**成绩**



**自然辩证法**

**结课论文**

**日 期： 2023年10月25日**

**题　　目：人工智能的伦理困境及其治理研究**

**学 院：计算机学院**

**专业名称：计算机科学技术**

**学 号：3220231268**

**姓 名：张鹏杰**

**任课教师：李世新**

**评阅人：李世新**

人工智能的伦理困境及其治理研究

——基于马克思主义科技观分析

张鹏杰

（北京理工大学，计算机学院，计算机科学技术 3220231268）

摘 要：本文基于马克思主义科技观，探讨了人工智能发展所面临的伦理困境及其治理问题。人工智能的迅速发展有效推动了社会进步，然而随着人工智能技术的实用化，人工智能正面临着愈加严重的伦理困境，包括人权困境、道德困境、责任困境等，本文以马克思主义科技观为理论指导，分析了人工智能技术的本质和及其与社会发展的关系，提出坚持正确伦理观念的价值引领，通过提高全社会的文化道德素养、建立健全人工智能的法律制度和监督体系、引导人工智能技术的正确发展的治理对策，正确引导人工智能的发展，确保人工智能更好地发挥其积极的一面，推动人类社会的发展。

关键词：人工智能；伦理困境；治理方案；马克思主义科技观

引言

自20世纪50年代以来，人工智能经历了从初步探索到蓬勃发展的历程，近十年来人工智能技术前所未有地被广泛应用于人类社会，给许多领域带来了突破性的发展，然而任何技术都有其两面性，人工智能技术高速发展的背后是一系列伦理问题的逐渐显现，如数据隐私、算法偏见、责任归属等，这些问题逐渐引起社会各界的广泛关注，值得相关领域学者进行深入研究。

依据马克主义科技观对人工智能科学技术进行分析，可以帮助我们更深入地理解人工智能的本质，从而加深对人工智能伦理问题的认知，为相应的治理措施提供有力的理论支撑，为人工智能的可持续发展提供有效指导。

因此，基于马克思主义科技观对人工智能的伦理困境及其治理进行研究分析，具有十分重要的理论价值和现实意义。

一、人工智能面临的伦理困境

人工智能的产生和发展为人们带来了前所未有的便利和经济效益，但人工智能在解构传统伦理关系的过程中，也来带来了新的伦理矛盾。站在马克思主义的立场上客观辩证地分析人工智能带来的伦理问题及其原因，正确处理科学技术和伦理道德间的关系，是本文的研究重点。

（一）人工智能面临的人权困境

以算法为核心的人工智能通过对人类思维过程的模拟，在推理、判断等方面的表现越来越类人化，甚至在某些领域已经超越人类，如图像识别、棋类博弈等。人工智能的发展促进了生产力的提高，有效促进了人类的自由全面发展，然而在实践中越来越独立的人工智能，对人类主体地位也提出了严峻挑战[1]，这可能会对人的基本权益构成一定威胁。

1.人的主体性消解

马克思主义认为在自然世界和社会生活中，人是活动的主体，人的主体地位来自劳动，劳动使得人与对象分属两个阵营，形成人类独有的改造世界的主体性。然而人的这种主体地位受到了人工智能的严峻挑战，人类的主体性正遭到不断地异化。

马克思在《1844年经济学哲学手稿》中完整提出了异化命题，即异化主要为原本由人所创造的、 对象化的客体， 本应该服务作为主体的人，但这种客体反而控制了作为主体的人，这种主客体关系的颠倒就是异化[1]。恩格斯曾以英国纺织工人为例，描绘了机器如何将原本乡村中精神健全的农民，转化为工厂里愚钝痴呆的工人。

因此，原本为人类所创造用于服务自身的人工智能，不能使其反过来成为控制人类的手段和工具，而这恰恰是当前人工智能发展的一大挑战。从马克思主义实践观的角度来理解，现在的人工智能是人类认识客观世界、改造客观世界的一种工具，是人类劳动的对象化产物，是人类社会知识体系的凝结[5]，而当前社会舆论对人工智能的期待往往过高，这又引发了群众对人工智能技术的恐惧，随着科技的发展，若不能将人工智能放在一个合适的地位，人的主体性将会在社会层面上被动摇。

2.人的隐私安全

国内学者李开复将当前的人工智能解释为“深度学习+大数据=人工智能”，深度学习与大数据的组合为人工智能的技术发展赋予了强劲的发展动力，同时也不可避免地带来了关于隐私安全的问题。

当前主流的人工智能模型是深度学习模型，并且近期的发展趋势是模型规模的不断扩大，如ChatGPT4就是大模型的一种应用成就，因此人工智能的发展对数据的要求越来越高，模型训练数据的收集有时不可避免地侵犯人的隐私，此外在人工智能应用过程中，已经可以做到利用大数据收集汇总的人的日常隐私数据，对人的想法、行踪进行解读，这都对人的隐私安全构造了极大的威胁。

3.人的发展限制

随着人工智能技术的不断革新，为适应社会发展的多层次需求， 慢慢演化出小到扫地机器人、强到水下极限作业机器人、大到货运机器人等等，种类不同的人工智能，以帮助人们更快、更好地完成任务[4]。

然而，人工智能会导致人的主体能动性的削弱，而人类主体能动性是人类区别于其它物种的独特属性和专有能力，如认知能力、审美能力[1]，是人类发展的关键特质。人工智能既可以应用于底层简单的体力劳动，也可以应用于高层复杂的脑力劳动，这将对人的发展空间造成挤压，从经济层面上引发就业困难的问题，从社会层面上消解人的社会性，从发展层面上削弱人的主观创造力，这将导致人的发展受到极大约束。

（二）人工智能面临的道德困境

康德在《实践理性批判》中指出：“世上最使我们震撼的是头顶的星空和我们心中的道德”，道德地位是人类区别于其他生物最主要的特质，然而迅速发展的人工智能使得之前只是从事简单体力活动的智能机器人逐渐具有一定的感知能力、推理能力，是否应该赋予具备人性的机器人以人权，将其纳入人类伦理道德体系内，成为一个争议不断的问题[2][9]。

关于该问题的解决方法，有两种不同的观点：一种观点认为可以将人工智能作为道德主体，法国著名哲学家拉梅特里认为： “人是动物， 因而也是机器， 只不过是更复杂的机器”，将人与机器等同起来，因此赋予人工智能道德主体地位是合理的[4]；而另一种观点认为人工智能应该仅仅被视为工具，并且由于人工智能相对于人类单体的强大能力，必须时刻警惕其威胁到人类生存的可能性。

鉴于人们对于人工智能能否同人类一样拥有道德主体地位缺乏统一的认知，社会对于人工智能能否拥有“人权”的态度也具有差异性，随着人工智能的日渐发展，该问题将逐渐成为一个亟待全社会达成共识的命题。

（三）人工智能面临的责任困境

通常情况下，一项新技术的发展往往具有复杂性和不确定性，在应用过程中会出现难以避免的消极影响[1]，而当前人工智能作为一种强大的技术工具，其深度学习的架构使其技术黑箱的问题十分明显，强大的能力使其对人类社会产生巨大影响，而不可解释的特性使其输出结果具有不确定性，这使得人工智能在引发问题时，伦理责任的归属往往陷入困境。

例如，无人驾驶是当前人工智能最经典的应用领域之一，然而无人驾驶颠覆了传统驾驶的以司机为责任对象的追责体系，在人工智能拥有了一定的判断决策能力后，如何明确责任划分成为一个重要的现实问题，如2017 年 11 月美国拉斯维加斯一辆搭载 15 名乘客的无人驾驶巴士撞上一辆人力驾驶的货车，所幸最终无人员伤亡。

通过我们对责任归属的判定，主要依据法律条文中的相关规定，但人工智能的发展速度往往远超相关法律条文的更新速度，但在人工智能拥有自主判断能力的情况下，若简单将人工智能失控归咎于产品的设计开发者，责任分配或许有失衡的风险，因此如何在责任划分体系内正确对待人工智能，需要相关领域人员的实践探索。

二、马克思主义科技观视野下的人工智能

马克思主义科技观是马克思主义关于科学技术的系统性的观点和方法， 我国发展人工智能，必须坚持马克思主义科技观的基本立场和基本原则。

马克思主义认为，“科学技术是人的创造物，是人类认识世界、改造世界的手段和工具”[10]，同时科学技术是推动社会变革和时代进步的关键因素，但是需要指出的是，科学技术具有二重性，科技在给人类带来巨大收益的同时也会带来难以预料的风险[1]。

因此，在发展人工智能时，必须始终坚持人类主体地位，认识到任何科学技术都只是基于人类需求创造出来的客体，如果将主客体位置颠倒，不仅会阻碍人类对科技的正确认识和运用，还会误导人类对科技负面影响根源的把握。

（一）人工智能的本质

在马克思主义科技观中，“人本”意蕴是该思想的基本立场，马克思从“现实的人”出发，指出科学技术是人的本质力量的体现，因此现实的人的实践决定了科学的本质和内涵。马克思在阐述有关科学、技术和人类社会关系的基本观点时，一直强调人作为主体的重要性，并认为科技发展的最终目标是促进人的自由全面发展，人在科技活动中不仅仅是被动的接受者，更是科技的创造者和使用者。

作为当代科学技术的代表之一，人工智能为社会发展及人类发展带来了深刻的变革，但其本质仍是人类科学的产物，是人的造物。马克思主义强调认识来源于实践，而人工智能所拥有的知识并不是其自身实践的产物，而是来源于人类自身，因此关注人工智能更应该关注人工智能背后的人。

追本溯源，人工智能是一串数据[5]，人工智能技术与应用，包括计算机视觉、自然语言处理、深度学习等，无一能脱离数据本身而存在，人通过自身的知识和实践将无意义的数据组织为有意义的人工智能系统，因此人工智能的本质是人创造的新的生产工具，是人的一般智力的现代表现形式，任何脱离对人工智能的工具性的认知，都是脱离人和社会的非理性辩证的结果。

（二）人工智能与社会发展

生产力是由各种要素构成的，其发展是各种要素相互作用的结果，现代科学技术对其他生产要素的渗透和作用日益强烈，当它作为生产的内在要素渗透于生产要素的其他要素之中时，必然会导致其他要素随之变化，从而促进生产力整体质量和水平的提高[7]。

人工智能与传统科技所不同的是其对生产力的极大促进，尤其是这种促进是通过模拟人的智能以促进甚至代替人在某些领域的工作，牛津大学发布的《就业的未来》的研究报告指出，在未来 20 年内，人工智能将会威胁到收银员、厨师、服务员、建筑工人、安保人员等体力劳动者的工作。不仅如此，就连艺术家、教师、医生等脑力劳动者的工作也会受到人工智能的极大影响。

因此，人工智能把人们从体力劳动中解放出来，这极大缩短全社会的必要劳动时间，提高社会的劳动生产率，促进工业向高质量阶段发展，促进社会的发展进步，进而使人的活动空间不断扩大，促进人的发展[7]。

然而需要警惕人工智能在应用过程中被阶级利用，使其在某个阶段反过来制约生产力的发展、制约人的发展，例如目前凸显的人工智能引发的社会发展问题有[6]：

1.劳动时间的延长

随着越来越多的工作以电脑等智能设备为基础，社会的工作节奏越来越快，工作要求越来越细致，工作和生活的界限日渐模糊。

2.劳动强度的增强

人工智能作为信息化的生产工具，本应提高生产效率、减轻劳动者负担，但随着低端工作被自动化工具占据，劳动者反而需要将更多劳动时间分配到高强度工作中，体力和智力的付出反而有所提升。

3.自由权益的侵占

在资本主义的条件下，人工智能被用于监视用户的生产、生活，使人的隐私被完全暴露在算法之下，数据化和网络化的结果是使每个人处在自己的全景监狱中而不自知，使人的自由发展反而受到限制。

（三）人工智能的伦理问题

利用马克思主义科技观客观分析人工智能后，我们会发现人工智能作为科学技术的工具性是毋庸置疑的，正确认识这一点后，将人工智能发展带来的伦理问题一味归咎于科技本身，是有失偏波的。

人工智能伦理问题的产生，一方面是由于社会共识的缺乏和法律制度的缺席，另一方面是由于部分科研工作者、科技公司道德责任感的缺失，这都不应该成为迟滞人工智能技术发展的阻碍，近年来人工智能威胁论甚嚣尘上，加之部分媒体的夸大其词，使得群众对人工智能技术本身产生了一些恐惧心理，这是不应该的。

因此，讨论人工智能的伦理困境的治理方案，应坚持马克思主义科技观的基本原理，更多地关注于人的思想转变，以及社会制度的变革，从而实现人工智能技术的合理应用，由此人工智能的伦理问题将不攻自破。

三、人工智能伦理困境的治理方案

讨论人工智能伦理困境的治理方案，首先要坚持正确伦理观念的价值引领，要坚持马克思主义科技观的指导，坚持新时代中国特色社会主义伦理思想引领。

必须坚持以人为本、和谐共生的理念，牢记人是审视人工智能发展的价值尺度、实现人的解放是人工智能发展的根本诉求，为此人工智能技术的发展要始终强调人在社会发展中的主体地位[1]，坚持科技要为人民服务，构建正确的、和谐的人与人工智能的关系[3]，以人工智能的科技创新，推动社会生产力的提高。

下面本文将从三个方面给出治理人工智能伦理问题的具体方案，包括提高全社会的文化道德素养、建立健全人工智能的法律制度和监督体系、引导人工智能技术的正确发展。

（一）提高全社会的文化道德素养

提高全社会的文化道德素养，形成对人工智能技术广泛的、正确的态度共识，这是突破人工智能伦理困境的根本。

一方面，应提高普通民众的文化素养，使其正确认识人工智能。库兹韦尔将人们对待新技术的态度分为三个阶段，首先，人们会惊叹新技术为人类所带来的便利；然后，人们会恐惧隐藏在新技术背后的危险和弊端；最后，人们将认识到真正理性的道路是设计出一条科学的发展路线，真正实现科技为人类服务的初衷。因此我们必须理解群众对于人工智能这一新兴强大技术的担忧，加强宣传教育，引导人们树立客观的评价标准。

另一方面，应着重加强科研人员的道德责任感。人类是一切科学技术的创造者和推动者，人工智能的发展与相关科研人员的工作密不可分，必须加强对科研人员正确的思想引领，培养其责任意识，提高其道德水平，对人工智能技术负责，对社会、对未来负责。

（二）建立健全人工智能的法律制度和监督体系

当前人工智能发展过程中所产生的很多伦理问题，其主要原因就是缺乏统一的规章制度和监督体系，尤其是缺乏在法律层面上的约束，如人工智能的隐私泄露问题、责任归属问题等，这需要国家科研管理部门和法律部门以及社会各界的共同努力。

法律作为国家强制性手段，能最大程度地限制人工智能发展过程中的无序性，降低人工智能伦理问题给社会带来的危害。首先要从人工智能的法律主体关系入手，厘清人工智能在可能发生的伦理事故中的责任归属，同时建立人工智能的法律问责机制，规范人工智能系统的设计和生产，限制人工智能的非法应用[1]。

同时，应当构建人工智能技术监督体系，降低技术本身产生伦理问题的风险。在人工智能技术的开发环节，确保其研发设计理念不违反社会道德伦理，在人工智能产品的生产使用环节，建立起完善的售后服务体系[2]，同时政府、社会建立各级各层的人工智能监管机构，确保人工智能技术的安全可控性。

（三）引导人工智能技术的正确发展

人工智能技术的不成熟是引发其伦理风险的重要原因之一，当前的深度学习模型可解释性较差，许多人工智能技术是一个科技的“黑箱”，这使人工智能的发展具有一定的不可控性，因此正确引导人工智能技术未来的发展，降低其发展的不确定性，是规避伦理问题的一种有效措施。

一方面，从科技发展层面上，学术界和工业界需要注重对人工智能基础理论的研究，对人工智能尤其是深度学习背后的原理构建良好的解释模型，降低人工智能产品的不确定性；另一方面，学术界和工业界需要恪守法律道德原则，避免开发使用违反社会主流价值观的人工智能产品，避免某些容易被不法分子利用的人工智能技术流出而对社会造成冲击。

参考文献：

[1]杨双彪.人工智能的伦理困境及其规避进路——基于马克思主义科技观分析[J].理论观察,2023(06):52-56.

[2]董子涵,任凤琴.困境与超越：人工智能的伦理审视[J].科技智囊,2023(06):65-70.DOI:10.19881/j.cnki.1006-3676.2023.06.09.

[3]杨丽萍,刘华,杨晓萍.人工智能技术的伦理问题及其治理路径研究[J].互联网周刊,2022(05):30-32.

[4]郭志龙,李晓红.人工智能技术应用中的伦理问题及治理研究[J].赤峰学院学报(汉文哲学社会科学版),2020,41(12):34-38.DOI:10.13398/j.cnki.issn1673-2596.2020.12.007.

[5]侯苏皖. 人工智能时代人的主体地位研究[D].南京信息工程大学,2023.DOI:10.27248/d.cnki.gnjqc.2023.001359.

[6]程萌.人工智能会使无产阶级变成“无用阶级”吗？——技术图景中的马克思阶级政治学说[J].理论月刊,2023(04):26-35.DOI:10.14180/j.cnki.1004-0544.2023.04.003.

[7]张鑫欣.马克思主义哲学视域下人工智能的利与弊[J].现代商贸工业,2023,44(16):148-151.DOI:10.19311/j.cnki.1672-3198.2023.16.052.

[8]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(05):128-136.DOI:10.16290/j.cnki.1674-5205.2017.05.012.

[9]陈晋. 人工智能技术发展的伦理困境研究[D].吉林大学,2016.

[10]李桂花.科技的人化[M].长春：吉林人民出版社，2004.