

生成式人工智能的政府监管体系构建研究

苟圣杰, 阳卓伶

四川省社会科学院法学所, 四川 成都

收稿日期: 2024年4月22日; 录用日期: 2024年5月9日; 发布日期: 2024年6月19日

摘要

随着生成式人工智能继续搅动科技行业, 世界各国政府开始考虑对其予以监管, 以打击其滋生犯罪和助长偏见的潜力。人工智能并不是中立的、价值无涉的, 不仅应将生成式人工智能视为社会发展的创新技术, 也应将其视作社会治理的规制对象。本文分别从理论和实践层面探究生成式人工智能监管的必要性, 通过研究生成式人工智能国内外监管趋势, 进一步提出规范我国生成式人工智能发展的基本规则并构建我国生成式人工智能治理的基本框架。应当坚持伦理先行并持续关注算法与信息安全, 以数据质量充分保障为基础, 以赋能型监管为核心, 促进法律规制与数字技术融合发展, 开创多元协同共治、科技协同监管的良好局面, 最终实现监管与创新的动态平衡。

关键词

必要性分析, 监管动向, 原则要求, 制度建构

Research on the Construction of Government Regulatory System for Generative Artificial Intelligence

Shengjie Gou, Zhuoling Yang

Institute of Law, Sichuan Academy of Social Sciences, Chengdu Sichuan

Received: Apr. 22nd, 2024; accepted: May 9th, 2024; published: Jun. 19th, 2024

Abstract

As generative artificial intelligence continues to influence the technology industry, governments around the world are beginning to consider regulating it to combat its potential to breed crime and promote prejudice. Artificial intelligence is not neutral and value-free, and generative artificial intelligence should be viewed not only as an innovative technology for social development, but

also as an object of regulation for social governance. This paper explores the necessity of regulating generative artificial intelligence from the theoretical and practical levels respectively, and further proposes basic rules to regulate the development of generative artificial intelligence in China and builds a basic framework for the governance of generative artificial intelligence in China by studying the domestic and international regulatory trends of generative artificial intelligence. We should adhere to the ethics first and continue to pay attention to algorithm and information security, based on adequate protection of data quality, with enabling supervision as the core, promote the integration and development of legal regulation and digital technology, create a good situation of multiple collaborative governance and scientific and technological collaborative supervision, and ultimately achieve a dynamic balance between supervision and innovation.

Keywords

Analysis of Necessity, Regulatory Movement, Principles of Supervision, Institutional Infrastructure

Copyright © 2024 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

国家互联网信息办公室、国家发改委、教育部、科学技术部、工业和信息化部、公安部、广播电视总局七部门于7月15日联合发布《生成式人工智能服务管理暂行办法》，迈出了我国针对生成式人工智能监管的第一步。面对生成式人工智能爆炸发展所带来的数据泄露、著作权侵权、非法使用、歧视偏见等现实法律困境，如何有效引导人工智能有序健康发展，进而充分释放技术价值红利成为了各国政府面对的共同课题。解读《办法》所释放的政策信息，从法律视角剖析生成式人工智能的技术特性和治理挑战，进一步探索与之适配的监管模式和治理框架是缓解技术恐慌、应对秩序冲击的有效手段，也是我国在构建世界通行人工治理规则和体系中掌握更多话语权的必然要求。

2. 生成式人工智能监管的必要性探究

(一) 现实层面：潜在风险

1) 国家安全风险

鉴于生成式人工智能的复杂性、技术失控的可能性以及其被恶意利用的可能性，可能会对国家数据安全、金融安全等多重领域构成严重威胁。首先，由于生成式人工智能系统的深度神经网络结构，其操作和决策过程得高度复杂性可能导致技术失控，即系统产生的意外结果或行为可能对国家安全产生负面影响[1]。其次，生成式人工智能训练所需的大量数据如果未受到适当的保护从而导致数据泄露或滥用，可能导致国家重要信息或技术泄露。此外，生成式人工智能技术的双用途特性显著增加了国家安全管理复杂性，例如，恶意国家或非国家行为体可能使用生成式人工智能制造虚假信息、进行恶意网络攻击或社会工程学攻击等，以达到政治、经济或军事方面的目标，如何有效监管和控制AI技术的军事应用，防止其成为新型武器竞赛的导火索，是国际社会面临的共同挑战。

2) 公共安全风险

生成式模型的强大文本生成能力使其易于产生具有欺骗性质的文本，催生了虚假新闻和信息操纵风险，进而可能对社会政治稳定和公共利益与公共安全造成深远的影响[2]。恶意行为者可以利用生成式人

工智能的强大文本生成能力, 如基于 Transformer 模型的 GPT 系列, 以自动化方式制造虚假新闻、不实言论或误导性信息, 然后迅速扩散至社交媒体和新闻平台, 引发公众恐慌、社会不安[3]。生成式人工智能还可以自动生成多样性的信息, 适应不同的语境和受众。这意味着虚假信息可以根据受众的特定背景和信仰进行个性化定制, 以更有效地操纵受众的情感, 这种信息的高度个性化可能导致信息“过滤泡沫”效应, 加剧社交媒体上的信息孤岛化, 削弱人们对多元化观点的接触, 从而加大了信息操纵的效果。

3) 其他社会隐忧

生成式人工智能还可能会对个人隐私带来侵犯风险, 如未经许可地从互联网上收集和使用权个人或敏感信息[4], 以及修改、屏蔽和删除个人信息的救济机制不完善, 以及从人工智能数据库和机器学习模型中共同删除数据的诉求机制缺失。除此之外, 生成式人工智能可能影响人们的意见形成、决策行为和社会互动方式, 导致信息过载和社会动荡, 人类的辨识能力和认知能力会不会因为生成式人工智能的广泛应用而出现反塑危机? 而且生成式人工智能的发展可能威胁到某些创造性领域, 如文学、艺术等, 引发对人类创造力的担忧。在一些领域, 自动化生成的内容可能会取代人工创作, 可能导致某些职业的减少, 引发更深层次的失业问题。

(二) 理论层面: 多维考量

对生成式人工智能(人工智能)的监管需求源于多个层面的法律、伦理和社会考量, 这些考量共同构成了对该新兴技术的原则性、框架性制约。

1) 从法理学角度审视

首先, 自然法理论强调道德原则在法律制定中的中心地位, 认为技术发展不应偏离增进人类共同福祉的基本目标。生成式人工智能技术, 具备模拟人类创造力的潜能, 其不受约束的应用可能导致前述风险, 进而威胁到社会的公平正义和秩序安宁。其次, 从正义理论的视角来看, 约翰·罗尔斯认为社会基本结构的安排应当能够为所有人带来最大的自由和机会[5], 生成式人工智能在无适当监管的情况下可能加剧社会不平等, 例如通过加强信息控制的垄断或通过替代劳动力而导致就业市场的分化。因此, 监管生成式人工智能不仅是技术层面的必要, 也是实现社会正义的必需。此外, 从责任伦理学的角度, 技术发展应预见并防范可能的负面后果。汉斯·乔纳斯在其作品《责任原理》中强调了“预见的责任”, 主张在科技进步面前应采取谨慎原则, “人类要对他自己、对他遥远的后代乃至他所统治的全部地球生命负责”[6]。生成式人工智能的发展同样需要在道德责任框架内进行审慎考量, 确保技术进步不会损害公共利益或引发伦理危机。

2) 从伦理学角度考究

从伦理角度出发探讨对生成式人工智能的监管必要性根植于几个关键伦理学原则: 责任归属、公正性与偏见、伦理创新的推动, 这些原则共同构筑了一种框架, 旨在确保技术发展不仅符合技术逻辑, 而且遵循广泛认可的道德准则[7]。

首先, 责任归属问题在生成式人工智能的伦理讨论中占据核心地位, 挑战着传统的责任和自由意志观念。如何在开发者、使用者和人工智能实体本身之间分配责任成为一个棘手的道德问题, 这一问题的探讨要求我们重审人机交互中责任和自由意志的概念, 并为新兴技术环境下的责任分配提供道德指导[8]。其次, 人工智能系统处理数据和生成决策的方式可能不自觉地复制社会中现有的偏见, 尤其在种族、性别和社会经济地位等方面, 监管措施必须确保在人工智能系统的整个生命周期中主动避免、识别并纠正这些偏见。最后, 伦理创新的推动强调监管不仅是对现有问题的回应, 更是对未来社会愿景的积极塑造, 通过和技术开发的早期阶段融入伦理考量, 可以指导人工智能技术的发展方向, 确保其贡献于解决而非加剧社会问题, 推进一个更公平、包容和可持续的未来。

3) 从社会学角度分析

根植于对技术进步与社会发展动态相互作用的全面理解, 对生成式人工智能进行监管的迫切需求凸显了对一系列深层次的社会结构、行为模式及公共福祉方面的考虑。从社会学角度出发, 技术发展不仅是推动经济增长和创新的动力, 更被视为塑造社会结构、文化认同和公共福祉的关键因素, 技术发展应当支撑社会的公平正义原则, 确保技术创新在推动经济增长和提高生活质量的同时, 不加剧现有的不平等现象或社会分层风险[9]。因此, 对生成式人工智能的监管需求强调确保技术进步与增强社会凝聚力、促进社会公正以及保护社会弱势群体之间的平衡。具体而言: 第一, 生成式人工智能技术的监管框架需考虑到技术对社会凝聚力的潜在影响, 防止技术应用加剧社会分裂, 促使技术发展有助于促进不同社会群体之间的沟通与理解以加强社会纽带。第二, 必须通过细致的监管机制确保人工智能技术的应用支持信息的真实性以维护公共空间诚信, 防止对公共讨论的质量、媒体信任度以及民主制度本身造成破坏[10]。第三, 监管政策应旨在保护那些可能受到人工智能技术负面影响最大的群体, 如工作被自动化取代的工人阶层, 或者是那些容易被数据驱动的决策过程边缘化的社会群体, 通过在监管中融入社会保护的措施, 确保技术进步不以牺牲社会公正和福祉为代价, 促进技术创新与社会发展相协调、技术进步与社会进步和谐共生。

3. 生成式人工智能的国内外监管动向

(一) 欧盟的监管策略

在当前的全球科技法律框架中, 欧盟是推进对人工智能监管的先行者, 欧盟的监管动向和策略体现了其对技术创新与个人权利保护之间平衡的追求。欧盟前期通过《数据治理法案》(Data Governance Act)和《数字服务法案》(Digital Services Act)等立法, 试图建立一个安全可靠的数据分享环境和有力监管下的在线平台, 其中包括对生成式人工智能使用的数据来源和透明度的明确要求。2024年3月13日, 欧盟议会审议通过《人工智能法案》, 该法案是全球首个针对人工智能系统全面监管的法律提案, 尝试根据人工智能系统的风险等级对其进行分类, 并据此制定相应的合规要求[11]。对于被认为风险最高的应用, 如用于重要基础设施的人工智能或可能侵犯基本权利的应用, 将施加最为严格的限制和要求, 其中将 Chat GPT 进一步归类至有限风险人工智能系统, 需要履行透明义务, 在批准其进入市场之前, 应用研发机构外部的伦理审查委员会进行, 确保用户的自主性、公平对待和隐私。

欧盟在人工智能应用领域始终强调人文主义, 主张人工智能应当符合人类伦理, 不应当背于人类道德和价值观。与一般人工智能相比, 欧盟对于以 Chat GPT 为代表的大语言模型的伦理监管更加谨慎, 主要对于其基本伦理秩序的维护和 Chat GPT 的应用过程当中公民基本权利的保护予以特别的关注。

(二) 美国的监管策略

2022年10月, 美国白宫科技政策办公室发布《人工智能权力法案蓝图》, 提出安全、公平、隐私保护、及时通知和选择退出等5项基本原则。在 Chat GPT 发布以后, 美国更加关注大语言模型可能引发的重要的一些伦理问题, 2023年1月公布了人工智能风险管理框架, 指导相关机构组织降低在开发和部署人工智能系统过程当中的安全风险, 避免产生偏见、歧视等负面后果, 提高人工智能可信度, 保护公民的自由权利; 2023年2月, 美国总统拜登也签署了一项行政命令, 要求联邦政府设计和使用人工智能时, 要避免受到偏见和算法歧视的威胁; 2023年6月, 美国政府对国家人工智能战略征求意见, 部分内容聚焦保护人权, 促进公平公正等方面。总体来看, 在 Chat GPT 问世以后, 美国虽然鼓励重点发展生成式人工智能等大语言模型人工智能, 但也对大语言模型带来的伦理问题高度关注, 从其密集出台的相关政策文件可以看出端倪, 多份文件都提出了确保人工智能的合理性、可信赖性, 以及对公民的个人信息、隐私、公平自由等合法权益的保障, 体现了美国在应对大语言模型人工智能发展问题当中的人本主义立场, 并在联邦政策和法律的框架当中有所表达[12]。

(三) 英国的监管策略

英国针对生成式人工智能主张多方参与的监管框架, 包括政府机构、行业、学术界和社会公众。这种协作旨在促进技术创新和发展的同时, 充分考虑社会影响和伦理问题。英国注重预防原则, 强调在新技术应用早期阶段就进行监管, 以避免潜在问题的出现[13]。英国政府于 2021 年发布了《人工智能产业战略》, 该战略提出了人工智能发展的愿景、优先领域和指导原则, 其中包括对生成式人工智能的关注。英国在生成式人工智能领域的监管制度主要关注数据隐私和安全、算法公平性和透明性、可解释性和稳健性等方面, 相关法规包括《数据保护法》《人工智能和数据伦理指南》等。这些法规和指南要求生成式人工智能系统的开发和使用必须符合一定的道德和法律标准。

(四) 我国的监管策略

国家互联网信息办公室首先于 2023 年 4 月 11 日发布《生成式人工智能服务管理办法(征求意见稿)》, 紧接着七部门联合在同年 7 月 15 日发布《生成式人工智能服务管理暂行办法》, 我国进入生成式人工智能政府监管的新阶段。生成式人工智能的监管理念、政策和制度主要是以促进创新和规范发展为主导, 同时注重对生成式人工智能实行包容审慎和分类分级监管。同时, 中国也重视人工智能的安全和可控性, 强调发展与安全并重。在政策方面, 中国已经有了一系列与人工智能相关的政策, 包括“新一代人工智能发展规划”、“促进新一代人工智能产业发展三年行动计划”等。总的来说, 国内外对生成式人工智能的监管理念、政策和制度存在一定差异。国内更注重创新和规范发展, 而国外则更加注重分类分级监管和伦理法律问题。

4. 生成式人工智能发展的原则性要求

为确保生成式人工智能的良性发展, 平衡道德原则与发展需求, 以社会主义核心价值观为指引, 从现有人工智能治理制度中提炼出生成式人工智能发展治理元规则, 作为普遍性指导原则, 用于生成式人工智能的开发、应用和监管中[14]。

(一) 伦理先行

生成式人工智能的发展和应用, 应深刻植根于“以人为本”的核心理念。地域、历史、文化、经济、性别等因素的差异决定多元主体间存在不同的伦理价值判断标准, 生成式人工智能不应设立同质化的伦理价值判断标准, 特别不能形成或强加某一特定的价值立场, 因此伦理第一是前提规则。

在“以人为本”这一理念下, 伦理第一原则成为生成式人工智能治理的前提规则, 旨在确保技术创新和应用不仅遵守法律法规, 更重要的是符合伦理道德的基本要求[15]。治理实践中, 应强化人类尊严的保护, 确保生成式人工智能的每一项应用和创新都体现对人的尊重和促进人的福祉。政府建立的监管框架应提供明确的伦理指引, 技术开发者在设计和实施过程中应融入伦理考量, 引导使用者识别并防范可能的伦理风险[15]。除了通过建立完善的、合乎人类利益与道德标准的法律规范, 并将其贯穿于公法与私法实践中从而形成强制性约束外, 还应通过教育和宣传增强公众对生成式人工智能伦理问题的认识, 激发用户的伦理自觉, 帮助他们在使用过程中识别潜在的伦理陷阱和风险, 从而在保障人类尊严和伦理价值的基础上推动技术的健康发展。

(二) 算法安全

算法是生成式人工智能的决策原理, 为防止存心不良的主体利用生成式人工智能的神经网络技术向用户强制进行意识形态输出, 要保证算法的运行方式和决策逻辑应该是清晰、可理解和可解释的[16]。算法的决策过程和数据流理应被审查和验证, 其技术的合规性能进行评估, 确保生成式人工智能技术能够在合法、道德、可控的框架内运作, 且应确保算法的决策和生成过程是可追溯、可解释和可追责的, 确保技术开发者 and 提供者对其后果负有适当的责任。

(三) 信息安全

信息安全要求兼顾个人利益和国家安全。生成式人工智能的一大技术特点就是可获取性,即通过在各地区获取信息来进行深度学习。因此要对信息挖掘的深度以及数据流动的范围进行限制。在信息挖掘深度方面,在具备合法前提下,要求达到最小满足的原则,确保信息主体的权益在信息挖掘过程中不被侵犯,采取措施保护用户和相关个体的隐私权利,防止他们的个人信息被滥用、泄露。在数据流动范围方面,应加强信息区分敏感度,进行信息安全风险评估,择出一般信息和敏感信息,对敏感性程度不同的信息设定不同流动审批要求,确保数据的机密性、完整性和可用性,防止数据泄露、篡改、丢失或未经授权的访问。

5. 生成式人工智能监管的制度性建构

(一) 基本观点

1) 建立创新与监管平衡机制。政府监管应在保护公共利益的同时,保持对生成式人工智能技术创新的支持。监管政策和实践应考虑到技术的动态性,在合理范围内保持技术的灵活性并鼓励创新实践,同时防止滥用和潜在风险的发生^[17]。政府与科技界、学术界和社会各方应建立紧密的合作关系,共同制定监管框架和指导方针。

2) 构建监管能力动态提升机制。政府监管部门应加强生成式人工智能的监管能力和专业知识。通过项目研究和人才培养,使监管部门人员动态学习和更新,关于人工智能技术、算法、应用和监管领域的专业知识,以便有效监控和评估生成式人工智能系统的运行和影响。

3) 探索科技协同监管机制。通过专门的人工智能安全管理系统来管理生成式人工智能,在不同人工智能之间形成相互监督模式,从而实现有效和精准监管,弥补人类在人工智能监管方面难以克服的认知局限,使监管更为专业科学化。

(二) 制度建构

建立一个综合性的框架,涵盖技术开发、应用、监管和社会参与等多个方面,以确保生成式人工智能技术的合法可信赖及可持续发展。

1) 保障数据质量

生成式人工智能的成功,得益于算法优、算力强和上千亿参数的模型、数以千亿计的训练语料数据,数据作为训练和优化人工智能模型的核心元素,具有不可替代的作用。建议国家层面加速数据要素市场的基础性制度建立,促进数据开放共享和数据要素在全国范围内合理流动,并着力建立多模态公共数据集,打造高质量中文语料数据库,为我国的生成式人工智能技术发展提供丰富的“原料”,确保数据的最大价值实现,并且在统一的制度框架下最大化保障数据的安全和合规性。

2) 开创赋能型监管

重视监管对于生成式人工智能发展的赋能作用,克服静态的、命令式的、仅关注结果的监管模式的局限性,强调监管的动态性、灵活性和目标导向,以监管作为枢纽,坚持全面监管和分领域重点监管相结合,发挥科技监管优势以形成相互监督模式,促进合作共治、促进合规、促进科技向善。

具体而言,需要在全面监管与分领域重点监管之间找到恰当的平衡。全面监管确保了对生成式人工智能技术广泛应用的基本规则和原则设立,形成统一的监管框架和标准;而分领域重点监管则针对特定行业或应用场景中的特殊需求和风险进行细致的规制,如医疗健康、金融服务等关键领域,以确保监管措施的针对性和有效性。同时,科技监管的优势应充分发挥,通过引入创新的监督机制和技术手段,比如算法审计、数据透明度要求、伦理评估等,构建一个多层次、多主体参与的相互监督模式。在这一模式下,不仅政府监管机构,还包括行业协会、科研机构、企业以及公众等多方利益相关者共同参与

成式人工智能的伦理审视、安全评估和监管执行中来, 形成合作共治的局面。这种合作共治不仅促进了监管的透明度和公众参与度, 还有助于分享最佳实践、提升行业整体的合规水平, 进而推动科技向善。赋能型监管还应注重监管技术和方法的创新, 例如运用区块链、智能合约等技术提高监管效率和准确性, 通过大数据分析预测潜在风险, 实现对生成式 AI 应用的实时监测和动态调整。这样的监管不仅更加灵敏和有效, 也能更好地适应技术快速发展的需求, 为生成式人工智能技术的健康发展和广泛应用提供有力支持。

3) 促进法律规制与数字技术融合发展

生成式人工智能对于数字技术具有高度的依赖性, 数字技术发展直接影响着生成式人工智能发展, 立法要保持对数字技术不断发展的兼容性, 并且要把数字技术的成果融入到立法规制当中去, 利用数字技术实现立法规制内容。

这要求立法过程中不仅要关注当前的技术状态, 更要有预见未来技术演变趋势的能力, 确保法律框架的弹性和适应性。以前瞻性的法律设计思路, 通过充分利用数字技术在法律实践中的潜能来适应快速变化的技术环境, 避免法律规定在具体技术应用面前迅速过时。而将数字技术的成果融入到立法规制中, 不仅意味着在法律文本中反映技术发展的成果和经验, 更重要的是利用数字技术提高法律规制的执行效率和准确性。例如, 通过大数据分析识别和预测技术发展趋势, 辅助立法预判可能的社会影响; 利用区块链技术提升法律执行的透明度和可信度; 运用人工智能辅助判断和分析, 提高司法和执法的精确性和公正性。

4) 促进协同(多元)共治

在治理模式上, 坚持软硬法协同共治, 优化包括不同部门、不同主体, 也包括国际多元主体参与的协同共治模式。确立政府和社会多方联动的监管模式, 明确要求企业进行内部的基本制度建设和技术措施建设, 定期检查漏洞并及时修复; 鼓励团体标准先行, 并进一步制定国家标准和行业标准, 通过制定技术标准和技术规范对个人信息处理、数据的可信性、数据安全、数据产品、数据服务、算法模型等加以规制。

在宏观层面, 政府应发挥引导和协调作用, 建立起一个开放、透明、高效的政策环境, 为生成式人工智能的发展提供清晰的法律和伦理指南。政府在政策制定过程中应坚持开放性和透明性原则, 积极吸纳社会各界的意见和建议, 通过公开征求意见、举办政策论坛、建立多方利益相关者协商机制等方式增强政策的公众参与度和社会接受度[18]。

对于企业而言, 除了遵守国家法律法规和行业标准外, 还应主动承担起社会责任, 加强内部管理和自我规范。包括建立健全的伦理审查机制, 确保技术开发和应用项目在启动前进行充分的伦理风险评估; 建立持续的技术安全评估和漏洞修复机制, 保障产品和服务的安全性和可靠性; 同时, 企业应加强与利益相关者的沟通和协作, 听取消费者、社会组织和公众的意见和建议, 不断优化和调整其产品和服务, 以更好地适应社会需求和伦理期待。

在技术标准和规范的制定上, 应倡导开放的标准制定过程, 鼓励包括学术界、产业界、政府机构和公民社会在内的多方参与和贡献。这种开放和协作的标准制定机制, 不仅能够提高标准的科学性和实用性, 还能增强标准的社会接受度和执行效力。通过这样一套全面、协同、开放的治理模式, 可以确保生成式 AI 技术的发展既符合技术创新的内在要求, 又遵循社会伦理的外在指引, 最终实现技术发展与社会福祉的双赢。

基金项目

本文是四川省社会科学院第七届学术新苗课题《我国构建生成式人工智能的政府监管体系研究》的

阶段性成果。

参考文献

- [1] 张广胜. 生成式人工智能的国家安全风险及其对策[J]. 人民论坛·学术前沿, 2023(14): 76-85.
- [2] 杜娟. 生成式人工智能的扩展风险与社会规制——基于 ChatGPT 社会效应的探讨[J]. 东北师大学报(哲学社会科学版), 2024(2): 127-133.
- [3] 漆晨航. 生成式人工智能的虚假信息风险特征及其治理路径[J]. 情报理论与实践, 2024, 47(3): 112-120.
- [4] 王东方. 生成式人工智能对个人信息权益的侵害风险及其法律规制[J]. 征信, 2024, 42(2): 31-37.
- [5] 王进. 实践性正义观与立法实践——阿马蒂亚·森对约翰·罗尔斯理论的重构与启迪[J]. 理论与改革, 2015(6): 49-52.
- [6] 张旭. 技术时代的责任伦理学: 论汉斯·约纳斯[J]. 中国人民大学学报, 2003(2): 66-71.
- [7] 冯子轩. 生成式人工智能应用的伦理立场与治理之道: 以 ChatGPT 为例[J]. 华东政法大学学报, 2024, 27(1): 61-71.
- [8] 袁曾. 生成式人工智能的责任能力研究[J]. 东方法学, 2023(3): 18-33.
- [9] 容志, 任晨宇. 人工智能的社会安全风险及其治理路径[J]. 广州大学学报(社会科学版), 2023, 22(6): 93-104.
- [10] 李韬, 周瑞春. 生成式人工智能的社会伦理风险及其治理——基于行动者网络理论的探讨[J]. 中国特色社会主义研究, 2023(6): 58-66+75.
- [11] 江海洋, 魏书敏. 基于风险的通用人工智能监管——从欧盟《人工智能法案》视角展开[J]. 科技与法律(中英文), 2024(2): 88-97.
- [12] 董杨, 张蕴萌, 张雪然. 基于公共责任的人工智能监管: 美国的关键做法及启示[J]. 中国行政管理, 2024(1): 112-120.
- [13] 周辉, 金僖艾. 英国人工智能监管实践、创新与借鉴[J]. 数字法治, 2023(5): 194-206.
- [14] 商建刚. 生成式人工智能风险治理元规则研究[J]. 东方法学, 2023(3): 4-17.
- [15] 张林. 生成式人工智能应用的风险研判及伦理规制[J]. 人民论坛, 2023(24): 136-138.
- [16] 陈兵, 董思琰. 生成式人工智能的算法风险及治理基点[J]. 学习与实践, 2023(10): 22-31.
- [17] 徐磊. 生成式人工智能内容风险的一体化规制[J]. 出版发行研究, 2024(1): 59-66.
- [18] 郭小东. 生成式人工智能的风险及其包容性法律治理[J]. 北京理工大学学报(社会科学版), 2023, 25(6): 93-105+117.