

从伦理风险到韧性重构：人工智能治理的 四维驱动机制与实践进路

宋雁宾

北京建筑大学城市经济与管理学院，北京

收稿日期：2025年5月29日；录用日期：2025年6月13日；发布日期：2025年7月30日

摘要

人工智能技术的快速发展在提升产业效率的同时，也引发了隐私泄露、算法歧视等伦理风险，如何通过伦理治理增强产业韧性成为数字经济时代的重要命题。本研究使用文献综述法和逻辑推理法，构建“风险识别-韧性响应-治理迭代”的动态模型，运用跨学科分析框架和产业集群实证案例，系统解析人工智能伦理风险对产业韧性的三重侵蚀路径：技术滥用引发的社会信任崩塌、合规成本激增导致的创新资源挤占、技术迭代受阻形成的路径锁定效应。研究发现，伦理治理通过技术透明化重塑信任基础、敏捷治理提升风险免疫能力、伦理嵌入驱动创新范式转型、多元共治构建韧性网络生态四维机制，形成“治理投入-韧性增益”的价值闭环。本研究创新性地揭示了伦理风险与产业韧性的相互作用机制，提出的分层治理工具为差异化场景治理提供解决方案，通过开发技术适配性工具与建设国际标准话语权，为全球人工智能治理贡献理论范式与实践路径。

关键词

人工智能伦理治理，产业韧性，理论模型，作用机制，实践进路

From Ethical Risk to Resilience Reconstruction: The Four-Dimensional Driving Mechanism and Practical Approach of Artificial Intelligence Governance

Yanbin Song

School of Urban Economics and Management, Beijing University of Civil Engineering and Architecture, Beijing

Received: May 29th, 2025; accepted: Jun. 13th, 2025; published: Jul. 30th, 2025

文章引用：宋雁宾. 从伦理风险到韧性重构：人工智能治理的四维驱动机制与实践进路[J]. 现代管理, 2025, 15(7): 207-214. DOI: 10.12677/mm.2025.157204

Abstract

The rapid development of artificial intelligence technology has not only improved industrial efficiency, but also triggered ethical risks such as privacy breaches and algorithm discrimination. Enhancing industrial resilience through ethical governance has become an important proposition in the digital economy era. This study uses literature review and logical reasoning methods to construct a dynamic model of “risk identification-resilience response-governance iteration”. Using an interdisciplinary analysis framework and empirical cases of industrial clusters, it systematically analyzes the triple erosion path of artificial intelligence ethical risks on industrial resilience: the collapse of social trust caused by technology abuse, the crowding out of innovation resources caused by the sharp increase in compliance costs, and the path locking effect formed by the obstruction of technology iteration. Research has found that ethical governance reshapes trust foundations through technological transparency, enhances risk immunity through agile governance, drives innovation paradigm transformation through ethical embedding, and constructs a four-dimensional mechanism of resilience network ecology through diverse co-governance, forming a value loop of “governance investment-resilience gain”. This study innovatively reveals the interaction mechanism between ethical risks and industrial resilience and proposes a hierarchical governance tool to provide solutions for differentiated scenario governance. By developing technology adaptation tools and building international standard discourse power, it contributes theoretical paradigms and practical paths to global artificial intelligence governance.

Keywords

AI Ethical Governance, Industrial Resilience, Theoretical Model, Mechanism of Action, Practical Approach

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

人工智能技术的快速迭代正深刻重构全球产业发展格局，其在提升效率、优化决策的同时，也引发了隐私泄露、算法歧视、责任归属模糊等伦理风险[1]。尤其在产业领域，技术应用的先行性与场景复杂性放大了伦理问题的系统性影响。与此同时，全球化进程中的供应链波动、地缘政治冲突及技术壁垒加剧了产业发展的不确定性，产业韧性成为应对复杂环境的核心能力。如何平衡 AI 技术驱动的效率优势与伦理风险衍生的脆弱性，成为学界与业界共同关注的焦点。现有研究多聚焦 AI 伦理风险的分类与治理路径，但对其与产业韧性间的作用机制仍缺乏系统性解析。具体而言，需回答两个问题：其一，AI 伦理风险如何通过技术滥用、信任缺失等路径削弱产业系统的适应性、抗压性与恢复力？其二，如何通过伦理治理将风险转化为韧性增强因子？本研究使用文献综述法和逻辑推理法，构建“风险识别-韧性响应-治理迭代”的动态模型，揭示伦理风险与产业韧性的相互作用关系，对现有研究进行有益补充。基于产业集群、园区治理等场景的实证分析，提出分级分类的伦理治理工具箱，为政策制定者优化 AI 产业监管框架、为企业构建“伦理合规-风险缓冲”机制提供决策依据，助力实现“技术向善”与“韧性增长”的协同目标。

2. 文献综述

(一) 人工智能伦理风险

近年来,人工智能技术的快速发展在体育、医疗、教育、环境治理等领域带来显著效益,但也引发了复杂的伦理风险。研究表明,AI伦理风险具有跨领域共性,表现为数据隐私泄露[2][3]、算法歧视与透明度缺失[4][5]、责任归属困境[6]以及技术滥用导致的主体性削弱[7]等问题。不同应用场景存在异质性风险,体育领域关注竞赛公平性,医疗领域强调患者自主权与安全性,而教育领域则侧重数据伦理与教育公平[8]。研究揭示,技术加速主义下AI伦理风险的生成源于技术缺陷、利益驱动与制度滞后[9],需构建多层次治理体系。当前国际共识聚焦透明度、公平性、隐私保护等七大原则,但人工智能伦理风险评估标准存在地域性差异。针对治理路径,学者提出了多元框架。体育领域建立“防-建-管”三层体系与跨学科平台,教育领域基于4R危机管理模型设计多方协同防范机制,算法审计制度因兼具约束力与灵活性成为新兴解决方案。共性治理策略包括加强顶层设计、推动风险评估标准化、强化伦理教育,以及平衡技术创新与伦理规制。

(二) 产业韧性研究的多维拓展

近年来产业韧性研究呈现多维化发展趋势,学者们通过差异化的研究视角与方法论构建推动了该领域的理论深化。在研究对象维度,既有研究覆盖了文化产业[10]、生猪产业[11]、农业产业[12]等具体产业门类,形成了“特殊产业韧性-共性韧性机理”的研究路径。其中,田晓仪等创新性地揭示了平台经济与产业共聚对文化产业的韧性赋能机制;郑瑞强团队则通过构建四维韧性评价体系,实证揭示了生猪产业韧性提升的结构性障碍。在研究范式层面,学界呈现出定量测度与质性分析的混合研究态势。尤亮等[13]通过系统梳理产业韧性理论谱系,构建了涵盖经济、社会、生态系统的三维韧性框架,为后续研究提供了基础性分析工具。而孙良顺[14]提出的“资源能力进阶”模型,则为乡村产业韧性发展提供了动态演化视角下的理论解释。数字技术赋能与产业韧性关联性研究成为新兴热点。舒长江等[15]基于新结构经济学框架,揭示了数字经济与高新技术产业韧性间的非线性关系;朱建民等[16]则创新性地将全球价值链协调视角引入制造业韧性研究,构建了“产业韧性-绿色创新-耦合发展”的作用链模型。这些研究在方法论上融合熵权法、耦合协调模型等多元工具,推动产业韧性研究向纵深发展。

既有文献在产业韧性领域形成了丰富的理论积淀,但存在三个显著的研究盲区:其一,现有研究多聚焦传统风险类型(如自然灾害、市场波动等)对产业韧性的冲击,对人工智能伦理风险这一新型系统性风险缺乏必要关注;其二,研究视角局限于技术创新、组织变革等传统驱动因素,尚未构建伦理治理与产业韧性提升的理论关联;其三,现有分析框架多采用单一维度或静态视角,缺乏对人工智能伦理风险与产业韧性动态演化机制的系统性解构。这为本研究突破传统分析范式,从人工智能伦理风险治理视角构建产业韧性提升的创新性理论框架提供了学术创新空间。

3. 人工智能伦理风险冲击产业韧性的理论模型

若某类系统性风险会通过破坏信任、增加合规成本或引发技术滥用对产业韧性产生负面影响,则该类风险必然导致产业韧性受损。人工智能伦理风险是一类系统性风险,其具体表现为算法歧视、数据隐私泄露、技术失控等。因此,人工智能伦理风险会通过上述机制对产业韧性产生负面影响。

(一) 社会信任崩塌引发产业防御工事溃退

人工智能伦理风险引发社会信任崩塌对产业韧性的影响是多方面的。从数据维度来看,隐私泄露等问题会严重损害公众对人工智能技术的信任。当个人敏感信息被滥用或泄露时,人们会担心自己的隐私安全,从而对使用人工智能产品产生抵触情绪。这种信任的缺失会导致市场需求下降,进而影响产业的抵抗力。例如,在医疗领域,如果患者担心自己的健康数据被泄露,他们可能会拒绝使用人工智能辅助诊断系统,这将限制医疗产业的发展。从算法维度来看,算法黑箱和算法偏见等问题也会削弱社会信任。当算法决策过程不透明且存在歧视性时,人们会质疑其公正性和可靠性。这种不信任会进一步影响产业

的恢复力。例如，在招聘领域，如果人工智能招聘系统存在性别歧视，企业可能会面临法律诉讼和声誉损失，从而影响其恢复能力。**从系统维度来看**，系统漏洞和价值冲突等问题会引发公众对人工智能系统的担忧。当系统失效或违背人类价值观时，人们会失去对人工智能技术的信心。这种信任的崩塌会影响产业的转型力。例如，在交通领域，如果自动驾驶车辆因系统漏洞而失控，公众可能会对自动驾驶技术失去信任，从而阻碍交通产业的转型。**从人因维度来看**，设计者偏见和用户交互偏见等问题也会导致社会信任的下降。当算法设计者将个人偏见嵌入算法或用户与 AI 系统的交互数据包含偏见时，人们会认为人工智能技术存在不公平和不公正。这种不信任会影响产业的抵抗力、恢复力和转型力。例如，在金融领域，如果人工智能信贷系统存在种族歧视，银行可能会面临客户流失和声誉损失。¹

(二) 合规成本激增致使产业修复引擎阻滞

人工智能伦理风险导致合规成本激增对产业韧性的影响也是显著的。**从数据维度来看**，数据偏见等问题会引发监管机构的关注，导致企业需要投入更多资源来确保数据合规。例如，企业可能需要加强数据加密、访问控制和审计等措施，以防止数据泄露和滥用。这些措施会增加企业的运营成本，从而影响其抵抗力。**从算法维度来看**，算法黑箱和算法偏见等问题也会导致合规成本的增加。企业需要投入更多资源来解释算法决策过程、消除算法偏见和确保算法公平性。例如，企业可能需要开发新的算法解释工具、进行算法审计和培训员工等。这些措施会增加企业的研发成本和管理成本，从而影响其恢复力。**从系统维度来看**，系统漏洞和价值冲突等问题会引发监管机构的严格审查，导致企业需要投入更多资源来确保系统合规。例如，企业可能需要加强系统安全测试、修复漏洞和调整系统设计等。这些措施会增加企业的技术成本和时间成本，从而影响其转型能力。**从人因维度来看**，设计者偏见和用户交互偏见等问题也会导致合规成本的增加。企业需要投入更多资源来提高员工的伦理意识和确保用户交互数据的公正性。例如，企业可能需要开展伦理培训、建立用户反馈机制和改进算法设计等。这些措施会增加企业的人力成本和管理成本，从而影响其抵抗力、恢复力和转型力。

(三) 技术迭代受阻导致产业进化路径僵化

人工智能伦理风险导致技术迭代受阻对产业韧性的影响也是不可忽视的。**从数据维度来看**，数据安全会限制数据的收集和使用，从而影响技术迭代的速度。例如，企业可能因为担心数据隐私问题而不敢收集更多的用户数据，这将限制人工智能模型的训练和优化。这种技术迭代受阻会影响产业的抵抗力，因为企业无法及时推出更优质的产品和服务来满足市场需求。**从算法维度来看**，算法黑箱和算法偏见等问题会阻碍算法的创新和改进。当算法决策过程不透明且存在歧视性时，研究人员可能会面临更多的挑战来改进算法。例如，开发新的算法解释工具和消除算法偏见需要大量的时间和资源。这种技术迭代受阻会影响产业的恢复力，因为企业无法快速响应市场变化和技术进步。**从系统维度来看**，系统漏洞和价值冲突等问题会限制系统的升级和优化。当系统存在漏洞或违背人类价值观时，企业需要投入更多资源来修复漏洞和调整系统设计。例如，自动驾驶车辆的安全性和可靠性问题需要大量的测试和改进。这种技术迭代受阻会影响产业的转型能力，因为企业无法及时推出更先进的系统来满足市场需求。**从人因维度来看**，设计者偏见和用户交互偏见等问题会限制技术创新和应用。当算法设计者将个人偏见嵌入算法或用户与 AI 系统的交互数据包含偏见时，研究人员可能会面临更多的挑战来改进算法和系统。例如，开发新的算法设计方法和用户交互机制需要大量的时间和资源。这种技术迭代受阻会影响产业的抵抗力、恢复力和转型力，因为企业无法及时推出更优质的产品和服务来满足市场需求。

4. 人工智能伦理治理驱动产业韧性提升的作用机制

本部分基于负责任创新理论、敏捷治理理论及复杂适应系统理论，从技术、制度、价值、生态四个

¹ 本研究中的四种分类参考了《人工智能伦理风险分析报告 2019》等资料。

层面，系统阐述人工智能伦理治理如何通过重塑信任、优化资源配置、激发创新活力与构建韧性网络，驱动产业抵抗力的强化、恢复力的提升与转型力的激活。

（一）技术透明化与信任强化机制：增强产业抵抗力

技术透明化与信任强化机制在增强产业抵抗力方面发挥着至关重要的作用。技术透明化通过提升决策可解释性，如算法白盒化，重建社会信任。根据社会资本理论，社会信任是社会关系和经济活动的润滑剂，能够降低交易成本，促进合作和创新。当技术透明化时，消费者和投资者的风险感知会降低，从而稳定市场需求与资本投入。信号传递理论进一步指出，企业主动披露伦理治理信息，如数据使用协议、算法审计报告，向市场传递可信信号，增强品牌声誉与市场竞争力。在需求端，可解释性 AI 工具能够降低用户疑虑，提升产品采纳率。在供给端，透明化能够减少法律纠纷与监管处罚。欧盟的《AI 法案》预认证机制降低了企业的合规不确定性。政府也会基于企业的透明度评级提供税收优惠，以 IBM Watson Health 为例，通过公开医疗 AI 训练数据来源与决策逻辑，获得了 HIPAA 合规认证以及市场份额提升。技术透明化不仅强化了社会信任，还通过优化资源配置和政策环境改善，增强了产业的抵抗力。

（二）敏捷治理与风险免疫机制：提升产业恢复力

人工智能技术的快速发展和广泛应用，给社会带来了巨大的机遇和挑战。在面对这些挑战时，敏捷治理和风险免疫机制成为了提升产业恢复力的重要手段。敏捷治理理论强调动态调整监管框架，如监管沙盒、伦理豁免期，在风险可控的前提下加速技术迭代，缩短危机恢复周期。冗余理论则指出，通过系统冗余设计，如多模态数据备份、算法容错机制，能够增强技术鲁棒性，降低突发冲击的破坏性。在实际应用中，监管沙盒允许企业在封闭环境中测试伦理争议技术，如上海临港自动驾驶伦理试验区，缩短了产品上市周期，测试效率提升了 40%。区块链赋能的伦理审计链，如腾讯的“伦理链”，能够实现问题追溯与责任厘清，减少纠纷处理时间。此外，分级治理体系引导资源向高优先级领域倾斜，如核心领域的“红线清单”与创新领域的“豁免期”。欧盟的 AI 伦理认证体系通过动态更新的高风险 AI 清单，使企业能够针对性优化合规策略。敏捷治理和风险免疫机制通过快速响应能力、危机修复效率和资源弹性配置，提升了产业的恢复力，使产业能够在面对突发冲击时迅速恢复。

（三）伦理嵌入与创新协同机制：激活产业转型力

在人工智能技术的发展过程中，伦理嵌入和创新协同机制对于激活产业转型力具有重要意义。负责任创新理论强调将伦理约束内嵌至技术研发流程，如隐私保护设计，推动技术向“社会可接受”方向演进，避免路径锁定。协同演化理论指出，伦理标准与技术创新协同演化，如可解释性算法驱动医疗 AI 升级，能够开辟新兴市场空间。在技术层面，隐私增强技术催生了联邦学习、差分隐私等新范式，重塑了数据驱动型产业。例如，医疗联邦学习平台支撑了跨机构模型训练。在商业模式方面，伦理合规产品形成了差异化竞争优势。在全球范围内，主导国际伦理标准能够减少技术贸易壁垒，拓展海外市场。例如，中国的大疆无人机通过 ECOA 认证重返了欧洲市场。伦理嵌入和创新协同机制通过技术路线转型、商业模式创新和全球标准主导，激活了产业的转型动力，推动了产业的转型升级。

（四）多元共治与生态韧性机制：构建产业韧性网络

在人工智能技术的发展过程中，多元共治和生态韧性机制对于构建产业韧性网络具有重要意义。生态系统理论强调政府、企业、公众、国际组织协同参与治理，形成风险分散与能力互补的韧性网络。集体行动理论指出，通过行业联盟制定自律规范，能够降低个体治理成本。在实际应用中，跨产业链伦理数据共享能够降低单点失效风险。例如，汽车行业联合建立自动驾驶事故数据库。政企合作开发公共伦理工具能够减少中小企业技术门槛。例如，国家 AI 伦理开放平台为中小企业提供了技术支持。在国际层面，区域伦理联盟增强了全球供应链稳定性。例如，RCEP 框架下的 AI 伦理互认机制促进了区域内的合

作。多元共治和生态韧性机制通过风险分散、能力互补和国际协同，构建了产业的韧性网络，使产业能够在面对复杂多变的环境时保持稳定和持续发展。

5. 人工智能伦理治理驱动产业韧性提升的实践进路

(一) 制度层面：构建分层治理体系

1) 核心领域实施伦理红线清单

在人工智能技术飞速发展的当下，核心领域如国防、能源等关乎国家安全与社会稳定的行业，面临着前所未有的机遇与挑战。为确保这些领域的人工智能应用不偏离正轨，实施伦理红线清单至关重要。伦理红线清单应明确列出在这些领域中绝对禁止的人工智能应用行为，如在国防领域，禁止使用人工智能进行违反国际人道法的自主武器攻击；在能源领域，禁止利用人工智能进行恶意破坏能源设施等行为。通过制定并严格执行伦理红线清单，可以为这些核心领域的人工智能应用设定明确的边界，防止因技术滥用而引发的严重后果，保障国家和社会的核心利益，为产业的稳定发展提供坚实的基础。

2) 民生领域推行伦理影响星级评估

民生领域的 AI 应用，如医疗、教育等，直接关系到民众的切身利益。推行伦理影响星级评估，能够有效衡量这些应用的伦理风险与价值。以医疗领域为例，对于一款 AI 诊断系统，评估其在数据隐私保护、诊断准确性、对患者权益的影响等方面的表现，给予相应的星级评定。星级评估结果不仅为医疗机构和教育机构选择合适的 AI 产品提供了参考依据，也促使企业更加注重产品的伦理合规性。教育领域的 AI 教学软件，通过星级评估，可以确保其内容符合教育伦理，不会对学生产生不良影响。这种评估机制有助于推动民生领域 AI 应用的健康发展，提升民众对 AI 技术的信任度，增强产业的韧性。

3) 创新领域设立伦理豁免期

创新领域如元宇宙、AIGC 等，充满了无限的可能性，但也面临着诸多不确定的伦理风险。设立伦理豁免期，为这些领域的创新提供一定的发展空间，同时引导其在探索过程中逐步建立伦理规范。在伦理豁免期内，企业可以在遵守基本法律法规的前提下，大胆尝试新的技术和应用模式。例如，元宇宙中的虚拟社交平台，可以在豁免期内探索如何平衡用户隐私保护与社交互动的需求。AIGC 领域的创作者，可以尝试利用 AI 生成各种创意内容，同时思考如何避免侵权和虚假信息传播等问题。政府和相关机构应在这个过程中密切跟踪和评估，及时总结经验教训，为后续制定更为完善的伦理规范提供依据，促进创新领域的可持续发展。

(二) 技术层面：开发适配性伦理工具

1) 可解释性算法与隐私增强技术的应用

在人工智能技术的应用中，可解释性算法和隐私增强技术的应用至关重要。可解释性算法能够为 AI 决策提供清晰的逻辑和依据，使人们能够理解 AI 系统是如何得出结论的。例如，在医疗诊断中，医生可以通过可解释性算法了解 AI 诊断系统的判断依据，从而更好地与患者沟通和做出决策。隐私增强技术则能够有效保护用户的个人数据，防止数据泄露和滥用。在金融领域，银行可以利用隐私增强技术确保客户的账户信息和交易记录的安全。通过应用这些技术，不仅能够提升用户对 AI 系统的信任度，还能够降低企业因数据隐私问题而面临的法律风险，为产业的健康发展提供技术支持。

2) 区块链赋能的伦理审计与追溯机制

区块链技术以其不可篡改、可追溯的特点，为人工智能伦理审计与追溯提供了强大的工具。在 AI 系统的开发和应用过程中，区块链可以记录每一个环节的数据和操作，确保数据的真实性和完整性。例如，在供应链管理中，利用区块链技术可以追溯产品的生产、运输和销售过程，确保 AI 系统在各个环节的决策都符合伦理规范。同时，区块链赋能的伦理审计机制可以对 AI 系统进行定期审计，及时发现和纠正潜

在的伦理问题。在智能合约的应用中，区块链可以确保合约的执行过程透明、公正，防止因人为干预或技术漏洞而导致的道德行为，提升产业的公信力。

(三) 国际层面：提升标准话语权

1) 主导区域标准

在经济全球化的背景下，区域标准的制定对于人工智能产业的发展具有重要意义。以 RCEP 框架下的 AI 伦理准则为例，通过主导该准则的制定，区域内各国可以在人工智能伦理方面达成共识，促进区域内的人工智能产业协同发展。主导区域标准的制定，不仅可以提升本国在国际舞台上的话语权，还可以引导区域内其他国家遵循更加合理的伦理规范。例如，在跨境数据流动方面，RCEP 框架下的 AI 伦理准则可以规定数据的收集、存储和使用必须符合一定的伦理标准，保护用户隐私和数据安全。同时，区域标准的制定也有助于消除贸易壁垒，促进区域内的人工智能产品和服务的自由流通，提升产业的国际竞争力。

2) 推动领域突破

在特定领域，如新能源领域，推动人工智能伦理的 ISO/IEC 标准制定，能够为该领域的技术创新和市场拓展提供有力支持。ISO/IEC 标准具有国际权威性，能够在全球范围内得到广泛认可和应用。在新能源领域，人工智能技术的应用可以提高能源利用效率、优化能源管理系统等。通过制定相关的 ISO/IEC 标准，可以确保这些应用符合伦理规范，促进新能源产业的可持续发展。例如，在智能电网的建设中，ISO/IEC 标准可以规定人工智能系统在电网调度、故障预测等方面的操作规范，保障电网的安全稳定运行。同时，领域标准的制定也有助于提升本国企业在国际市场的竞争力，推动相关技术和产品的出口。

3) 建立南方国家 AI 伦理联盟

南方国家在人工智能领域的发展面临着诸多挑战和机遇。建立南方国家 AI 伦理联盟，可以加强南方国家之间的合作与交流，共同应对人工智能伦理问题。联盟可以组织南方国家的专家学者、企业代表等，共同研究和制定适合南方国家的人工智能伦理规范和标准。例如，在数据隐私保护方面，南方国家 AI 伦理联盟可以制定统一的数据保护准则，促进南方国家之间的数据共享和合作。同时，联盟还可以开展人才培养和技术交流活动，提升南方国家在人工智能领域的整体水平。通过建立南方国家 AI 伦理联盟，可以增强南方国家在国际人工智能伦理领域的话语权，推动全球人工智能产业的均衡发展。

6. 结论与展望

本研究通过构建“风险识别-韧性响应-治理迭代”的动态模型，揭示了人工智能伦理风险与产业韧性间的非线性作用机制，系统论证了伦理治理对产业韧性提升的赋能路径，主要形成以下结论：首先，人工智能伦理风险通过三重路径形成产业韧性侵蚀效应。技术滥用导致的信任崩塌会削弱产业系统的社会资本积累，合规成本激增将挤占企业创新资源投入，而技术迭代受阻则造成产业升级的路径锁定。这种系统性风险冲击呈现出跨领域传导与跨周期叠加特征，在数字技术深度嵌入产业生态的背景下，伦理风险的负外部性较传统风险更具隐蔽性与扩散性。其次，伦理治理通过四维机制实现韧性重构。技术透明化重塑产业信任基础，敏捷治理提升风险免疫能力，伦理嵌入驱动创新范式转型，多元共治构建韧性网络生态。研究证实，伦理治理并非单纯的成本约束，而是通过风险转化机制形成“治理投入-韧性增益”的价值闭环。

展望未来，建议从三个维度深化研究：其一，构建动态化的伦理风险韧性影响评估体系，开发多源数据融合的监测预警平台；其二，探索区块链、元宇宙等新兴技术对“伦理-韧性”作用机制的再塑造；其三，加强全球治理协同研究，重点关注数字主权博弈下的伦理标准互认机制。这些研究方向的突破将有助于完善人工智能时代的产业韧性理论体系，为全球数字经济治理提供智力支持。

参考文献

- [1] 李梦薇, 徐峰, 晏奇, 等. 服务机器人领域人工智能伦理风险评估方法的设计与实践[J]. 中国科技论坛, 2023(10): 74-84.
- [2] 覃立, 刘青, 尹志华. 体育人工智能伦理风险检视及其纾解探径[J]. 成都体育学院学报, 2025, 51(1): 34-41.
- [3] 张艾一, 余中光. 基于国内外政策分析的医疗人工智能伦理风险治理研究[J]. 中国医学伦理学, 2024, 37(9): 1061-1067.
- [4] 林曦. 技术加速主义人工智能伦理风险的类型学分析[J]. 江海学刊, 2024(4): 164-172+256.
- [5] 宋艳, 陈琳, 李琴, 等. 人工智能伦理风险感知、信任与公众参与[J]. 科学学研究, 2022, 40(7): 1153-1162+1171.
- [6] 陈雄燊. 人工智能伦理风险及其治理——基于算法审计制度的路径[J]. 自然辩证法研究, 2023, 39(10): 138-141.
- [7] 杜严勇. 人工智能伦理风险防范研究中的若干基础性问题探析[J]. 云南社会科学, 2022(3): 12-19.
- [8] 王佑镁, 房斯萌, 柳晨晨. 风险社会视角下教育人工智能伦理风险分类框架研究[J]. 现代远程教育, 2024(3): 28-37.
- [9] 赵志耘, 徐峰, 高芳, 等. 关于人工智能伦理风险的若干认识[J]. 中国软科学, 2021(6): 1-12.
- [10] 田晓仪, 顾江. 平台经济、产业共聚与文化产业韧性[J]. 现代经济探讨, 2025(1): 73-85.
- [11] 郑瑞强, 章清, 黄璜, 等. 我国生猪产业韧性水平测度、影响因素及提升策略[J]. 黑龙江畜牧兽医, 2024(24): 9-15.
- [12] 平卫英, 张谊瑞. 政策性金融赋能农业产业韧性的机制与效应研究[J]. 农村经济, 2024(10): 100-111.
- [13] 尤亮, 任晴. 产业韧性: 内涵、影响因素与展望[J]. 江西财经大学学报, 2024(4): 33-44.
- [14] 孙良顺. 资源能力进阶、多层次治理与乡村产业韧性发展[J]. 南通大学学报(社会科学版), 2024, 40(6): 81-89.
- [15] 舒长江, 郭瑞硕. 新结构经济学视角下数字经济对高新技术产业韧性影响[J]. 江西财经大学学报, 2025(3): 37-49.
- [16] 朱建民, 张慧妹. 产业韧性、绿色创新对制造业耦合发展的影响——基于全球价值链协调视角[J]. 科技进步与对策, 2024, 41(15): 44-54.