

量子人工智能伦理治理的模式探索与实践路径

Investigations into the Models and Practical Paths for the Ethical Governance of Quantum Artificial Intelligence

刘燊 /LIU Shen¹ 汪澜 /WANG Lan²

(1. 安徽农业大学人文社会科学学院, 安徽合肥, 230036; 2. 安徽农业大学林学与园林学院, 安徽合肥, 230036)

(1. School of Humanities and Social Sciences, Anhui Agricultural University, Hefei, Anhui, 230036;

2. School of Forestry and Landscape, Anhui Agricultural University, Hefei, Anhui, 230036)

摘要: 随着量子计算的发展, 将量子计算与人工智能结合的量子人工智能受到关注, 但是它也带来了伦理争论与风险。本文首先对量子人工智能的概念和本质加以界定, 并探讨了它的历史历程。接着介绍了量子人工智能带来的隐私和数据泄露风险、量子算法决策和责任归属问题、算法偏见和歧视、认知自主性和人类主体性问题。最后, 提出算法透明与过程跟踪、情境伦理决策与偏见消解以及伦理对话与共识构建等伦理治理建议以寻求科技进步和伦理规范之间的平衡。

关键词: 量子人工智能 量子计算 人工智能 伦理治理

Abstract: The progress in quantum computing technology has led to significant attention towards quantum artificial intelligence, an innovative field that merges quantum computing with artificial intelligence. However, it also raises ethical controversies and potential risks. This paper initially defines the concept and essence of quantum artificial intelligence and then delves into its development history. It then elaborates the concept and data leakage risks linked to quantum artificial intelligence, the debates over decision-making and accountability in quantum algorithms, issues of algorithmic bias and discrimination, and the challenges tied to cognitive autonomy and human subjectivity. Ultimately, it provides recommendations for ethical governance, such as enhancing algorithm transparency and process traceability, facilitating context-aware ethical decision-making and bias reduction, and promoting ethical dialogue and consensus-building, with a purpose to strike a balance between technological progress and ethical standards.

Key Words: Quantum artificial intelligence; Quantum computing; Artificial intelligence; Ethical governance

中图分类号: O413; N031 DOI: 10.15994/j.1000-0763.2026.01.011 CSTR: 32281.14.jdn.2026.01.011

近年来, 人工智能技术快速成为科技领域的焦点, 深刻改变了社会各个行业。传统的人工智能已经广泛运用到模式识别、自然语言处理、决策支持这些领域当中, 极大地改善了效率和创新水平。伴随互联网用户以及数据数量

的逐步增加, 处理复杂状况及大量信息的需求也在同步增长, 于是就产生了对更为先进的计算技术的要求, 以便改善信息传递与处理流程。量子科技属于前沿技术, 给人工智能带来新的机会和挑战。

基金项目: 国家社会科学基金重大项目“当代量子诠释学研究”(项目编号: 19ZDA038); 安徽省高等学校哲学社会科学优秀青年项目“具身社会心智的神经哲学进阶研究”(项目编号: 2022AH030089)。

收稿日期: 2024年8月1日; **返修日期:** 2025年10月18日

作者简介: 刘燊(1992-)男, 安徽无为, 安徽农业大学人文社会科学学院教授, 研究方向为量子认知与社会认知神经科学。Email: liushen@ahau.edu.cn

汪澜(2005-)女, 安徽池州, 安徽农业大学林学与园林学院本科生, 研究方向为量子认知与社会认知神经科学。Email: 2199069524@qq.com

量子人工智能 (Quantum Artificial Intelligence) 融合了量子计算与人工智能, 提供了更强的计算能力与新算法解决方案。^[1] 但是, 它的发展也带来了许多伦理问题。量子计算机的强处理能力会给现有的数据保护方式带来极大的威胁, 我们需要重新考虑如何保证数据安全、隐私安全等问题。其次, 量子人工智能在算法的设计与应用中存在偏见, 会造成不公正的决策结果, 从而对社会的公平性和稳定性造成影响。

本文将深入探究量子人工智能的本质, 剖析其应用时出现的伦理风险和难题, 探寻怎样在技术发展与伦理准则之间达成调和。本文希望能够给出有效的伦理框架以及监督手段, 保证这项前沿科技既能够推动创新, 又不会违背社会道德和伦理准则。

一、量子人工智能的研究历程与概念界定

1. 量子人工智能的研究历程

量子人工智能产生于量子计算与人工智能的交界处。人工智能发端于20世纪中期, 1943年, 沃伦·麦卡洛克 (Warren McCulloch) 和沃尔特·皮茨 (Walter Pitts) 提出了神经网络模型, 奠定了神经元模型的基础。^[2] 1950年, 艾伦·图灵 (Alan Mathison Turing) 提出了图灵测试来衡量机器的智能程度。^[3] 1956年达特茅斯会议使人工智能成为一门学科, 确定了逻辑推理、知识表征、机器学习等方向。^[4] 20世纪70-80年代, 知识表征、专家系统成为重点, 专家系统模拟专家决策显示了人工智能在某些特定领域应用的潜力, 但是存在知识获取及适应性等方面的难题。^[5]

20世纪80-90年代, 量子计算走进学术视野成为量子人工智能的理论研究阶段。理查德·费曼 (Richard Feynman) 于1982年提出量子计算概念, 表示能够对量子物理系统实行模拟从而突破传统计算机处理复杂量子系统的不足之处。^[6] 后来彼得·肖 (Peter Shor) 提出Shor算法、Grover算法在因式分解和数据库搜索中占据明显优势, 使量子计算理论有了突破。

^{[7], [8]} 同时, 人工智能领域也在机器学习以及神经网络技术的推进下发展。^[9]

进入21世纪, 以深度学习为代表的人工智能技术以其强大的表达能力和大数据训练方法获得广泛应用, 在语音、语言和图像等处理中被作为核心技术, 尤其是深度卷积神经网络非常擅长于图像数据的处理。^[10] 21世纪以来, 由于量子技术的进步, 小型化的量子计算机研发成功, IBM、Google的量子平台以及Sycamore处理器等均展现了量子计算的潜力。^[11] 这一时期是量子人工智能的实验阶段和技术发展期, 量子机器学习和量子优化算法成为研究热点, 显示了量子计算在大数据、复杂优化问题方面的优势。^[12]

2. 量子人工智能的概念界定及其与传统人工智能的本质区分

随着科技的进步, 量子计算和人工智能的融合逐渐受到关注。学者们认为, 这种结合将形成新的技术形式, 即量子人工智能。量子人工智能的核心在于通过量子计算增强人工智能的处理能力, 尤其是在传统计算难以应对的复杂问题上展现独特的优势。^[13] 量子认知的成功为量子人工智能的未来应用提供了启示, 显示出量子技术和人工智能之间的互补性。^[14] 一方面, 人工智能可以增强量子系统的自动控制, 解决量子计算中存在的实际问题; 另一方面, 量子计算可以加速神经网络的训练, 推动机器学习的创新。^[15]

量子人工智能与传统人工智能的核心区别在于前者利用量子叠加和纠缠等特性, 极大提升了处理复杂问题的效率。量子位的叠加态和量子纠缠使量子计算机能够并行处理多种状态, 大幅提升计算精度。^[11] 相比之下, 传统人工智能依赖经典计算模型, 常在复杂问题上遭遇瓶颈, 而量子算法则在量子化学、优化问题和大数据分析等领域展现出传统方法难以匹敌的优势。^{[12], [16]-[18]}

二、量子人工智能面临的伦理困境

传统人工智能存在很多伦理争议, 涉及算

法公平、数据隐私保护以及自主决策的透明度等方面。虽然人工智能能够给经济增长、社会福利带来裨益,但是低可解释性、数据偏见以及安全隐患也使用户面临风险。^[19]量子人工智能发展过程中也存在类似的难题,并且量子计算能够实现并行处理,会对数据隐私造成潜在威胁,因此伦理方面的讨论也就显得更加困难紧迫。因此,量子人工智能的特殊性需要我们研究并解决这些伦理困境。

1. 隐私与数据泄露风险

随着量子人工智能的迅猛发展,隐私以及数据泄露成为学界所担忧的议题,尤其是对量子计算带来的对于传统的数据加密手段的威胁。传统的加密算法,如RSA(Rivest-Shamir-Adleman)和DSA(Digital Signature Algorithm),依赖于大整数分解和离散对数问题,虽然这些问题对于经典计算机来说是难以破解的,但是量子计算由于其叠加和纠缠的特性可以大大缩短破解时间。^[8]例如,Shor算法已经证实量子计算机可以在多项式时间内破解RSA加密,而经典计算需要指数级时间。^[8]2019年Google的量子霸权实验显示了量子计算的能力,给未来也许会冲破传统加密技术给予了实际可能。^[20]一旦量子计算成熟,现有的加密算法很快就会被破解,这就涉及隐私、商业和国家安全风险。^[21]

为了应对并化解这一威胁,后量子密码学的研究正在加速,并且主要聚焦于格密码学、多变量方程密码学以及哈希函数密码学。^[22]其中,格密码学因其结构上的复杂性被认为是目前抵御量子计算攻击潜力最大的一种方案。美国国家标准与技术研究院推动了对后量子加密算法的标准化过程,有一些算法如Crystals-Kyber和FrodoKEM已经接近完成标准化工作。^[23]但后量子加密技术还是存在密钥过大、计算资源耗费等方面的限制。

2. 量子算法决策与责任归属的争议

量子人工智能的复杂性不仅带来了技术上的革新,也催生了新的伦理争议,尤其在金融和医疗等高风险领域,量子算法决策的透明性和责任归属问题由此变得尤为关键。量子算法

依赖量子力学原理,能够高效处理复杂问题,显著加速决策过程,但其运作方式不同于传统算法,往往呈现“黑箱”特性,使得外界难以理解其具体决策逻辑。^[12]

在金融领域,量子算法已经被应用于优化投资组合、风险评估以及市场预测。例如,一些金融机构已经采用量子人工智能技术进行超高速交易,利用其强大的计算能力快速分析市场数据并自动生成交易决策。^[24]然而,量子算法的高复杂性使其决策逻辑晦涩难懂。例如,在量子算法交易中,研究者观察到该算法在处理金融衍生品时,可能会导致部分群体遭受不公平的利率或者融资条件,进而引发金融歧视。^[25]这种现象的背后是量子算法无意中放大了数据中的偏见,而由于量子计算的“黑箱”特性,外部审计和监管机构难以追踪这些决策的具体因果链,从而导致责任模糊的情况。又如,量子算法在处理金融数据时,由于过度依赖特定的历史数据集,导致某些低收入群体被标记为高风险客户,进而面临更高的贷款拒绝率。量子算法借助量子纠缠和叠加的计算方式放大了数据集中已有的偏见,使得部分决策具备歧视性。^[26]在这种情况下,责任的归属就变得尤为复杂:银行、程序员还是算法设计者,谁应当对于这一不公正结果承担责任?

在医疗领域,量子人工智能同样被用于提升诊断精确度、优化治疗方案以及开发个性化药物。量子算法能够分析大量的医疗数据,包括基因信息、影像数据和临床记录,从而提供更精确的诊疗方案。然而,复杂的量子计算过程也增加了责任归属的不确定性。某些量子算法在实际医疗应用中显示出对于少数群体的诊断精度较低的问题,例如,量子算法在分析癌症患者的基因组数据时,过度依赖来自某些特定人群的数据集,从而导致非代表性人群的诊断准确率降低。^[27]这种偏差可能直接影响病人的治疗效果,甚至威胁到生命安全。这种情况下,如果患者由于量子人工智能的错误诊断导致病情恶化,追究责任变得困难。很难判断是负责开发算法的程序员应当承担责任,还是负责部署该技术的医院或者医疗机构承担责任。

由于量子算法的高度复杂性，医疗机构和监管部门在面对这类技术时，难以明确量化责任归属。

在实际应用中，量子算法的“黑箱”特性常常导致歧视现象。无论是在金融还是医疗领域，量子算法通过对大数据的快速处理，可能无意中放大了数据中隐含的偏见和差异。例如，在信用评级系统中，量子算法可能由于训练数据集中的偏见而给某些社会群体打上“高风险”标签，导致他们在贷款或者保险申请中被不公平地拒绝。^[25]此外，在医疗诊断领域，不同种族和性别的病患由于训练数据的不均衡，可能面临算法偏差导致的错误诊断或者治疗方案不当的风险。^[27]在这种情况下，算法决策过程的不可解释性与复杂性引发了明显的伦理困境：当量子人工智能做出有偏见或者错误的决策时，谁应当对此负责？程序员无法完全理解算法的决策路径，医疗机构难以审查算法背后的逻辑，而患者和消费者则可能成为受害者。在一个责任归属尚未明晰的法律和伦理框架下，量子人工智能的广泛应用可能进一步加剧这一困境。

3. 算法偏见与歧视问题

量子人工智能的发展可能会加剧技术偏见，这些偏见源于数据集的不充分代表性、算法误差传播、优化目标设定或者算法设计中的偏差。这些偏见可能导致算法在实际应用中对于某些群体产生偏向，影响资源分配、机会获取或者社会互动，进而造成不公平的社会影响。

数据集选择和收集受到多种因素的影响，如数据渠道、样本代表性和数据采集者的文化背景。例如，在医疗领域的量子人工智能应用中，由于训练数据集中少数族裔样本较少，算法的诊断效果在不同人群中表现出显著差异。^[27]一项研究显示，量子算法在分析癌症患者的基因组数据时，对于白人患者的诊断准确率要显著高于其他族裔群体。基于量子计算的诊断系统在少数族裔患者中的错误率较高，尤其是在处理癌症病变和心血管疾病时，这种误差可能导致严重的医疗后果。^[28]这表明，由于数据集代表性不足，量子算法可能放大原本存在的

社会不平等，导致少数群体获得的医疗服务质量低于其他群体。

量子算法在金融领域的应用也暴露出类似的问题。量子人工智能在金融决策中被广泛应用于信用评级、贷款审批和风险评估等任务。然而，量子算法在处理历史金融数据时，可能无意中强化了数据集中存在的社会偏见。例如，一项针对量子算法信用评分的实验中发现，来自特定社会经济背景的人群（如低收入或者少数族裔）被标记为高风险客户的概率显著上升，尽管他们的实际信用行为并不支持这一结论。^[25]这种现象表明，量子算法可能会继承并且放大数据中的历史性偏见，使得这些群体在金融系统中的机会受到限制。

由于量子算法能够通过强大的计算能力提取数据中的隐藏模式，它们有可能放大现有的社会偏见。研究表明，量子计算不仅可以高效处理大规模数据，还可能在无意中强化已有的社会分层结构。^[25]在金融领域，某些算法被用于预测信用风险，这些预测基于历史数据进行优化。然而，如果历史数据本身就存在对于某些群体的不公平对待，量子算法可能会继续沿用这些不公平的模式，进而进一步加剧社会不平等。

因此，要加强伦理准则对算法设计、数据收集以及执行方面的应用。尽管已经取得了一些进展，量子人工智能的算法仍然需要优化，研究者需要不断探索有效解决方案，减少科技带来的偏见和歧视，实现科技进步与社会伦理的和谐共生。

4. 认知自主性与人类主体性面临的挑战

伴随着量子人工智能的发展，量子人工智能会不会替代人本身在认知过程中的主体性，成为了一个值得思考的问题。当下，弱人工智能已经在生活里普遍应用起来，在某些领域表现很好，可是它所拥有的能力被预设规则以及少量的数据所限制，无法执行广泛的任务。相比传统的人工智能，量子人工智能借助于量子非局域性、叠加和纠缠等特性，或许可以在模拟人类智能、模仿人脑认知现象等方面实现比传统人工智能更胜一筹的表现。^[29]

量子计算技术上的优势使得量子人工智能可以处理一些复杂的、大数据的问题等超越人类的情况。例如,量子机器学习对医疗数据处理的时候可以找到传统方法找不到的规律,进而带动医学发展。^[30]不过,量子人工智能究竟能否真正超越人类的认知水平也引起了不少的讨论。

公众担心量子人工智能也许会拥有超越人类的自主学习和决策能力,进而影响社会结构和价值体系。认知自主性在传统上是人类智能所特有的,也包括了情感、道德和伦理的成分。^[31]量子人工智能虽然在某些方面展现出了高效、准确的能力,但是是否可以完全代替或者超越人类对情感、道德、伦理等方面的判断仍需进一步探究。未来量子人工智能发展会怎样影响到这些方面,仍旧值得深究与审视。

三、量子人工智能伦理治理

量子人工智能属于前沿科技,拥有巨大的潜能,发展可能胜过传统人工智能,开拓崭新的应用范畴。不过,量子计算能力和量子算法逐步提升之后,量子人工智能就在技术上冲破了以前的边界,也带来新的伦理和社会问题。这些挑战涉及数据隐私、算法偏见、责任归属等方面,因此,研究和建立有效的伦理治理模式显得尤为重要。

虽然量子人工智能在理论上和实验上都有非常广阔的发展前景,但是它真正应用于生活还是一个探索的过程。量子人工智能的发展进步,将会给技术、社会结构、经济运行以及个人生活各方面带来巨大的影响。例如,它会改变现有的工作场所,改变人们之间的互动方式,并且会重新定义隐私和数据保护的标准。为了保证量子人工智能能够长久地发展下去,并且尽可能发挥出它的积极作用,就需要学者以及相关的专家共同努力去创建起一个科学、严谨的伦理治理框架。

这个框架要包含对量子人工智能技术全方位的评估和监管,数据隐私保护、算法透明度、公平性以及责任归属这些方面都要被纳入考虑

范围当中。制订并执行这些伦理规范之后,就能够保证量子人工智能的技术发展与社会的道德准则保持一致,并且还能够促进社会发展的过程中起到积极作用。

1. 伦理审查与监督

量子人工智能技术可能会引发隐私泄露和数据安全方面的问题,传统的伦理规范对于这种新的技术而言显得不足。所以,特别要创建起专门的伦理审查与监督体系,从而保证量子人工智能的发展符合社会道德和价值观念。根据《通用数据保护条例》(General Data Protection Regulation, GDPR)等已有法规,审查机制要着重注意如下几方面:

第一,数据最小化原则的审查与监督。GDPR明确规定了数据最小化原则,即量子人工智能系统只能收集和使用必要的数。^[32]但量子计算强大的处理能力使得系统可能会倾向于收集过多的数据并进行分析,所以应当定期对量子计算机进行单独的审查,确保其仅针对必需的数据进行收集。每次数据更新和用途变更都需要重新审查,而且要使用自动化工具实时监测数据处理过程,防止超出了合法正当的范围。

第二,目的限定及透明性机制。GDPR强调对数据使用的限制需合法,量子时代的数据处理速度与复杂程度可能导致数据被误用到非预期的目的上。所以要强化目的限定审查机制,添加技术手段以改善透明性,创建用途记录体系,使量子人工智能系统处理用户数据的时候把用途随时记录下来,并且备案到监管机构可以审查的数据库里。用户和监管者都可以访问此数据库,从而使得数据的处理过程能够透明化。同时设有用途变更通知,在用途发生变更时必须马上告知使用者并重新取得同意。^[32]

第三,使用以数据安全原则为基础的量子增强方案。量子计算有可能损害传统的加密手段,因此必须应用后量子加密技术,保障数据在加工和储存期间能够抵制量子攻击。在法律范围内强制采用这些加密技术,并且定期对量子人工智能的安全手段做动态检查,特别是针对加密和存储的保护部分要随时处于最新安全

防护状态。

第四，加强用户的同意取得和管理。在量子人工智能体系里，对用户同意外部信息的获得以及管理会更为复杂。由于量子系统的数据处理方法很复杂，用户很难知道自己的数据会如何被使用，所以需要加强同意。系统要给出清晰易懂的解释，告诉用户数据有什么用处、是怎么处理的、会带来哪些影响，还要利用人工智能改进的界面把复杂的技术信息简单化，让用户可以在充分了解之后作出同意。而且用户应该可以随时撤销同意并要求删除数据，量子系统应提供一个方便的接口来及时处理用户的对数据访问、修改或者删除的要求。^[32]

2. 算法透明与过程跟踪

量子人工智能具有复杂性以及“黑箱”特征，因此算法偏见、过程跟踪及责任归属成为了重要话题。面对这些挑战，想要提升决策透明度就非常必要，设计出透明的算法才是破解这类难题的主要方法。

首先，设计透明算法。量子人工智能系统要首先追求透明，让决策过程对用户和利益相关者具有可解读性。系统要对决策所依据的数据、算法以及推理过程做出详尽的展现，这既可改善用户的信任度，又能够在问题浮现时给予明晰的诠释和依照。^[33]可解释性人工智能（Explainable Artificial Intelligence）技术对此至关重要，它让算法推理逻辑显露出来，从而使用户明白量子人工智能的决策途径。在医疗领域，可解释性人工智能能够帮助医生追踪并检验诊断建议，提高治疗精确度。^[34]在金融领域，可解释性人工智能则可以显示风险评估模型的细节，使得监管者检查量子金融决策是否公平且合法。^[35]

其次，区块链技术的应用。区块链技术还可以进一步提高量子人工智能的透明度和可追溯性。通过分布式账本特性，每次决策的记录可以以不可篡改的方式存储下来，为整个决策过程提供安全透明的操作记录。在医疗领域，区块链技术可保证病患病历以及治疗情况不可被篡改，并加强其安全性和隐私度。^[36]在金融领域，区块链技术可以对每一次交易决策记录

详细的信息，并保证操作透明化，从而有据可循地做到合规。另外，在复杂的量子计算环境中，区块链还可以以去中心化的方式分发计算资源，并提供安全的数据存储和验证。^[37]

第三，日志记录与审计。为了进一步提高透明度和责任归属，量子人工智能系统应记录每次决策所用的数据输入、算法执行过程以及输出结果。这些日志有利于定期审视系统运作状况，尽早觉察到算法偏见或者错误。在医疗、金融这样的高风险行业里，日志记录给监管机构以及外部审计给予了追踪线索，保证系统能够在合法的范围之内运作。例如，在医疗领域，日志记录可以保存每次诊断意见的完整路径，给以后可能发生的医疗事故提供决策时重要的证据支撑；再者在金融行业，日志记录能够保留每一次的投资以及风险评估决策的细节情况，给审计和合规评估提供依据。

最后，审计机制的引入。有效的审计机制是保证量子人工智能系统的透明和公正的重要手段。为保障评估的客观性及公平性，审计应由内部队伍或者独立第三方机构按时进行。审查量子系统各步骤，涵盖数据处理、算法执行、模型训练、决策生成等环节，从而证实其符合性与透明度。在金融领域，独立审计机构根据审计日志对量子模型的公正性、透明度进行评价；在医疗领域，审计能够保证量子算法对数据的处理遵守隐私保护方面的法规要求。

3. 情境伦理决策与偏见消解

量子人工智能在不同的领域的使用时会遇到不同的伦理难题，所以要解决这些问题必须结合具体的应用场景，而不能只是依据一般的道德规范。

在医疗方面，量子人工智能包含敏感的个人健康数据，这就需要在隐私保护、数据安全以及病患自主权之间找到恰当的调和点。一方面算法偏见常常是由于数据集缺乏代表性或者不平衡造成的，比如医学影像以及基因数据往往来源于特定的族群，从而使得该算法在诊断少数族裔或者是某一性别的患者的时候会产生不准确的情况。^[27]为了减轻这种偏见，用不同种族、性别、年龄等数据集的多样性特征可以

加强算法的应用型以及诊断准确性。此外,还可以引入公平性算法,比如差异性隐私、公平性损失函数等,使得量子人工智能在不同群体的诊断精度更均衡,避免出现系统性的不正。^[33]在癌症诊断和个性化治疗时,量子人工智能凭借对海量基因数据的分析而给出准确的治疗。医疗机构可以采用对抗性训练技术“对抗”不同的数据集,从而减少潜在的算法歧视,提升诊断准确度和算法鲁棒性。^[27]

在金融领域,量子人工智能被用来做信用评分、贷款审批和风险评估,不过过去的数据库里藏着偏见,可能让系统无意识地加大种族、性别和社会经济地位之类的不公。^[25]对抗性训练可以发现并纠正偏见,如系统可以引入虚拟对手模型同算法的偏见“对抗”以减少决策中的偏见。公平性算法可以保证贷款审批这类金融决策在不同的申请人群体之间的一致性。可解释性模型(比如“白箱”模型)可以提供一个透明的决策途径,让金融监管机构审查其是否遵守了反歧视法律框架下的公平性要求。^[33]量子人工智能在信用评估系统中的应用可以利用大量的信用记录,使用多种多样、具有公允的数据集和算法,使历史信用数据中的社会偏见不再成为问题,不会因为数据中有歧视,就拒绝某些人的贷款申请。金融机构可以设置独立的算法检查制度,在部署算法之前执行全面的偏见检测与公平性检验,以此保证符合反歧视法规。

另外,数据采集策略的选择多样与工具设计也对消除算法偏见有着至关重要的影响。采用代表性强且多样化的数据集可以为量子人工智能系统提供全面的学习材料,减少因数据偏差而产生的不公平决策。采取多样化数据收集策略以保证不同社会经济背景、种族、性别等因素在数据集里具有平衡的代表性,从而防止少部分群体被排除于数据之外。以医疗系统为例,利用分层采样技术保证各个患者的群体数据均能够被公平采集。^[27]同理,偏见检测工具也可以实时监测到数据的采集过程,防止因为地域、性别或者经济原因导致的数据采集有偏差的情况出现。金融机构也能够经由这些工具

来监测贷款申请数据,从而保证在搜集并处理数据时各类群体都能够受到公平对待。

4. 伦理对话与共识构建

随着技术发展突飞猛进,量子人工智能的伦理问题引起关注。要应对这些问题,需要跨学科专家之间展开对话和合作,在如下几个方面思考:第一,跨学科对话。量子物理学家能够去探究量子技术的原理和潜在风险,计算机科学家就可针对量子算法的达成和效果展开研究。法律学者可以研究量子人工智能涉及的法律问题,社会学家也可以分析它对社会结构、经济活动以及人际关系产生的影响。这方面的跨学科合作能够帮助发现量子人工智能所遇到的伦理难题,进而找到解决方法。^[38]第二,形成共识。形成伦理共识是有效治理的前提,经由跨学科的深入探讨,专家们可以一起界定隐私保护、数据利用以及算法公平性这些关键伦理问题。这种共识可以促进形成一种标准化的伦理评估框架,从而给技术应用以明确的方向性指导,在帮助确立起统一的伦理标准的同时也能够为量子人工智能的实际运用提供相应的规范指引。第三,伦理评估与认证。制定包括透明性、隐私保护、公平性以及责任在内的伦理评估标准,并对符合这些标准的系统进行认证,这是执行伦理治理的关键步骤。认证机构应该凭借技术审查和道德审查,判定量子人工智能系统的合规情况,对符合标准的系统予以认证。这样既提高了该系统的可信度,又提升了社会对量子人工智能技术的信任程度。总的来说,量子人工智能技术的伦理研究与应用,要依靠跨学科合作、达成伦理共识并制订全面评估准则。这些举措可以保证技术负责任、可持续地发展,在促进技术发展的过程中,尽可能少地产生一些伦理问题和社会风险。

余 论

量子人工智能是一种新兴技术,具有潜在的革命性影响,可能会超过人类现有的智能水平,从而带来社会、法律和伦理方面的重大问题。为了应对这些挑战,需要各方合作来创建

科学的伦理框架以及评估机制，并且推动透明而负责任的技术实践活动以确保可持续发展。虽然量子人工智能的研究还处于摸索当中，但预防相关的伦理问题也很重要。未来的相关研究也需要重视这一方面，并且要尽力将量子人工智能技术所带来的正面效果发挥到最大程度并尽可能减少负面的影响。这就包含建立应对复杂伦理难题的体系，确保量子人工智能安全可靠、公平公正并且持久持续发展。这样做的目的既能够让科技不断进步，又可以保证科技符合社会道德标准，促进社会发展。

[参考文献]

- [1] 陈向群、郭晓. 人工智能的量子方法及其辨析[J]. 科学技术哲学研究, 2019, 36(2): 8-13.
- [2] McCulloch, W. S., Pitts, W. 'A Logical Calculus of the Ideas Immanent in Nervous Activity'[J]. *Bulletin of Mathematical Biophysics*, 1943, 5(4): 115-133.
- [3] Turing, A. M. 'Computing Machinery and Intelligence'[J]. *Mind*, 1950, 59(236): 433-460.
- [4] McCarthy, J., Minsky, M. L., Rochester, N., et. al. 'A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence'[J]. *AI Magazine*, 1955, 27(4): 12-14.
- [5] Clancey, W. J. 'The Epistemology of a Rule-Based Expert System—A Framework for Explanation'[J]. *Artificial Intelligence*, 1983, 20(3): 215-251.
- [6] Feynman, R. P. 'Simulating Physics with Computers'[J]. *International Journal of Theoretical Physics*, 1982, 21(6/7): 467-488.
- [7] Shor, P. W. 'Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer'[J]. *SIAM Review*, 1999, 41(2): 303-332.
- [8] Long, G. L. 'Grover Algorithm with Zero Theoretical Failure Rate'[J]. *Physical Review A*, 2001, 64(2): 022307.
- [9] Mitchell, T. M. *Machine Learning*[M]. New York: McGraw Hill, 1997, 8-23.
- [10] LeCun, Y., Bengio, Y., Hinton, G. 'Deep Learning'[J]. *Nature*, 2015, 521(7553): 436-444.
- [11] Kelly, J., Barends, R., Fowler, A. G., et. al. 'State Preservation by Repetitive Error Detection in a Superconducting Quantum Circuit'[J]. *Nature*, 2015, 519(7541): 66-69.
- [12] Biamonte, J., Wittek, P., Pancotti, N., et. al. 'Quantum Machine Learning'[J]. *Nature*, 2017, 549(7671): 195-202.
- [13] Taylor, R. D. 'Quantum Artificial Intelligence: A "Precautionary" US Approach'[J]. *Telecommunications Policy*, 2020, 44(6): 101909.
- [14] 刘燊、汪澜. 量子认知对意识的“难问题”的突破及其展望[J]. 自然辩证法研究, 2023, 39(9): 54-58.
- [15] Zhu, Y., Yu, K. 'Artificial Intelligence (AI) for Quantum and Quantum for AI'[J]. *Optical and Quantum Electronics*, 2023, 559(8): 697.
- [16] 刘燊. 量子认知的研究回溯、核心应用领域与未来展望[J]. 自然辩证法通讯, 2024, 46(4): 41-48.
- [17] Aspuru-Guzik, A., Walther, P. 'Photonic Quantum Simulators'[J]. *Nature Physics*, 2012, 8(4): 285-291.
- [18] Cao, Y., Romero, J., Olson, J. P., et. al. 'Quantum Chemistry in the Age of Quantum Computing'[J]. *Chemical Reviews*, 2019, 119(19): 10856-10915.
- [19] Floridi, L., Sanders, J. W. 'Artificial Evil and the Foundation of Computer Ethics'[J]. *Ethics and Information Technology*, 2004, 6(3): 203-214.
- [20] Arute, F., Arya, K., Babbush, R., et. al. 'Quantum Supremacy Using a Programmable Superconducting Processor'[J]. *Nature*, 2019, 574(7779): 505-510.
- [21] Preskill, J. 'Quantum Computing in the NISQ Era and Beyond'[J]. *Quantum*, 2018, 2: 79.
- [22] Dong, X., Li, Z., Wang, X. 'Quantum Cryptanalysis on Some Generalized Feistel Schemes'[J]. *Science China Information Sciences*, 2019, 62(2): 22501.
- [23] Bernstein, D. J., Lange, T. 'Post-Quantum Cryptography'[J]. *Nature*, 2017, 549(7671): 188-194.
- [24] Egger, D. J., Gambella, C., Marecek, J., et. al. 'Quantum Computing for Finance: State-of-the-art and Future Prospects'[J]. *IEEE Transactions on Quantum Engineering*, 2020, 1: 1-24.
- [25] Lu, Y., Yang, J. X. 'Quantum Financing System: A Survey on Quantum Algorithms, Potential Scenarios and Open Research Issues'[J]. *Journal of Industrial Information Integration*, 2024, 41: 100663.
- [26] Shaydulin, R., Li, C., Chakrabarti, S., et. al. 'Evidence of Scaling Advantage for the Quantum Approximate Optimization Algorithm on a Classically Intractable Problem'[J]. *Science Advances*, 2024, 10(22): eadm6761.
- [27] Dlamini, Z., Skepu, A., Kim, N., et. al. 'AI and Precision Oncology in Clinical Cancer Genomics: From Prevention to Targeted Cancer Therapies-an Outcomes Based Patient

- Care'[J]. *Informatics in Medicine Unlocked*, 2022, 31: 100965.
- [28] Ramesh, S., Tomesh, T., Riesenfeld, S. J., et. al. 'Quantum Computing for Oncology'[J]. *Nature Cancer*, 2024, 5: 811-816.
- [29] Mulgan, T. 'Superintelligence: Paths, Dangers, Strategies'[J]. *The Philosophical Quarterly*, 2016, 66(262): 196-203.
- [30] Giordani, T., Koch-Janusz, M., Eisert, J. 'Quantum-Enhanced Machine Learning'[J]. *Nature Reviews Physics*, 2019, 1(9): 625-638.
- [31] Lauscher, A. 'Life 3.0: Being Human in the Age of Artificial Intelligence'[J]. *Internet Histories*, 2019, 3(1): 101-103.
- [32] Hoofnagle, C. J., Van Der Sloot, B., Borgesius, F. Z. 'The European Union General Data Protection Regulation: What It is and What It Means'[J]. *Information & Communications Technology Law*, 2019, 28(1): 65-98.
- [33] 孟令宇. 从算法偏见到算法歧视: 算法歧视的责任问题探究[J]. *东北大学学报(社会科学版)*, 2022, 24(1): 1-9.
- [34] Zhang, Y., Weng, Y., Lund, J. 'Applications of Explainable Artificial Intelligence in Diagnosis and Surgery'[J]. *Diagnostics*, 2022, 12(2): 237.
- [35] Chen, X. Q., Ma, C. Q., Ren, Y. S., et. al. 'Explainable Artificial Intelligence in Finance: A Bibliometric Review'[J]. *Finance Research Letters*, 2023, 56: 104145.
- [36] Chen, Y., Ding, S., Xu, Z., et. al. 'Blockchain-based Medical Records Secure Storage and Medical Service Framework'[J]. *Journal of Medical Systems*, 2019, 43: 1-9.
- [37] Varma, J. R. 'Blockchain in Finance'[J]. *Vikalpa*, 2019, 44(1): 1-11.
- [38] Lin, P., Lewis, B., Cogburn, D. 'Examining the Potential of Autonomous Vehicles with a Focus on Equity'[J]. *Transportation Research Part C: Emerging Technologies*, 2020, 113: 477-489.

[责任编辑 李斌]