

人工智影响政府监管效能的耦合原理、问题与解决路径研究

张雅琪 姜万勇 姜亚妮
齐鲁理工学院

摘要 人工智能技术的深度演进正深刻重塑政府监管的底层逻辑，其影响已超越传统工具优化范畴，形成数据驱动、算法支配、智能融合三重技术逻辑与监管体系的多维耦合。数字时代监管环境持续变革，技术应用推动监管从被动响应转向主动预判，从分散管理迈向协同治理，从经验决策升级为数据支撑决策。本文系统阐释人工智能作用于监管效能的内在原理，揭示技术与制度的多层次耦合机制，分析人工智能应用于政府监管时面临的问题，并提出针对性的对策建议。旨在为智能监管体系构建提供理论支撑与实践参考，助力破解技术应用与监管变革的适配难题，推动形成兼具效能与公平、创新与规范的新型监管范式，适应数字文明发展需求。

关键词 人工智能；政府监管效能；技术赋能；制度重构；耦合机制

DOI <https://doi.org/10.6914/tpss.070502> **文章编号** 2664-1127.2025.0705.10-19

一、引言与文献综述

人工智能技术快速发展，在政府监管领域应用范围持续扩大，从根本上改变了监管实施方式、运行效率与实际成效。人工智能可以为监管实践提供先进技术支持，助力监管在精准性、前瞻性与协同性方面实现提升，更推动监管体系在组织结构、流程机制与价值取向等维度发生适应性变革。但这一变革进程中，技术应用与制度建设融合产生算法透明度不足、数据安全风险、责任界定不清等新问题，将制约政府监管效能进一步提升。

在此背景下，系统探究人工智能影响政府监管效能的内在原理与耦合机制，剖析实践推进中存在的具体问题并提出针对性对策建议，具有重要的理论价值与现实意义。理论层面，这一研究有助于丰富数字时代政府监管理论体系，构建人工智能与监管体系深度融合的分析框架；实践层面，能够为政府部门优化智能监管策略、提升监管实效提供可操作的路径指引，推动人工智能在监管领域的规范健康应用，从而更好

©The Author(s) 2025. **Article History** Received: 25 June 2025; Accepted: 18 July 2025; Published: 31 October 2025. 社会科学理论与实践(Shèhuì kēxué lǐlùn yǔ shíjiàn) *Theory and Practice of Social Science*, ISSN 2664-1127 (Print), ISSN 2664-1720 (Online). Email: wtocom@gmail.com, <https://ssci.cc>, <https://cpcl.hk>. Published by **Creative Publishing Co., Limited** .CPCL®. This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License, which permits non-commercial use, sharing, adaptation, distribution, and reproduction in any medium or format, provided the original author(s) and source are credited, and derivative works are licensed under the same terms.

地服务于经济社会高质量发展。目前,国内外学术界已围绕人工智能对政府监管效能的影响展开多维度、多视域、多领域的研究,形成了一系列具有参考价值的研究成果,主要包括以下方面:

(一) 人工智能应用的风险类型与生成机制研究

人工智能在推动经济社会发展的同时,其风险问题成为学界关注的核心议题,相关研究从技术特性、系统运行及应用场景等层面揭示了风险的具体类型与生成逻辑。

唐要家(2022)研究得出人工智能系统的黑箱模式及其不可解释性是风险产生的主要原因,具体表现为技术不可解释、结果不可解释、运行动态性与不确定性,以及使用者人为封锁决策过程。人工智能应用风险是由系统运行不同环节的风险所共同构成的,在开发设计环节存在算法结构缺陷、代码错误及伦理缺失;数据投入环节受数据不完整、有偏性影响;运行环节则面临任务配置偏差、安全漏洞等问题;产出应用环节可能会由于结果误读、过度依赖引发风险^①。

卢安文(2025)从工具理性与价值理性失衡的视角,聚焦生成式人工智能,将风险分为虚假信息传播、隐私侵犯、数据偏倚、不可解释性^②。郭朝先(2025)强调人工智能已进入“奥本海默时刻”,其风险不仅涉及技术层面,更延伸至社会伦理与国家安全,如大规模失业、系统性经济风险、军事应用失控等风险,且风险具有跨领域传导性,因此需要进行系统性防控^③。赵志君(2023)则从中国实践出发,指出人工智能应用还面临数据质量低、核心技术受限及伦理争议等现实风险,这些风险与发展阶段、技术储备不足直接相关^④。

(二) 人工智能监管的原则、模式与国际比较研究

针对人工智能风险,学术界与实务界围绕监管原则、模式及国际差异展开了深入探讨,形成了多维度的治理框架。

在监管原则方面,唐要家(2022)提出“以人为中心”的核心准则,具体包括安全性、透明度、公平性、问责性。OECD、澳大利亚等提出的如包容性、隐私保护等原则也被纳入这一框架,强调技术服务于人的根本目标^⑤。

在监管模式与国际比较上,吴洁(2024)系统对比了中美欧的差异,指出欧盟采取“强硬监管”,以《人工智能法案》为核心,按照不可接受-高-有限-低风险的风险分级实施精准管控,禁止社会评分等高危应用,强调法律强制力;美国推行“包容审慎监管”,以行业自律为主,通过总统令、法案引导创新,兼顾安全与技术领先;中国则坚持“鼓励创新与安全监管并行”,既通过《生成式人工智能服务管理暂行办法》等规范应用,又通过国际倡议推动协同治理^⑥。

卢安文(2025)补充了新加坡关于理性自愿、自由发展的监管模式,其通过自愿性“人工智能验证”体系,给予企业更大的自主权和决策权,同时也激发了企业在人工智能研发过程中的创新活力,有力推动了企业自身的创新发展以及市场的自由竞争与繁荣^⑦。唐要家(2024)将各国进行对比,美国凭借基础研究实力与高端人才储备形成技术优势,欧盟以伦理规范与法律框架构建制度特色,中国则依托庞大市场规模与丰富应用场景走出差异化路径,但在核心技术自主可控与监管精细化实施方面仍需持续突破^⑧。

(三) 人工智能监管的实践路径与优化对策研究

基于风险特性与国际经验,学术界已形成针对性的监管实施路径与优化策略体系,强调技术、制度与多元主体的协同。

①唐要家、唐春晖:《基于风险的人工智能监管治理》,《社会科学辑刊》2022年第1期。

②卢安文:《生成式人工智能:风险、监管与治理模式探究》,《重庆邮电大学学报》(社会科学版)2025年第3期。

③郭朝先:《加强人工智能监管与治理》,《北京工业大学学报》(社会科学版)2025年第3期。

④赵志君、庄馨予:《中国人工智能高质量发展:现状、问题与方略》,《改革》2023年第9期。

⑤唐要家、唐春晖:《基于风险的人工智能监管治理》,《社会科学辑刊》2022年第1期。

⑥吴洁:《中美欧人工智能监管治理比较》,《中国经济报告》2024年第3期。

⑦卢安文:《生成式人工智能:风险、监管与治理模式探究》,《重庆邮电大学学报》(社会科学版)2025年第3期。

⑧唐要家、王钰:《创新友好人工智能监管:目标导向与实施体制》,《人文杂志》2024年第9期。

在技术与制度适配层面,段永彪(2025)以浙江省“市场监管数字化改革”为例,提出“技术-制度双向适配”机制,通过“一网协同”整合监管场景与数据,“数字执法”提升流程效率,“精准画像”实现风险动态监测,同时配套法规与标准,构建“场景驱动、资源整合、流程优化、平台赋能”的智能监管体系,为地方实践提供参考^①。

在多元协同治理方面,王若璇(2024)从协同治理视角提出,需明确上游开发者关于数据合规与模型质量的责任、下游应用者保证内容安全的义务、以及担任社会监督角色的用户终端的权责,通过“技术之治”与“制度之治”,形成政府、企业、社会联动的治理网络^②。

关于我国人工智能监管的优化路径,学术界已有针对性研究成果。赵志君(2023)提出,应构建覆盖广泛的人工智能教育体系,以此破解人才供给短缺问题;通过完善《数据安全法》《个人信息保护法》等法律制度强化安全治理框架;并深化国际合作交流以突破技术应用壁垒^③。唐要家(2024)则着重指出,需加强基础研究与开源生态建设,优化算力设施与数据平台等基础支撑条件,培育制造业、公共服务等领域的智能化应用场景,同时建立兼具“创新友好”特质的监管体制机制^④。

(四)文献述评

综上所述,现有研究围绕人工智能风险、监管原则、国际模式及实践对策构建起较完整分析框架,相关成果体现在三个维度。风险识别从技术层面拓展至全生命周期与多领域,深入揭示风险复杂性与系统性特征;监管模式比较兼顾国际差异与本土特色,为中国监管策略制定提供有益参考;实践路径探索强调“技术-制度”协同与多元治理理念,凸显动态适配在监管实践中价值。

但现有研究存在明显不足。一是对生成式人工智能等新兴领域对政府监管的影响研究深度不足,大模型伦理风险具体防控机制需进一步深化;国际比较多聚焦宏观模式,规则互认、跨境监管协作等微观机制探讨薄弱;技术治理与制度设计结合缺乏具体可操作方案。未来研究可结合技术迭代与治理实践,强化跨学科整合,提升监管策略针对性。

本文的主要贡献体现在三个方面。系统阐释了人工智能影响政府监管效能的多重原理,将技术赋能、制度重构与风险制衡纳入统一分析框架,揭示了三者相互支撑、动态平衡的内在逻辑;通过构建“核心驱动-耦合机制-效能输出”的三层分析模型,清晰呈现了人工智能与监管体系从基础机制互动到生态系统适配的深度耦合过程,为理解智能时代监管变革提供了系统性研究视角;结合实践中的具体问题,提出了技术优化、制度完善、实践推进与国际合作相结合的对策体系,既回应了算法黑箱、数据安全等现实挑战,又为构建兼具效能与公平、创新与规范的智能监管新范式提供了可操作的路径参考,丰富了数字时代政府监管理论的实践内涵。

二、人工智能对政府监管效能影响的原理与耦合机制

人工智能技术的深度演进正深刻重构政府监管的底层逻辑,这种变革已超越传统工具优化的范畴,形成了数据驱动、算法支配、智能融合三重技术逻辑与监管体系的多维耦合。其效能提升机制建立在技术赋能、制度重构与风险制衡的动态平衡之上,共同塑造着适应数字文明的新型监管范式。以下将系统阐释人工智能作用于监管效能的内在原理,揭示技术与制度的多层次耦合机制,为智能监管体系的构建提供理论支撑。

(一)技术赋能机制:数据-算法-算力的协同进化

^①段永彪、董新宇:《“技术-制度”双向适配:数智化转型重构政府监管机制的逻辑——基于浙江省“市场监管数字化改革”的案例研究》,《电子政务》2025年。

^②王若璇:《协同治理视角下生成式人工智能的监管研究》,《国际公关》2024年第21期。

^③赵志君、庄馨予:《中国人工智能高质量发展:现状、问题与方略》,《改革》2023年第9期。

^④唐要家、王钰:《创新友好人工智能监管:目标导向与实施体制》,《人文杂志》2024年第9期。

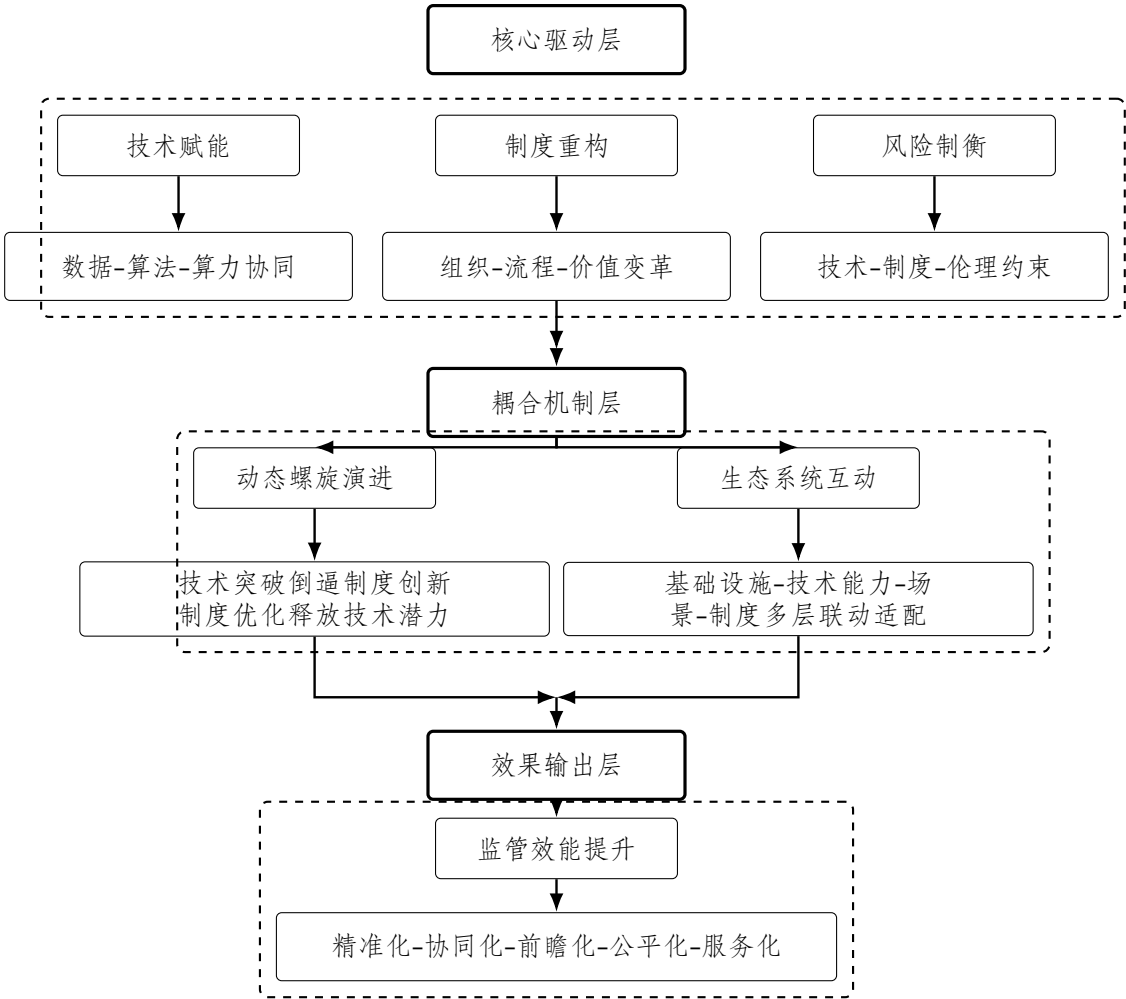


图 1: 人工智能对政府监管效能影响的耦合原理与机制图

在现代监管体系中，技术赋能的核心在于数据、算法与算力的深度协同，三者构成相互强化的“技术三角”，共同推动监管范式从传统模式向智能化转型。

数据作为基础性生产要素，其价值已超越单纯的信息载体范畴，成为驱动监管变革的核心动力。物联网设备的全域部署实现了物理世界状态的实时采集，区块链技术确保了数据流转过程中的可信性与完整性，云存储技术则突破物理空间限制，支撑起海量数据的规模化存储与高效访问，三者的有机结合，构建起了覆盖监管全场景的感知网络。数据融合产生的协同效应，推动监管决策实现了三重跃迁，一是从抽样检查转向全量分析，从而提升监管的全面性；二是从静态档案管理转向动态画像更新，增强对监管对象的实时感知；三是从被动响应转向主动预警，强化风险防控的前瞻性。当数据维度突破临界值时，其量变引发质变，使监管者获得透视复杂系统运行规律的“数据显微镜”，从而助其精准识别潜在风险。

算法作为智能监管体系的核心中枢，推动监管决策从传统的规则执行模式向认知层面跃升。机器学习技术通过深度神经网络解析海量数据中的非线性关联，为复杂监管场景提供精准分析支撑；计算机视觉技术的应用则推动非接触式执法模式的普及，显著提升监管执行效率。生成式技术的发展进一步拓展了算法应用边界，自然语言处理构建动态法律知识体系，智能体技术实现监管场景模拟推演，多模态融合算法增强对新型监管对象的识别能力。算法技术的持续进化使监管系统逐步形成“预测-模拟-决策”的完整闭环，为风险防控与公共安全保障提供了技术支撑。

算力体系为智能监管提供坚实的基础设施支撑，其分布式架构呈现三层协同特征。第一层为云端超算

中心,其负责处理PB级历史数据,支撑复杂模型的训练与优化;第二层为边缘计算节点实现毫秒级现场响应,满足实时监管需求;终端设备承担即时数据预处理任务,以提升整体处理效率。

数据、算法、算力的协同形成动态强化的技术三角,高质量数据为算法进化提供“养料”,优化的算法提升算力使用效率,强大的算力则释放出数据的潜在价值。在智慧城市建设中,这种协同催生“城市数字孪生”系统—物联网数据构建虚拟城市模型,AI算法模拟交通流量变化,超算平台优化信号控制策略,最终形成“感知-决策-执行-反馈”的智能闭环。技术三角的持续运转,会推动监管效能呈指数级提升,使传统监管难以应对的跨域性、隐蔽性、突发性风险得到有效地管控,从而为现代治理体系提供坚实保障。

(二) 制度重构机制:监管体系的适应性变革

人工智能的深度渗透正推动监管体系发生系统性重构,这种变革并非单纯的组织结构调整,而是延伸至流程机制与价值取向的深层转型,最终形成适配数字时代的治理新形态。

在组织架构层面,监管机构呈现三维结构性变革特征。纵向维度上,中间管理层级得以精简压缩,省级监管平台可直接调度县级执法资源,实现指令传达与执行的高效贯通,减少层级传导中的信息损耗;横向维度上,“数据中台”的构建有效打破部门间的信息壁垒,市场监管、税务、海关等部门的监管数据实现实时交互共享,跨域协同能力显著增强;空间维度上,传统实体监管所向虚拟化指挥中心转型,通过数字技术重构监管场景的空间连接方式。

监管流程的范式重构同步推进。事前环节,预测性监管模式逐步替代被动响应机制,基于机器学习的风险评估模型将企业划分为四级,低风险企业免检比例达到60%,在提升监管精准度的同时减少对市场主体正常经营的干扰;事中环节,非接触式监管成为主流,远程监测设备对重点污染源的覆盖比例达80%,实时监管精度得到实质性提升;事后环节,智能归因系统实现事件根源的自动化追溯,为责任追究提供快速技术支撑。流程重构推动监管价值取向发生三重转变,从“一刀切”的粗放式管理转向精准化施策,从事项导向的静态管理转向风险导向的动态管理,从权力本位的管制思维转向服务本位的治理理念。信用监管体系作为这种转变的核心载体,通过跨领域信用画像的构建,使“守信者无事不扰,失信者寸步难行”的监管目标成为现实,市场环境得到持续优化。

人工智能进一步拓展了监管体系的能力边界。在精度维度,微观监管可达到分子级检测与毫秒级响应水平;从覆盖范围来看,平台经济算法监管等新领域持续拓展监管范围;关于干预深度方面,监管干预从单纯的行为约束延伸至对系统运行的动态调节。更具革新意义的是“监管即服务”新范式的逐步形成。智能咨询系统整合多领域合规要求为企业提供精准指引;政策仿真平台通过动态模型推演帮助企业预判监管政策调整带来的影响;区块链存证技术则进行流程优化与成本控制降低企业合规证明的获取与维护成本。

生成式技术的开源特性重塑了监管技术的供给模式。开源模型使地方政府能够以更低成本构建专属监管系统,有效降低技术应用门槛;分布式协作机制吸引全球开发者参与算法优化,形成多元主体共同推动技术创新的格局。这种开放创新实践正催生“公共部门主导-企业技术支撑-社会力量参与”的三元治理结构,可有效弥补政府技术能力短板,推动智能监管技术向普惠化方向发展。

(三) 风险制衡机制:技术伦理的约束框架

应对算法异化需要构建多层次制衡体系。技术层面通过可解释性技术实现算法决策追溯,增强决策过程的透明度;制度层面建立算法备案审查制度,要求关键监管算法提交测试数据与评估报告,强化对算法应用的规范管理;伦理层面则构建算法影响评估框架,对教育、医疗等敏感领域的算法应用开展偏见检测,保障算法应用的公平性与公正性。

数据风险防控需聚焦采集、使用、共享三个关键环节。采集环节严格遵循最小必要原则,通过技术方案优化强化个人隐私保护;使用环节建立分级授权机制,按数据敏感度实施差异化管理,确保数据安全可控;共享环节依托隐私计算技术,通过跨部门联合建模实现数据“可用不可见”的安全协同。人工智能应

用带来的责任边界模糊问题需要通过制度创新加以厘清。对自动化决策实施“人类监督员”机制，在金融监管等领域明确要求保留人工复核环节，保障决策的合理性与准确性；建立算法责任保险制度，将系统错误导致的损害赔偿纳入承保范围，降低技术应用风险；推行“算法责任追溯码”制度，使每项监管决策附带可验证的数字凭证，为责任追溯提供技术支撑。

司法领域首现算法责任案例，法院明确监管机构需对 AI 系统错误决策承担主体责任，为责任认定提供实践参照。化解技术排斥需多元主体参与，依托算法听证制度保障利益相关方表达权；推行基层监管人员数字素养提升计划，通过年度常态化培训强化技术应用能力；搭建公众监督平台并开通问题反馈通道，强化社会对算法应用监督。

同时，针对数字弱势群体保留传统服务渠道，确保技术进步红利全民共享，促进社会公平正义实现。

（四）深度耦合机制：三位一体的演进系统

人工智能与监管体系的耦合呈现动态螺旋式演进特征，其核心表现为技术突破与制度创新的双向互动，技术应用对现有规则边界的突破形成持续性倒逼，推动制度框架适应性调整；而制度创新通过构建规范环境，为技术潜力释放提供保障。这种互动在智慧医疗监管领域尤为显著，AI 诊断技术发展催生新型审批制度，制度优化又加速技术临床转化。当前处于第三轮上升阶段，大模型技术推动监管知识管理系统重构，通过功能升级强化技术与制度的协同效应，促进二者深度融合。

“善智”理念指导下的技术发展呈现出了新特征。公平性算法在社保监管中自动校正区域差异，保障了社保政策的公平实施；可解释模型在司法监管中生成裁判说理依据，提高了司法的透明度和公信力；人类监督机制确保关键决策价值校准，避免了技术的盲目应用。某市开发的“良知”算法框架，将 24 项伦理准则转化为技术参数，使 AI 系统自动规避歧视性决策，实现了技术与伦理的有机结合。这种融合推动监管价值从效率优先转向为效率与公平动态平衡，更好地满足了社会发展的需求。

智能监管生态系统由四层结构构成。基础设施层以 5G 与算力网络为支撑，为智能监管提供基础的网络传输与计算能力；技术能力层依托 AI 平台与区块链等核心技术，构建智能监管的技术支撑体系；应用场景层聚焦智慧市监、环境大脑等具体领域，实现技术与监管需求的精准对接；制度规范层通过算法治理规则的建立，为整个生态系统运行提供制度保障。

各层次需要持续互动，技术突破扩展应用场景，应用实践检验制度效能，制度创新引导技术发展。如“区块链+监管”生态是一个典型案例：分布式账本技术实现全程可追溯，智能合约自动执行监管规则，通证机制激励企业合规，最终形成技术赋能、制度规范、市场激励的协同体系，提高了监管的整体效能。

三、人工智能在政府监管效能提升中面临的现实挑战

（一）技术层面的固有局限

算法的“黑箱效应”与不可解释性构成首要技术瓶颈。人工智能算法，尤其是深度学习模型，其内部参数交互与特征提取过程具有高度复杂性，加之自主学习能力带来的动态演化特性，使得决策逻辑与运行机制常超出人类直观认知范畴，形成难以穿透的“算法黑箱”。例如，在金融风险监管中，智能系统通过多维度数据关联分析识别的异常交易模式，监管人员往往无法完整追溯其风险识别的具体特征权重、关联规则与推理链条，这种不可解释性不仅削弱了监管过程的透明度，降低了公众对监管决策的信任度，更给事后责任界定与问题追溯设置了技术障碍，难以明确错误决策的根源所在。

数据质量瑕疵与安全风险构成另一重挑战。作为智能监管的“燃料”，数据的质量直接决定算法模型的可靠性。实践中，数据采集的片面性、标注过程的人为偏差、样本分布的代表性不足等问题普遍存在，可能导致模型输出带有系统性偏见，影响决策的公正性。同时，数据全生命周期管理中仍面临多重安全风险：在采集环节，过度索权与非必要信息收集可能侵犯个人隐私；在存储环节，中心化数据仓库易成为黑

客攻击的目标；在共享环节，跨部门数据流转中的权限管理疏漏可能导致数据泄露或滥用。例如，在跨区域市场监管数据共享中，部分敏感信息因加密技术不足或访问控制不严，存在被非法截取与滥用的风险，对数据安全与个人权益构成威胁。

技术单一依赖可能引发系统性风险。监管部门对人工智能技术的过度依赖，可能使监管体系陷入“技术路径锁定”，削弱人工判断与应急处置能力。一旦系统遭遇技术故障、极端场景下的模型失效，或遭受针对性网络攻击，都有可能导致监管数据被篡改、决策失准的后果，更甚者会引发监管流程的中断。

（二）制度层面的适配滞后

法律法规体系对人工智能监管的适应性尚且不足。现行法律框架多基于传统监管模式设计，尚未充分回应智能监管的特殊需求，在算法备案审查的具体标准、数据权属的清晰界定、开发者与使用者的责任划分等关键环节尚未形成系统性规范。特别是针对生成式人工智能等新兴形态，其内容生成的责任归属、训练数据的知识产权保护、输出内容的合规性判定等问题，仍存在立法空白与模糊地带，导致监管实践中“无法可依”或“法条适用冲突”的现象时有发生，难以有效应对技术快速迭代带来的监管挑战。跨部门与跨区域监管协同机制梗阻尚未彻底破除。智能监管高效运转依赖监管资源有效整合与协同联动，实践推进中部门利益壁垒、技术标准不一、数据接口兼容不足等问题，造成监管信息共享受阻，业务协同效率不高。不同地区智能监管系统建设缺乏统筹规划，技术架构与数据标准差异显著，进一步加大跨区域监管协同实施难度。

人工智能监管领域伦理规范体系仍需完善。技术应用伦理边界缺乏明确界定，未能为算法设计与运行提供清晰价值指引。信用评分、公共服务资源分配等场景中，算法可能因训练数据隐含历史偏见，对特定社会群体造成不公平影响，现有监管体系尚未建立常态化伦理审查机制与偏见纠偏工具。某地区社会救助智能评估系统曾遇类似情况，算法设计未充分考量区域发展差异，导致欠发达地区申请者评分普遍偏低，形成隐性歧视，此类伦理争议解决缺乏明确规范依据，一定程度上会影响监管工作社会认同度。

（三）实践层面的执行困境

复合型人才供给不足制约着技术应用效能。智能监管的落地需要既精通人工智能技术，又熟悉监管业务流程与法律规范的复合型人才，而当前这类人才的供给与实际需求存在显著缺口。这种现象在基层监管部门尤为突出，一方面，部分人员对智能系统的操作熟练度不足，难以充分发挥系统的高级功能；另一方面，对算法潜在风险的研判能力欠缺，无法及时识别模型输出中的不合理结果，导致智能工具的应用效果未达预期。

应用场景拓展与技术成熟度仍存在局限。人工智能技术在监管领域的应用仍集中于标准化、流程化场景，而在公共卫生应急、复杂社会治理等非结构化场景中的渗透深度仍旧不足。同时，部分应用场景的技术成熟度与实际需求存在差距，例如，基于计算机视觉的非现场执法系统在复杂天气条件下识别准确率明显下降；用于预测经济风险的宏观模型因变量选取不全面，对突发事件的响应灵敏度不足，难以满足复杂监管任务的要求。

企业合规成本攀升也带来了新的治理矛盾。智能监管对企业的数据报送格式、系统对接标准、实时监测能力等提出了更高要求，而中小企业受限于技术储备与资金规模，往往难以负担相应的系统改造与维护成本。如一个行业的智能监管平台要求企业接入实时生产数据监测系统，而一套标准化的终端设备与数据传输模块成本超过中小微企业的年度技术投入预算，部分企业因无法达标而面临监管处罚，客观上抬高了经营门槛，甚至可能抑制市场主体的创新活力，与“放管服”改革中“激发市场活力”的目标形成张力。

四、优化人工智能赋能政府监管效能的路径探索

为应对人工智能在政府监管应用中面临的技术局限、制度滞后与实践困境，需从技术优化、制度完善与实践创新三个维度构建系统性解决方案，推动智能监管从“技术赋能”向“效能跃升”转化，实现创新活力与监管秩序的动态平衡。

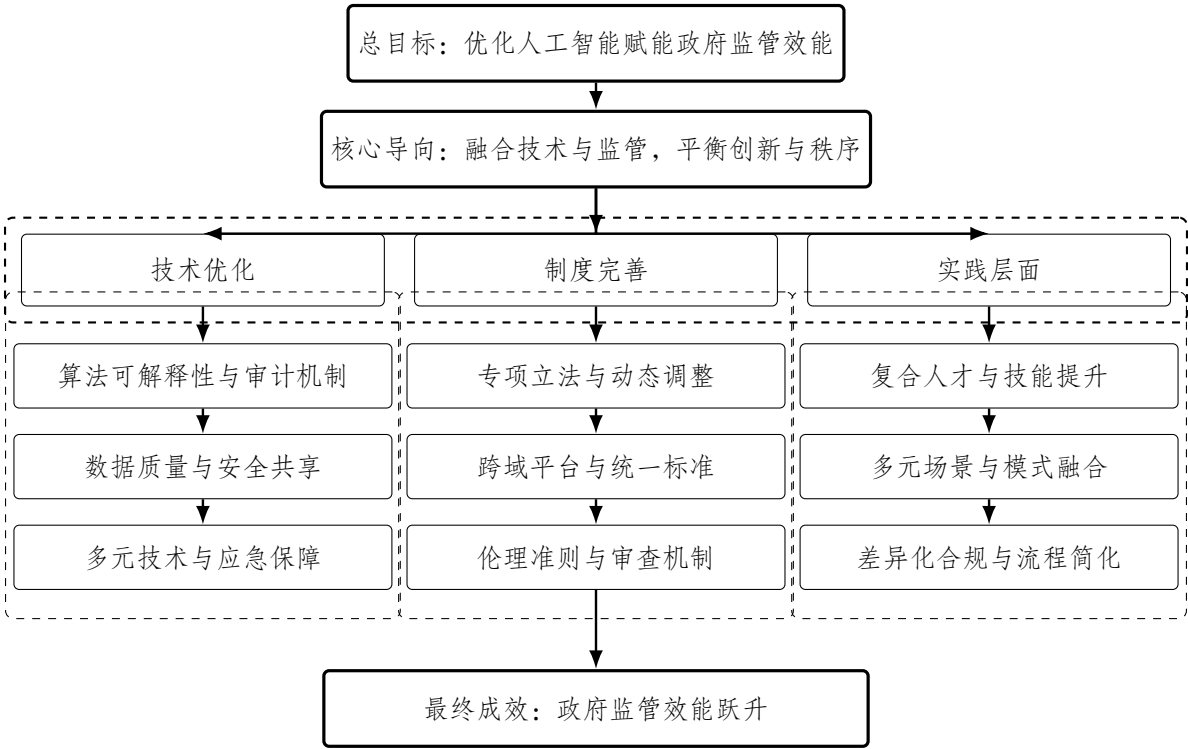


图 2: 人工智能赋能政府监管效能的路径探索图

（一）技术层面的优化路径

技术层面的突破应聚焦算法透明化、数据高质量与风险可控性，为智能监管提供可靠的技术底座。强化算法的可解释性与透明化建设，需要加大对可解释性人工智能技术的研发投入，突破深度学习模型的逻辑可视化技术，推动算法决策过程从“黑箱”向“白箱”转化。针对金融风险研判、公共资源分配等敏感领域，应建立算法决策追溯标准，要求开发主体提供技术说明文档，明确特征变量权重与关联规则。同时引入第三方独立审计机制，对算法公平性、准确性开展常态化评估，形成全生命周期技术监督闭环。

构建高质量数据治理体系，需建立覆盖数据采集、清洗、存储、共享全流程的标准化管理规范，明确不同场景下的数据质量指标。通过自动化校验与人工复核相结合的方式，减少数据缺失、标注偏差等问题。在数据安全保障方面，采用端到端加密、权限分级、隐私计算等技术，构建“采集-存储-共享”全链条安全屏障。依托区块链实现监管数据分布式存证与溯源，能够在保障数据主权的前提下提升跨域协同效率，为综合监管提供高质量数据支撑。

完善技术风险防控机制，应推动监管技术体系多元化布局，避免对单一技术的过度依赖，构建“人机协同”的双重校验模式，确保人工在关键环节的最终把控权。建立智能监管系统安全应急响应体系，定期开展压力测试与漏洞排查，制定算法失效、系统遭袭等场景的应急预案。通过冗余设计、异地灾备等方式保障业务连续性，可有效降低技术风险对监管效能的冲击。

（二）制度层面的完善方向

制度层面的创新需着眼于法律适配、协同机制与伦理规范，为智能监管提供稳定的规则框架。健全适

配智能监管的法律规范体系,需要加快推进专项立法,在现行法律基础上补充算法备案、数据权属、责任划分等条款,明确智能监管的法律边界。针对生成式人工智能等新兴形态,应规范内容生成责任、训练数据知识产权保护等内容,填补新兴领域的立法空白。建立法律法规动态调整机制,定期评估技术发展对监管规则的影响,确保法律与技术迭代保持动态适配。

优化跨域协同监管机制,要打破部门与区域行政壁垒,以“数据互通、业务协同、效能提升”为目标,构建统一的智能监管协同平台,推动数据标准化对接与实时共享。完善跨部门联席会议制度,明确职责分工与协作流程,形成“风险共防、问题共治、结果互认”的协同格局。由牵头部门制定统一的技术架构、数据接口等标准,推动不同地区、领域监管系统的兼容对接,可提升跨域协同的系统性与高效性。构建人工智能监管的伦理框架,应确立公平、透明、普惠的伦理准则,作为智能监管技术应用的基本遵循。建立由技术专家、法律学者、社会代表组成的算法伦理审查委员会,对信用评分、资源分配等涉及公共利益的应用开展伦理评估,排查潜在的歧视因素。将伦理要求嵌入算法开发环节,开发伦理风险识别工具,形成“设计-应用-评估”的管控闭环,确保技术应用符合公序良俗与公平正义,增强公众对智能监管的认同感。

(三) 实践层面的推进措施

深化推进实践层面需聚焦人才培养、场景拓展与成本优化三大方向,为落地实施智能监管提供坚实保障。复合型监管人才队伍建设需构建多层次培育体系,推动高校增设人工智能与公共管理、法学等学科交叉专业,定向培养兼具技术应用能力与监管专业素 Vernacular of the 复合型人才。基层监管人员技能提升可通过设计专项培训课程实现,课程涵盖算法基础原理、系统实际操作、风险识别研判等内容,依托理论教学与实操训练结合模式,强化其运用智能监管工具的实际能力。与此同时,应建立健全人才引进与激励机制,吸引技术领域专家参与监管实践,畅通“政产学研”人才流动渠道,为智能监管实践注入专业力量。

拓展智能监管的应用场景与深度,要以实际需求为导向,在公共卫生应急、社会治理等复杂领域开展试点,通过“试点-评估-优化-推广”的路径逐步拓展应用边界。加大研发投入,鼓励面向复杂场景的技术创新,提升模型对非结构化数据的处理能力、动态风险预测精度与极端事件响应灵敏度。推动智能监管与传统模式融合创新,在保留人工监管优势的基础上,通过智能工具提升监管精准性与效率,形成“人机协同”的新型监管模式

降低企业合规成本与制度性交易成本,需实行差异化合规要求,根据企业规模、行业特性制定分级标准,对中小企业实施轻量化合规方案。搭建公益性技术服务平台,为企业数据对接、系统改造等支持,通过政府购买服务、补贴等方式降低企业合规投入。依托数据共享与智能分析减少重复检查、多头检查,在提升监管效能的同时,最大限度减少对市场主体的经营活动的干扰,激发市场活力。

基金项目 1. 教育部中华优秀传统文化专项课题(A类)重点项目(尼山世界儒学中心/中国孔子基金会课题基金项目)(23JDTCA037)“弘扬中华优秀传统文化与提升国家文化软实力、中华文化影响力研究”; 2. 山东省社科联 2025 年一般课题“人工智能对政府数字监管效能的影响及提升路径研究”; 3. 齐鲁理工学院校级科研课题“传统制造企业数字化转型关键成功因素及作用机理研究”(QIT24SN012)。

作者简介 张雅琪,女,1998年出生,山东临邑人,助教,硕士,研究方向为智能会计、财务管理。姜万勇(通讯作者),1970年出生,男,汉族,山东海阳人,博士,教授,研究方向为数字经济与数字社会治理、软实力、人工智能与企业管理。姜亚妮,女,1997年出生,汉族,山东菏泽人,助教,硕士,研究方向为数字经济、企业管理。

引用本文 张雅琪,姜万勇,姜亚妮.人工智影响政府监管效能的耦合原理、问题与解决路径研究[J].社会科学理论与实践,2025,7(5):10-19. <https://doi.org/10.6914/tpss.070502>.

Cite This Article Yaqi ZHANG,Wanyong JIANG.(2025). Research on the Coupling Principle, Problems, and Solutions of Artificial Intelligence Affecting Government Regulatory Effectiveness . *Theory and Practice of*

Research on the Coupling Principle, Problems, and Solutions of Artificial Intelligence Affecting Government Regulatory Effectiveness

Yaqi ZHANG, Wanyong JIANG, Yani JIANG

Qilu Institute of Technology

Abstract The profound evolution of artificial intelligence (AI) technology is deeply reshaping the underlying logic of government regulation. Its impact has transcended the scope of traditional tool optimization, forming a multidimensional coupling of three technical logics and regulatory systems: data-driven, algorithm-dominated, and intelligent fusion. In the digital era, the regulatory environment continues to change, and technological applications are driving regulation from passive response to proactive anticipation, from decentralized management to collaborative governance, and from empirical decision-making to data-supported decision-making. This article systematically explains the inherent principles of AI acting on regulatory effectiveness, reveals the multi-level coupling mechanism between technology and institutions, analyzes the problems faced by AI when applied to government regulation, and proposes targeted countermeasures and suggestions. It aims to provide theoretical support and practical reference for the construction of an intelligent regulatory system, help solve the adaptation problem between technological application and regulatory reform, and promote the formation of a new regulatory paradigm that combines efficiency and fairness, innovation and regulation, to meet the development needs of digital civilization.

keywords Artificial intelligence; Government regulatory effectiveness; Technology empowerment; Institutional reconstruction; Coupling mechanism

(责任编辑: 黄欣 Email: wtocom@gmail.com)