

学 号 112015301000014

# 延衛大學 硕士学位论文

# 人工智能发展的伦理问题研究

论文作者: 谷雨

指导教师: 邹顺康 副教授

学科专业: 伦理学

研究方向: 应用伦理学

提交论文日期: 2018 年 4 月 20 日

论文答辩日期: 2018 年 6 月 2 日

学位授予单位: 西南大学

中国•重庆 2018年6月

# 独创性声明

学.	位论文题	E		人	工智	能发	展	的作	理	问	题研	究

本人提交的学位论文是在导师指导下进行的研究工作及取得的研究成果。论文中引用他人已经发表或出版过的研究成果,文中已加了特别标注。对本研究及学位论文撰写曾做出贡献的老师、朋友、同仁在文中作了明确说明并表示衷心感谢。

学位论文作者: 谷雨

签字日期: 2018 年6 月7日

# 学位论文版权使用授权书

本学位论文作者完全了解西南大学有关保留、使用学位论文的规定,有权保留并向国家有关部门或机构送交论文的复印件和磁盘,允许论文被查阅和借阅。本人授权西南大学研究生院(筹)可以将学位论文的全部或部分内容编入有关数据库进行检索,可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

(保密的学位论文在解密后适用本授权书,本论文:□不保密,□保密期限至 年 月止)。

学位论文作者签名: 给雨

签字日期: 2018 年 6月 7 日

导师签名:

签字日期:20/8年6月7日

# 目 录

摘要	要	I
Abs	stract	iIII
第-	一章	绪论1
	1.1	选题的背景和问题的提出1
	1.2	研究现状的综述1
		1.2.1 国外研究现状1
		1.2.2 国内研究现状
		1.2.3 国内外研究评述4
	1.3	研究的目标与意义4
		1.3.1 研究的目标4
		1.3.2 研究的意义5
	1.4	研究的思路与方法5
		1.4.1 研究的思路5
		1.4.2 研究方法
		(1) 分析与综合的研究方法5
		(2) 系统分析方法6
	1.5	创新与不足6
	1.5.	1 创新之处6
	1.5.	2 不足之处6
第二	章	人工智能概述7
		人工智能技术简介7
	2.2	人工智能的发展8
	2.3	人工智能的应用领域以及对人类社会的影响9
		2.3.1 人工智能的主要应用领域9
		2.3.2 人工智能的发展对人类社会的影响11
第三	E章	人工智能发展引发的伦理问题及成因分析13
	3.1	人工智能发展引发的伦理问题13
		3.1.1 人权伦理问题
		3.1.2 道德地位问题
		3.1.3 责任伦理问题16
		3.1.4 环境伦理问题

	3.1.5	代际伦理问题	. 20		
3.2	人工智	智能伦理问题的成因分析	. 21		
	3.2.1	技术上的局限性	. 21		
	3.2.2	伦理规则的缺失	. 22		
	3.2.3	监督管理体系的不完善	. 23		
第四章	人工智	智能伦理问题的应对策略	. 25		
4.1	研发原	层面的应对策略	. 25		
	4.1.1	用"以人为本"的科学发展观指导人工智能发展	.25		
	4.1.2	让程序语言融入哲学思想	. 26		
	4.1.3	构建人工智能技术的伦理标准	. 27		
4.2	人类自	自身层面的应对策略	. 27		
	4.2.1	增强科学家的道德责任感	. 27		
	4.2.2	提高民众的科学素养,正确认识人工智能	. 28		
	4.2.3	加强人工智能的国际交流合作	. 29		
4.3	制度规	观范层面的应对策略	. 30		
	4.3.1	制定法律法规来规范人工智能的发展	. 30		
	4.3.2	科学管理人工智能产品的使用	. 30		
	4.3.3	严格监督人工智能产品的研发和运用	. 31		
第五章	结论		. 33		
参考文献					
致谢					

# 摘要

自 1956 年人工智能第一次作为一个学科概念出现在世人面前以来,人工智能科技的发展日新月异,目前弱人工智能已经被全面应用到人类生活的方方面面。现代的互联网大数据检索、智能无人驾驶汽车的研究、智能家居用品的应用等让人工智能不仅仅停留在理论层面,也让我们的生活充满了人工智能的成果。同时,"爱丽丝"高仿真聊天机器人在脱口秀上的精彩表现、"阿尔法狗"在围棋大赛中战胜韩国选手李世石,让"人工智能威胁论"开始充斥整个舆论界。对人工智能伦理问题的研究刻不容缓,笔者意图通过自身的伦理学知识的运用,站在辩证的哲学角度上分析人工智能伦理问题出现的原因,给出一些解决办法,在人工智能的蓬勃发展过程中贡献一点力量。

本文共分为五章:第一章是绪论;第二章是人工智能技术概论,包括人工智能的技术简介、人工智能技术的发展、人工智能的主要应用领域以及人工智能对人类社会的影响;第三章详细描述了人工智能发展中产生的伦理问题,其中包括人工智能的人权伦理问题、道德地位问题、责任伦理问题、环境伦理问题以及代际伦理问题。并且探讨了人工智能伦理问题的成因,主要包括以下几个方面:人工智能技术上的局限性、伦理规则的缺失、监督体系的不完善以及公众科学素养不高导致的对人工智能的误解;第四章笔者通过对前文中伦理问题的分析,从三个层面给出应对的策略:在技术开发的层面,有用以人为本的科学发展观思想引导人工智能的发展、让程序语言融入哲学思想以及构建人工智能的伦理标准。在人类自身的层面给出的应对策略有,提高科学家的道德素养、提高民众的科学素养和文化素养、加强国间的交流合作。在政府规范政策的层面,通过立法规范人工智能技术的发展、科学管理人工智能产品的开发、严格监督人工智能产业的使用;第五章为结论。总之,面对这一次人工智能技术的飞跃,人类要时刻做好技术反思工作,不断地总结问题,并采取相应的策略,这样才能够使人工智能技术趋利避害,更好地造福于人类。

关键词:人工智能;伦理问题;伦理规则;对策

#### **Abstract**

Since 1956, artificial intelligence has emerged as a discipline for the first time, artificial intelligence technology is developing rapidly. At present, weak AI has been applied to every aspect of our life.Modern Internet big data retrieval, intelligent driverless vehicle research, smart home appliance application and so on, make AI not only stay at the theoretical level, but also make our life full of AI achievements.Meanwhile, the "Alice" high simulation chat robot and the "Alfa dog" defeated the Korean chess player Li Shishi, also let the "Ai threat theory" start to fill the whole public opinion. There should be no delay on researching AI problems, In order to find solutions, to contribute in the development of artificial intelligence, the author attempts to use the knowledge of its own ethics to analysis reasons of ethical problems in artificial intelligence perspective and dialectical philosophy.

This paper include five chapters: the first chapter is the introduction; the second chapter describes the emergence, development and classification(strong AI and weak AI) of AI technology, and the impact of AI application on human society. The third chapter fully describes the ethical problems arising from the development of AI, including the three suspense of AI: can AI really be endowed with human thinking? Is there a limit for artificial intelligence? And whether artificial intelligence is really controllable by human? At the same time, researchling the main reasons for the three major suspenses which are: whether artificial intelligence should be given "human rights"; Whether artificial intelligence has moral character, employment conflict brought by labor substitution of artificial intelligence. The fourth chapter discusses the causes and essence of the ethical problems of artificial intelligence according to the problems raised in the previous article. The causes include: The limitations of the artificial intelligence technology, the lack of ethical rules, the imperfect supervision system and the misunderstanding of artificial intelligence. The essence of the ethical problems of artificial intelligence is the ethical status of artificial intelligence, the problem of responsibility, the environmental ethics and the interpretation of the philosophy of science and technology. In the fifth chapter, the author gives the corresponding countermeasures through the analysis of the ethical problems in the previous article. The development of philosophy, the ideological guidance of artificial intelligence to improve scientists' moral quality, improve people's scientific literacy to correct understanding of artificial intelligence, through the development of legislation, the development and application of artificial intelligence

#### 西南大学硕士学位论文

technology, scientific management of artificial intelligence products to construct artificial intelligence technology ethics and development principles; the sixth chapter is the conclusion.

In conclution, in the face of the leap of AI technology, we should always do the work of technical introspection, constantly summarize problems and take corresponding strategies, so that we can skillfully make artificial intelligence technology take advantages and avoid disadvantages, and better benefit mankind.

Key words: artificial intelligence, ethical problems, ethical rules

# 第一章 绪论

#### 1.1 选题的背景和问题的提出

最早出现于上世纪五十年代的人工智能是一门研究机器是否能开发出智能的高科技学科。这门学科由计算机科学、控制论、信息论、心理学等等多种学科互相融合共同发展而来。人类通过人工智能进行这样的尝试:用机器的思考来替代人力的思考,以赋予机器思维的方式来分担人类的脑力劳动,可以说人工智能是探讨人脑奥秘的科学的一个重要的分支,同时也是计算机应用领域的一个重要研究方向。它同原子能和空间技术一起,并称为人类社会的三大尖端科技。不断发展的人工智能已经渗透到了人类生活的方方面面,例如军事、服务、娱乐、建筑、经济等等。但人工智能的高速发展一方面为人类提供了更加便利的生活方式,另一方面也带来了许多的伦理问题。人类在享受人工智能带来的便利的同时,也开始反思和探讨人工智能发展背后出现的伦理问题。本文则试图通过了解人工智能的本质,列数人工智能已经出现和潜在的伦理问题,并给出具有针对性的建议。

## 1.2 研究现状的综述

#### 1.2.1 国外研究现状

自 20 世纪 60 年代开始,一个专业色彩非常浓厚的词汇——"人工智能"出现在人们的生活中。随着时代的发展,人们深刻地感受到仿真机器人的发展之迅速,它们将在许多方面远远超越人类智能。"佳佳"作为我国首台特有的体验交互机器人在记者的提问下也可对答如流,人类为机器人的智力水平惊叹不已,并且期待着未来机器人智能达到与人类的智能比肩的程度。但是人们在享用人工智能为人类社会发展带来的方便、快捷的同时,那些与人工智能发展并行的人工智能发展的问题始终牵挂着人们不安和探究的心。人工智能的问题研究涉及哲学、信息论、控制论、计算机科学、系统科学、信息科学、心理学、社会学、生物学等等众多学科领域,学者们从各自领域、各自视角展开研究。

国外有关人工智能的发展问题研究颇丰。例如安德鲁的《人工智能》<sup>①</sup>、门泽尔、阿卢伊西奥的《机器人的未来:类人机器人访谈录》<sup>②</sup>、玛格特 A 博登的著作《人工智能哲学》<sup>③</sup>等。总结提炼国外的人工智能相关文献之后发现,大致可分为四大类。

(1) 对人工智能本质的研究。目前学者们的主流看法是,人工智能是系统地

① 【英】安德鲁.人工智能【M】.刘新民译.山西科学技术出版社,1987

② 【美】门泽尔,阿卢伊西奥.机器人的未来: 类人机器人访谈录【M】.张帆译.上海辞书出版社.2002 年

③ 【英】玛格丽特·A·博登.人工智能哲学【M】.刘希瑞,王涵琪译.上海译文出版社.2005

对信息进行处理的结果,深度地对人脑进行模仿,并在此基础上对人脑进行拓展的一项智能计算机技术。英国计算机科学家阿兰•图灵(Alan Turing)曾明确表示:一台机器在与人交流过程中让人察觉不到自己正在与一台机器交流,才能真正意义上被称作智能。麻省理工学院温斯顿(Wiston)教授则认为人工智能就是一种替代人力劳动的工具,表示"人工智能就是探索计算机怎样才能做过去只有人才可以完成的工作"。 坦福大学费根鲍姆(Feigenbaum)教授对人工智能给出的定义是:"对知识和信息进行规则处理。"人工智能的本质讨论不止出现于严肃的学术领域,在一些科幻的文学作品中也有所呈现。例如,1886年法国维里耶德利尔•亚当作家早在小说《未来的夏娃》中就通过仿人机器"安德罗丁"对人工智能的种种可能性做出设想。

- (2)对人工智能发展未来的展望,大部分学者都认为人工智能有望达到或超过人类智慧。作为人工智能创始人之一的赫伯特·亚历山大·西蒙(Herbert Alexander Simon)在研究过程中认为人工智能能达到人类智能的水平。尝试通过仿照人脑活动思维,将机器大脑活动运行解码的亨利·马克莱姆(Blue Brain)表示,人类大脑在计算机中进行复制运行只是时间问题。美国的雷·库兹韦尔(Ray Kurzweil)认为,人工智能和人类智慧可以同时使用,在未来某时刻,除智能水平之外,在情感和意志上机器也能够具备与人类相近的水平,并认为"21世纪结束之前,人类将不再是地球上唯一具有智慧的生命实体。"<sup>②</sup>
- (3)对人类未来如何应对人工智能不断发展的威胁的研究。在研究人工智能的过程中,我们发现人工智能的发展是一次科技的爆炸性突破,发展速度超出人们的想象。因此一些学者认为人类智慧终将会被人工智能超越。被称之为"人工大脑之父"雨果·德·加里斯就曾在《智能简史一一谁会取代人类成为主导物种》<sup>®</sup>一书中提过这个问题,并且在清华大学中发表讲演时表示:通过数据分析发现,人的大脑的转换速度在低于人工智能的情况下,机器的运算速度是人大脑运算速度的好几十万倍(10 的 24 次方倍),因而雨果得出结论:"未来二三十年人工智能机器可能会与人类成为朋友,50 年后,人工智能会是人类最大的威胁。人类与机器之间的竞争是无法避免的,甚至智能机器会对人类的生命产生威胁"。
- (4)对人工智能伦理问题的研究。国际上对人工智能伦理问题的研究都处于初级阶段,从人工智能应用到生产、生活的各个方面开始,人工智能的伦理问题才逐渐浮出水面。苹果公司高级工程师雷德哈尔曾在国际人工智能研讨会上提出: "人工智能对人类劳动力的替代将大幅提高人类社会的生产力,但我们更应该重

① 【美】费跟鲍勃.人工智能【M】侯怡然译.上海科学技术文献出版社.2008

② 【美】弗里德曼.制闹者: 创造堪比人脑的智能【M】.张陌,王芳波译.三联书店,2001

③ 【美】雨果•德•加里斯.智能简史一一谁会取代人类成为主导物种【M】刘长阳译.清华大学出版社,2007

视这背后对我们可能会失去工作的伦理问题。"针对人工智能已经出现或即将出现的伦理问题,2013 年未来生命研究院正式提出"23 条人工智能原则",第一次提出完整的人工智能伦理准则框架。2017年,欧盟提出要制定"机器人宪章"<sup>①</sup>,其中包含对机器人的民事立法。同时,当代哲学家们也对人工智能的伦理问题进行了热切的讨论,其中被探讨最多的问题是关于人工智能发展的环境伦理问题。例如在艾博尔发表的《Robot Ethics: The Ethical and Social Implications of Robotics》(机器人伦理:机器人的伦理地位及其社会影响)一书中就提到有关于机器人的制造垃圾将会对环境造成辐射影响,这在未来将会是巨大的人工智能环境伦理问题。"这除了会影响我们的健康状况,甚至影响到了蜜蜂的蜂蜜、花粉的授粉。农业的生产率也会受到影响,农业污染将是一个巨大的灾难。"<sup>②</sup>

#### 1.2.2 国内研究现状

与外国学者相比,我国学者对人工智能的关注时间较晚。但是在我国生产力水平与日俱增,经济发展迅速的背景之下,我国学者对人工智能的研究也取得了一定的成就,在一些研究领域成果甚至超过了国外学者。学者们对人工智能领域中哲学问题的分析也是相当活跃。综观国内的人工智能研究,可以大致分为四个方向:

- (1)对人工智能发展史的研究。董军关于人工智能所发表了作品《人工智能哲学》,书中对人工智能的发展历史以及发展过程中面临的各种问题都统一做出了概述,就人工智能所涉及到的哲学概念和问题结论做出归纳,针对思维的科学性和智能工程所掺杂的思想结合中国所特有的禅宗思维进行阐述,详细阐述了智能设备在模拟运作中的作用。赵泽林的《人工智能的基础哲学问题探秘》也对同类的问题进行了研究。
- (2)从语言哲学的角度解读人工智能的研究。徐英瑾的代表作《心智、语言和机器:维特根斯坦哲学和人工智能科学的对话》,从语言的角度以及哲学的立场分析了人工智能科学的发展状况以及目前人类社会的技术掌握情况。使人们领略到维氏哲学在知识表象特征、语言解读、机器人、逻辑推理等方面的贡献,为人工智能的发展与其他学科之间搭建起了一座桥梁。
- (3)对人工智能伦理规则和法律制度的初探。就目前来看,由于人工智能非常具有时代特色,所以在相关的伦理规则和法律制度方面并不太完善,目前我国有许多学者就这一问题进行了深刻的反思和初步的探讨。比如王晓楠的《机器人

①【法】机器人宪章 【N】时代科技报,2018-3-18

②【美】艾博尔.机器人伦理: 机器人伦理地位及其社会影响,周仁华议【M】上海:上海先锋出版社,2012

技术发展中的矛盾问题研究》<sup>©</sup>、张宝生的《人工智能法律系统的法理学思考》<sup>©</sup>、姜潔的《给机器人做规矩,要赶紧了》<sup>®</sup>、唐昊淶的《人工智能与法律问题初探》 <sup>®</sup>等都对这一问题进行了研究分析。

(4)对人工智能伦理问题的初步探讨。近年来,国内伦理学学者尤其是科技哲学学者们开展了一系列关于人工智能伦理问题的研究,发表了丰厚的研究论文成果。例如:郝勇胜的《对人工智能的哲学反思》(科技哲学博士论文),江昕教授的《哲学视阈中人工智能发展的问题研究》等等。这些对人工智能伦理问题研究的论文都有一个共同的基点,那就是用哲学的思维解答、缓解甚至解决人工智能的伦理冲突和问题。

#### 1.2.3 国内外研究评述

在人工智能的技术研究方面,由于我国学者入手研究的时间较晚,所以与国外学者相较,掌握的素材以及研究的现实基础较为薄弱。迄今为止,在人工智能学术研究方面,全世界共同面对的与人工智能有关的问题是:理论体系的不完整、产品使用的不规范。迄今为止没有任何一个有关人工智能的完善的理论体系;关于人工智能产品的使用法律法规的出台也无法同步跟进。究其本质,并不是因为这些伦理问题不够突出,而是人工智能技术本就尚处在初级阶段,对于其出现的伦理困境都还是需要试探性的假设和预判。所以,我们有理由相信即使现阶段关于人工智能的研究完整性不足,但是根据现阶段该技术研究的发展情况来看,研究的成熟化和系统化在将来是能够实现的。

#### 1.3 研究的目标与意义

#### 1.3.1 研究的目标

本文试图通过对人工智能的技术研究为依据进行探讨,而后再一一列举出人工智能对人类社会产生的影响,接着通过对人工智能的哲学反思,总结出人工智能现有阶段所出现的伦理问题以及未来可能会出现的伦理困境,最后通过与我们所处的现实生活进行磨合接洽,创造性的提出一些能够解决问题的对策。总的来说,文章的主旨意在让越来越多地人类有意识的去了解人工智能的本质,消除有关人工智能的不真实言论,同时让人工智能的发展受到道德伦理的约束,坚持"以人为本"的科学发展精神,使得人工智能能够始终走在向善的、福泽人类的发展

①王晓楠.机器人技术发展中的矛盾问题研究【D】大连理工大学, 2011

②张保生.人工智能法律系统的法理学研究【J】,法学评论, 2011

③姜潘.给机器人做规矩,要赶紧了【N】文汇报,2011-6-7

④唐昊涞.人工智能与法律问题初探【J】,哈尔滨学院学报,2007;1

道路上。

#### 1.3.2 研究的意义

#### (1) 理论意义

哲学家恩格斯曾表示: "无论自然科学家持有怎样的立场,最终都要受到哲学支配。事实上: 不论是否受哲学的支配,他们仍旧想在现有理论基础上进行研究。"人工智能技术为社会的发展做出了巨大贡献,但是也给人类带来了很多伦理问题。而社会的发展需要人工智能技术在类的高新技术的支持,这对人工智能技术提出了更高的要求,也使人工智能技术面临更多新的挑战。本文通过对学者们的结论进行借鉴总结,尝试从伦理学的角度对现有阶段的人工智能研究进行深入了解,力求解开人工智能发展中的某些问题。旨在明晰伦理学在人工智能发展中所起的作用,明确伦理规则在人工智能以及科学研究中的指导意义,弥补伦理学对人工智能发展思考中的缺失,丰富人工智能发展的伦理思考,明确人工智能发展的关键所在,对我国人工智能发展战略以及实施提供一定的理论依据。

#### (2) 实践意义

尽管人工智能技术的快速发展引发了诸多伦理阏题,但对于这项正在快速发展中的新技术所带来的伦理问题不必过度担心或恐慌。任何事物都有它的两面性,人类只要能够对人工智能技术可能引发的伦理问题做好防范工作,对于己经造成的伦理问题及时采取拯救措施,做到在人工智能技术发展的过程中趋利避害,这将不仅有助于人工智能技术的顺利发展,也会促进人类社会的不断发展与进步。可见,清醒的认识人工智能技术的伦理问题并施以相应的对策,这对人工智能技术更好的为人类造福具有非常重要的实践意义。

## 1.4 研究的思路与方法

#### 1.4.1 研究的思路

自 19 世纪 60 年代开始,短短 50 年时间,人工智能的发展水平突飞猛进,人类在与人工智能相关应用领域的改革创新也是层出不穷。本课题首先从人工智能技术的概况入手,接着对人工智能技术的主要伦理问题进行了阐述,然后对人工智能技术伦理问题的产生进行了比较客观的成因分析,最后针对这些伦理问题产生的主要原因进行了相应的对策探讨。

#### 1.4.2 研究方法

#### (1) 分析与综合的研究方法

分析与综合的方法指的是在认识的过程中将整体分解为部分以及将部分重新

整合为整体的研究方法。本文利用分析与综合的方法,对与人工智能研究和发展相关的大量资料书籍进行了搜索,搜索领域囊括人工智能技术的研究以及与人工智能有关的学科的发展,然后在这些书籍和资料文献的基础上进行归纳总结,为本文的撰写奠定坚实的理论基础。

#### (2) 系统分析方法

系统分析方法是将一个项目放到系统的运行中进行考察,通过对系统中各个 要素的分析找出各个要素的联系与规律,从而解决项目问题。本文在对人工智能 伦理问题的研究中,将人工智能技术的伦理问题与相应的应对策略看成是一个系统,通过分析系统内部各个要素之间的关系来找出能够解决或减轻这些伦理问题 的应对策略。

# 1.5 创新与不足

#### 1.5.1 创新之处

论文在课题的研究中,意图进行以下几个方面的创新:其一,从研究对象上来说,国内外关于人工智能的发展问题和对该项技术的哲学反思研究成果都很多,但对于人工智能的伦理问题的研究,基本上都是对某个问题进行研究,很少有像本文这样对人工智能的伦理问题作为一个系统来研究,从整体上进行相应的对策研究;其二,从研究内容上来说,由于人工智能是一项正在不断发展中的新兴科技,伦理问题日益突出,相关的政策、法规却很少,造成目前解决人工智能伦理冲突的现实对策不足,本文给出了很多具有现实意义的应对策略,相信这些方法、对策都具有一定的指导意义,能够有效的缓解、解决现有的伦理冲突问题;其三,从研究方法上来说,采用多学科研究方法,将人工智能发展带来的伦理问题放到一个系统的视阈中进行研究,运用哲学思维结合相关学科理论,深入研究哲学给人工智能带来的启示和对人工智能发展的指导作用。在人工智能的技术符号融入哲学思想,用"以人为本"的科学发展观指导人工智能发展。

#### 1.5.2 不足之处

现阶段在人工智能伦理学领域相关的著作和期刊文献等研究成果已经非常丰富了,在对这些文献进行查阅过程当中,笔者发现自身的理论水平非常有限,对于学术界的很多观点无法综合消化与吸收,这就造成了论文的深度不足。同时,迄今为止无论是人工智能技术本身还是人工智能的哲学反思研究都还处于起步阶段,关于人工智能的一些学术研究尚未成型,所以本文给出的一些应对人工智能伦理问题的对策属于初级宏观的层面。

# 第二章 人工智能概述

## 2.1 人工智能技术简介

人工智能,是计算机科学中一门新兴的边缘学科,是研究如何用计算机模拟人脑所从事的论证、显示、鉴定、理解、设计、思考、计划和解决问题。费根鲍勃(E.A.Feigenbaum)教授定义人工智能为"计算机科学中的一个分支,它涉及计算机的智能化,是未来计算机的发展方向。"

从工程的角度来看,人工智能是利用人工开发使机器具有与人类智能有关的功能,如论证、显示、鉴定、理解、设计、思考、计划和解决问题。它是人类的思维和智慧在机器上的反应;从学科的角度看,人工智能是一门研究智能机器建设或实现机器智能的学科,它是人类智能的模拟、延伸和扩展的科学;从学科地位和发展水平角度来说,人工智能是现代科学技术的前沿阵地,它在计算机科学、信息处理、控制处理、心理学、数学、生物学、哲学等等多个领域的基础上发展起来的,是一门新思想、新理论、新技术的新兴学科,是一门综合性极强的边缘学科。

可以说,人工智能就是对人的思维过程的模拟,虽然人工智能不是人脑智能,但能像人一样的思考,甚至有可能超越人脑的思考。人工智能是一门极具挑战性的科学,需要有计算机知识、心理学甚至是哲学知识才能从事这项工作的研究。广义地讲,人工智能技术是关于人造物的智能行为,这种智能行为一般包括知觉、学习、推理、交流以及在复杂环境中的行为,也可以说,人工智能就是让机器能够胜任一些复杂的、人类才能够完成的工作。

目前对人工智能的研究,可以分为两大种类:强人工智能与弱人工智。二者的区分标准在于是否具有"学习能力"。强人工智能一般指的是在接收到外界的刺激影响之后,机器可以产生自己的思维判断,进而产生自我意识,进行一种类似于学习的活动,拥有自我推理和判断能力。也就是说强人工智能机器不仅仅能够完成人类输入的指令,并且可以进行与人类大脑相似的思考与判断活动。例如,对于"美"的思考,强人工智能能够自主发现一幅画的美丽之处,而不仅仅依靠人类向其输入的美的概念来判断,可以说强人工智能拥有属于自己的对美的区分标准。而弱人工智能则是在接触到外界环境刺激之后做出相应的反应,无法产生自我意识,只能按照人们输入的指令做出相应的行动。"根本不存在意向状态,只是受到电路和程序支配机械性地运动而已。也就是说,弱人工智能并没有学习思考的能力,无法拥有自己的判断思维能力。""迄今为止,弱人工智能的应用已

① 【英国】玛格丽特·A·博登著作.人工智能哲学【M】.刘西瑞,王汉琦译.上海译文出版社,2005;83

经比较广泛地运用于人们的大部分指令当中,但如果涉及到一些含有较多自我意识的指令,如音乐、文笔、颜色等,则需要强人工智能来完成,但是目前强人工智能领域暂无较大的突破。

#### 2.2 人工智能的发展

人工智能的真正形成的标志可以说是1956 年在美国的达特茅斯召开的关于利用计算机模拟人类智能行为的学术会议,会议由数十位数学、心理学和信息论方面的学者参加,也就是在这一场会议上第一次出现了"人工智能"一词,标志着人工智能第一次作为一个新兴学科正式诞生在人们面前。而在1969年举办的国际人工智能联合会议,标志着人工智能这一学科已在全世界范围内得到了公认。但在这之后,人工智能由于学科过于新颖并且无重大科技成果问世,经历了一段时间的低潮。直到1997年,IBM 公司研制的"深蓝"计算机在国际象棋大赛中击败了国际级象棋大师卡斯帕罗夫,这是人工智能第一次在棋类博弈中赢过人类智能,世界再次开始热切关注人工智能,人工智能自此开始全面走向人们的生活当中。

总的来说,人工智能的发展阶段可以分为五个大阶段:

#### (1) 萌芽期(1956年之前)

在1956年之前,人类处于对人工智能的展望时期。人们企图通过工具替代自身的脑力劳动,但限于当时的技术和生产水平,人类仅仅处于对智能的初级探讨与展望,并未触及到人工智能的研发。

#### (2) 发展期(1956年-1978年)

1956年达特茅斯会议之后的几十年,是人工智能发展的黄金时期。在这段时间内,计算机出现并且开始被应用于解决代数应用题、证明几何定理、学习和使用英语等,这些应用使人们对计算机的能力赞叹不已的同时也让人们对人工智能的发展前景充满信心。当时学者们普遍认为,"二十年内人类就可以研发出可以完全替代人类工作的机器人"、"人类在四到八年内就能得到一台和人类智能完全相同的机器人"。期间有许多重大研究成果问世,例如,1958年,约翰麦卡锡(John McCarthy)开发出编程语言Lisp,之后这种编程语言就成为人工智能研发中的主流编程语言。1959年他又提出了"Advice Taker"(假想程序)概念,这也是第一个完整的人工智能系统。

#### (3) 低潮期(1966年-1974年)

这段时期,人工智能的发展遇到技术瓶颈,导致人们当初对人工智能发展的 乐观期望遭到严重的打击。当时由于计算机的复杂指数指数级上涨,数据量缺失 等问题,一些关于人工智能的难题好像完全找不到答案,因而人工智能开始广受 批评,研究经费也开始转移到其他项目上。马文明斯基就曾说到: "我们认为研究已经完全停滞,因为人工智能的基本理论几乎是一片空白,我们对感知器做出了大量的研究和实验,但我们完全摸不清它的工作原理。"<sup>①</sup>

#### (4) 第二波发展浪潮(1975年-1985年)

这段时期内,日本关于人工智能的研究迎来了新的小浪潮,期间日本经济产业省拨款八亿五千万美元用以研发第五代计算机,目标是研发出能够像人类一样处理图像、文字信息并且能够进行逻辑推理的计算机。随后,英美等国也开始纷纷效仿,在AI(人工智能)的研发中投入大量资金。1981年,"专家系统",即一种知识处理、信息筛选的智能系统被研发,人工智能开始迎来短暂的春天。但在1985年到1989年,由于知识处理的发展再遇瓶颈,人们对人工智能的研发再度失去热情,研发一度停滞。

#### (5) 高速发展期(1989年-至今)

1989年以后,随着计算机技术、人工神经网络以及行为主义等研究的深入开展,人工智能技术不断得到突破,也越来越广泛的应用到人类社会当中,人工智能开始进入黄金告诉发展期。在此之后,人工智能的研发成果日益丰硕:1989年贝尔实验室的研究人员开始将反向传播算法应用到多元神经网络,人工神经网络的研发得到重大突破;1995年理查德研发的聊天机器人"爱丽丝"让机器人第一次拥有了自己的自然语言样本;1997年,IBM公司研发的"深蓝"计算机在国际象棋大赛上打败了人类象棋冠军等等。

而我国人工智能技术起步较晚,1978年人工智能技术才开始纳入国家研究计划,我国人工智能技术自此受到重视并开始了大量研究: 1984年我国第一次召开全国人工智能学术讨论会; 1986年,我国将一系列智能化重大项目纳入国家高技术研究计划; 1993年,将计算机系统与人工智能列入国家科技攀登计划。随着我国对人工智能技术的国家政策的逐步引导,人工智能取得前所未有的发展成果,创造了许多具有国际领先水平的创新成果。现在,国内已经有成千上万的研究人员和大学教授学者从事人工智能技术的研究工作,人工智能技术将为我国的现代化建设做出重大贡献。

# 2.3 人工智能的应用领域以及对人类社会的影响

#### 2.3.1 人工智能的主要应用领域

人工智能作为一项新兴的科学技术,在近代以来许多领域中上取得了优秀的 研究成果,建立了具有一定机器智能的计算机系统。同其他的学科一样,不同的

①【美】神经网络的局限性 马文明斯基,刘清明译【D】北京:人民大学出版社,47

学者对于人工智能的应用领域有着不同的分类方法,但尽管分类方法不同,也有 着一些被广泛公认的分类,本文将介绍一些人工智能的主要应用领域。

#### (1) 专家系统

专家系统是一种比较智能的计算机系统,可以运用知识和逻辑来解决一些具体领域内的重大问题,因为这些问题只有学识经验丰厚的专家才有可能解决,因而将其称为"专家系统"。例如在医疗、金融、统计领域都已广泛应用到专家系统。专家系统的特点在于拥有着海量的专家知识,可以通过模拟专家的知识思维来解决复杂的问题,得出同专家同样的答案。

#### (2) 普通语言理解与机器系统翻译

人工智能最早期的研究领域是语言的处理。人与人之间的交流和沟通是非常容易的,要建立一个能够产生和"理解"一个单词的计算机系统却是非常困难的。因为通过思维的"结构"点需要与思维结构的结构相关的巨大的公共思维,这就像一个人,需要对其背景进行梳理,才能够对其有深入的理解。现如今,在各大领域,计算机翻译已经有所运用,但机器翻译依旧无法完全摆脱人工的复查和整合。主要存在两个方面的问题:首先是翻译的准确性问题,第二是翻译的速度问题,翻译需要一个非常庞大的系统库,对于该系统库的信息要及时地予以更新,确保翻译的速度以及翻译的效率。在系统库的信息准确性和有效性得到确保的情况下翻译的效率和准确性也能够得到保证。

#### (3) 人工神经网络

人工神经网络的研究起源于上世纪40年代,网络神经指的是通过多个处理单元在同一种方式下单元之间相互连接而构成计算机系统。人工神经网络通常被用于对人类的大脑神经活动进行模仿,实际上这是一种对生物神经活动通过计算机进行模拟的方式,对于很多人类神经才能处理的问题进行处理,例如对语言的识别、对金融业经济学的理解等等。如果从智能的角度进行区分,智能的不同类型是人工神经网络和专家系统之间最显著的区别。专家系统偏重于对人类智能信息处理的模拟,而人工神经网络则偏重于对人脑神经结构的模拟。

#### (4) 模式识别

模式识别指的是对于事物的种类或者模式,通过计算机进行判断和分类征同类模式,或者种类归入相同的,或者相似的模式当中去,在很多领域,这一功能均有所运用,例如文字的处理、音乐的抽象归类、生物物理实体的归类等等。迄今为止,在对图像和声音的识别归类领域当中,模式识别的应用最为普遍。举例来讲,在手写文体的识别归类领域当中通过模式识别,将手写文体直接归类转化为印刷体,又例如将声音直接识别转化为字体。很多通讯软件使用的语音转化为文字功能就使用到了这一模式。

#### (5) 机器人智能化

在人工智能应用领域当中,智能机器人是一个非常关键的领域。这一领域主要包括对机器人设备的研究以及对智能机器人本身的研究。目前人类已经建立起了一套复杂的机器人系统,机器人会根据事先编程好的指令,对一些重复性工作,机械性的进行执行。大多数机器人只能工作,不能思考,如何使机器人获得思考能力,是机器人研究领域以及实践领域中的一个重要的发展方向。因此今后人工智能研究方向,可以说是以人工智能的个性化思维为主,怎样赋予智能机器人个性是今后研究的重点所在。

#### 2.3.2 人工智能的发展对人类社会的影响

2014年以来,人工智能领域爆发了一场巨大的科技改革,AI(人工智能)对人们生活的影响可已经到了"无孔不入"的程度。人工智能的成果被广泛的应用到人们的生活、生产当中,例如安保系统、图像声音的识别系统、无人驾驶车辆、无人驾驶飞机、金融分析等。总的来说,人工智能不仅改变了人们的生活方式,而且改变了人们的生产方式,进而影响社会结构以及经济生产。笔者将从人工智能对人类社会的影响、人类社会结构的变化以及人类思维方式的变化、对人类经济效益的提高和对人类生活方式的改变几个方面来思考人工智能对现代社会的影响。

#### (1) 人工智能改变了人类的社会面貌

人工智能发展对人类社会面貌的改变可以从以下几个方面说明:第一,人工智能使财富快速增长。与过去相比,人工智能的劳动力能够生产出原本人类劳动力数十倍甚至数百倍的产品;第二,人与人之间的联系越来越紧密。人工智能对国家的社会制度产生了非常大的影响,人际接触频率明显增高,人与人之间的距离缩短。整个世界被互联网连接起来,在这种情况下,人们的生产不再是孤立的,而是密不可分的。贸易过程当中相关主体都是紧密的相关方,金融交易变得更加频繁,教育国际化水平越来越高,国家之间、地区之间的政治交流与沟通变得更加频繁;第三,信息快速增长、快速更新。一方面,人们每天通过计算机输入的信息以几何级数的形式增长,而另一方面人们也通过电脑快速获得大量的信息。人们期待信息传播,为了满足这种需求,越来越多的人通过人工智能收集和输入新信息。

#### (2) 人工智能改变了人类的社会结构

近十多年来,社会结构因人工智能的广泛应用已经悄悄的起了变化,人一机器的社会结构正在向人一智能机器一机器发展。智能机器人就是人工智能的一种。 从发展的角度来说,医院的护士医生、商店餐馆的"服务员"、办公室的秘书、 指挥交通的交警等直观重复的劳动工种都将被人工智能替代。因此,人们必须学 会与智能机器和谐相处,并适应与智能机器打交道的人的社会结构。

#### (3) 人工智能增加了社会的经济效益

人工智能创造了巨大的利益,专家系统就是一个典型的例子。伴随着计算机 零部件的市场化与技术的相对成熟,市场推陈出新的速度不断加快,计算机的价 格变得越来越便宜了,而对于人工智能来说,计算机是其组成部分,计算机零件 价格的下降,必然会降低人工智能的成本,从而产生更大的效益。成功的专家系 统可以为其建设者、所有者和用户带来显著的经济利益。由于软件有很多共通的 地方,专家系统可以广泛传播专业知识和经验,并推广有限且昂贵的专业人员的 知识。如果得到适当的维护,软件的寿命可以维持上百年。因此,人类专家的经 验可以在没有人类专家生命的限制的情况下得到扩展,这是另一个巨大的财富。

## (4) 人工智能改变了人类的思维方式

人工智能使人类的思维活动被延长,有利于人们应用正确的思维方法。人工智能的运作遵循客观规律,需要人们在应用和发展中采取一种务实的态度,而不是感性的,容不得一点感情用事。人们传统的思维方式已经固定了,新产生的人工智能将会影响到传统的思维,并且最终改变人们的思考方式。例如,传统知识通常印刷在固定的书籍、报纸或杂志上,因而这些知识都是固定的、一成不变的。而人工智能系统知识库的知识可以不断地修改、扩充和更新。知识不再一成不变,而是时刻在更新与进步的,这将打破人们以往的僵化思维。作为另一个例子,一旦专家系统用户开始相信系统(智能机器)的判断和决定,他们就可能不想多思考,过于依赖专家系统,失去了解决问题的能力。那些过分依赖计算器的学生,也会大大减少他们的思维和计算能力,这也是一种依赖性。在设计和开发智能系统时,应考虑上述问题,并鼓励用户积极参与解决问题的过程。

#### (5) 人工智能改变了人类的生活方式

人工智能应用主要体现在两方面: 其一,遵循内涵深入; 其二,外延拓展。现在全世界每年有几十万个成熟软件推向市场,而无论是商务的还是娱乐的,都改变了人们单调的工作和生活模式; 其次,外延拓展主要是人工智能与其他技术结合, 光子、超导和激光计算机是人工智能发展的新出路。在社会广阔领域中,人工智能借助机电光声技术,为社会提供了电子排版系统、电视节目编辑器、复印机、学习机、家庭影院、音乐喷泉、B超检查、CT检查和机器人,给人们生活带来了新的气象。

# 第三章 人工智能发展引发的伦理问题及成因分析

#### 3.1 人工智能发展引发的伦理问题

人工智能给人类带来了很多便利,这是无法否定的。但同时也带来了很多问题,例如,随着人工智能数量的不断增加,是否会完全替代人类完成各种工作?那些被人工智能替代的人会失去工作无法供养家庭。而且人工智能的智能水平达到一定标准之后,人类对人工智能技术的过度依赖可能会导致的身心康及精神上的伤害;人工智能技术引起的所有权问题、法律纠纷等。例如在劳动替代的问题上,随着越来越多的机器人被应用于各个行业,造成了劳动力需求的大量下降,很多工人被迫下岗,对劳动力市场产生非常大的影响。亚马逊网站最近购买了机器人来从事仓储搬运工作,这就大大降低了对工人的需求。此外,在非制造业当中,也能看到人工智能的影子。例如,律师能够使用智能程序来代替案例研究,使人工智能成为了自己办案的助手。《福布斯》杂志发表了一片关于智能软件程序来取代记者的报道,大量的数据表明,现在税务、财务、旅游网站都用人工智能软件取代了大量的员工,人工智能能够取代大部分的统计、分析工作。所以,像其他新兴技术一样,人工智能技术也有两面性,人工智能的伦理问题是无法回避的,是不应当被人类所忽视的。

#### 3.1.1 人权伦理问题

自从人工智能一词出现,便引发了关于人权问题的探讨。要探讨人工智能中的人权问题,首先应该了解人权的概念以及人权适用的范围。对于人权的概念,目前还没有一个统一的说法,人们对于人权的界定分为三大类:人权是法律的、政治的、道德的<sup>©</sup>。关于人权的概念也可以引用我国著名学家甘少平的定义:"人权是指每个人在一定的社会和历史条件下享有或应该享有的基本权利。"所谓人权,指的是无论在何种历史条件下,每个公民都应该按照他的尊严和本质享有的基本权利。或者说,人权指的就是每个人都拥有平等生存和自由发展的权利"<sup>©</sup>。14无论种族、肤色、国籍、性别、年龄,每个人都拥有人权,没有任何国家或者组织可以无视人权的存在。如果从伦理学的角度来解释,人权则指的是:"人权伦理是基本人权,它包含伦理道德,以及人权体系中的一切事物,体现道德、价值观、伦理关系和伦理原则的道德规范,是伦理道德的总和。"<sup>®</sup>人权以及人权伦理的主体性表明了人权的主体必须为"人",必须要承认人的主体地位,人权的

①张志.当代中国人权伦理研究.【D】.广西师范大学,2004,4-7

②甘少平.人权伦理学.【M】.中国发展出版社,2009: 2

③张志兵.当代中国人权伦理研究.【D】.广西师范大学,2004:10

实践性和普遍性才能体现出来。

随着人工智能的发展,各种类型的人工智能相继诞生,然而这种人造的产品颠覆了人们的传统思维,使得"人权"遭受了挑战。当人工智能从仅仅是从事简单体力劳动的机器,在不同程度地被赋予了情感和"人性"后,智能机器人领域的应用更加广泛,机器人甚至可以起到"陪伴"的作用。但随着这些智能机器人不断被投入使用,智能机器人面临着日益严峻的人权伦理问题。

美国学者库兹韦尔曾经在书中写道,"可能在2029年会出现拥有自我意识的非生物体(也就是人工智能),他们可能会存在一定程度上模拟人类或者与人类类似的微妙的感情。"<sup>©</sup>现代社会的许多人工智能计算机,都被输入了类似于"表情"或者"情感"的信息。例如人类常见的苹果手机中的siri(人工智能助理软件)就被注入了感情的色彩。人类可以通过语音与siri对话,并且可以调节siri的性别,得到两种不同声线的回答。例如在问询siri"你叫什么名字"时,siri的回答是"我叫siri,幸会幸会"。这个回答就被注入了人类的情感色彩,使得该人工智能助理软件拟人化。这些拥有了人类情感色彩的拟人化人工智能,引发了许多关于人权的问题。既然人工智能可以拟人化,给人类带来与和人类交流相似或者相同的感受,那么人工智能是否可以被赋予"人权"?对此问题讨论非常激励,有人反对说:"人类不需要强大的人工智能,人工智能同时不应当被允许拥有与人平等的权利,如果人类让计算机拥有所谓的人权,那么它就违反了机器人的三条原则。"<sup>®</sup>而赞成者则认为,强人工智能的出现是不可阻止的,没有人应该去阻止科技的发展,如果人工智能可以拥有一定的道德修养,那么人工智能就可以被赋予和人类相似的权利。

2017年10月,在沙特阿拉伯举行的投资计划会议上,研发的机器人被授予沙特阿拉伯国籍。大会主持人对机器人说道,"你是首个被被授予沙特阿拉伯国籍的机器人"。机器人对此回答到:我感到莫大的荣幸,是阿拉伯政府赋予了我生命,感谢阿拉伯政府,我将会永远终于阿拉伯政府,为阿拉伯政府效力终生。这是首个被授予国籍的高拟人化机器人,也是第一个机器人被"赋予"人权的例子。沙特阿拉伯的公民每年从巨大的石油销售中得到一笔可观的收入。外国人很难获得公民身份,尤其是那些在恶劣条件下工作的外来务工人员获得公民身份的机会微乎其微。而"索菲亚"获取了沙特阿拉伯的公民身份,也就意味着"她"将享有作为沙特阿拉伯公民的所有权益。"索菲亚"是一个高拟人化的机器人,硅胶做成的脸部可以模拟62种人类表情,并且能够通过计算分析、识别人类的面目,

① 【美国】.库兹韦尔.《如何创造思维》.【M】.盛燕杨译.浙江人民出版社,2014:195

②黄建民.我们要给机器人以"人权"吗?【J】.读书与评论,2009:6

与人类做眼神和情感、语言的交流。"索菲亚"的设计师汉森说过,"索菲亚"的目标就是像人类一样的生活,拥有同样的意识、创造性和其他能力。那么如此高拟人化的机器人,是否应该拥有"人权"?从人权的基本要求来说,一个人拥有了全人属性才能拥有人权。全人属性包括人的自然属性和社会属性,即自然繁衍的生命属性和社会关系属性。例如克隆人就是拥有人类的基因、人类的所有听力、视力、思考能力等,但缺少了自然繁衍属性,即克隆人是通过基因的复制而生,并非自然两性繁殖,从这一角度来说,克隆人就无法拥有"全人属性",因而不能拥有人权。同样的,现在的机器人都还停留在弱人工智能的层面上,即使"索菲亚"拥有硅胶脸部,几乎与人类无异的表情,但"索菲亚"实际上是机器,并非人类的肉体,即使"她"在某一些逻辑层面与人类相似甚至相同,但"她"的本质还是机器,并不能拥有"全人属性",所以就不该被赋予人权。但未来人类社会是否能够真的出现强人工智能的机器人,甚至是拥有"全人属性"的机器人,我们不得而知,对于是否能够应当赋予机类人器人人权的问题也还需进一步探讨。

未来机器人会越来越多的与人类互动,类人机器人将穿梭在人类之间,甚至有一天以朋友的身份与人类共同生活。那么对机器人的人权问题就不能单纯地否定,对于一些真正像人类一样拥有全部人的属性的机器人,人类是否应该赋予它们一定的人权,笔者认为是有必要的,而给与多大程度的"人权",则需要人类进行"度"的思考。

#### 3.1.2 道德地位问题

智能化随着社会的进步已经成为了现代科技的发展方向,智能化是"互联网+"的产物,人工智能在很多领域都产生了影响。在生物领域人工智能能实现对人类身体结构的改造;在社会领域,能改变社会内在联系;在生态领域,还会改变人与自然的关系。同人工智能技术给人类带来的人权伦理问题类似,人工智能技术也给人类带来了道德地位伦理问题。生态哲学的代表人物利奥波德认为,整个生态系统都应该具有道德地位,即拥有生存和发展下去的权利,但不是每一个生物都拥有"神圣不可侵犯"的权利。<sup>©</sup>鉴于此,对于人工智能技术创造的智能机器人,人工智能的出现,我们就不得不为其进行道义上的权利申辩,尤其是当具有人类情感的"全人属性"机器人出现的时候,当他们具有相当的"人性"的时候,他们就该拥有相应的道德地位。道德地位随之而来的一方面是道德权利,一方面是道德责任。

如雷根所说道德权利的直接体现和根本实质就是尊重,这是由于尊重是人类

① 【美】.利奥波德,沙乡年鉴【M】侯文慧译.长春, 吉林人民出版社,1997,40

进行社交的起码礼仪和基本权利,是获得和行使其他权利的基础,同时尊重应当是双向的、彼此的<sup>©</sup>。如果一个社会成员被赋予了社会地位,那么他便拥有了道德权利,如果人类在与人工智能的相处过程当中,否定其道德地位,人类就可以肆无忌惮地对待人工智能。可以是友善的,也可以是恶劣的、残酷的。上文已经分析过对于拥有"全人属性"的人工智能应当赋予其一定的人类权利,也就是道德权利,对于人工智能的处境,人类应当设身处地地考虑。当跟人类一样拥有"感情"的类人机器人受到虐待,就如同动物被虐待一样,是人类道德观念、价值观所不允许的。身为人工智能的主人,人类可以利用人工智能完成很多工作,人类是否有足够正当的理由和资格来要求机器人一直服务人类,满足人类所有的要求?人类一直将自己定位为机器人的主人,但对这种定位从来没有反思过。在人工智能的发展越来越接近人类智能的同时,需要充分考虑人工智能的道德地位、诉求。综上所述,人类要尊重和重视人工智能的权利,特别是其道德权利。

如前文所述,当类人机器人真正出现时,机器人会拥有自己的思维而不仅仅是人类的符号输入,那么它们在进行思考时就不仅仅只有对信息的处理分析,可能会对事物拥有自己的判断,那么它们在判断对错时就应该建立起自己的道德准则,在拥有道德权利的同时,履行相应的道德义务。总的来讲,道德义务可以归纳为人们在特定的社会关系当中所应当选择的道德行为。也可以说是在社会关系中,其他人所承担的道德义务。对于他人的道德权利,任何人都不能侵犯。例如;人工智能的管家机器人在工作过程当中,会对人类的日常喜好进行记录,从而配合人类的日常生活。这种设计是非常人性化的,使人类的生活在智能机器人的管理下井井有条。但这并非是不存在风险的。例如,对于管家机器人使用者的隐私安全的保护,就存在一定风险。通过一个特定的数据盘,智能机器人将包括使用者私密信息在内的一切信息进行存储,对于拥有一定技术的人而言,访问并获取其中的信息是轻而易举的,这对于机器人用户而言,存在着非常大的安全隐患。所以,机器人不光有智能管理的功能,还具有盗窃能力。一旦有人恶意利用这一点,便能够盗窃人们的隐私。在这种情况下,就需要这些管家机器人具有一定的道德责任意识,在面对任何指令时,都应当以维护主人的权益为准则。

#### 3.1.3 责任伦理问题

责任一词包含着十分丰富、深刻的含义。"责任,是随着人类社会的产生而产生的,同时也随着人类社会的进步而日益的明确、具体。"<sup>②</sup>目前学术界在定义责任含义的过程中分为了两派。第一派以张文显、孙笑霞为代表,他们对责任进

①雷根.打开牢笼,面对动物权利的挑战【M】马天杰译.北京:中国政法大学出版社,2005,65

②程东峰.责任伦:关于当代中国责任理论与实践的思考.【M】.北京,中国林业出版社,1995,13

行了法律解读,他们更加注重责任的综合意义。张文显学者就曾这样定义责任: "第一,责任应该理解为分内应做的事,这种责任的核心意思应该是义务; 第二,责任是特定的人对特定的事的发生、发展产生的积极作用,也就是承担事务中的责任; 第三,责任是在未履行法定责任的情况下应当承担的不利后果,因此,责任可以理解为违约责任。" <sup>®</sup>另一派将法律和政治、道德结合在一起,在此基础上具体阐释了责任的定义,可以理解为"责任就是理解和维护人的绝对尊严、绝对权力、绝对责任的学问。"从哲学角度来看,因果关系是责任的前提条件和产生基础,而责任又是因果联系的具体反映。有行为就会产生相应的结果,对社会产生影响。同时这些行为结果和影响是可以预见的,因为人类对自己的行为具有控制力。<sup>®</sup>

机器人发明于20世纪50年代末,经过几十年的发展已经取得了巨大进展,并迅速席卷了人类社会生活的各领域,且相关智能技术更新也越来越快。在功能分类设计方面也日益细化,如专门陪伴人的伴侣机器人、用于餐厅服务的餐厅型机器人,目前可应用于医院医疗工作的辅助机器人也已取得巨大成果。但人类在自豪的同时,也应当就机器人的高度智能化进行思考和警惕。随着技术更新速度的不断加快,机器人的智能化和自学能力不断加强,在达一定程度后是否会发生"质变"?若具备人类思维的机器人不再完全服从命令,对人类生产、生活而言可能会造成严重影响<sup>®</sup>。

综上所述,确定机器人负责人至关重要,确定机器人的责任能够确保日后在有问题出现时,有人能承担责任和及时处理、解决问题。在技术更新日新月异的现代社会,技术伦理制度的制定要跟上技术的发展步伐。如果制度建设滞后于科技发展速度,将会引发一系列列社会问题和伦理问题,对人类来说也将面临新的危机和压力。人工智能的基础作用就是代替人类劳动力,从事一些体力活动,这就必然会导致很多人失业,谁将对此负责?再如,医疗行业的人工智能,人类无法断定其治疗操作是否绝对不会出现问题,如果对患者造成伤害,谁将对事故后果负责?因此大规模推广人工智能前,要先要界定好责任,否则将为人工智能的后续发展埋下巨大的隐患。电影《人工智能》在开头借科学家的角色向人类发问:"人类具有双向的'爱'的本能和向往,在智能机器人的设计中"爱人类"是其基本设定,这种"爱人类"是否需要双向地去回应?还是理所当然地单方面接受?"但是比起是否要回应对方的爱,更重要的是人类要对自己的创造行为负责。而关于规范或改善人工智能技术伦理问题相关方面的伦理规制与伦理原则的制定及更

①张文显.当代社会责任理论【M】.北京人民出版社,2008,15

②曹南燕.科学家和工程师的伦理责任,《哲学研究》2000年第一期

③姜鹏.给机器人做规矩,要赶紧了?【N】.文汇报,2011.6.7.11

新的速度远远没有跟上人工智能技术发展的速度,这使得人工智能技术带给人类的责任伦理问题显而易见,人类正面临着各方面的责任伦理压力。特别是由谁来负责人工智能技术引发的责任伦理问题是争论的焦点之所在。

此外,责任伦理还强调了对于现在和将来的负责,也就是说人工智能不光是对于当下人们的使用不产生相关的危害,这种危害更不能延续到下一代人的身上。人工智能的负责者不仅包括创造者和市场使用者,还包括政府部门,全社会都应当自觉承担起对人工智能发展的责任。现代的技术是不可逆的,这就警惕我们要慎重的对待这一问题,不能将人类的未来当作"赌注"。人类不能以未来的可持续发展为筹码来忽视人工智能的潜在问题,技术和时间具有不可逆性,当下的忽视必将造成不可挽回的后果。当前的舆论普遍认为,人工智能科技的突飞猛进可能会对人类造成许多新的伦理问题,而传统伦理学对于人工智能技术中出现的伦理问题是无法涵盖和应对的,尤其是责任伦理问题。所以,为了人类的长远发展,人工智能科技术发展所引发的责任伦理问题不容忽视,我们要给予人工智能技术引发的责任伦理问题相应的关注,且进行深入的思考和探讨。

#### 3.1.4 环境伦理问题

在世界范围内,环境运动始于二十世纪中期。而引起这一运动的原因来自于 《寂静的春天》一书中对世人的警醒,书中对于被污染破坏后生灵涂炭的环境进 行了详细描写,使得人们第一次意识到环境保护问题。人们开始反思人类的发展 进程,谴责以往的工业发展以自然资源和环境的牺牲为代价,这种舆论氛围形成 了对工业发展的道德约束。也就是从那时候起,环境伦理开始被人们所关注。环 境伦理,实际上就是一种人权问题。随着科技的高速发展,各国间的经济、政治、 文化的交流不断深入,世界已经成为了一个巨大的"地球村"。环境的污染已经 不仅仅是局限在一个国家内, 例如核辐射、全球变暖等问题都是跨越地域的, 无 国界的。同时,热爱和保护自然是环境伦理的基本要求,追求人类与自然的友好 和平的相处,坚决反对发展经济的同时牺牲自然的做法。环境伦理包括可持续发 展、人与自然和谐共存和环境平等三个原则。第一个原则主导着环境伦理学的发 展,使其基础架构更加科学合理,具体体现在三方面:一是遵守资源公平分配原 则,这种公平覆盖代内、代际两个阶段;二是坚持持续发展为前提的原则,诸如 水、土壤、煤炭等资源的储量是有限的,人类需要牢记这一点并在发展中落实这 一原则,合理地安排发展规划并有效的利用这些资源;三是共生原则,任何危机 的出现都必然伴随着其他危机的出现,可持续原则要求人类要重视自然危机,确 保人类与资源、环境和谐共生。另一个人与自然和谐共存的原则认为物种之间存 在着即互益又相互制约的关系,这是维系生态平衡的关键,也是环境伦理理论的

本质要求和终极目的,该思想原则以是否对人类和自然有利的双重标准作为衡量准则<sup>①</sup>。此外,环境平等是科学环境伦理学建设的又一重要原则,它是人类道德的延续和延展,启示人类在发展和建设过程中关注并重视人与自然关系,在人与人之间的关系、人与动物的关系中都强调和谐平等、共同发展。

显而易见,任何一个新兴科技的出现,都会对环境造成一定的影响,人工智 能作为一种高科技出现,势必会对环境造成影响。高新技术能造福人类这是毋庸 置疑的,但在发展过程中人类似乎越来越远离了科技创造初衷,科技逐渐走向了 异化,引发了资源浪费、环境污染等全球性问题,而人工智能的发展也难逃这种 命运。当今时代人类的生活生产到处充斥着人工智能技术,对其依赖性还将不断 加强,这代表着未来人工智能技术对资源材料的需求还将增大。人工智能技术在 人类社会的应用不胜枚举,以工厂机器人为例,近几年全球范围内很多企业都计 划为其工厂购进了大量机器人,以此来替代人类劳动力,利用智能机器人将能为 公司发展节约成本并获取更多利益。若当前经济的发展速度持续下去,未来智能 机器人的需求会激增,将大大提高各方面产出效率,但带来的如资源紧缺、社会 矛盾激化、社会不稳定等负面影响也将日益扩大。另外,智能产品升级换代的周 期越来越短,由此产生的产品垃圾问题需要引起人类高度重视。以宇宙开发为例, 从功能来看智能技术在宇宙开发中发挥着不可或缺的作用,为科研提供了技术支 持;然而从环境保护角度来看,它在制造越来越多的太空垃圾,太空垃圾的威胁 是不可估量的,一旦和地球碰撞将造成巨大损失和生命威胁,对其他星体的正常 运转和飞行器的安全航行也会产生不良影响。

环境伦理是一门将伦理和自然相结合的新学科,其核心是尊重自然价值与权利。环境伦理在研究和掌握人与自然相互影响的规律基础上,对现代社会存在的环境问题进行了深入分析,倡导人类通过道德约束技术,从而实现人与自然环境的和谐共生。用智能技术关爱自然、为人类谋幸福是该学科的主要思想。它要求科研人员自觉将关爱和保护环境的概念贯穿于科技发明设计的全过程中,用道德约束技术<sup>②</sup>。不遵循环境伦理思想而进行科技创新,是产生科技与环境伦理矛盾的根本原因。另外一个重要因素也严重阻碍了人工智能技术的进一步发展,即关于人工智能的伦理思想发展很慢,远远滞后于技术的进展,两者的不协调限制了技术进步的空间。通过前文分析可知,人工智能技术需要相应的环境伦理来指导和约束,片面发展技术的想法和行为值得全人类警惕。

①余谋昌.环境伦理学【M】.高等教育出版社,2004,239

<sup>&</sup>lt;sup>②</sup>格光民,环境伦理与科技发展的关系【D】.西北大学, 2011, 23

#### 3.1.5 代际伦理问题

在现代化社会家庭结构发生巨大变化的背景下,代沟使得代与代之间的冲突显著加剧,辈分不同的人之间隔着一条深深的鸿沟,而导致这一现象出现的主要原因是科技的发展及其全球化的不断加速。人类的代际关系,随着社会结构和关系的不断变迁,也发生了翻天覆地的变化。代际伦理是一门新兴的交叉性学科,但是针对什么是代际伦理这一问题,迄今为止还没有一个明确的定义。但是还是有一些学者尝试着从不同角度对其进行界定。汪加堂教授认为世界范围内一些学者在对代际伦理研究过程当中主要是从家庭伦理的角度进行阐释,说明世代之间伦理关系以及在家庭结构当中代际伦理的调节功能。廖小平教授认为,代际伦理实际上就是人类代与代之间伦理关系和规范的总称。现阶段,很多学者都比较认同廖小平教授的论述。

在理论研究领域对于代际伦理的研究,直至近几十年才取得了较为显著的成就,但是理论会随着社会的发展而不断变化,代际伦理关系随着市场化、全球化、信息化的不断加速,而产生相应的变化。因此,新的代际伦理随着社会的不断变迁,必将会接踵而至,尤其是在社会文化和社会变革的加速会使得代际伦理问题讨论的激烈程度日益加剧背景之下。相对于其他伦理形态和道德规范,代际伦理同样也有自己的基本准则:"道义与功利相统一、和谐与整体相一致、生存与发展相协调是代际伦理最鲜明的三大基本原则。"<sup>®</sup>很显然,代际伦理所具有的这些原则可以作为分析或界定人工智能技术代际伦理问题理论上的依据和支撑。

现代社会主要面临的代际伦理主要体现在以下几点:第一,道德、价值观的代沟,以及道德价值观的沟通问题。第二,现代家庭代际伦理。第三,可持续发展过程当中的代际伦理支持问题,以及代际伦理公平问题。第四,全球化背景下的代际伦理对于社会和谐稳定以及人类可持续发展而言,以上代际伦理问题与其息息相关。所以对于这些代际伦理进行严格的审视是非常重要的。人工智能在发展过程当中,也带来了较为复杂的代际伦理问题,这是无法避免的。通过上文分析能够看出,在某些方面,人工智能已经违反了一些代际伦理原则,例如,"和谐与整体相一致原则"以及"生存与发展相协调原则"。首先,当今世界发展过程当中,和谐是主题,也是人类社会共同追求的目标。其次,对于人类而言,可持续发展也是其重要目标,而当看到因人工智能技术而导致的各种生态的或是给人类造成的精神上的恐慌等这些这明显破坏代内和代际间的公平的事实时,人工智能技术代际伦理问题不捅自破。

当然,人工智能技术带来的代际间的伦理问题远远不止这些,比如,如果有

①廖小平.论代际伦理及其关涉视域和基本原则【M】.复旦学报(社会科学版), 2004, (02):101-107.

一天当智能机器人同人一样具有生存权或是人权的时候,那么人工智能机器将有可能给人类带来"代际"的困惑。智能机器人本身是模仿人的智能,那么不可否认,他们也会如人类一样拥有自己的后代。如今,智能型机器人正在进入"类人机器人"的高级发展阶段。比如有些科学家正在从事一项研究:他们设想给智能机器人植入一种"人造染色体",这样机器人将与人类一样拥有自己的"基因代码",这将会使这类机器人同人类一样拥有喜怒哀乐的情感。另外再假设,如果将来可以为这类机器人设计出类似人类的X和Y染色体,那么机器人也许会有"性别"之分,如此看来,这类机器人的自我复制也就成为可能。如果这些雌雄机器人能够"相爱"的话,通过进行"染色体"之间遗传物质的交换来产生出一套新的"染色体",如果把这套新的"染色体"直遗传物质的交换来产生出一套新的"染色体",如果把这套新的"染色体"也会成为可能。当然,这些都只是猜测。但是,也不可否认这些设想在将来有可能会实现。如果从代际伦理的视角而言,假设智能机器人变得很普遍,而且又已无法进行必要控制的,那么是否会出现代际混乱呢?不管这样的担忧是否是杞人忧天,人工智能技术代际伦理问题都是不容忽视的。

#### 3.2 人工智能伦理问题的成因分析

#### 3.2.1 技术上的局限性

人工智能现阶段的发展是不完善的,许多人工智能都仅仅停留在弱人工智能 的层面上,准确地来说迄今为止人类并没有任何一个强人工智能的成果问世,技 术上的局限也就导致了人类在研究人工智能伦理问题的过程中缺乏成熟的研究对 象,从而无法得出确切实际的解决对策。

人类最接近强人工智能的研究就是深度神经网络,谷歌的围棋软件"阿尔法狗"就是以这项技术作为核心。但"阿尔法狗"并不仅仅只能用于下围棋,它也在很多其他的领域做出了超乎想象的卓越成就。例如说"阿尔法狗"还可以完成视觉分类的任务,它可以很直接的分辨出图片中是一只企鹅还是猫,这一结果相当精确,却仍不能达到人脑思考的水平。研究者在对图片稍作修改之后发现,人们可以轻易的骗过"阿尔法狗",即使它在围棋上可以进行精妙的计算,却无法逃过人类小小的欺骗。人们被图片欺骗这件事本身没有什么大不了的,人脑也会经常被骗,但骗过"阿尔法狗"的图片,是被普遍认为相当正常容易被辨认的图。这些现象都说明了一个问题,人类的强人工智能并未真正产生,人工智能在技术上有相当大的局限性。这种局限性可以分为几个部分说明。

#### (1) 认识的局限性

人们对于思维认识的过程是比较片面的,实际上思维过程是不可能通过物理

符号所运算出来的,无论运算结果有多么相似,人脑思考和计算机思考的方式都是不同的,抽象思维是无法通过物理符号运算。

#### (2) 智能优化方法与途径的局限性

从机械角度出发,智能优化方法主要分为结构派和功能派。结构派研究的是人的大脑神经结构,从而模拟出人的神经网络。但是人的神经网络及其复杂,神经元数量巨大,这就使得结构派的研究非常困难,目前来说研究局限性很大。功能派则研究人的思维活动的智能能力和心理特征,但是这归根结底还是使用的图灵模型,没有逃出物理符号,对于强人工智能来说具有很大的局限性。

#### (3) 计算机模型的局限性

在大多数情况下,人类在解决一个社会问题时,可能并不需要精确地答案,而是需要满意的答案。这是以图灵模型研制出的计算系统所不能实现的。现在的人工智能对待问题只能得到唯一的精确答案,并且在计算过程中需要人为的输入数学意义上的解析式,这也就是强人工智能在计算方式上的局限性。

#### (4) 实现技术方面的局限性

毕竟人脑和和人工智能还是有很大区别的,人脑的逻辑思维能力是人工智能所无法实现的。人脑可以从浅显的知识中运用概念的抽象思维提炼出更高层次的核心知识,在解决问题时能够快速从多种解决办法中找到最优办法,在这过程中,可能根本不需要大量的运算,这也可以称之为"经验"。但这恰恰是人工智能无法做到的。

#### 3.2.2 伦理规则的缺失

众所周知,人工智能是一项新兴的科技,在此出现之前人们对于人工智能的伦理规则构建几乎是一片空白。早期虽然有一些哲学家和学者对未来的人工智能可能出现的伦理困境做过预设,却因为科技进步的太快而无法做出真正符合现实情况、准确的构建。只有当科技真正发展到相应程度,与之相对应的社会伦理规范才能够产生。同时,只有当科技的进步与社会伦理规则相一致时,合乎伦理道德的科学发明才能朝着正确的方向发展。在中国古代的科技发展中,就将追求"大道"作为伦理规范,"大道"在传统的技术发展中一般被理解为"平衡的、和谐地照顾到各方之间的冲突和制约。""但这其中就存在一个矛盾,人类无法预知科技的发展方向和发展程度,从而也就无法提前预设与其相对应的伦理规范。这一点在人工智能的发展上体现得尤为突出,人工智能的发展明显快于人们对于人工智能伦理规范的构建速度,从而导致人工智能伦理规范的缺失。

人工智能的许多伦理问题的产生,都可以归于伦理规则的缺失。例如最近美

①王乾."道""技"中国文化背景的技术哲学【M】.北京人民出版社,2008: 147

国一家公司生产了一款超仿真机器人索菲亚,"她"在脱口秀节目中与人类对答 如流,对话中,索菲亚流露出的喜悦、惊讶、厌恶等情绪令人惊叹。有参与对话 的人表示,和索菲亚建立了很好的友谊。这不禁让人思考,长时间的与聊天机器 人对话,是否会使人产生对机器人的情感依赖?这种类似于柏拉图式的感情交流 其实并不能满足人们对情感的最终诉求,这就可能使得人们在情感关系的后期产 生失望、焦虑等负面情绪,缺乏对现实的真实感知甚至出现更为可怕的后果。现 在聊天机器人已经在民众生活中比较普及,前文中说过的Siri(苹果人工智能软件) 也具有一定的聊天能力,但还没有任何规范与人"交往"的人工智能的伦理规则。 不过,随着能与人互动的人工智能的普及,这一问题也得到了越来越多地关注。 2017年4月13日,中国社科院和腾讯研究院在北京联合举办了"2017年度人工智 能技术与伦理研讨会",在此次研讨会中,学者们针对超级人工智能(ASI)给人 类社会带来的影响,讨论拟制定了一系列的伦理规则与基本法律原则。提出了人 工智能发展的六个原则,这些原则覆盖了"自由、正义、福利、伦理、安全、责 任"等方面。在人工智能高速发展阶段,人工智能在造福人类生活的同时,也带 来了许多来自伦理方面的问题,人类不能只看人工智能的有利一面,而忽略其发 展过程中存在的突出问题。这就需要社会各界的共同努力,共同构建合乎人工智 能发展的各项规范和伦理规则,对人工智能的发展方向进行正确的引导。而且机 器人的应用范围在将来必将会越来越广泛,机器人使用量的稳定增加也会带来一 些预料之外的风险和道德问题。英国设菲尔德大学人工智能和计算机技术专家诺 埃尔•夏基教授也因此呼吁, 应立即制定机器人道德准则。 ①任何科技的出现都会伴 随着一定的风险,但只要人类构建相应的法律法规和伦理规范,相信科技带给人 类更多的一定是社会的进步。因此,我们不能在人工智能的问题上因噎废食,因 为可能出现的问题就止步不前,而是用更加完善的规范来规避损失,促使科技更 好地发展。

#### 3.2.3 监督管理体系的不完善

人工智能是一种新兴技术,目前并没有与之相配套的法制体系,现阶段人类立法的速度无法跟上科技发展的速度。例如人工智能带来的人权问题、环境问题和道德问题等问题出现时,更多地是借助于科学家自身的道德素养和社会的舆论来解决问题,而很难利用法律来有效缓解或是解决这些伦理问题。不管是科学家的自觉性还是舆论的压力都不具备强制性,因而约束效果并不显著。遗憾的是,目前还没有一部完整的、专门的关于人工智能技术伦理方面的法律体系,尽管有些文件或者规范或许有所涉及,但是也不全面,不足以应对如今人工智能技术所

①胡瑶.机器人也要有"德性"【N】.经济参考报,2008-12-24

产生的一些伦理问题。历史的经验告诉我们,如果一项科技在发展和使用的过程中缺乏有力的监管,那么这项科技就可能会被滥用,大量的伦理问题会随之产生。科学的发展是一把"双刃剑",一面是科学提高了人类的主体地位,促使人类更加高效地利用自然资源、成为自然的支配者;另一方面,在科学技术的运用方式不得当的情况下,会使我们的社会陷入一种道德伦理困境。包括人工智能在内的科学技术本身是没有好坏之分的,人工智能的未来究竟是为人类造福还是带来灾难,都取决于人类自己。因此,如果人类不构建完整的监管体系,在不久的将来,人工智能可能真的会沦成全人类的灾难。

# 第四章 人工智能伦理问题的应对策略

目前人工智能的研究重点就是希望机器人能够具有伦理判断与行为的能力,使机器人的行为对人类更有利。要想研究出有思想的机器人,最基本的手段就是对机器人进行伦理设计。试想一下,拥有高度自由权的机器人如何为人类服务?如果机器人拥有高度自由选择的权利,他们选择与人为敌岂不是让很多科幻电影中的情景变成了现实?如果人工智能发展到比人类智能更强大的地步,是否还会同意被人类统治?是否人类会被机器人统治?不管这些问题是否会成真,笔者认为都要防患于未然,对人工智能伦理问题的研究刻不容缓。针对这些情况的出现人类要从伦理学的角度出发,从研发层面、人类自身层面和政府政策层面出发解决机器人研发带来的问题,引导智能机器人向着正确的方向发展,而不是一味的打压。

### 4.1 研发层面的应对策略

#### 4.1.1 用"以人为本"的科学发展观指导人工智能发展

"追求技术的高端并非是人类科研活动的最终目的,科研活动的最终目标是是为了更好地服务人类,为人类的发展做出贡献。"<sup>①</sup>机器人的使命是服务于人,为人类提供更加安全、健康、便利、愉悦的生活。人工智能技术的本质是机器通过信息处理做出最优决策,帮助、替代人们完成任务,因此,人工智能机器人可以说是人类触角的延伸,能够满足人们对安全、健康、便利和愉悦更高的要求。在人工智能的科学发展之中,不管是开发者还是使用者都要准确认识到人工智能发展的首要准则:"以人为本"的科学发展观。

"以人为本"的科学发展观,在应用到人工智能产品的开发领域中时,笔者认为可以概括为以下三个要求:第一,在人工智能的发展中,必须以保障人的安全为前提。在1978年的9月6日,日本一家工厂的切割机器人由于出现故障,在切割钢板工作时误将一名工人当成钢板,这当时成为了闻名世界的"机器人杀人事件"。最后专家认为,事故的原因可以归结为软件系统的故障。这一事件在当时引起了大范围民众的恐慌,也引起了人们对使用机器时的安全问题的思考。任何的科技发展,都必须保障人的安全,人工智能的系统具有复杂和高智能化的特征,因此更应全面地以保证人的安全。第二,在人工智能的发展中,要以保护人的尊严和权利为核心。任何人工智能都必须尊重和严格执行操作人下达的指令,人类在人工智能面前的主体性地位不可动摇。科幻小说家阿西莫夫提出的机器人三大定理中指出,机器人在不伤害人类的前提下,要严格遵循人类的命令。人工智能

①林德洪.科技哲学十五讲【M】.北京大学出版社,2004,280

的发展在不侵害人类利益的前提之下,必须确保人应享有的权利和尊严不被侵犯。 第三,人工智能的发展中,要以保持人类社会的健康发展为目标。人工智能的快速发展,使得人类以前许多繁杂的工作都变得轻松,比如说智能联想输入法让人类的打字时间大大缩短,因而很多青年人都更加倾向于使用电脑打字,造成了很多青年人出现"提笔忘字"的情况,人类的书写能力在逐步退化。为了避免类似的因便利而导致的人的行为能力退化的现象的出现,人类必须确保在人工智能的发展中,不能以削弱人类自身的能力为代价。比如人类可以将人工智能的研发领域予以限制,使其只能应用于简单的辅助工作上,禁止其进行大范围的人际、决策的工作。人机合作的初衷是为了促进工作的效率,而一些具有创造性思维的工作还是需要由人自己完成,这样就可以保证人的能力不会因机器的"代工"而减弱,从而保证人类社会的健康发展。人工智能的发展的目的始终都是为人服务。只有把握好"以人为本"的宗旨,才能促使人工智能向善的方向发展。

#### 4.1.2 让程序语言融入哲学思想

机器人的迅速发展,是好是坏还需要人类的不断引导。如果控制不好,机器人的发展将会给人类带来毁灭性的破坏。如果技术没有得到良好的运用,不仅没有给人类的生活带来便利,还会增加人类的生活风险。这种风险通常有两种情况:首先是机器人控制整个地球,对人类进行屠杀;另一种是一些不法分子控制机器人杀手。所以为了减少风险,需要加强对机器人的控制。一些研究者提出对于机器人的使用者人类没有办法进行甄选,所以即可能出现以下情况:好人和坏人都可以使用机器人,这一情况增加了机器人的使用风险,如果不能进行有效地遏制,那么后果不堪设想。<sup>①</sup>同时,随着智能化技术的不能发展,一些研究者试图将人脑和机器人相结合,简单来说就是把芯片植入人脑。随着植入芯片的不断高阶化,人脑与芯片互相支持,谁主谁从很难说。这条路对于人工智能发展可能是条捷径,但也极具风险,伴随着一系列的问题:比如谁起主导作用?谁接受控制?这一技术堪比洗脑,如果是将好的思想植入人脑还好说;如果将负面的的程序植入人脑,可能会带来灾难性的破坏。

对于这些技术滥用问题的出现,我们把它们交付给科学家德道素养和法律来解决。而技术缺陷造成的风险则需要我们从目前的人工智能底层技术做起,提高科学技术,弥补在技术运用上的漏洞,最重要的是计算机语言程序的完善。现阶段的计算机语言在输入上还存在局限性,主要是机械的输入各种指令,没有思想的输入,这就显得机器人在各个方面都显得比较生硬。这就需要我们在进行人工

①李桂花.科技的人化[M].长春: 吉林人民出版社, 200:210.

智能技术研发的过程中不断更新我们的编程语言,让程序语言融入哲学思想。所以未来的研究方向就是:让程序语言融入哲学思想,使机器人更加有思想、有个性,从而可以满足人类的个性化的服务的要求,同时机器人也更加有尊严的服务人类。

#### 4.1.3 构建人工智能技术的伦理标准

随着现代科学与社会关系联系的紧密性日益加强,特别是科学应用方面的负面效应逐渐显著,科学研究早已不再纯粹是客观中立的认知活动,而是逐渐与社会伦理学建立起紧密的联系,科学研究的伦理性特征日益显著。<sup>©</sup>在人工智能技术发展过程当中,人工智能伦理标准的建立刻不容缓。

第一,客观公正性标准。在人工智能的传播和应用上,应该强调客观公正性。 人工智能的发展不仅要坚持客观性,确保认知过程中信念的真实性,而且要具有 公正性,保证科学知识和信念得到了公正的传播和使用。这既是认知进步的需要, 也是人道主义的要求。对于人工智能,目前大多数民众不甚了解,甚至有些媒体 为了博人眼球夸夸其谈,造成错误的舆论发酵,加深了民众对人工智能的误解。 这种现象产生的根本原因在于,人工智能在传播和应用过程当中,没有建立起一 套客观公正的伦理准则,如果政府在宣传过程当中,根据相关的客观标准严格的 予以把控,相信对于人工智能,人们的误解会逐渐减少,进而营造出一种健康有 序的人工智能技术研发和应用环境。

第二,公共利益优先准则。人类在进行科学研究过程当中,必须要以公共利益作为出发点,不能只为了某一集体或者某个部门的私利,而不顾社会公共利益。因此,科学研究过程当中,公共利益是其最高准则,也是对科研者的最高要求,如果人类研究的科学技术会危害到当代人的利益或者对后代人的公共福祉,损害生态环境的可持续发展的科学研发活动显然是违背了伦理道德的<sup>②</sup>。在人工智能发展过程当中,也应将人类的利益放在首位,不仅要考虑当代人的利益,对于后代人的影响以及对生态环境产生的影响,也要充分地纳入考量范畴,确保任何的人工智能活动,都不损害公众利益,这样人工智能才能取得健康持续的发展。

# 4.2 人类自身层面的应对策略

#### 4.2.1 增强科学家的道德责任感

人工智能的发展方向与科学家息息相关,任何新科技的研发都是一把双刃剑 有利有弊。科学家在对人工智能技术的研发过程当中,如果出于个人利益而不顾

①薛桂波.科学共同体的伦理精神.【M】北京.中国社会科学出版社,2014,7,139

②刘大榕.科学伦理: 从规范研究到价值反思【N】.南昌大学学报,2001,02

社会整体利益,就会对人类的发展产生负面影响。如果科学家从社会整体利益出发,那么科研成果将会更好的服务于人类造福后代。爱因斯坦曾经说过:"怎么利用科学技术,科学技术带给人们的,究竟是灾难还是幸福,全部取决于人类自身,而不是取决于工具"<sup>①</sup>。现代人工智能发展过程当中,产生的种种问题很大程度上是由于科学家缺乏道德责任感。在人工智能的研发过程当中,有些科学家秉持一种中性的态度,认为科学的发展不应当受到任何道德伦理的制约,他们将自己封闭在狭隘的专业领域,对人工智能对外界产生的各种影响充耳不闻。在这种情况下,非常容易造成一些别有用心之人利用人工智能作恶,从而导致人工智能技术的滥用。科学家必须对自己行为导致的后果予以关注,在研发初始就应当对科技某些影响力进行评估,同时在研发过程当中尽量趋利避害,认真履行科学家自身的社会道德责任。

就人工智能技术可能带来的消极后果而言,科学家们要从以下三方面做出表率:首先,科学家能够预测科学结果的社会和政治影响比任何人更科学,因为他们有在科技领域更为专业的知识。他们有责任向公众和政界说明这些结果,这是一种社会责任;其次,由于科学家的社会地位对公众具有特殊的重要性,所以他们可以在一定程度上参与并影响到政治家的决策。因此,这一领域的科学家对人工智能的应用具有一定的社会责任感。例如在一些推理技术对社会的影响,科学家们对不唇服务的后果的闲话;最后,当人工智能的某些科研项目对人类有害时,科学家们在一定程度上是有权利参加或退出这个项目的。知识的力量与之相应的责任是成正比的,科学家们拥改变世界的力量,就应当肩负起对人类社会健康发展的重大责任。在人工智能的发展当中,一定要增强科学家们的道德责任感,承担起科研中的道德责任,让人工智能产品的研发始终保证为大多数人服务,促使人工智能的发展始终走在向善的道路之上。

#### 4.2.2 提高民众的科学素养,正确认识人工智能

在中国乃至世界范围内,人工智能都属于一种尖端科技。对于人工智能的工作原理、工作机制等内容,人类了解并不是非常充分<sup>®</sup>。对于未知的事物,人们总是会产生恐惧之情,进而衍生出很多误解。近年来很多科幻电影上映,例如《黑客帝国》这些电影中的内容,颠覆了人们关于人机的传统认识。人工智能留给人们的印象并不好,电影的结局是机器人最终掌握了主动权,这种结局可能会引起人类的恐慌,人们担心机器人最终会取代人类的地位。尽管这些内容仅是源自于科幻电影,但对于不了解人工智能的普通民众而言,就会演变成一种恐慌心理,

①爱因斯坦文集:第三卷【M】.许文良译,北京商务印书馆,1979,78

②王大洲.技术哲学、技术实践和技术理性【J】.哲学研究,2004,11

有些民众开始从心底里排斥人工智能。同时,某些媒体为了吸引眼球,对人工智 能的报道过于夸张,对于人工智能发展过程当中出现的一些问题进行过度诠释, 对民众产生误导。以上种种因素使得人工智能的研发无法得到大众的认可,这在 无形当中加剧了人工智能的伦理问题的严重程度。但实际上,对于人类发展而言, 人工智能的贡献是不言而喻的。人工智能只是计算机技术当中一个分支,其不论 是研发还是应用,始终是以服务人类作为导向的。所以,目前应当通过各种途径 来向民众普及人工智能的相关知识,提高民众的科学素养,使普通大众能够正确 的认识人工智能,从而扭转民众对人工智能的错误印象。在这个过程当中,政府 媒体、科研机构需要共同携手努力。例如;建立更多的人工智能科普阵地,在社 区、高校等地定期的开展讲座,不断深入人工智能的科普工作。此外,要充分的 利用媒体的舆论导向功能,对人工智能技术进行正确的宣传,严格打击不符合现 实的虚假报道。具体而言,首先要让民众认识到人工智能是一项中性技术,与其 他科学技术之间没有任何差别,只要人类在人工智能的开发当中,融入更多的伦 理观念就能够对一些问题予以避免, 民众是无需对人工智能感到恐惧的。其次, 要让民众相信政府和科研学者,进而相信人工智能始终是为了服务人类而存在的, 其目的不是取代人,更不是让机器控制人,我们不会允许人工智能"反客为主" 的事情发生。最后,在面对人工智能伦理问题的社会舆论共识要发挥出媒体的正 确舆论导向功能,民众在可能出现的伦理问题面前保持冷静,理性的取舍。对于 一些谣言不宣传,减轻民众对人工智能的恐慌,让智能机器人更好的为人类服务。

#### 4.2.3 加强人工智能的国际交流合作

作为一门新兴学科,人工智能还有非常大的发展空间,但人工智能目前还有很多难以攻克的难题成为了人工智能研发的瓶颈,这其中有很大一部分来自于各国之间研发部门沟通不足,错过了相互借鉴学习的机会。世界范围内人工智能技术研发水平都存在着不平衡状态,各国的人工智能技术的发展水平参差不齐,从而导致解决人工智能问题的效率低下。随着全球化的不断深入发展,在面对人工智能伦理问题上各国要展开相互合作,共同解决问题。合作共赢是唯一的选择<sup>①</sup>,世界各国科学家们携手进行头脑风暴,会对相关问题的解决大有裨益,人工智能的研发在各国学者合作交流之下,也会迎来新的发展突破。笔者认为可以通过发起系列科学计划和科学活动的方式,加强国际间的人工智能交流合作。例如;交流论坛、学术会议、外籍人工智能青年科学家的培养等途径均能够予以尝试。

①陈肇雄.人工智能发展迫切需要国际协作.【N】新京报 2017,12,05

### 4.3 制度规范层面的应对策略

#### 4.3.1 制定法律法规来规范人工智能的发展

人工智能飞速发展,引发了一系列的社会问题。但是迄今为止,我国尚未出台任何与人工智能有关的专门性法律法规。仅有的相关规定散见于各种法律文书当中,并且存在着许多的漏洞,没有解决实际问题的能力。事实上,国际上也没有关于人工智能发展的规范法条,这已经引起了世界各国法学家的注意。2018年1月,美国国会就人工智能的发展问题提出了《人工智能未来法案》,如果众议院和参议院通过这一法案,他将会成为世界范围内首部针对人工智能的系统法案<sup>©</sup>。

"无规矩不成方圆",对于人工智能规范化发展而言,相应的政策法规判定标准都有利于其健康发展。在人工智能发展过程当中,产生了很多社会问题,比如由其引发的就业问题、安全问题、隐私问题、知识产权问题、算法合谋问题等等。这些问题都使人们对于人工智能的发展产生了一定的担忧。究竟如何应对这些问题带来的挑战,是摆在人们面前的一道难题,而要破解这一难题,就要依靠法律和公共政策的有机配合。现在各国加紧立法,就是为了填补在人工智能领域的法律空白,让遇到相关问题时有法可依、有章可循。完善人工智能的法律法规和公共政策建设,是国际各国共同的责任,这也将成为规范人工智能发展的重要准则。这需要各国政府的支持与配合,需要全人类的共同努力。相信在得到国际上的高度重视之后,我们能够尽快出台一部国际通用的人工智能发展规范的法案,从而引导人工智能朝着更加光明、去弊的方向发展。

#### 4.3.2 科学管理人工智能产品的使用

在人类生活当中,人工智能技术应用早已崭露头角。在人类社会发展过程当中,人工智能所扮演的角色越来越重要,对于人类生活的影响作用也越来越显著。但是人工智能产品的不科学、不规范的使用为人类带来了一些潜在的问题。例如,人工智能专家系统用于医学领域和金融领域,信息一旦更新不准确,或者信息被不法分子所篡改,可能会造成错误诊断,以及金融危机现象的出现。这对人们将是灾难性的结果。所以科学管理人工智能产品是人类安全高效地对人工智能产品使用的基础和前提。

笔者认为,可以通过对人工智能产品进行编号、分类的方式来有效的管理人工智能产品的使用。比如说,给每一个投入使用的专家系统独立的条形码,这些条形码的独特性可以类似于人们的"身份证",在每一次的产品维护、维修时都能有效的顾及到所有的专家系统,避免造成对专家系统管理的疏忽。同时,可以通

①千家智.人工智能发展,需要立法规范吗?【N】.新京报.2018,1,29,第二版

过限制专家系统的操作员来避免有人恶意利用专家系统。例如医学专家系统的条形码可以设计成只允许某些医学专家操作,在条形码的信息中就输入对操作人员的识别,除去这些医学专家其他人都不具备操作资格。其次,这不仅仅是可以限制专家系统不被滥用,也可以在专家系统出现问题时有效的追责,人工智能产品出现任何闪失,都可以直接追索到操作负责人。

总而言之,人工智能产品是否能对人类造福取决于我们如何使用它们。科学的管理人工智能产品的使用,有效的产品维护检查、落实到个人的追责系统都有利于我们正确的使用人工智能产品。

#### 4.3.3 严格监督人工智能产品的研发和运用

当前,对人工智能技术的监管不到位产生了一个很大的社会问题,那就是人工智面临的社会伦理问题。要想解决这一问题,就要求人类从多角度、多方面对人工智能进行监督和制约。不仅如此,还要针对人工智能制定详细的规章制度,甚至成立专门的法律来规范它。对于违反规定和法律的责任人,应给予其严厉的处罚。人工智能是一场新兴技术革命,作为一种新型的信息技术,人工智能的传递信息速度非常快,如果被一些别有用的人利用,将会产生不可估量的危害。因此,从一开始,人类就要时刻警惕这种潜在威胁,要通过一系列的措施来降低或者消除这种风险。具体的措施可以从以下几方面进行开展:

首先,人类在人工智能产品的设计之初,就应该在人性化理念的指导下进行设置,确保在设计层面不能存在反人类的因素,要使程序简单明了,避免使程序具有伪装性,最重要的是必须保证其在安全性的基础上,会为人类带来便利性,不能对人类社会产生威胁和破坏。

其次,在人工智能产品的完成之时,人类应该通过足够的测试来确保其安全性,就像我国研制的载人航天飞机过程中必须经过无数次的实验,在确保百分之百的成功率后再投入市场,并且设置试用期,在试用期内发生任何问题必须第一时间解决。

再次,在人工智能进入市场之际,生产者要附上一份详细的使用指南,使用指南除了常规的产品说明、使用注意事项、以及使用安装步骤之外,还需要将产品的设计者、组装者、出库检者、安全负责人等信息予以标明。以防在产品出现问题的情况下,能够及时地联系到相关责任人,通过这种方式能够将损失降到最小,而且还能确保产品的品质。有了一目了然的生产责任人信息,购买者在消费和使用过程中更加放心,进而实现人工智能产品的销量的增加。

总的来说,随着新的技术革命出现而产生的人工智能,从其最开始的设计环节到其进入市场环节,人类要做好每一环节的监督工作,从而来防范新的技术给

人类带来不利的影响,缓解甚至减少人工智能技术的伦理问题。例如,对于智能机器人,若被当作工具,则不能设计有伤害人类的功能;机器人永远不得侵犯人权、人类隐私等法律规定的人类的基本权利;机器人需要设定非常明显的机器属性,禁止那些存在明显伪装性和欺骗性机器人的生产,而且在机器人投入使用前要保证它的功能确实能够按照预先设定的程序执行;另外,必须要让使用该机器人的公众提前熟悉该机器人的各种特效,包括可能造成的危害等,以此切实维护公众的知情权和选择权,从而减少一些可以避免的麻烦。

# 第五章 结论

今天,人工智能技术已经与我们的生活密切相关,许多人工智能产品已经应用到我们生活的方方面面。未来,人工智能的发展还会对我们的生活、学习产生更大的影响。人类社会是个不断发展进步的社会,促进科学技术的发展,是当代人的必然选择。尽管人工智能技术引发了不少伦理问题,但历史告诉我们,只有科技的不断发展和进步,人类才不会倒退,人类社会文明才能够不断进步。人工智能技术作为现代科学技术的最前沿代表之一应该不断被发展。

虽然目前人工智能的发展尚且存在诸多的伦理学问题,但我们坚信人类的 认识是有限和无限的辩证统一,世界上没有无法被认知的事物,只有还没被认知 的事物!面对这一次人类科技的飞跃,我们要时刻做好技术反思的工作,不断地 总结问题并采取相应的策略,这样才能够巧妙地使人工智能技术趋利避害,更好 地造福于人类。

# 参考文献

- [1] 陈升, 孙雪. 国内外军用机器人的现状、 伦理困境及研究方向. 制造业自动化,2015(11):27-28.
  - [2] 蔡自兴. 人工智能辞典. 化学工业出版社, 2008.:50-54
  - [3] 蔡自兴, 徐光祐. 人工智能及其应用.第4版. 清华大学出版社, 2010.:12-66
  - [4] 单明. 科技发展要在乎伦理. 当代工人, 2015(9):19-19.
  - [5] 杜严勇. 机器伦理刍议. 科学技术哲学研究, 2016, 33(1):96-101.
  - [6] 杜严勇. 人工智能安全问题及其解决进路. 哲学动态, 2016(9):99-104.
  - [7] 杜严勇. 现代军用机器人的伦理困境. 伦理学研究, 2014(5):98-102.
  - [8] 杜严勇. 机器人伦理研究方法论原则. 中国社会科学报, 2015:1-2.
- [9] 杜森, DuSen. 人工智能的法律与伦理意识形态问题研究. 黄冈职业技术学院学报, 2016,18(2):64-67.
  - [10] 冯定. 神经网络专家系统. 科学出版社, 2006:78-80
  - [11] 冯继宣. 计算机伦理学. 清华大学出版社, 2011:44-46
- [12] 陆冰章, 陆丙甫. 神奇的数字 7±2:人类信息加工能力的某些局限. 心理科学进展, 1983, 1(4):53-65.
  - [13] 龚园. 关于人工智能的哲学思考. 武汉科技大学, 2010.:94-99
  - [14] 胡增顺. 人工智能的伦理问题. 开封大学学报, 1998, 12(3): 57-59.
  - [15] 霍金斯. 人工智能的未来. 陕西科学技术出版社, 2006.22-26
  - [16] 胡增顺. 计算机的伦理学困境与出路. 开封大学学报, 2013, 27(4):86-88.
- [17] 贺欣晔. 科幻文学中人工智能与人类智能的关系. 沈阳师范大学学报(社会科学版), 2016,40(2):111-115.
  - [18] 何华灿, 李太航,等. 人工智能导论. 西北工业大学出版社, 1988.:102-108
- [19] 计海庆. "机器人"观念的形成及其影响的哲学考察——以分析技术的本质和人与机器(人)关系为视角. 复旦大学, 2005.3-7
- [20] 景英娟, 董育宁. 生物特征识别技术综述. 桂林电子科技大学学报, 2005, 25(2):27-32.
  - [21] 廉师友. 人工智能技术导论. 西安电子科技大学出版社, 2007.189-190
- [22] 林德宏. 人与机器:高科技的本质与人文精神的复兴. 江苏教育出版社, 1999.40-42
- [23] 梁路宏, 艾海舟, 徐光档,等. 人脸检测研究综述. 计算机学报, 2002, 25(5):449-458.
  - [24] 李武军, 王崇骏, 张炜,等. 人脸识别研究综述. 模式识别与人工智能,

#### 2006, 19(1):58-66.

- [25] 李俊平. 人工智能技术的伦理问题及其对策研究. 武汉理工大学, 2013::21-24
  - [26] 麦好. 机器学习实践指南:案例应用解析. 机械工业出版社, 2014.138-141
  - [27] 毛磊. 保护隐私权:个人尊严和价值的体现. 紫光阁, 2003(5):58-67
  - [28] 芮延年. 机器人技术及其应用. 化学工业出版社,33-35
  - [29] 唐凯麟, 李诗悦. 大数据隐私伦理问题研究. 伦理学研究, 2016(6):103.
  - [30] 陶悦宁. 论西方科幻电影中的新型伴侣关系. 影视长廊, 2016, (3): 39-44.
  - [31] 王悠然. 预防人工智能伦理缺失. 北京:中国社会科学报, 2015:1-2.
  - [32] 王树林. 人工智能辞典. 人民邮电出版社, 1992:77
  - [33] 王万良. 人工智能及其应用.第 2 版. 高等教育出版社, 2008:50
  - [34] 王万森. 人工智能原理及其应用.第 3 版. 电子工业出版社, 2012:80
  - [35] 王前, 杨慧明. 科技伦理案例分析. 北京:高等教育出版社, 2009.词典:34
- [36] 王东浩. 人工智能体引发的道德冲突和困境初探. 伦理学研究, 2014(2):68-73.
- [37] 王绍源. 论瓦拉赫与艾伦的 AMAs 的伦理设计思想--兼评 《 机器伦理: 教导机器人区分善恶》. 洛阳师范学院学报, 2014(1):30-33.
- [39] 王绍源. 应用伦理学的新兴领域:国外机器人伦理学研究述评. 自然辩证法通讯, 2016,38(4):147-151.
  - [40] 王飞跃. 从 AlphaGo 到平行智能:启示与展望. 科技导报, 2016(7):72-74.
  - [41] 魏鸿. 科幻电影中人机关系认同研究. 华东师范大学, 2016:53-57
  - [42] 王晓楠. 机器人技术发展中的矛盾问题研究. 大连理工大学:2011:7
  - [43] 吴胜, 王书芹. 人工智能基础与应用. 电子工业出版社, 2007.89-94
  - [44] 姚洪阳. 试论人机关系的历史发展及其文化考量. 长沙理工大学, 2010.
- [45] 徐豪, 邹锡兰. 人工智能"风口",医疗与金融先起飞?. 中国经济周刊, 2016(35):31-33.
  - [46] 杨宪泽. 人工智能与机器翻译. 西南交通大学出版社, 2006:64-67
  - [47] 佚名. 电脑会永远正确吗. 中国技术监督, 1996(2):40-41.
  - [48] 朱贻庭. 应用伦理学.上海辞书出版社, 2013.
  - [49] 朱勤, 王前. 欧美工程风险伦理评价研究述评. 哲学动态, 2010(9):41-47.
  - [50] 周昌乐. 无心的机器. 湖南科学技术出版社, 2016.41:8-12
- [51] 张翠平, 苏光大. 人脸识别技术综述. 中国图象图形学报, 2000, 5(11):885-894.
  - [52] 张一南. 人工智能技术的伦理问题及其对策研究. 吉林广播电视大学学

- 报, 2006(11):123-124.
- [53] 翟振明. "强人工智能"将如何改变世界——人工智能的技术飞跃与应用伦理前瞻. 人民论坛·学术前沿, 2016(7):22-33.
- [54] 张蕊, 张佳帆, 江灏. 可穿戴式柔性外骨骼人机智能系统可靠性及应用伦理问题研究. 机电产品开发与创新, 2008, 21(5):19-21.
- [55]玛丽·雪莱著, 胡春兰, 侯明古译. 弗兰肯斯坦. 北京:人民文学出版社, 2004:158-162
  - [56]RAYKURAWEIL. 奇点临近. 机械工业出版社, 2015:71-75
- [57]布律诺·雅科米, 雅科米, Jacomy,等. PLIP 时代:技术革新编年史. 中国人民大学出版社,2007:30-33
- [58] Arkin R C. Governing Lethal Behavior in Autonomous Robots. Crc Press, 2009, 37(2):19-22
- [59] Brown W S. Ontologcal Security, Existential Anxiety and Workplace Privacy. Journal of BusinessEthics, 2000 (23):61-65.
- [60] Danielson P. Artificial Morality : Virtuous Robots for Virtual Games. Routledge, 1992:212-218
- [61] Lisa Damm. Moral Machines: Teaching Robots Right from Wrong. Oxford University Press,Inc. 2008:12-16
- [62] Moor J H. The Nature, Importance, and Difficulty of Machine Ethics. Intelligent Systems IEEE,2006, 21(4):18-21.
  - [63] Minsky M. The society of mind. Simon & Schuster, Inc. 1986:371-396.
- [64] Nick, Bostrom. Ethical Issues in Advanced Artificial Intelligence. 2003, (12): 1-5.
- [65] R, V, Yampolskiy. Artificial Intelligence Safety Engineering: Why Machine Ethics Is a WrongApproach. Roman V. Yampolskiy: 389-393.

# 致谢

在本论文即将完成之际,谨此向我的导师致以衷心的感谢和崇高的敬意,提供给我较好的实习的机会,使我同时在思想上、人生态度和意志品更深层次的高了自己的专业视野,也在生活、思想方面给予了我难忘的谆谆教诲,这些教益必将激励着我在今后的人生道路上奋勇向前。在生活上我结实了很多好朋友,感谢你们的陪伴。是你们让我的生活变得更加充实而多彩。亲情、恩情和友情一个也不能少。我更需要感谢我的师兄师姐们,感谢你们在学习上给我最精准的领航,感谢你们让我深深的感触到我们组建的大家庭中的团结和温馨。

由衷的感谢各位师友的陪伴,是缘分让我们一路陪伴,一起在图书馆里翻查 资料,一起上讨论我们共同关注的领域的事件和改革,一起谦虚谨慎的做好科学 研究的讨论为我拓宽了研究的思路加快了进展。