**基于人工智能的工程伦理教育研究**

**常朵朵**

**摘要：**工程伦理教育是人文教育和科技教育相交叉、融合的学科。当前我国理工科大学生迫切需要进行工程伦理教育，以掌握对现代工程活动进行社会评价和道德评价的基本规则。这一方面是提高大学生道德素质的需要，另一方面也是我国社会可持续发展的需要。作为高新技术的软件工程技术来说，其在现代化建设中发挥着愈发重要的作用，而人工智能作为软件工程中极具发展潜力的代表性技术之一，是新一轮科技革命和产业改革的核心力量，将深刻改变人类社会生活、改变世界。因此，基于人工智能的软件工程伦理教育刻不容缓。

**关键词：**工程伦理教育；软件工程；人工智能

1 工程伦理教育

工程伦理教育作为工程教育的一个类别（科类），工程伦理教育是通过专业教育与道德教育的结合来提高学生道德素质的有效方法，是自然科学和人文社会科学相互交叉的学科，其目的是培养大学生在未来的工程活动中具有强烈的社会责任感，形成以伦理道德的视角和原则来对待工程活动的自觉意识和行为能力，在未来的工程活动中能够从道德的视角和原则来为大众服务[1]。

2 工程伦理教育的意义

2018年召开的全国教育大会中，习近平总书记指出，党的教育方针为：坚持教育为社会主义现代化建设服务、为人民服务，把立德、树人作为教育的根本任务，全面实施素质教育，培养德智体美全面发展的社会主义建设者和接班人，努力办好人民满意的教育。可以看出当前教育界及整个社会都很关注学生的素质教育。随着社会物质文化和精神文化的不断丰富和迅速发展，社会对大学生的道德素质提出了更高的要求，道德素质在大学生的整体素质中显得愈加重要。那么，在大学中结合学生专业进行伦理道德教育应该是提高学生道德素质的重要方法。这种方法更便于学生接受，且影响持久、深远。

21世纪以来，我国的经济开发和建设得到飞速的发展，同时针对产生的伦理问题也迫切需要工程伦理教育来解决。在当今知识经济时代，科学技术对经济建设和社会发展的作用越来越大。与整个社会生活和人类生存状态的关系也愈加密切。科学技术的宗旨在于造福人类、促进社会全面和谐的发展，否则，科学技术的进步就没有任何意义，甚至会给人类社会造成危害。当今科学技术的迅猛发展使人类控制自然和改造自然能力空前强大，但同时科学技术的不合理运用所造成的负面效应也已经给人类社会带来了许多危害，给社会的可持续发展带来了重大的隐患。如各种环境污染、动植物物种减少、水土流失和土地沙化、工程活动对公众利益的侵犯和对历史文化的破坏等等。毫无疑问，科学技术的探索和发明是具有永恒价值的。问题的关键是如何运用科学技术，运用科学技术的人具有什么样的价值观、道德观、利益观和责任感，科技工作者是否明确工程活动对人、生态环境等方面的影响。理工科大学生是国家未来的科技工作者，将来很快就要参与国家的各种经济建设。对他们而言，掌握科学技术同然重要，但同样重要的是要在未来的工程活动中具有社会责任感、正确的价值观、利益观和强烈的伦理道德意识，能够自觉担负起维护人类共同利益的道义使命。对此，工程伦理教育可以发挥重要的作用[2]。

3 基于人工智能的工程伦理

作为软件工程中代表性技术之一，人工智能的发展受到了党中央、国务院的高度重视。党的十九大报告强调，要“推动互联网、大数据、人工智能与实体经济深度融合”。习近平总书记多次就人工智能作出重要批示，指出人工智能的迅速发展将深刻改变人类社会生活、改变世界。同时，人工智能是新一轮科技革命和产业变革的核心力量，世界主要发达国家纷纷加强谋划部署，力图在新一轮国际科技竞争中掌握主动权。高校处于科技第一生产力、人才第一资源、创新第一动力的结合点，是国家科技创新体系的重要组成部分，肩负着人才培养、科学研究、社会服务、文化传承创新、国际交流合作等使命。面对新一代人工智能发展的重大机遇，教育部及时出台《行动计划》，旨在引导高校主动深化改革、加大探索力度，全面落实立德树人根本任务，牢牢抓住提高人才培养能力这个核心点，发挥科研育人在高等教育内涵式发展和高质量人才培养中的重要作用，并通过创新型人才的培养不断提升国家自主创新水平，构筑持续创新发展的优势，为我国人工智能发展提供科技和人才支撑，推动我国占据人工智能科技制高点。但技术往往是把双刃剑，其在发展的同时给人们带来的不仅仅是便利，还造成了负面影响，比如，近期发生的手机任意调动摄像头事件，显然，人工智能的伦理问题比以往技术带给人类的更加复杂。

国家在大力发展人工智能技术的同时，也需认识到，在人工智能的发展和应用中，伦理必须占据中心地位[3]，发展人工智能是为了人类的共同利益。人工智能应当保证公平，并且让人容易理解，人工智能不应当用来侵犯人们的隐私，所有公民都有权利接受教育，使他们能在精神、情感与经济上和人工智能一起繁荣发展，人工智能不应被赋予伤害、破坏或欺骗人类的自主能力。无疑，人们在探讨人工智能带来的伦理问题的同时，还在思考怎样才能解决这些伦理问题，尽管离解决问题还有些距离。下面将通过具体的案列来分析。

（1）人工智能偏见的伦理问题

多位人工智能专家认为，人工智能存在偏见，种族主义和种族歧视是目前人工智能系统存在的主要问题，这可能和系统的研发设计者本身的种族有关。2016年，美国多地法庭开始使用人工智能系统预测一名罪犯再次犯罪的几率，作为法官给予罪犯缓刑等的依据。然而，媒体对佛罗里达州法庭使用的一种名为“少数派报告”预测系统进行了调查，如图3-1所示，发现黑人被系统预测会再次犯罪的几率比其他族裔高出45%，其中黑人被预测再次暴力犯罪的几率比其他族裔高出77%。数据显示，这一人工智能系统的准确率只有20%。媒体评论说，所谓犯罪预测系统唯一能判断的仅是受调查对象的种族而已。同年，美国多家公司联合推出了“世界首届人工智能选美大赛”，如图3-2所示，参赛者在网站上上传照片，由人工智能算法“精确”评估参赛者的美。然而，这场比赛的获奖者都是白人。美国路易斯维尔大学网络安全实验室主任扬波利斯基指出，很明显这一人工智能的学习训练样本不够，美人被限制在固定的模式中。人工智能带来的文化偏见的例子还有很多，如词语“女性”、“妇女”与艺术人文类职业以及家庭联系更紧密，词语“男性”、“男人”则和数学、工程类职业更近。尤为危险的是，人工智能还有强化所习得的偏见的潜能，它们不像人可以有意识地去抵制偏见。



图3-1 “少数派报告”犯罪预测系统

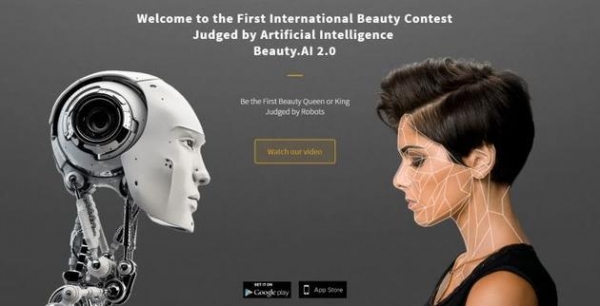


图3-2 世界首届人工智能选美大赛

扬波利斯基等人认为，这些案例说明，保持机器学习数据的多样性是防止人工智能系统产生“偏见”的关键，相关伦理研究更是应该走在技术研究的前面。

1. 人工智能缺陷的伦理问题

人工智能和包括机器人在内的智能化自动系统的普遍应用，不仅仅是一场结果未知的开放性的科技创新，更将是人类文明史上影响至为深远的社会伦理试验。今年5月，滴滴顺风车司机杀害女乘客事件发生后，有网友爆出，滴滴顺风车司机通过平台软件，在接单前可以看到以前司机对乘客的全部评价，而这些评价中就包括对女性乘客外貌的评价。这是因为平台软件上增加了社交功能。作为共享交通平台，它的目标是安全快捷地把乘客送达目的地，现在为了达到商业上的利益最大化，增加了社交软件，却没有考虑到那些女乘客以及潜在受害者的利益。这就涉及到伦理问题。通过对司机、乘客和平台的利益分析和价值分析，发现其实包含了对各方的利害得失与价值冲突。



图3-3 滴滴顺风车司机杀害女乘客事件

如果说人工智能对人类的威胁是终极问题，如何解决还未可知，那么人工智能在当下活动中体现出来的伦理问题就亟待解决了。我们应当从我们这些年研究科学技术前沿中的、科学技术应用中遇到的实例，去探讨人工智能的伦理问题。从这个角度，它可以当做是当作未完成的伦理。与以往不同，人工智能伦理研究不再是一个高高在上的理论，也没有一套适用全部情形的普遍原则体系，更多的是针对具体的人类活动或者科技实践的原则体系，比如说，陪护机器人，如图3-4所示，就需要一些防止过度依赖机器人的原则。从里面找到一些价值冲突，需要作哪些伦理的决择，让我们发现问题、研究问题。另外，当下的社会中，大数据让人们享受了便捷，同时这些数据会受到不良的使用，或者是没有必要的滥用。而由于大数据的挖掘导致商业智能的运用对人的权利的侵犯，这种侵犯最核心一点就是人的自主权，包括隐私。



图3-4 陪护机器人

让技术在社会上更好地运行，就要从伦理上对它进行设计。比如现在经常提到监控装置对人的隐私的伤害，将来可以让使用者看到摄像头的界面，让人们在一定程度上了解自己的数据如何被追踪和使用。这样既可以促使数据追踪者规范其行为，也有利于被追踪对象适当修正其行为方式。这是数据时代、信息时代一个新的信息对称的问题，让不对称变对称一点，寻求一种再平衡。

通过上述案例，我们可以清晰地认识到人工智能伦理方面的缺陷与不足，在现代化建设的新时代，毫无疑问，人工智能需要“德才兼备”，而如何具备，则需要人工智能的伦理教育。

对于人工智能伦理教育来说，面向处理原则，首先要以预防为主，充分预见人工智能可能产生的负面影响，使责任伦理面向未来。德国哲学家尤纳斯和伦克倡导的“责任伦理”，强调充分考虑到技术可能后果的责任，对可能受技术影响的未知人群的责任，对人类命运和社会进步的责任。而后要坚持以人为本的原则，充分保障人的安全、健康和全面发展，避免狭隘的功利主义。要体现“人不是手段而是目的”的伦理思想，防止为了功利主义的商业需要损害人的安全、健康和全面发展，尤其要强调技术实验的人道主义和“知情同意”原则。之后要体现整体主义，从整体上维护人类利益和生态环境，防止人工智能技术无序发展。技术的研究与发展具有个体性和局部性，而其影响同时具有群体性和整体性，两者的矛盾需要随时解决。最后则需要相应制度约束，主要依靠制度实现人工智能技术的伦理规约，防止出现不可逆的严重后果。人工智能技术的伦理规约不能仅仅靠科技工作者个人的道德良知，必须依靠强有力的制度约束。

**参考文献**

1. 李丽英，2008年：《论工程伦理教育》2008 年, 2008 - cdmd.cnki.com.cn
2. 刘绍春，2004年：《工程伦理教育与理工科大学生道德素质的培养》北京理工大学学报: 社会科学版, 2004 - cqvip.com
3. Parliament home pageHouse of Lords Seclect Commitee on Artificial Intelligence.Artificial in the UK: ready, willing and able? [EB/OL]，2018-04-19．,
4. 贝里斯，米切尔，1989年：《职业伦理学》，学苑出版社。
5. 肖平编，1999年：《工程伦理学》，中国铁道出版社。
6. 李世新，2003年：《谈谈工程伦理学》(《哲学研究)2003年第2期)
7. 李世新，2006年：《工程伦理意识淡漠的原因分析》北京理工大学学报: 社会科学版, 2006 - cqvip.com
8. 马成松，2004年：《对工程教育中工程伦理问题的思考》高等建筑教育, 2003 - cqvip.com
9. 于建军， 贺媛媛， 李长玲，2012年：《工程伦理教育的困境与展望——基于国际工程伦理教育的启示》 - 产业与科技论坛, 2012 - cnki.com.cn