Advances in Social Sciences 社会科学前沿, 2026, 15(6), 817-824
Published Online June 2026 in Hans. https://www.hanspub.org/journal/ass
https://doi.org/10.12677/ass.2026.156537

Hans汉斯

生成式人工智能引入数字政府建设的风险与规制路径

——基于TOE理论框架

刘晓睿

西南民族大学管理学院，四川 成都

收稿日期：2026年3月30日；录用日期：2026年6月22日；发布日期：2026年6月30日

摘　要

生成式人工智能凭借数据融合、智能生成与自迭代能力，正深刻重塑数字政府的治理逻辑与运行范式，推动政府治理从人工决策加速转型为人工智能决策。然而，在生成式人工智能应用的过程中引发了技术、组织、环境多重风险，这些风险对公民权利保障和政府公信力都构成了挑战。基于TOE (技术－组织－环境)理论分析框架，本文系统剖析生成式人工智能嵌入数字政府过程中存在的风险困境并提出规制路径，有助于推动生成式人工智能与数字政府建设的良性互动，助力提升政府治理效能与治理现代化水平。

关键词

数字政府，人工智能，生成式人工智能，风险治理，TOE理论

The Risks and Regulatory Paths of Introducing Generative Artificial Intelligence in Digital Government Construction

—Based on the TOE Theory Framework

Xiaorui Liu

School of Management, Southwest Minzu University, Chengdu Sichuan

Received: March 30, 2026; accepted: June 22, 2026; published: June 30, 2026

文章引用: 刘晓睿. 生成式人工智能引入数字政府建设的风险与规制路径[J]. 社会科学前沿, 2026, 15(6): 817-824.
DOI: 10.12677/ass.2026.156537

Abstract

Generative artificial intelligence, through its capabilities of data integration, intelligent generation, and self-iteration, is profoundly reshaping the governance logic and operational paradigm of digital governments, accelerating the transformation of government governance from manual decision-making to artificial intelligence decision-making. However, during the application of generative artificial intelligence, multiple risks such as technical, organizational, and environmental risks have emerged, posing challenges to the protection of citizens' rights and the credibility of the government. Based on the TOE (Technology-Organization-Environment) theoretical analysis framework, this paper systematically analyzes the risk dilemmas existing in the process of embedding generative artificial intelligence in digital governments and proposes regulatory paths, which is conducive to promoting the positive interaction between generative artificial intelligence and digital government construction, and helping to enhance the governance efficiency and modernization level of the government.

Keywords

Digital Government, Artificial Intelligence, Generative Artificial Intelligence, Risk Governance, TOE Theory

Copyright © 2026 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

数字化转型背景下，我国数字政府改革创新步入深水区，数字政府建设已成为推进国家治理体系和治理能力现代化的核心抓手。生成式人工智能作为人工智能领域的重点突破，凭借其本身具有的数据融合、智能生成、自迭代功能等方面的独特优势，以及在数字政府建设中凭借其在公文处理、政策解读等多场景的高效应用，正逐渐成为数字政府建设的核心驱动力。这一技术浪潮正以前所未有的速度和深度嵌入社会各领域，对于追求高效、透明、以公民为中心的数字政府建设而言，既是历史性机遇，亦构成系统性挑战。《中华人民共和国国民经济和社会发展第十四个五年规划和 2035 年远景目标纲要》[1]以及2025 年《国务院关于深入实施“人工智能+”行动的意见》[2]都提到要建设数字政府，规范人工智能在政务服务、公共治理等领域的应用要求。我国现已颁布了《中华人民共和国网络安全法》[3]《中华人民共和国数据安全法》[4]《中华人民共和国个人信息保护法》[5]《生成式人工智能服务管理暂行办法》[6]《国家人工智能产业综合标准化体系建设指南(2024 版)》[7]等多部法律规范。在此背景下，系统梳理生成式人工智

[1]https://www.gov.cn/xinwen/2021-03/13/content_5592681.htm
[2]https://www.gov.cn/zhengce/content/202508/content_7037861.htm
[3]https://flk.npc.gov.cn/detail?id=021e7d7684474107b8f3febbb1c4f8b5&fileId=&type=&title=%E4%B8%AD%E5%8D%8E%E4%BA%BA%E6%B0%91%E5%85%B1%E5%92%8C%E5%9B%BD%E7%BD%91%E7%BB%9C%E5%AE%89%E5%85%A8%E6%B3%95
[4]https://flk.npc.gov.cn/detail?id=021e7d7684474107b8f3febbb1c4f8b5&title=%E4%B8%AD%E5%8D%8E%E4%BA%BA%E6%B0%91%E5%85%B1%E5%92%8C%E5%9B%BD%E7%BD%91%E7%BB%9C%E5%AE%89%E5%85%A8%E6%B3%95
[5]https://flk.npc.gov.cn/detail?id=ff8081817b6472a3017b656cc2040044&fileId=&type=&title=%E4%B8%AD%E5%8D%8E%E4%BA%BA%E6%B0%91%E5%85%B1%E5%92%8C%E5%9B%BD%E4%B8%AA%E4%BA%BA%E4%BF%A1%E6%81%AF%E4%BF%9D%E6%8A%A4%E6%B3%95
[6]https://www.gov.cn/zhengce/zhengceku/202307/content_6891752.htm
[7]https://www.gov.cn/zhengce/zhengceku/202407/content_6960720.htm

能在数字政府建设中的应用风险，构建科学完善的法律规制体系，既是防范技术风险、保障公民权利的现实需要，也是推动数字政府高质量发展、实现治理现代化的必然要求。

2. 理论基础

学者贾开、蒋余浩认为“人工智能是建立在现代算法基础上，以历史数据为支撑，而形成的具有感知、推理、学习、决策等思维活动并能够按照一定目标完成相应行为的计算系统。”[1]，学者张鑫、王明辉认为“通过赋予机器感知和模拟人类思维的能力，人工智能使机器达到乃至超越人类的智能。人工智能不同于常规计算机技术依据既定程序执行计算或控制等任务，而是具有生物智能的自学习、自组织、自适应、自行动等特征”[2]。基于此，我们认为人工智能(Artificial Intelligence)，是一门融合计算机科学、数学、逻辑学、社会学等多学科的交叉前沿学科与技术体系。其通过人工的方法与技术，模拟、延伸和扩展人类的智能行为，令其具备类人的感知、推理、决策、学习、交流、创造等多元能力，最终能够自主处理复杂问题的技术系统。

TOE 框架为 Tornatizky 和 Fleischer (1990)所提出，最初应用于企业行为分析，技术(Technology)、组织(Organization)、环境(Environment)三个维度共同影响创新采纳过程。随着公共管理研究的发展，TOE 框架被逐渐应用于公共部门创新、政策执行、基层治理、政府数据治理等领域，其中，技术维度聚焦于技术本身特性，组织维度侧重治理主体的权责配置、协同机制，环境维度则涵盖组织应用技术时所处的内部与外部环境。TOE 框架的核心优势在于能够有效整合技术、组织与环境因素，规避了单一维度分析的片面性，构建起一个逻辑严密、层次清晰的系统性分析框架。

3. 生成式人工智能引入数字政府建设的多维度风险

3.1. 技术(T)维度

存在生成偏差风险。生成式人工智能的输出高度依赖高频率、长时间的训练数据，若政务领域训练数据存在偏差或不完整，易产生人工智能幻觉问题，即生成虚假、错误的政务信息或决策建议。例如，在应急模拟场景中，模型可能因参数设置偏差，造成场景模拟失误，影响决策科学性。同时，生成式人工智能的自迭代能力可能会导致模型参数变化，进而导致模型行为偏离预设目标，而政务系统的实时性要求又加剧了失控风险的传导速度。

存在算法黑箱风险。生成式人工智能的算法具有高度复杂性，其决策过程难以被人类理解与追溯，形成算法黑箱。在行政审批、行政裁量权等权力应用场景中，算法黑箱可能导致政府行为的合法性与合理性难以验证，例如，算法提速时，黑箱属性致使知情权难以客观行使，“算法黑箱”可能会因“商业秘密”获得法律保护，知情权不仅需要对抗无法探知符码化“算法黑箱”，而且需要与“商业秘密”进行权利衡平，削弱了传统行政赋予的知情权[3]。算法从来就不是客观中立的，至少不是像人们所想象的那样“公正无私”，算法在改善治理、提升效率和创新服务的同时，也带来了算法歧视、技术霸凌和社会信任侵蚀等问题[4]。即数字行政过程中，出于对某些秘密的保护与对效率的追求，会使得程序上的公正公平受到损害。

存在数据泄露风险。出于国家机密、商业秘密、公民个人隐私权的考虑，公共数据开放程度是有限制的，而生成式人工智能的数据处理需求必然涉及大规模数据。若数据安全防护技术不足，易引发数据泄露事件，如 AI 训练过程中未对敏感数据进行有效脱敏，可能导致个人隐私信息被非法获取，数字政府与互联网接轨，数据泄露风险的出现会侵害到公民个人权利。同时，数据共享机制不健全可能导致数据滥用，若 AI 技术突破数据使用权限，过度采集与分析公民信息，则会侵犯公民基本权利。

技术规制工具不足。算法透明与可解释性要求难以落地。生成式人工智能的复杂性导致其可解释性难度极大，难以满足公众的知情权。同时，缺乏针对政务大模型的技术标准与测评体系，对模型“幻觉”、

算法偏差等问题的技术防控手段不足，难以从源头防范风险。

3.2. 组织(O)维度

监管机制不健全。基于全生命周期理论，当前监管多集中于技术上线前的安全评估与备案，缺乏对模型运行过程、迭代更新、废止注销的全周期监管机制。对于算法的动态调整、数据的实时处理等环节，监管手段相对滞后，难以及时发现并处置风险。同时，监管主体不明确，缺乏明确统一的协调机制，容易出现数据监管重叠或数据监管空白，造成行政效率过低。

权力与责任归属模糊。法律对权力与责任划分的规定尚存在模糊地带。生成式人工智能模糊了传统行政责任链条。数字政府中“人”与“数字技术”的争夺，导致了权力的归属问题。在《社会契约论》中，政府行使的权力来源于人民主动让渡的权利[5]。这说明权力源自于民，也将为了保护人民的权利。而现代数字技术的介入，使得主体关系涉及到政府、人工智能模型开发者、人工智能技术供应商等多方问题。此外，算法对治理的介入或许打破了传统“边发现边修复”二元治理，陷入技术崇拜的唯一“正确”治理[3]。同时，傅建平提出“算法卸责”，即各主体相互推诿责任，导致受害者难以获得有效救济。应在技术开发和应用前，考虑其可能带来的社会影响和伦理问题，通过法律制度嵌入伦理要求，形成一套明晰的责任体系，以实现权利救济[6]。对生成式人工智能应用中的责任划分的法律规范尚在探索阶段，便捷权利救济渠道建设不完善、行政复议、行政诉讼等传统救济方式如何适用算法时代的维权需求仍待思考。

组织协同汇聚困难。数字政府的传统组织形态以科层制为核心，数据资源分散掌握在各职能部门，缺乏强有力的、常态化的跨部门数据协调机制。公共数据开放与共享、授权与运营等往往来源于上级压力，而非基于实际场景需求，易导致数据上传不及时、共享响应缓慢，而生成式人工智能应用与数字政府建设需跨部门扁平化的资源调配机制，组织结构矛盾易导致的“数据孤岛”使生成式人工智能难以获取全面的训练数据；同时，权力分配模式导致技术决策权集中，基层部门缺乏对模型的调整权限，难以适应本地化服务需求。

人才支撑保障薄弱。生成式人工智能的应用需要既掌握政务流程又具备人工智能技术能力的复合型人才，而当前政府部门人才结构存在明显短板。各部门在招录考试时大多录取的为所属类别的专业人才，多数部门依赖外部技术供应商提供服务，技术应用自主性不足，还因外部供应商对政务逻辑的理解有限，易造成模型与实际需求脱节。以 ChatGPT 为代表的新一代人工智能技术在数字政府治理领域彰显了巨大潜力，同时随着技术与治理活动的深度融合，数字政府治理主体将更加频繁地依赖这些技术。当决策权由人类向技术或机器转移时，数字政府治理主体的主导地位将面临挑战，尤其是技术实力较弱的政府与公众，可能陷入“智能无权”的困境[7]。

3.3. 环境(E)维度

法律体系碎片化。生成式人工智能的发展速度远超制度更新速度，现有法律法规难以覆盖生成式人工智能应用的新场景、新问题。在算法监管方面，算法透明度、可解释性、公平性等要求尚未转化为可操作的法律标准，不同地区、部门的监管规则存在碎片化问题。现有规制措施分散于不同法律法规与政策文件中，难以覆盖技术应用的全流程风险。不同政策与法律规范之间衔接不畅、标准不一，导致规制效果大打折扣，部分条款存在重复或冲突，且对生成式人工智能特有的技术特性与应用场景考虑不足，导致法律规制出现“真空地带”。此外，监管规则与法律制度的建立需要经过复杂的调研、论证和程序，这一过程难以与人工智能创新发展保持同步。时间上的迟滞使刚性的监管规则在出台时可能就已经落后于技术服务的实际发展状况，从而无法满足监管需求[8]。

社会信任不足。社会信任不足影响技术应用推广，相关法律保障机制尚不健全。公众对生成式人工

智能的认知存在偏差，一方面担心技术替代导致政务服务，认为人工智能缺乏情感能力，处理问题过于刚性；另一方面对数据安全与隐私保护存在疑虑，需防止信任赤字导致技术应用面临落地难，此外，数字鸿沟现象加剧了信任分化，农村居民、老年人等数字弱势群体因数字素养不足，难以有效使用智能政务服务，造成公共服务供给的不平衡。现有法律规范对公众数据权利的保障机制不够完善，个人信息泄露后的救济渠道不畅通，责任主体与实施路径不明确，难以有效化解社会信任危机。

工具理性与价值理性的冲突。工具理性导向下可能造成价值导向偏离，该风险源于技术逻辑对公共价值的冲击，法律对公共价值的平衡规范不够完善。生成式人工智能主要以效率为核心导向，而数字政府建设需要兼顾效率与公平、发展与安全等多元价值平衡。因此，需要维护国家意识形态安全，如国外生成式人工智能模型的训练数据蕴含其价值观，若需采用国外生成式人工智能直接应用于政务服务，其潜藏的风险可能间接影响公共价值导向的偏离。应当警惕生成式人工智能利用其敏感信息抓取能力对国家安全形成的威胁，数字政府的核心数据库中往往包含了与国家安全和国家利益密切相关的敏感信息，这些信息若未经严格筛选和脱敏处理便被直接用作人工智能的训练数据，其泄露风险将显著增加[6]。同时，国际数据治理差异导致跨境数据流动面临违规风险，数字政府建设中涉及的跨境协作、国际政务服务等场景，易因规制体系差异引发法律纠纷。我国现有法律规范对涉外政务人工智能应用的审查、跨境数据流动安全评估标准等尚缺乏明确规定，可能导致技术应用结果偏离政府核心价值，因而亟待完善生成式人工智能应用价值平衡机制。

3.4. 技术(T)-组织(O)-环境(E)维度

生成式人工智能嵌入数字政府的风险并非技术、组织、环境三个维度的独立叠加，而是相互交织、彼此强化，形成动态的风险传导链条、反馈回路与风险放大。

正向传导路径：环境缺陷－组织障碍－技术风险。法律体系的碎片化特征易引发监管主体权责界定模糊、跨部门协同效能不足等问题，进一步加剧“数据孤岛”与决策响应滞后，进而逐步诱发算法生成偏差、算法透明度不足等技术风险。

反向反馈机制：技术风险－社会信任损耗－制度压力提升－组织行为调适。算法黑箱等技术风险会逐步削弱公众对技术应用与治理主体的信任，社会信任的损耗会转化为对政府及立法层面的制度优化压力。在此压力下，监管主体可能采取阶段性集中治理或强化规制力度等举措，若此类规制要求与技术现实可行性存在偏差，则可能制约技术应用推进，影响技术赋能治理的效能，进而形成规制供给与技术应用适配失衡的治理困境。

风险放大效应：时间维度的制度滞后性与空间维度的责任分散化。时间维度上，生成式人工智能技术迭代速度与法律制度渐进完善进程存在结构性时差，制度滞后性使得风险在规制完善周期内持续累积与放大。同时，法律规范从顶层设计到落地执行，需依托配套实施细则与实践探索，制度滞后问题短期内难以彻底消解。空间维度上，多元治理主体参与格局易引发责任边界模糊问题，风险发生后的责任分散现象会弱化各主体的风险防控主动性，从而提升风险发生与扩散的概率。

4. 生成式人工智能引入数字政府建设的三维法律规制优化路径

4.1. 技术(T)维度

构建全流程技术法律规范，维护透明与安全底线。一是增强人工智能技术可解释性，破解算法黑箱、偏差、歧视等困境。生成式人工智能的算法模型在数字法治政府建设中，具备实现透明性的技术潜力。理论上可形成包含行为特征提取、法律条文匹配、裁量标准适用等关键节点信息的结构化决策日志，并且其逻辑链条可通过可视化界面呈现，使行政相对人理论上能直观知悉“为何作出此决定”[9]。将解释

人工智能技术应用于数字政府建设中，即要能够向监管部门与公众说明决策依据与生成逻辑，对输出内容进行强制性标识，明确其为人工智能生成，提醒用户注意内容的参考性。对涉及公民权利、公共利益的核心决策，需提供可视化的逻辑链条与数据支撑。二是建立算法安全防护与公平审查体系，要求政府部门在引入生成式人工智能模型前，委托第三方机构对算法训练数据、决策逻辑进行安全与公平测评，重点审查是否存在性别、地域等歧视性因素。三是构建算法实时动态监测法律机制，“尝试在自动化行政机制中有机嵌入人工干预机制，‘人工干预是贯穿于所有自动化行政阶段的一种必要机制’”[10]。四是要求政府部门与技术供应商建立算法运行日志留存制度，实现算法决策全过程可追溯，对违法违规、敏感内容进行实时预警，发现问题及时启动技术拦截与人工干预双重保障。

完善数据治理法律规则，保障数据合规与安全。一是细化数据分类分级法律制度，依据《中华人民共和国宪法》《中华人民共和国数据安全法》《中华人民共和国个人信息保护法》等法律规范，制定政务数据分类分级实施细则，明确不同级别数据的处理规则与安全保障措施。二是规范数据处理全流程法律责任，明确政府部门在数据收集、存储、使用、共享等环节的义务与责任，要求数据收集必须遵循合法、正当、必要原则，获得数据主体明确授权；数据共享必须建立安全评估机制，签订数据共享安全协议；数据使用必须符合政务服务目的，不得超出授权范围。三是建立数据质量保障法律制度，要求政府部门建立政务数据质量评估体系，对用于生成式人工智能训练的数据进行事前审核、事中监测、事后评估，确保数据的真实性、完整性、准确性，对数据污染导致的决策失误，明确相关主体的法律责任。

健全系统化法律保障，促进技术规制工具适配。一是制定政务人工智能技术标准法规，统一数据格式与模型适配要求，建立跨部门数据共享接口标准，实现不同政务系统的数据互通与功能协同。二是明确技术整合中的法律责任划分，规定因系统不兼容导致数据丢失、服务中断等问题的，由技术供应商承担主要责任；因政府部门未按标准实施导致的，由政府部门承担相应责任。中国信息通信研究院发布的《数字政府一体化建设白皮书》提出，推进数字政府一体化建设有利于推动政府履职协同化，持续增进民生福祉。建设一体化政务大数据体系，建立跨部门的信息共享和业务协同机制，可以打破信息壁垒，加快政务数据在不同层级、地域、系统、部门、业务之间的流通互认共享，提升全链条、全过程、全区域的协同化履职[11]。

4.2. 组织(O)维度

构建全周期监管实施流程，覆盖技术应用全链条。实施数据全生命周期质量管理，数据全生命周期即建立数据生命周期是对数据从产生到流转、利用、维护的全流程分阶段管理体系，核心作用是明确数据管理步骤、优化政府数据开放的流程。建立“事前评估、事中监测、事后审计”的全周期监管机制。事前，实行应用备案与安全评估制度，政府部门引入生成式人工智能系统前，需向监管部门备案，提交应用方案、风险评估报告、安全保障措施等前期材料，监管部门对高风险应用进行严格审查。事中，建立实时监测与动态预警机制，对所运用的人工智能模型实时对其数据处理流程、生成内容质量进行动态监测，对模型“幻觉”、数据异常、算法偏差等风险进行自动预警，及时进行人工干预。事后，开展定期审计与效能评估，评估后要有所反馈，并进行长期跟踪。

创新多元协同监管机制，夯实数据共享安全保障。一是明确监管主体与监管内容，避免监管权责重叠与监管内容空白，在全国统一要求下，同时鼓励依据地方差异建立区域性监管平台，因地制宜细化措施。同时可以有条件地汲取国内“互联网 + 政务服务”领域的典型经验，聚焦线上线下协同的靶向路径，协同应对技术带来的自动化行政难题[10]。二是建立数据共享安全保障法律制度，要求跨部门数据共享必须签订安全协议，明确数据共享的目的、范围、期限与安全责任，建立数据共享追溯机制，对数据共享全过程进行记录，确保数据安全可控。

健全责任追究法律体系，明确各方主体责任边界。一是明确主体责任，区分政府部门、技术供应商、模型开发者的责任边界。政府部门承担应用主导责任与监督责任，对技术应用的合法性、合理性进行审查监督；技术供应商承担技术质量与数据安全责任，对算法缺陷、数据泄露等问题承担赔偿责任；模型开发者承担算法合规责任，对算法设计中的歧视性因素、安全漏洞等承担责任。二是建立多元问责法律机制，对数据泄露、算法歧视、虚假信息等问题实行双重追责，既追究直接责任主体，也追究监管责任主体。三是完善权利救济法律途径，明确公民、法人因生成式人工智能应用遭受权益损害时，可通过行政复议、行政诉讼、民事诉讼等方式获得救济，例如可适用举证责任倒置原则，由政府部门或技术供应商证明其行为的合规性。

完善人才保障法律制度，构建复合型人才支撑体系。一是建立政企校多方协同的人才培养机制，鼓励政府部门、高校、科研机构、企业等各方合作开设政务人工智能相关课程来定向培养复合型人才。二是政府内部健全人才激励与保障法律制度，将数字素养纳入公务员考核体系，对掌握人工智能技术的公务员给予激励；建立公务员数字技能培训制度，定期开展人工智能技术应用培训，提高公务员数字技术应用能力。三是规范技术外包法律关系，生成式人工智能凭借其数据处理的规模性与逻辑推理的稳定性，为数字法治政府的决策支持系统注入了技术动能。以 DeepSeek 为代表的技术范式，通过 FP8 混合精度训练框架优化计算资源配置，显著提升了模型训练效率与精度，能够在短时间内整合法律文本、执法记录、经济指标等多源数据[9]。制定政务人工智能技术外包管理办法，明确政府部门与技术供应商的权利与责任，要求技术供应商协助政府部门掌握技术应用技能，降低技术依赖风险。

4.3. 环境(E)维度

完善层级衔接法制法律体系，铸牢专项法治根基。制定专项立法，衔接现有法律。我国现在已有《中华人民共和国网络安全法》《中华人民共和国数据安全法》《中华人民共和国个人信息保护法》《生成式人工智能服务管理暂行办法》《国家人工智能产业综合标准化体系建设指南(2024 版)》等法律规范、政策文件。一是要细化配套制度，填补规制空白。在现有法律法规基础上，针对算法治理、责任划分、权利救济等重点问题，制定配套实施细则。二是确立辅助决策原则，明确生成式人工智能不得替代人类作出重大政务决策，在公共服务与公共产品供给方面等涉及公民权益的场景中，必须保留人工审核与干预环节，形成“人工 + 人工智能”双重保障模式。三是明确数据处理的安全优先原则，严格落实“涉密不上网、上网不涉密”要求，规范政务数据的采集、脱敏、训练、存储全生命周期流程管理。

培育社会信任法律保障机制，提升公众技术接受度。一是建立技术应用公开透明法律制度，要求政府部门公开生成式人工智能应用的范围、目的、技术原理、决策逻辑等信息，保障公众的知情权、参与权、监督权；对涉及公共利益的重大技术应用，可通过线上平台信息公开征集、线下听证会等方式广泛征求公众意见。二是完善数字包容法律制度，要求政府部门在引入生成式人工智能技术时，充分考虑老年人等弱势群体的需求，保留人工服务兜底保障；建立数字技能培训法律制度，要求政府部门与社会组织合作开展数字技能培训，提升公众数字素养。三是健全权利救济与赔偿法律机制，明确生成式人工智能应用导致公民权益损害时各责任主体应承担相应赔偿责任，建立快速反馈机制，确保受害者及时获得救济。

构建国际协同法律应对机制，应对外部风险冲击挑战。一是加强核心技术自主研发法律保障，通过支持国产大模型发展，鼓励政府部门优先采用国产生成式人工智能技术与产品，降低技术依赖风险；建立技术创新激励制度，例如，对在政务人工智能领域取得重大突破的企业、科研机构可给予税收优惠等支持。二是参与国际规则制定与协同治理，全球治理倡议中也提出了奉行主权平等、遵守国际法治、倡导以人为本等核心理念，基于全球治理倡议积极参与全球人工智能治理规则制定，推动形成公平合理、包容开放的国际规则体系；在数据跨境流动、算法治理等领域加强国际合作，促进规则互认。三是建立

涉外政务人工智能应用审查法律制度，对引入的国外生成式人工智能模型进行安全评估与伦理审查，重点审查是否存在意识形态领域安全风险、数据安全风险等，确保符合我国法律规定与公共利益。

4.4. 技术－组织－环境维度

要破解技术－组织－环境三维度的风险强化效应，必须在三维规制基础上，构建专门应对风险强化效应的系统性策略，阻断风险传导链条、打破反馈回路、消除放大机制。

阻断正向传导链条，以环境补强与组织再造前置技术风险防控。一是通过专项立法明确跨部门协同义务，破解因环境界定模糊引发的责任推诿问题；二是以组织流程再造为抓手，强制建立“技术引入前组织能力评估”制度，从源头防范技术应用风险。

切断反向反馈回路，以透明度建设与信任修复机制化解规制难题。一是建立法定化的算法透明度分级披露制度，制定可操作的透明度标准，着力修复公众信任；二是创设“技术适应性规制”原则，进行组织行为调适，实现规制与技术发展的动态适配。

消除风险放大效应，压缩“制度迟滞”窗口与破解“责任分散”困局。一是由政府部门监管，允许符合条件的政府部门与技术供应商在可控范围内先行先试，同时明确试点主体与监管机构共同制定实时监测及风险预警方案；二是建立“单一法律责任”制度，破解多主体“算法卸责”与“责任分散”；三是建立风险定期评估与压力测试制度，常态化防范风险放大。

5. 结论

综上，生成式人工智能与数字政府建设的深度融合，是技术进步与治理现代化的必然趋势，既为数字政府建设、政府治理效能提升提供了全新动能，也带来了算法黑箱等多重风险。生成式人工智能引入数字政府建设既不是技术决定论的盲目推崇，也不是风险规避论的消极排斥，而是要在技术赋能与风险防控之间找到平衡点。通过构建 TOE (技术－组织－环境)三维协同的法律规制体系，能够最大限度地发挥生成式人工智能的技术优势，防范潜在风险，推动数字政府建设朝着高效、公平、透明、安全的方向发展，为国家治理体系和治理能力现代化提供有力的法治支撑。

参考文献

[1] 贾开, 蒋余浩. 人工智能治理的三个基本问题: 技术逻辑、风险挑战与公共政策选择[J]. 中国行政管理, 2017(10): 40-45.

[2] 张鑫, 王明辉. 中国人工智能发展态势及其促进策略[J]. 改革, 2019(9): 31-44.

[3] 高莹, 叶茂. 颉颃与再造: 数字政府的行政程序简化与趋向[J]. 2024(6): 35-45+124. https://link.cnki.net/doi/10.13903/j.cnki.cn51-1575/d.2024.06.006

[4] 董青岭. 人工智能时代的算法黑箱与信任重建[J]. 人民论坛·学术前沿, 2024(16): 76-82.

[5] 约翰·洛克. 政府论[M]. 北京: 商务印书馆, 2020: 14.

[6] 傅建平. 生成式人工智能引入数字政府建设的规制路径——基于风险社会的视阈[J]. 湖南大学学报(社会科学版), 2025, 39(1): 124-132.

[7] 洪富艳, 伍绘洋. 人工智能推进数字政府治理转型的内在逻辑、困境与实现路径[J]. 长春大学学报, 2025, 35(11): 73-78.

[8] 金太军, 杨芷仪. 生成式人工智能嵌入数字政府建设的多重风险与规制路径[J]. 行政论坛, 2025, 32(6): 142-148.

[9] 魏薪郦. 生成式人工智能下数字法治政府的治理[J]. 广西政法管理干部学院学报, 2025, 40(6): 13-23.

[10] 祁志伟. “技术-组织-环境”视角下的数字政府建设[J]. 理论探索, 2025(2): 85-93.

[11] 中国信息通信研究院. 数字政府一体化建设白皮书(2024 年) [EB/OL]. https://www.caict.ac.cn/sytj/202404/t20240401_474827.htm, 2024-04-01.