



南京邮电大学
Nanjing University of Posts and Telecommunications

(2021-2022 学年 第 1 学期)

《工程与社会》课程论文

题 目 人工智能技术背后的社会问题

所在学院 计算机学院

专 业 信息安全

年级班级 B190307

学 号 B19031614

姓 名 任远哲

授课教师 谢长征

社会与人口学院

2022 年 4 月 20 日

人工智能技术背后的社会问题

1. 摘要

2016 年 Alphago[1]战胜世界围棋冠军李世石这一历史性事件之后，人工智能这个概念传遍大街小巷，进入人们的视野。近年在大数据和强大算力的加持下，人工智能更是得以迅速发展。然而，现在的人工智能，只是弱人工智能，大部分只是用神经网络对数据的拟合，和在电影里看到的强人工智能还有非常大的差距，本质上并非智能。另外人工智能所带来的问题也层出不穷，直接或间接地引发了诸多社会问题。如果不解决好人工智能在技术上和社会上的问题，很可能会引发这个行业的第三次寒冬。作为《工程与社会》课程的大作业，本文主要针对人工智能社会层面的问题展开论述，从人权、安全、责任界定三个维度对人工智能困境进行梳理并分析这些问题的产生原因。最后总结，并就人工智能的发展，对社会各界提出倡议。

2. 背景和文献综述

历史上人工智能技术的发展经历了三次高潮和两次低谷:50 年代末，由神经网络主导了第一次高潮，利用感知器数学模型模拟人的神经元反应，完成简单学习及分类任务;70 年代初期，由于计算能力不足，无法使机器完成大规模数据训练和复杂任务而陷入低谷;而在 80 年代中期反向传播算法获得广泛关注，专家系统也投入应用，人工智能技术第二个黄金时期，但好景不长，90 年代初，由于没有严格的数学理论支撑人工神经网络的设计，专家系统也暴露出应用领域狭窄、知识获取困难等问题，导致政府投入缩减，人工智能再次陷入低谷;而 2016 年以来，云计算、大数据、物联网等技术的群体性突破，深度学习、对抗学习和强化学习等思想的提出和运用也促使 AI 领域实现重大突破，以 AlphaGo 战胜围棋冠军李世石为起点，人工智能开始了第三次高潮，并以前所未有的态势快速渗透进人类生活的方方面面。

在技术发展的同时，针对人工智能所带来的社会问题的研究也在持续。在国内，李伦主编的《人工智能与大数据伦理》[2]中对人工智能的道德决策、主体地位等也进行了探讨，闫佳翻译的《人工智能会取代人类吗?》叙述了人工智能的发展、现状、缺陷和对人的影响，并对未来进行展望。简小烜《人工智能体的道德地位研究——基于责任论视角》[3]，较为全面总结伦理问题的，针对主体性及责任进行深入研究；邬桑的《人工智能的自主性与责任》[4]保持审慎立场，并提出寻求算法决策与算法权力的公正性，呼唤更加透明、可理解和可追责的人工智能技术，反思智能化的合理性及可能出现的伦理冲突等价值诉

求。在人工智能法制建设方面也有不少研究，袁曾的《人工智能有限法律人格审视》[5]，郑志峰的《人工智能时代的隐私保护》[6]等，探讨人工智能的主体问题、隐私侵犯问题，并提出加强法律完善的建设性意见。赵云，张煌在《智能化战争的伦理审视》[7]中对智能化武器进行伦理审视，主张致命性自主武器系统在没有或无法植入社会正义理念的情景下，其研发、交易、部署和使用等每个环节都应受到国际规制严格约束和监管。王银春在《人工智能的道德判断及其伦理建议》[8]一文中剖析了人们对于人工智能技术持有三种不同态度：乐观、中立以及悲观，并且他认为应该坚持“人类中心主义原则”用审慎的态度去对待人工智能。蔡自兴在《中国人工智能 40 年》[9]中详细论了国外及中国的人工智能技术的发展历程，由迷雾重重发展到拨云见日最后上升到国家战略，发展之路不可谓不艰难。同样韩晔彤在撰写的论文《人工智能技术发展及应用研究综述》[10]中也阐明了人工智能如今的重要作用，其重要作用主要是指人工智能技术在解放劳动力方面的重要作用，并且在此过程中可以提高产量及效率使得数据更加真实有效。段伟文在《人工智能时代的价值审度与伦理调适》[11]中认为要从有限自主与交互智能体出发，校正人工智能的价值与伦理。

在国外，国外人工智能技术发展早于中国，相关的研究著述成果颇丰。Kazim Emre 在《AI 伦理的高级概述》[12]中认为人工智能的伦理问题是基于关于人类、人类及其在世界特别是技术世界中的相互作用的论题提出来的，我们应当主动应对具有伦理互动的主体，并将人类的生命作为最高的道德价值和优先事项。Turt Russell 在《人工智能武器的立场》认为自主武器系统应当被限制在法律规范的范围内，并形成国际公约，确保遵守国际法的程序与规定。Sabine Hauert 认为社交媒体应该成为客观、正确信息的传播者，避免夸大和炒作人工智能技术，政府、研究机构、商业发展机构，以及研究和工业协会都有责任公开对话、推广和科学传播人工智能技术及其影响。

当下无论在生产，生活，还是军政领域，都能看到人工智能的身影。人工智能技术的发展和运用必然会对当下社会伦理关系产生冲击。我接下来将从人权、安全、责任界定三个方面分别阐述。

3. 人工智能带来的社会问题

3.1 人权问题

人权这一问题是在近代才提出来，强调的是人区别于动物的尊严与价值，人是有思想的人，追求自我价值实现，马斯洛需求层次理论中，人的最高级需

求便是自我实现，成为一个完善人格、全面发展具有价值的人。然而，人类正慢慢丧失决策权力，这使我们的主体地位受到前所未有的挑战。人工智能所具有的高效信息收集、分析、决策能力是人类所难以企及的，通过人工智能对数据进行分析，得出“最优解”是专家系统等人工智能产品的目的，人们利用“最优解”来做出决策，便是丧失了完全的自主决策权的先验条件。无论是人类借助于人工智能做出判断还是直接让人工智能做出判断，都有人工智能介入，影响人的自主判断。除了根据专家系统给出的治疗方案救治病人、利用金融投资软件确定投资比重与数额等应用之外，人工智能还用来帮助人们寻找伴侣：婚恋网站上通过兴趣爱好、性格特点推荐你的合适伴侣，茫茫人海利用人工智能进行筛选不失为一种提高效率、增加牵手成功率的方法，但是却处于被人工智能“安排”的被动状态，只能在既定的范围内做选择。因此人丧失了应当具备的主体性认知，丧失了自主自由选择的权利。

另一方面，人工智能导致人类能动性、创造性的丧失，产生“无用的阶层”。“无用的阶层”的最早是由尤瓦尔·赫拉利在《未来简史》中提出，《人类简史》中再一次提到“人工智能技术的快速发展必然导致无用的阶层的兴起”，并引发了学界的探讨，而与之相对的是少数精英阶层。日常被人工智能辅助下的人类则如温水中的青蛙，丧失了避免成为无用的阶层的唯一方法就是终身学习，但事实是真正可以终身学习的只是少部分人，难以不断自我提升的人会在人工智能时代被机器所取代。

3.2 安全问题

人工智能的安全问题是业界最关注的问题之一，安全是人类最基本的生理需求，只有在保证信息、人身安全的前提下，一切便利才有意义。而如何证明人工智能的安全性一直是难题。因为人工智能本身具有“不可解释性”，“只要输入既定程序和规定算法，人工智能便可以在极短时间内完成计算、推论和策划，人们无需也无从知道运算过程的始终，更无法验证其运算结果的真伪。”

从我的专业“信息安全”的角度看，人工智能越是“智能”，就越需要获取、分析越多的数据信息，进而通过学习、模拟以提高自身在某个领域的能力。大家可能都有这样一种体验，有时候可能就是随口提了某个东西，接下来在各大 app 就能精确地收到相关推送。“推荐系统”作为人工智能领域的一大分支，它的数据来源于哪儿呢？有些 app 在运行时会申请各种权限，比如读取剪贴板，调用录音机，使用者稍有不慎就会面临被人工智能看光的风险。不仅是个人，国家的安全也可能被侵犯，作为自动驾驶领域的先驱，特斯拉汽车被限制驶入政府军事部门，因为它所携带的摄像头很可能泄露一些机密。

在战争安全方面，人工智能时代下的军事领域的竞争也从简单的火力对抗转为信息科技的竞争，战争的制胜机理将完全改变。2020年初，伊朗军事骨干领导人苏莱曼尼、莫哈丹接连被自动瞄准机枪暗杀，表明基于人工智能与无人化作战系统的充分融合。智能化军事提高了暗杀、刺杀等军事活动的成功率，大大促进战争的智能化。然而智能化降低伤亡成本的同时是否会增加军备竞赛的风险？自动化武器或自主武器是否有权利决定杀死谁？如何确保不会杀死平民、伤及无辜？给用来杀戮的武器限定人道主义原则，本身就是对程度进行把控的难题。

3.3 责任界定问题

由于人工智能本身具有安全风险，当其出现事故、对人造成伤害时，其具有的自主性、不可预测性等特征可能给责任承担提出挑战，使得受害人的损害无法获得弥补。弱人工智能不具有自主性、独立意识，只是算法模型的工具，无力承担运行错误或失误导致的损失。当出现事故，需要分析是由于出厂设计过失、外界数据挟持还是错误使用所致，由产品设计公司、供给商还是使用者自身承担责任，如自动驾驶汽车、陪伴机器人等对人造成伤害，是否应当让人工智能体承担法律责任或道德责任？人工智能体又如何承担责任？然而，人工智能算法本身具有的不透明性、不可解释等特性，导致定责问题困难重重。

从法律角度来说，法律主体分为自然人与非自然人，对于自然人的法律主体界定根据自然人的行为能力、认知水平分为三种类型：18周岁以上公民完全行为能力人，14-18岁限制行为能力人，14岁以下则是无行为能力人，不承担法律责任。有观点主张将人工智能进行能力评估，例如类似图灵测试等专业评估与分级，判断人工智能应该定无行为能力人还是限制行为能力人。

还有一种观点是与自然人不同，法律主体还可以是法人，这里的“人”本身就是一种虚拟概念，指当前为了满足对经济活动与市场运行的管理，法律规定能够独承担民事责任的企业组织。人工智能完全可以视为法人，俄罗斯最早进行了最早尝试，公布该国首部也是世界上最早的关于智能机器人的法律地位的法律草案《在完善机器人领域关系法律调整部分修改俄罗斯联邦民法典的联邦法律》，也叫“格里申法案”。法案提出的“机器人—代理人”的概念，提出了机器人在不同法律关系中的发展阶段和定位，将不需要人类控制、能采取完全行动并能够评估后果的机器人进行登记，作为人类占有者的代理人。

4. 人工智能诸多社会问题的成因与解决办法

4.1 人工智能技术的局限性

人工智能并不是一门十分成熟的学科，它自身是存在着一定程度的局限性。现在的人工智能，只是弱人工智能，大部分只是用神经网络对数据的拟合，和在电影里看到的强人工智能还有非常大的差距，本质上并非智能。深度学习说是模仿人脑。但是却处处进行矩阵运算。然而人的大脑很不适合矩阵运算的。所以本质上来说深度学习对于人脑的模仿上是不成功的，他们完全两套机制，很难将人工智能真正看作是人。

4.2 监督管理不及时、政策法规不到位

在人工智能技术发展的早期，美国科幻作家阿西莫夫就曾给出了著名的“机器人三定律”。首先，机器人绝对不被允许做出威胁人类安全的事情，机器人在面临人类被伤害的处境时不能冷眼旁观；其次，机器人要坚决执行人类通过计算机系统而下达的指令，但即使这样也不能够违背第一条定律；最后，机器人在执行人类指令时必须维护自身安全，但不能违背以上两条定律。然而当下我们必须承认的是，关于规范人工智能技术发展的相关法律法规尚且十分不完备，需要我们在不断的不断发展中与时俱进，不断增添新条款。最先应该有明确解释和规范的法律条文的，是研发人工智能产品的相关人员、人工智能产品的相关运营商、购买人工智能产品的普通消费者等的相关权利和义务。而后，对于因人工智能产品被滥用而造成的恶劣后果和影响，需要来明确的法律处罚条例以及法律案例的适用，做到有法可依，有法必依，而不能使相关责任人有待无。

4.3 公众对于相关技术理解的偏差

有相当一部分人对人工智能技术抱有错误的理解，其中一部分人错误的把人工智能当作是神一样，无所不能，偏执地认为人工智能技术可以适用到任何一种条件中。但就目前的发展状况，技术远远不能达到无所不及的程度，就像真理的适用是有条件的一样，人工智能技术的使用也是有其范围的。一部分人认为凭借如今人工智能技术的发展现有状况来看，尚且不能将其进行大范围的应用。持有这种观点的人们则是认为以该技术的现有水平来看，人工智能不得不依赖于人工标注，故而其认为该技术不能被大范围的应用。另外的一部分人则认为在未来的 20-30 年中，瓶颈会到来，但在其后，人工智能将完全不受人

类掌控，人类只能任其自由发展，完全不能加以束缚。故由此而引发相当一部分群众误解人工智能，以为其将会为人类带来灭顶之灾，当一些人将人工智能引发的负面例子中的负面信息不断放大后，人们更加恐惧人工智能。所以，人们更加害怕人工智能技术，甚至对其抱有排斥的心理。

这一事实使得人工智能技术在人们头脑中的认可度越来越低。很明显，公众不认可也是加剧人工智能伦理问题形成的外部推力之一。要想解决由于人类的理性和道德问题造成的人工智能技术伦理问题，必须要逐渐提高人的文化素养和道德修养，这注定是一个缓慢的过程，过程中的每一个人，无论是专家学者还是普通大众，都应该以身作则，不以讹传讹，促进人工智能良性发展。

5. 总结和倡议

本文通过对人工智能技术的发展和社會问题进行汇总和梳理，以人工智能对社会造成的问题为探讨核心，并分析其背后原因。我们必须承认，人类的福利必须置于一切之上。科学绝不是一种自私自利的享乐。只有人工智能所带来的诸多社会问题得到解决，人工智能技术才能大规模地落地并投入使用，这是未来的计算机学人才和社会学人才应该合作攻坚的方向。

人工智能伦理困境的破解应该立足实践，针对不断出现的具体问题加强管理和治理，国家政府及时出台方针、政策及法律法规，企业在推出人工智能产品时，明确使用规则与情景，并在技术上不断改进和完善，规则和技术共同推进，而公民个人也应当梳理正确道德观念，提高自我保护意识，促进人工智能技术的合规使用。总之，人工智能各问题的破解需要各方面携手并进，才能共创美好智能新代。

作为一名计算机学院的学生，在未来的研究生生涯中我也打算选择人工智能方向。感谢《工程与社会》这门课程，让我以一个全新的视角重审自己研究的领域，我在今后也会不断努力，让技术造福人类，人与人工智能和谐相处。

参考文献

- [1] Dong X, Wu J, Zhou L. Demystifying AlphaGo Zero as AlphaGo GAN[J]. 2017.
- [2] 李伦主编. 人工智能与大数据伦理[M]. 北京: 科学出版社, 2018.
- [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):6.
- [4] 郭桑. 人工智能的自主性与责任[J]. 哲学分析, 2018, 9(4):10.
- [5] 袁曾. 人工智能有限法律人格审视[J]. 东方法学, 2017(5):8.
- [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):10.
- [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):029.
- [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] 段伟文. 人工智能时代的价值审度与伦理调适[J]. 人民大学学报.
- [12] Kazim Emre,Koshiyama Adriano Soares. A high-level overview of AI ethics[J]. Patterns,2021,2(9).