乃至全球和平。

人工智能技术的安全性问题涉及面极为广泛,不仅关乎个人权益,也关乎国家安全和社会秩序。其中蕴含的不确定性风险、隐私泄露隐患、军事应用危险、犯罪利用等诸多挑战,不容忽视。我们急需全面审视人工智能安全性问题,寻找有效的预防和应对之策。

### 3.3 责任界定

当探讨人工智能技术的责任界定问题时,我们面临的挑战主要来源于 AI 的复杂性和其在决策中的自主性。尽管 AI 系统在许多方面提高了效率,它们在出现故障或导致伤害时,责任归属的不明确性却给法律制度带来了重大挑战。

在当前的法律体系中,AI 既不是自然人也不是传统意义上的法人,这使得它无法完全适用现有的自然人或法人责任主体规定。当 AI 系统造成损害时,受害者很难获得有效的法律救济。问题可能源自 AI 系统的设计缺陷、输入数据的问题或是用户操作错误,而 AI 算法的"黑箱"特性又使得追踪具体责任变得更加困难。

AI 系统本身具有自主性、不可预测性以及不透明性等特征,这给事故责任的认定带来了前所未有的困难。我们很难直接将其等同于人为失误,去简单地将责任归咎于开发商或使用者。AI 系统的决策过程常常是一个复杂的"黑箱操作",外界无法完全看透内部的运算规则和决策逻辑,就无法判断到底是设计缺陷还是数据输入问题所致。

另一个棘手问题是,AI 系统在某些场景下的行为甚至可能已经超出了设计者的预期范畴,具有一定的"无法控制"性。这种情况下,如何认定责任主体就更加模糊了。比如一款自主作战的 AI 武器系统,一旦失控并对平民造成伤害,究竟是武器本身的程序缺陷还是意外情况下的不可预测行为,很难一概而论。

AI 技术的责任界定问题牵涉到诸多复杂因素,不仅需要对 AI 系统本身的特性有深入认识,还需要研究现有法律制度下如何定位 AI 的身份地位,以确定其应当承担何种程度的法律责任。这是一个亟需技术、法律和伦理多学科探讨的全新领域,也是未来社会发展必须正视和解决的重大挑战。

## 4. 解决方法

人工智能技术的发展将深刻影响和改变社会的方方面面,因此妥善解决人权、安全、责任等问题,事关社会长治久安和可持续发展。

首先,法律法规体系的完善,不仅关乎人工智能行业秩序,更是维护整个社会正常运转的重要保障。明确的权利义务界定和行为规范,能够为社会各主体的活动划定边界,避免无序竞争和无序使用,促进人工智能健康有序发展。

其次,加强技术审查和伦理监督,是维护社会公平正义的应有之义。人工智能一旦被滥用于违法犯罪行为,将严重破坏社会秩序和人们的安全感。通过严格审查,确保人工智能应用符合伦理道德底线,切实保护公众利益不受侵犯,这也是人工智能赢得社会认同的前提。再者,加大科普宣传力度,有助于消除偏见和谣言,营造人工智能技术在公众中的良好接受度。公众对新技术的认知理解程度,直接

关乎技术在社会上的传播和应用空间。只有通过科普教育,让全社会形成正确认知,人工智能技术才能顺利融入人们生活。

同时,强化伦理教育对于建设人工智能时代的社会价值观也至关重要。科技是中性的,最终是否造福社会,取决于操纵者的价值观导向。我们需要在全社会范围内加强人工智能伦理教育,培育积极向上的价值观,确保人工智能的应用有利于社会和谐发展。

最后,人工智能发展需要全球携手合作。人工智能技术的影响是全人类共同面对的挑战,任何一个国家或地区的单打独斗都是徒劳的。国际社会必须加强对话协作,就相关规则达成一致,化解分歧,共同促进人工智能在全球范围内的健康发展,造福全人类。

总之,人工智能技术与社会发展存在着千丝万缕的关联,解决人权、安全、责任等问题,不仅关乎行业发展,更事关整个社会的可持续和谐发展。我们需要从法治、技术、教育、国际合作等多个层面入手,多管齐下,构筑人工智能技术健康发展的坚实基础。这是一个漫长而艰巨的系统工程,需要全社会的高度重视和不懈努力。

## 5. 参考文献

- [1] 李伦. 人工智能与大数据伦理[M]. 北京: 科学出版社, 2018: 13-14.
- [2] 托比·沃尔什. 人工智能会取代人类吗? [M]. 阎佳, 译. 北京: 北京联合出版公司, 2018: 122-123.
- [3] 简小烜. 人工智能体的道德地位研究——基于责任论视角[J]. 湘潭大学学报: 哲学社会科学版, 2020, 44(5): 136-137.
- [4] 邬桑. 人工智能的自主性与责任[J]. 哲学分析, 2018, 9(4): 42-43.
- [5] 袁曾. 人工智能有限法律人格审视[J]. 东方法学, 2017(5): 50-57.
- [6] 郑志峰. 人工智能时代的隐私保护[J]. 法律科学: 西北政法大学学报, 2019(2): 55-56.
- [7] 赵云. 中国社会科学杂志社[J]. 军事学. 中国社会科学报, 2018.
- [8] 王银春. 人工智能的道德判断及其伦理建议[J]. 南京师大学报(社会科学版), 2018, (04): 34-35.
- [9] 蔡自兴. 中国人工智能 40 年[J]. 科技导报, 2016, 34(15): 12-32.
- [10] 韩晔彤. 人工智能技术发展及应用研究综述[J]. 电子制作, 2016(06X): 1.
- [11] Emre Kazim. A high-level overview of AI ethics[J]. Patterns, 2021, 2(9).