武漢大學學報(哲学社会科学版)
WUHAN UNIVERSITY JOURNAL (Philosophy & Social Science)
第77卷第6期 2024年11月
Vol. 77 No. 6 Nov. 2024 034~047
DOI:10.14086/j.cnki.wujss.2024.06.004

数字政府建设中应用生成式人工智能的风险及其克服

傅建平

摘 要 在数字政府建设过程中,生成式人工智能技术在强化信息整理和信息驱动能力、优化政府运作架构等方面能够发挥重要功能;同时,生成式人工智能技术在嵌入数字政府建设时也面临风险。旨在系统性应对现代化所致不安和危害的风险社会理论,可以为识别和克服这些风险提供帮助。在风险识别层面,数字政府建设中应用生成式人工智能可能导致数据安全风险、责任虚化风险以及隐私信息侵权风险。在风险治理层面,借助敏捷治理思路,应采取综合规制路径,建立数据分级分类规制框架,明确政府主体责任,有效激活隐私信息保护的私法规范,同时注意适时评估实效并及时修正。

关键词 风险社会;生成式人工智能技术;数字政府;隐私
中图分类号 C931.9 **文献标识码** A **文章编号** 1672-7320(2024)06-0034-14
基金项目 湖南省教育科学"十四五"规划2023年专项重点项目(XJK23AKS004)

党的二十大提出,我国要构建包括人工智能等技术在内的新增长引擎,完善以精细化服务、信息化支撑为目标的基层治理平台,提高社会治理能力[1](P4-28)。这表明,人工智能技术与数字政府建设具有紧密关联。生成式人工智能是人工智能的新近发展成果,它的快速发展为社会各领域的变革提供了契机,也为数字政府建设带来了机遇和挑战。根据2023年国家互联网信息办公室、中华人民共和国国家发展和改革委员会等部门颁行的《生成式人工智能服务管理暂行办法》的表述,生成式人工智能指的是具有文本、图片、音频、视频等内容生成能力的模型及相关技术。与传统人工智能相比,生成式人工智能在自主学习能力、人机交互和数据处理方面实现了显著提升,其技术赋能效果达到了新高度。由于具备广泛的适用性、发展潜力和创新生成能力等关键特征,生成式人工智能被视为强人工智能发展的初级阶段。凭借深厚的技术逻辑和卓越的算力算法,生成式人工智能能够为数字政府建设的输入输出流程提供数字化、信息化及协同化等全方位的工具支撑,从而推动形成一种协同高效、智能化的数字政府新模式。数字政府指的是在现代信息技术支撑下对组织架构、运作程序和管理服务进行重塑的政府形式。一方面,生成式人工智能的独特技术逻辑和显著算法优势,为数字政府的建设提供了坚实的技术基础和重要的资源支持;另一方面,数字政府的构建与持续发展离不开信息技术的科学驱动,而生成式人工智能凭借自身广泛的应用潜力和前景,为在数字政府建设的过程中实现深度融合与实际应用开辟了现实可行的道路,并提供了丰富的可能性。

不可否认的是,生成式人工智能的技术更新和迭代在带来巨大科技发展红利的同时,也会衍生出一系列的技术风险和制度风险。随着ChatGPT技术的升级和广泛应用,理论界对生成式人工智能介入数字政府建设的发展前景进行了激烈的讨论。从当前的学术研究成果来看,有学者认为,生成式人工智能在嵌入数字政府建设的过程中可能面临算法漏洞、数据篡改、系统瘫痪等技术风险[2](P91-102);还有学者指出,由于生成式人工智能的深度嵌入,政府的权威和主导地位可能被削弱,公共权力面临解构风险[3]

(P64-74)；也有学者提出，在参与数字政府建设的过程中，生成式人工智能的创新风险、安全风险和价值风险可能会对决策过程及决策结果产生冲击[4](P69-76)。尽管学界具有很强的风险研究意识，但由于缺乏风险社会理论的引导，这些研究呈现出了明显的碎片化特征。这种浅层化和碎片化的研究现状不仅让数字政府建设面临的风险机理处于雾里看花的不明确状态，也在很大程度上限制了政府风险规制策略的针对性和契合性。

在风险社会视域下，风险控制和风险预防成为社会治理的重要靶向。与早期的工业风险相比，现代社会中的风险类型具有高度复杂性、不可预测性和难以控制性的特征。立足风险社会理论，从个体与组织的双重视角对风险展开类型化分析，更能够从一个整体性、系统化视角对生成式人工智能参与数字政府建设的现实风险和可能风险进行阐述，从而为相关的理论研究和实践探索提供可视化方向。基于此，为进一步发挥生成式人工智能的技术优势，有必要从技术层和制度层对其底层技术逻辑和运行维护程序进行规制，以科技赋能和风险规避为逻辑链条，化解生成式人工智能介入数字政府建设的技术风险和制度风险，提升其在数字政府建设过程中的工具性价值，优化数字法治政府建设方案。

一、风险社会理论与生成式人工智能参与数字政府建设的契合性

许多学者都致力于通过风险社会理论解释现实变化，但无法准确界定风险社会的开始时间。利维塔斯主张，现代性与工业化的负面后果不再限于惩罚具体的群体而是扩展到所有个体是风险社会的起始标志[5](P300-319)。风险社会是工业社会发展的结果，是工具理性和科学进步观主导下的产物，是对工业社会的“自反”[6](P102-121)。伴随技术的迭代演进，风险的来源从最初的外部力量演变为内部决策，技术在不知不觉中影响了人们的决策内容和方式。因此，分析生成式人工智能参与数字政府建设的内部运作过程成为必要，要通过“技术—社会”的逻辑链条思考其可能招致的现实风险，以期达到预防风险和控制风险的目的。

（一）风险的社会性特征及风险社会理论的提出

伴随着工业社会和科学技术的发展，风险的存在场域不再局限于自然领域，其社会性特征逐步显现。1986年，德国社会学家乌尔里希·贝克在自己的学术研究中首次提出了“风险社会”的概念[7](P43-47)。他立足西方社会的具体背景，通过对现代社会发展的深入分析，敏锐地指出现代社会已经步入了风险社会的门槛。这一研究进一步激发了学术界对风险社会理论的浓厚兴趣，并引发了深入的研究探讨。随着全球一体化进程的加深，贝克于1999年发表了《世界风险社会》一书。在这部著作中，他将风险社会的研究视野从西方扩展到全球范围，强调风险社会正逐渐从一种地域性的社会现象转型为全球性的社会现象。在阶级社会向风险社会变迁的过程中，社会的价值体系发生了根本性转变。传统的基于阶级的不平等逐渐被基于风险的不安全取代，这标志着社会结构和价值体系的深刻变革。贝克对此给出了深刻的定义：“风险可以被视作一种系统性地应对现代化进程所带来的危险与不安全感的方式。”[8](P19)这一定义指出了风险社会与现代化进程的紧密关联。此后，西方学者吉登斯和卢曼也对风险社会保持了持续关注及探究。英国社会学家安东尼·吉登斯主张风险是现代性的必然产物，是一种由人类活动制造的风险。他认为，随着现代性的发展，人类社会的风险也在不断增加和复杂化。而尼可拉斯·卢曼则从技术层面出发，对现代社会的科技风险问题进行了深入的剖析。他指出，科技的发展虽然带来了社会的进步和繁荣，但同时也带来了新的风险和挑战。与此同时，道格拉斯、拉什等人则运用文化风险理论，对贝克、吉登斯的制度风险理论进行了批判性的分析。他们认为，除了制度层面的风险外，文化层面的风险也是现代社会不可忽视的重要方面。这一观点为风险社会理论的研究提供了新的视角和思路。尽管风险社会理论具有不同的研究视角，但总体而言，西方学术界对风险社会理论所具有的反思现代性的作用形成了高度共识。同时，随着全球风险社会理论的产生，西方学者将风险社会的风险主要归因于科技风险，“科学的进步驳斥了其最初的安全声明。正是科学的成功播种了对其风险预测的怀疑”[8]

(P78)。风险社会理论的核心观点主要包含三个层面:从风险分配上看,现代风险是全球性风险;从风险根源上看,现代风险源自工具理性导致的异化;从风险结构上讲,现代风险的典型特征是个体化[9](P12-14)。

具体来说,在风险的分配性方面,风险社会理论指出现代风险具备全球性的特性,其空间效应不受地理界限限制,时间维度上则展现出长期的持续性,可能对后代产生深远的影响。面对现代风险可能带来的灾难性后果,无论哪个群体或个人都无法置身事外,即便是那些拥有丰富财富和权势的人也难以避免其影响。在解析风险的形成机制时,风险社会理论将风险社会的出现归因于理性的内在分裂现象。具体来说,这种分裂一方面表现为工具理性的过度泛滥,随后科技理性与经济理性急剧扩张,导致了对科技进步和经济发展的盲目追求;另一方面,价值理性被忽视,导致道德观念淡化和社会发展的不平衡,这些因素共同为西方现代社会塑造了一个冷漠且极度物质化的环境,进而导致了社会关系的全面物化及异化,加剧了利益分化和冲突,使人与人、人与社会之间的关系变得紧张,社会结构面临巨大的挑战。另外,风险社会理论还强调,个体化是一个突出的结构性特征。这种个体化的基础在于福利国家推动的劳动力市场社会的普遍化进程,这一进程削弱了阶级社会和核心家庭的社会基础,而个体化则意味着个体获得了更多的自主性、权利、责任以及规范等属性。

关于应对风险的策略,风险社会理论主张塑造开放社会决策。开放社会决策意味着应当允许多元主体参与至决策结构之中,发展更具灵活性和保障能力的社会政策[10](P83-94)。传统政治中的一元决策体系无法及时识别现代社会下的新型风险,这注定造成了传统的风险应对体系在现代社会具有迟滞性及片面性,缺乏敏锐度。在开放的社会决策体系下,通过多元社会主体之间的共同协作,可以满足社会各方应对风险的需求。

在数字政府建设的语境下,开放社会决策的核心诉求在于实施“敏捷治理”模式。敏捷治理本质上是一套蕴含了高度灵活性、适应性及流动性的行动指南或方法论体系,它代表着一种自适应的、以人为本的、兼具包容性和可持续性的决策流程。与开放社会决策强调的灵活性不谋而合,敏捷治理的核心要义在于,面对快速变化的环境能够迅速感知并作出适应性调整,通过构建灵活多变的组织结构和流程机制,实现对业务需求和市场变动的即时响应。进一步地,针对开放社会决策对多元主体协同作业的内在要求,敏捷治理模式特别强调跨职能团队之间的紧密协作和无缝对接,倡导各方共同参与治理策略的规划和执行过程,确保所有相关方的利益诉求和观点都能得到充分考量和体现。因此,敏捷治理不仅具有高度的包容性,能够广泛吸纳并整合来自多元主体的声音和需求,而且还致力于追求治理实践的长期可持续性,力求在动态变化的社会环境中实现治理效能的不断优化及提升。

(二)风险社会、数字政府与生成式人工智能技术的关联性

生成式人工智能嵌入数字政府应用也带来了新的风险结构,这与风险社会理论紧密相关。复杂的算法模型对人类来说可能是不透明的,增加了人类理解其行为的难度,从而导致了不确定性。智能系统的错误或偏见可能会被放大并影响到整个社会,产生系统性风险。由于这种错误的发展速度往往快于相应监管措施的建立,可能会导致技术失控的风险。另外,由于伦理框架尚未完全建立,当算法作出错误决策时,责任归属问题变得模糊不清,因而自动化的广泛应用可能带来大规模失业,进而改变经济结构,对社会稳定产生影响[11](P72-74)。

人工智能虽然是算法驱动的,但其设计、开发、决策过程及最终的应用场景均受到人类行为的影响。数据集的选择、算法的架构以及训练过程中的人工干预都可能引入偏见。如果数据集不具代表性或含有偏见,模型的学习结果也会带有这些偏见,从而在实际应用中产生不公平或歧视性的结果。人类决策者在选择何时何地部署智能系统时可能会犯错,可能导致技术被错误地应用于不适合的场景,引发不必要的风险。开发者和使用者可能都没有充分考虑到应用的伦理后果,如隐私侵犯、自主权剥夺或是就业市场结构的变化。忽视伦理原则和价值观,可能导致技术发展偏离以人为本的方向,增加社会不稳定性

和个人权益受损的风险。通过风险社会理论的视角可以认识到，生成式人工智能的风险并非单纯的技术问题，而是深深植根于人类社会的复杂性之中。要有效管理这些风险，就需要跨学科的方法，结合技术、伦理、法律和社会学等领域的知识，以确保智能发展既能推动社会进步，又能维护人类福祉，减少不确定性和负面影响。

现代科技和互联网的特性打破了传统的地理边界，使信息、数据以及技术应用能够迅速地在全球范围内传播和交互。智能算法在数字政府中的应用不仅限于提升公共服务效率和质量，还涉及国家治理、公民隐私、数据安全等多个层面，其中的风险同样有跨国界的影响力[12](P161-173)。风险社会理论为理解全球性风险提供了深刻的洞见，其理论强调在全球化时代，风险的产生和传播不再局限于特定的地理区域，而是跨越国界，影响全球。为有效管理和减轻这些风险，需要采取全球视角，加强跨国合作，构建更加开放、透明、包容的全球治理体系。

风险社会理论提供了一种框架，用于理解和分析由技术进步带来的复杂风险。在高度技术化的现代社会中，风险往往超越了局部的、可预测的范畴，成为一种普遍的、全球性的现象，其影响广泛且深远。随着数字政府和生成式人工智能的深度融合，风险不再是孤立的事件，而是社会系统内部相互关联的产物。深入理解人工智能时代的风险社会理论，有助于我们从技术、制度、文化等多个维度识别潜在风险因素，区分技术风险、伦理风险、法律风险和社会风险等不同类别，量化风险可能导致的经济损失、社会不公、环境破坏和个人隐私侵犯等后果，我们可以基于历史数据和当前情境，预测未来可能发生的风险情景[13](P78-86)。风险社会理论提示，人们不仅要关注风险的表象，更要探究其背后的系统性原因以及不同风险之间的相互作用，这样有助于在人工智能时代制定精准的风险管理策略，促进数字政府的建设和发展，使其既高效又安全，既能推动科技进步，又能保障社会福祉。

风险社会理论强调了风险管理中政府、私营部门、非政府组织和公民社会等多元主体角色。通过促进这些多元主体之间的沟通、合作与协调，可以形成一种风险共担、利益共享的新型治理模式，增强整个社会对风险的抵御能力。要鼓励政府、企业、学术界和公民之间的对话与合作，形成风险共担、利益共享的治理模式，推动数字政府运作的公开透明，建立有效的监督和反馈机制，提高政府对公众的责任感；要采取前瞻性的措施，如开展风险评估、制定应急预案和培训专业人员，以减轻潜在风险的影响[14](P116-126)。为了有效应对复杂风险，各主体之间需要建立起常态化的协作机制。作为公共利益的守护者，政府应当主动承担领导角色，推动数字政府的公开透明运作，建立健全的信息披露机制，确保政策制定、数据使用、技术部署等关键环节的透明度，让公众能够了解和监督政府的行为，并应建立有效的监督和反馈渠道，鼓励公民参与，提升其对政府决策的信任度与满意度。私营部门，特别是科技企业和数据驱动型企业，是技术创新的主要推动力量，他们在风险管理中的角色不仅限于遵守法律法规，更应积极参与风险评估和管理实践，开发安全可靠的技术解决方案，促进产品和服务不会带来不可控的风险，同时承担社会责任，通过投资研发、参与公共讨论等方式，支持风险管理领域的创新与进步。非政府组织和公民社会作为社会监督的重要力量，对维护公共利益、倡导公民权利具有不可替代的作用。这些组织要通过提供独立的视角、开展公众教育、组织社会响应等方式，增强社会整体的风险意识和应对能力[15](P47-58)。学术机构和研究团体在风险管理中扮演着知识生产与传播的关键角色，他们可以通过科学研究提供风险评估的理论依据和技术支持，为政策制定者和实践者提供决策参考，推动风险管理方法的创新，促进跨学科交流，为解决复杂风险问题提供多元视角和解决方案。

二、数字政府建设中生成式人工智能的功能效用

政府数字化转型包含着三个层面的要素：一是运用数字技术进行治理，二是以信息数据作为治理对象，三是通过数字化实现政府架构和流程的重组[16](P5-14)。数字政府依托技术革新，将数据视为治理的核心资源，并以数字化实现政府组织架构和工作流程的根本性变革。因此，政府数字化转型是一个多

维度的系统工程,它不仅关乎技术的采用及创新,更涉及治理理念的转变与制度结构的重塑。这三个层面的要素相互依存、相互促进,共同构成了政府数字化转型的完整框架,推动政府治理向更加智慧、高效、开放的方向迈进。与之相对应,若要将新技术成功融入政府的行政系统,需要从技术、信息和结构三个视角进行观察。在将生成式人工智能运用于数字政府建设的过程中,技术连接、信息驱动、结构再造形成了三个环环相扣的阶段,其中,技术是治理手段,信息是治理对象,结构是治理效果,三方紧密结合,成为提升数字政府治理模式转型、治理结构优化与治理方式升级的变革之道。

(一)强化数字政府的信息整合能力

数字政府是以信息技术为基础,通过信息化、网络化和智能化的手段改进政府的运作模式,优化政府的组织结构和管理流程,提升政府的服务能力和治理效能[17](P94-105)。这必然涉及数据和信息的整合、共享和交换,同时也指向了数字政府建设所必须的技术联结能力。生成式人工智能可以通过自动化的方式进行数据整合,并进行智能分析,能从海量的数据中发现模式、趋势和见解。相比于传统的政府采用的信息整理技术而言,生成式人工智能更加整全化地整合了云计算、大数据、物联网、移动终端等技术,实现"云+网+端"的互联模式。这种互联互通的生成式人工智能模式为数字政府的技术联结能力建设提供了更具深度和广度的支持,能够帮助政府更好地实现各有关社会经济情况和公民需求等的数据和信息在各级政府结构之间的流转。此种技术联结能力确保了数字政府各个部门和单位能够实现信息共享和协同工作,从而使人工智能支持的实时决策更加高效和全面。通过生成式人工智能的接入,数字政府的技术联结能力得到明显提升,数字政府治理水平也随之提升。

除了数据整合和信息分析外,生成式人工智能还推动了自动化服务和个性化体验方面的技术联结能力的重塑。自 ChatGPT 面世以来,各地方政府愈发强调发挥生成式人工智能的技术联结功能,强调建设自己的专属政务大模型,通过类人化的对话交互模式,持续对各种政务应用系统进行训练和优化,对企业和公众需求进行层层递进、聚焦与逼近。以深圳市为例,通过与华为合作,在盘古政务大模型中植入了生成式人工智能算法,使得深圳市政务系统的深度学习、人机交互以及内容生成功能大幅提升,相当于为每一个提供服务的政府工作人员、为每一个被服务的企业甚至为每一位被服务的公众,都提供了一个智能助理。生成式人工智能在其中扮演的角色是"数字客服官"(digital customer service,DCS),即面向政府外部提供政务或服务知识的智能客服。此外,南宁市行政审批局研发的"邕易办"智能审批系统也是一个具有南宁特色的创新政务服务系统,该系统基于对事项申报场景的精细化梳理及业务模式改革,结合生成式人工智能技术,提供了包括智慧人机互动、个性化审批、全链化通办、自动化服务等业务模式,不仅推动了服务事项的高效化和智能化,同时也提升了群众办事体验感。

可以看出,各种形式的生成式人工智能从多个层面、系统性地提升了数字政府建设中技术赋能的效果。具体而言,首先,通过针对政务服务场景的预训练和微调策略,生成式人工智能在知识背景、语言表达及立场偏好等多个关键方面都更加贴近数字政府的治理需求;其次,生成式人工智能基于特化的政务语言大模型的构建,能够精确满足行政领域的语言表达需求,从而在处理政务数据和完成各类行政任务时表现出更高的效率;最后,在内容生成层面,生成式人工智能不仅展现出了更强的专业性,还特别注重行政主体与行政相对人之间的互动关系,增强了沟通的实效性。因此,相较于传统人工智能,生成式人工智能凭借丰富的政务知识和对政务场景的深刻理解,能够更好地适应并满足数字政府建设的实际需求。

(二)提升数字政府的信息驱动水平

在现代国家治理中,信息驱动已经成为基本的政府改革路径。数字政府本身就依赖信息技术驱动下的各类技术,通过人工智能、大数据等技术追求整体性的行政治理转型[18](P106-109)。通过运用先进的信息技术,政府可以更好地应对社会治理的复杂性和不确定性,提高政策制定和执行的效率,实现政府管理的智能化转型。生成式人工智能在深度学习、强化学习和大语言模型的支撑下,能够将信息驱动

功能发挥到极致。

在传统的政府管理框架下，信息的收集、处理及传播主要依赖政府工作人员的手动操作。这种依赖人工的方式不仅效率低下，而且容易出现错误，难以适应现代政府管理的高效和精确要求。相比之下，生成式人工智能的应用为政府管理带来了革命性的变革。生成式人工智能实现了信息决策的自动化与智能化传播，显著提升了政府的行政效率，同时大幅度降低了信息处理过程中出现错误的风险，展现出其在优化政府管理流程方面的巨大潜力。在实际政府管理场景中，生成式人工智能凭借强大的自然语言处理能力，能够准确理解公众的问题并作出精确回应，这不仅有效提升了政府服务的响应速度和服务质量，还为政府管理方式向数字化方向的深刻转型提供了有力的技术支持和创新动力。

在现代化国家中，政府作出决策需要的信息以及每天处理的信息，相较于以前而言都具有指数级的增长。除此之外，信息的碎片化程度也日益攀升。这都给数字政府的政务处理带来了巨大的挑战，增加了信息处理工作的复杂性。生成式人工智能具有强大的自然语言处理能力和大数据处理能力，能够理解和分析大量的文本信息，包括社交媒体上的评论、新闻报道以及公民在政府网站上留下的信息等。通过自然语言处理技术，生成式人工智能能够识别并提取出其中的关键信息和观点，将碎片化的信息整合成系统化的数据。例如，新加坡政府采用了一种名为Pair的公务员文书写作系统，该系统能够利用生成式人工智能的强大能力，快速整理大量资讯并撰写电邮及政府报告初稿。通过输入关键词和主题，系统能够自动生成初稿，并自动进行校对和修正，确保公文的规范性和准确性。这种应用不仅显著提升了公务员的工作效率，还保证了公文的质量，是生成式人工智能在数字政府建设过程中发挥信息驱动功能的典范。由此观之，生成式人工智能可以帮助政府机构克服传统文书写作中烦琐和耗时的问题，通过自动化智能化的方式，实现信息的快速生成和高效处理。

（三）优化数字政府的组织运作架构

生成式人工智能对数字政府组织运行结构的再造也体现出了生成式人工智能对数字政府的嵌入效能。政府组织运行结构优化能够使政府运行更加科学、规范和有序，从而减少甚至杜绝因政府的不规范运作而增加的行政成本和社会成本。此外，政府组织运行结构优化也有利于整合行政资源，充分发挥公共资源的效用[19]（P12-21）。因此，引入生成式人工智能再造数字政府的组织运行结构，有利于深入推进国家治理体系和治理能力现代化。具体而言，生成式人工智能通过深度融入数字政府的构建，从根本上实现了对组织运行结构的全面再造。这一变革主要体现在两个方面：一是组织结构的再造，它重新定义了政府部门的框架和布局，使机构设置更加科学合理，提升了管理效率；二是决策结构的再造，通过实时的数据分析及智能预测，使政府的决策过程更加动态灵活，能够迅速响应社会变化与民众需求。

一方面，生成式人工智能再造了数字政府的组织结构。通过生成式人工智能的引入，数字政府的组织结构得以重新设计和优化。生成式人工智能能够帮助数字政府减少冗余重复的工作，简化处理政务的流程，使任务自动化进行，加速决策过程，加强部门间的协作与协调，从而提高工作效率。此外，组织结构的及时变革能够使政府更加灵活地适应社会、经济和科技的快速变化。这意味着，政府能够更好地应对新挑战和问题，调整资源配置和功能设置。通过这种信息技术的应用，生成式人工智能成为破除不同政府组织之间连接壁垒的突破口，能够在同一政府组织内部、不同政府组织之间实现交流协作，形成一站式服务机制，通过深度交互系统来实现数字政府智能化。以“杭州城市大脑”为例，作为地方政府利用生成式人工智能实现协同转型的案例，旨在通过人工智能的应用嵌入减少信息不对称，促进行政流程的精简化，实现资源配置的智能化、社会协同的便捷化和公众互动的实时化[20]（P29-42）。这也验证了清华大学余凌云教授的论点：生成式人工智能促进了“人、环境、AI”三者间的深度融合与相互贯通，使其自身以及所赋能的数字政府构建成为一个跨越时间界限、打破区域壁垒、贯通不同层级的协同运作系统[21]（P95-97）。

另一方面，生成式人工智能再造了数字政府的决策结构。生成式人工智能模型还可以协助政府部

门进行政策制定和决策支持,通过对海量数据的挖掘和分析,为政府提供有针对性的建议和策略。科学行政决策的前提是大量信息和数据的收集,即运用技术对行政主体及其行为进行深度分析,由此形成“用户画像”并对基层行政治理活动进行事前预测,从而大幅提升行政决策的精确度。因此,通过“用户画像”等类似的创新数字机制的引入,有助于激发数字政府组织内部的创新能力,为新想法和新方法的实施创造条件。这将有助于推动数字政府治理水平的现代化,促进政府实现数字转型升级。这体现了人工智能对数字政府运行结构的再造,从而促进数据驱动的决策模式的建立,可以更准确地指导政策决策和资源配置,提高决策的科学性。

三、数字政府建设中应用生成式人工智能的风险类型

探讨生成式人工智能技术在数字政府建设中的技术增益和功能实现,必须正视其在应用过程中需要面对的风险与挑战。随着生成式人工智能技术在数字政府建设过程中的深度嵌入,数字政府的建设环境变得更加复杂多变,这为政府治理带来了一系列潜在的风险因素。在风险社会理论视角下,生成式人工智能嵌入数字政府建设面对的风险已超越传统工业社会风险的范畴,表现出高度复杂性、不可预测性和跨国界传播的特性。这些风险不仅涉及技术本身的局限和漏洞,更与数据安全、隐私保护、政府责任以及国家主权等重大问题紧密相连。因此,有必要对生成式人工智能在数字政府建设中的应用风险进行系统性、结构化的剖析,以揭示其潜在威胁,为后续的风险治理与防范提供坚实的理论基础。下文将从全球性风险、工具性风险以及个体性风险三个维度出发,详细探讨生成式人工智能在数字政府建设过程中可能引发的具体风险类型。

(一)全球性风险:国家数据主权安全风险

根据风险社会理论,随着科学技术的不断应用,风险已经跨越了地理性的界限,而具有全球性[22](P168-177)。在这一宏观背景下,特别是生成式人工智能已广泛应用的当下,风险的传播与影响机制发生了深刻变革。生成式人工智能的应用不仅促进了信息的快速流通和知识的深度整合,同时也为风险的跨国界传递提供了新渠道及可能。这意味着,风险不再局限于某一特定的国家或地区,而是可以通过技术网络迅速从一个国家传递至另一个国家,从而引发全球性的连锁反应[23](P133-146)。这一变化对传统国家主权的范畴构成了挑战,并促使我们重新审视和界定“主权”的概念。在生成式人工智能的推动下,数据的跨境流动与共享成为常态,数据作为新的战略资源,价值日益凸显。因此,数据主权风险应运而生,成为国家主权在数字时代的新延伸。数据主权不仅关乎国家对本国数据的控制与管理能力,还涉及在跨国数据流动中保护国家利益、维护数据安全以及应对潜在风险的问题。

现今,生成式人工智能模型的训练主要依赖西方科技公司和国家使用的英语数据集,这一现状使其他国家和地区的语言与文化面临边缘化的风险,从而在一定程度上揭示了技术霸权的存在[24](P17-31)。技术霸权(Technology Hegemony)国家将在国际关系中处于支配地位[25](P106-109),技术霸权还会引发数据权利分配不均的问题。一些科技巨头和发达国家凭借丰富的数据资源和强大的数据处理能力,能够研发出更先进的人工智能模型,如ChatGPT的不断更新和升级。相比之下,发展中国家由于技术上的短板,在全球竞争中处于不利地位,当技术公司(例如OpenAI)收集并利用全球数据来训练自己的人工智能模型(如ChatGPT)时,发展中国家的数据主权就会面临严峻挑战。

由于风险的社会性特征,其并不仅仅与个人命运相连接,社会政治组织也是风险社会的重要主体。基于此,为了化解后现代社会中的各类风险,人们被迫组成社会和政治的联盟,以组织中共同体成员的合力来弥补个体化过程的不足。为此,部分国家基于经济目标一致化的要求,可能会结成利益同盟,以技术霸权和技术垄断为进路强化自身的数字技术话语权。在数字政府建设场域,若公共权力被上述社会、政治霸权垄断联盟侵入,则政府的话语权和主导地位会被进一步削弱甚至剥夺。此时,公共权力的运行过程极易受到其他国家或地区技术资本的干预,公共服务的内容和质量可能会大打折扣。在“政府

对企业依赖性合作”的数字政府运行模式下，探讨生成式人工智能参与数字政府建设时，首先要解决的问题是谁能成为政府的合作者。尽管ChatGPT俨然成了生成式人工智能的代名词，但是其基本无缘我国的数字政府建设。一方面，技术原因限制了ChatGPT在我国数字政府建设中的应用。Open AI的研究表明，ChatGPT的性能与自然语言的语系存在正相关关系，ChatGPT的结构更擅长处理拉丁语语系的自然语言，不擅长处理汉语等更加复杂语言，相较于以英语提问，对以汉语提问的回答精准度不佳。另一方面，作为美国科技企业研发的人工智能产品，ChatGPT存在着国家安全风险不可控的问题，以及美国在高新技术领域对华“卡脖子”的问题。目前，ChatGPT未向中国内地及中国香港用户开放注册，这就意味着生成式人工智能参与我国数字政府建设，必须使用我国企业自主研发的、侧重于汉语语言处理的生成式人工智能产品。尽管我国企业也研发了众多生成式人工智能模型，但其应用的广泛度及成熟度尚无法与ChatGPT媲美，在我国数字政府建设中的应用尚需时日。

（二）工具性风险：政府主体责任失落风险

按照风险社会理论，风险的工具性意味着工具理性的泛滥以及价值理性的式微，从而造成了社会关系的异化，加剧了不同主体间的利益分化和冲突，使主体间的关系紧张化，社会结构面临冲击的危险。在政府层面，由于数字技术在政府得到广泛利用，本来作为工具中立的技术力量基于创新性能引发资本增殖形成一种资本力量，因而也存在通过工具性的异化而冲击政府结构的风险[26]（P135-148）。有数据显示，作为生成式人工智能模型的代表性产品，ChatGPT在公布后的5日内注册用户突破100万，预计Open AI公司的收入在2024年将达到10亿美元[25]（P106-109）。基于数据呈现的极高商业价值，资本看到了生成式人工智能的巨大潜力。目前，绝大多数生成式人工智能是由商业资本公司进行开发及运营维系，资本掌握着生成式人工智能的核心技术及数据资源。在生成式人工智能参与数字政府建设的过程中，资本很容易凭借对技术的垄断而对政府履职的独立性进行影响，甚至干预。

一是政府主体性遭到动摇。在数字技术更新升级的过程中，数字资本的异化加速了数字技术的“利维坦”过程，数字技术在运行过程中可能处于失控状态。在技术资本逻辑的影响下，数字政府建设被加入了逐利性色彩，并以数字技术为外在遮蔽。具体到实践过程中，数字技术可能会侵蚀数字政府的横向结构和纵向结构。在需要政企合作的公共性项目中，出于利益最大化需要，资本可能借助技术操纵公共决策和公共实践，以私权力的侵入消解公权力的主导地位。同时，在“层层上报”的公权力运行程序中，由于数字技术的嵌入，高层级政府与低层级政府之间的联结大大增强，信息传播速度和信息共享壁垒被大幅优化，这直接侵蚀着中间层级政府的作用[27]（P41-45）。换言之，资本对政府的影响无论好坏，在数字技术加持下都变得难以察觉，甚至“合理化”起来。具体到生成式人工智能的应用层面，由于其内嵌的拟人化表达及运行风格，在行政部门，相对于人利用此系统进行业务咨询和办理的过程中，都存在一种潜在的认知误区，即他们可能会误将系统视为真正的政府行政人员。这一现象实质上揭示了生成式人工智能技术在公共服务领域应用的一个深刻变革：政府部门的前台服务职能在某种程度上被智能系统替代，从而构建了一种虚拟的、由技术驱动的政务服务界面。此变革进一步意味着，智能系统背后的技术资本无形中掌握了部分原本属于政府的职能。这不仅仅是一个技术应用的转变，更涉及权力结构、公众认知以及行政责任等多维度的深刻变迁。

二是政府责任容易被架空。民主政治与民主行政本质上表现为责任政府。在社会契约论的框架下，政府是由人们让渡自然权利而形成的、代表社会共同体成员利益的公共性组织。政府若想积极、正面地回应和满足公众期待和公共需求，则需要承担一定的责任以强化政府的自我约束和自我控制。具体到生成式人工智能嵌入数字政府建设的过程中，技术遮蔽和私权力介入可能会导致政府责任关系陷入模糊的发展困境。一方面，由于生成式人工智能的高专业门槛，一般的行政工作人员无法完全解构其算法运行过程，再加上算法黑箱的存在，政府部门无法控制参数模型和算法指令的生成和操控环节，此时若出现决策失误，各方主体的责任范围和比例难以界定。另一方面，作为公权力机构，一旦在建设过

程被私权力主体侵入,不仅政府的公共属性极易被消解,其责任归属也难以清楚界定。由于政府的公权力运作以集体利益为导向,企业等私权力主体则以个体利益最大化为价值目标,因此,在数字政府建设过程中,若出现利益冲突情境下的追责情形,政府与私权力主体因立场差异,各方是否承担责任以及追责链条很难明确界定。

(三)个体性风险:个人隐私信息侵权风险

按照风险社会理论,现代性风险最终要由个体进行承担,风险的最终落脚点在于对个人基本权利的侵害。在生成式人工智能领域,这种个体性风险最终体现在对个人隐私的侵蚀。在新一代生成式人工智能快速发展的背景下,数据的重要性日益凸显,成为推动技术进步和应用创新的关键因素。生成式人工智能模型背后依赖的深度学习机制使其运行和生成新内容的能力高度依赖对海量、多样化数据的学习和处理,这种依赖关系不仅体现在模型训练阶段,也贯穿模型的不断优化及迭代过程。因此,数据的收集、整理、预处理以及质量控制成为构建和优化生成式人工智能模型不可或缺的基础环节,对提升模型的性能、增强生成内容的准确性和丰富度具有至关重要的作用。由于生成式人工智能本身的运作过程涉及大量信息的输入输出,由于算法而具有隐蔽性,加上信息来源与去向的不确定性,整个过程会造成对个人隐私的威胁。

首先,隐私信息收集不符合比例原则。信息收集是生成式人工智能开发和运行的前置环节。从其底层技术逻辑来看,无论是语言生成模块还是上下文学习模块,抑或参数模型的调适过程,都需要以强大的数据资源为基础。作为生成式人工智能的重要数据来源,WebText 数据从社交媒体平台 Reddit 所有出站链接的网络中爬取,是信息处理和内容输出的重要资源。在数字政府的建设与发展过程中,当公众(即行政相对人)有特定的公共服务需求,或是政府行政机关需根据实际情况来执行具体的行政行为时,生成式人工智能便扮演了关键角色。它要求收集并分析特定个体或群体的相关数据信息,以便更精准地理解需求、预测趋势,并辅助政府作出科学、合理的决策。为了提升输出内容的准确性,生成式人工智能需要以海量数据为依托。尽管很多个人数据与行政需求之间的联系并不紧密,生成式人工智能的算法也会收集这些数据来辅助验证,并通过知识蒸馏等方式进行深度学习以凝练结论[28](P29-43)。根据隐私政策,生成式人工智能模型有权随时收集用户的地址、浏览器类型及设置、用户与网站互动的数据,包括用户参与的内容类型和使用的功能,另外,它还会收集用户在不同时间及不同网站上的浏览活动信息。但目前,生成式人工智能模型的研发公司在收集用户个人信息(包括个人隐私以及敏感数据)时并未获得全部用户的同意,且由于生成式人工智能模型的先进性,其还可以不受时间限制跨平台地收集注册该模型用户在其他网络平台的信息。正是依靠公司收集的大量信息数据进行转型、升级以及创新,生成式人工智能模型引发的个人隐私侵犯可想而知。

其次,隐私信息的安全性保障较脆弱。当运用生成式人工智能处理信息之后,信息泄露风险也随之产生。现阶段,生成式人工智能所属公司与用户签署使用协议及隐私保护政策时,相关文件中并未包含任何明确赋予用户检查其个人信息在 AI 模型上留存状况权利的条款,也未发现有任何提及将存有用户个人信息的生成式人工智能数据库对外开放、以供社会进行监督的相关表述。另外,为了确保生成式人工智能中存储的治理对象数据的安全性,相关公司制定了一系列有利于自身的信息使用条款。令人遗憾的是,这些条款中对用户个人信息的存储保护及修复措施并未给出任何具体说明或承诺。2023 年 6 月,16 名匿名人士向美国加利福尼亚州旧金山联邦法院提起诉讼,指控 ChatGPT 在没有充分通知用户或获得用户同意的情况下,收集和泄露了包括详细的账户信息、姓名、联系方式、登录凭据、电子邮件、支付信息、交易记录、浏览器数据、社交媒体信息、聊天日志、cookie、搜索记录等在内的各类个人隐私信息。这引发了关于数字政府建设中个人信息安全问题的思考:一方面,生成式人工智能的底层技术逻辑使治理对象的个人信息极易被泄露。在行政机关实施行政行为、作出行政决策的过程中,生成式人工智能会将输入的个人信息和政务数据作为训练语料,通过反复调适,不断优化自身的表达能力。由于缺乏完善

的信息筛选和数据识别程序，以往用户的个人信息可能会暴露在下一位用户面前，导致个人隐私被泄露。另一方面，用户的使用习惯也加剧了个人信息泄露的风险。出于对生成式人工智能的信任，用户更容易袒露个人信息，自觉或不自觉地泄露了个人隐私。在此种情况下，生成式人工智能极易出现信息收集边界模糊、信息利用深度违规等风险，在诸如政府招标、舆情监控等活动中出现漏标、欺骗的违法违规后果，甚至可能被政治霸权、技术霸权国家利用，进而危害个体权益和国家利益。

四、数字政府建设中应用生成式人工智能的敏捷规制

在风险社会理论框架下，科技风险之防范是一个系统工程。风险社会理论主张通过开放社会决策来实现风险治理，以多元主体进行灵活性、针对性的风险防范，这可以对应为现代政府治理理念中的敏捷治理思维。敏捷治理致力于培养一种能够迅速且不断地察觉、适应并有效应对环境变化的治理能力，它特别注重促进多元主体之间的有效沟通和紧密协作，通过更加开放和灵活的策略满足各主体的多样化需求。具体来说，敏捷治理框架的核心要素体现在：既能保持稳定，又具备高度的灵活性；拥有卓越的动员能力和协作机制；展现出强大的自我组织和自我优化能力；具备有效分解并妥善处理复杂任务的能力；依托灵活多变的基础设施；积极倡导并实践拥抱变化的理念；确保所有相关主体之间的充分沟通；推动信息的全面开放与共享；鼓励试错行为，以促进持续学习和创新[29]（P141）。在生成式人工智能嵌入数字政府的过程中，亦应当以敏捷治理作为基本思维，从全球性、异化性以及个体性这三重风险视角进行敏捷性机制的构建。唯其如此，才能完成对科技赋能以及科技风险这一悖论的超越[30]（P96-99）。

（一）确立数据治理的精细化规制标准

敏捷治理意味着构建一套全面、精细的风险防范框架，以确保风险发生时能够迅速反应，将损失降至最低。具体到生成式人工智能参与数字政府建设可能引发的全球性风险，风险社会理论主张以法律手段对现代科技应用的风险进行规制[31]（P3-6）。《中华人民共和国数据安全法》第21条明确确立了“国家建立数据分类分级保护制度”这一基石性立法原则，该原则深刻体现了国家对数据保护领域的精细化管理理念，其核心精髓体现在，通过严谨而细致的流程，甄别并理解数据的多样性与复杂性特征，这涵盖了数据的敏感性、重要性、流通范围等多重关键属性。在此基础上，还需深入评估这些数据在流动与应用过程中可能触发的潜在风险与挑战。依据这些综合考量，科学地划分数据类别，并为其设定相应的保护等级，以确保数据的安全与合规使用，从而最大化地发挥数据的价值。

第一，纵横交错：数据类型与级别的标准化呈现。鉴于数字政府建设不可避免会涉及国家数据主权安全，因此，从数据分类和数据分级两方面出发对数据资源进行分流，或许可以避免生成式人工智能技术可能导致的国家数据主权安全风险。

一方面，从数据分类出发降低生成式人工智能引发的国家数据主权安全的风险。数据分类是根据数据的属性将数据划分为具有共同特征的不同种类，使数据按照一定的规律形成具有共同特征的集合[32]（P36-44）。国家数据主权安全之维护是一个系统工程，《数据安全法》对重要数据（第21条）、核心数据（第21条）以及属于管制物项的数据（第25条）这三类数据的内涵予以明确规定。这三类数据至关重要，生成式人工智能的不当使用易致数据主权安全风险。作为关键手段，数据分类能动态管理这些敏感数据，有效减少因处理不当引发的负面效应，保障国家数据安全。另一方面，从数据分级出发降低生成式人工智能引发的国家数据主权安全的风险。数据分级是将已经分类的数据按照不同价值识别标准划定保护级别，进而以此为根据确立恰当的数据保护策略和规则，以确保数据的完整性、保密性和可用性[33]（75-87）。根据主流观点，目前主要从数据安全、隐私保护和合规要求等角度对数据进行分级[34]（P933-940），分级标准为分类后的数据所承载的“利益分量”。数据分级后的生成式人工智能，不仅可以拥有丰富的资源，同时还可以形成一个安全可靠、操作性强的数据库，国家就是这一数据库的控制者，通过限制生成式人工智能访问该数据库的权限，避免了数据分级前生成式人工智能肆意使用重要敏感数

据、侵犯国家数据主权安全的风险,从而实现了分类后对个人敏感数据以及重要数据的特别保护。

第二,对症下药:不同数据类型和级别的具体规制。当前,政府数据分级制度的必要性及优势已获广泛认同。但在实施细节上,细化数据分类分级标准,强化国家数据主权安全,仍是一个亟须深入研讨的关键议题。

一方面,要建立面向数据权益的分类分级标准体系,旨在明确不同类别数据权属关系与保护等级,确保数据在收集、处理、存储、传输及利用等全生命周期中的合法合规性。具体而言,该体系须根据数据的敏感性、隐私性、商业价值等因素,将数据细分为个人信息、商业秘密、公共数据等多个类别,并为每一类别设定相应的保护级别。对涉及个人隐私的敏感数据,应实施最严格的访问控制和脱敏处理措施,确保数据不被非法泄露或滥用;对公开的、不具敏感性的数据,可在遵守相关法律法规的前提下,促进其在更大范围内的共享和应用。通过这样的分类分级,既能有效保护数据主体的合法权益,又能激发数据的流通活力,促进数据资源的优化配置。例如,《四川省政务数据、数据分类分级指南》重点明确了针对政务数据分类、分级的方法和流程等。该指南将数据划分为四级,按敏感等级由低至高分别为非敏感级、低敏感级、敏感级和极敏感级,以适应不同级别数据的共享开放原则和安全防护建议。另一方面,构建面向数据价值的分类分级标准体系。建立该体系是为了深入挖掘数据的潜在价值,推动数据要素的市场化配置及高效利用。这一体系要求从数据的质量、时效性、关联性等多个维度出发,对数据的价值进行量化评估和分级管理。对高价值的数据资源,如行业领先的研发数据、市场趋势分析数据等,应建立专门的数据管理平台,通过高级分析工具和技术手段实现数据的深度挖掘与精准分析,为企业的决策制定提供有力支持;对价值相对较低或时效性较短的数据,可通过数据交易平台进行快速流通,促进数据资源的循环利用和增值。通过这样的分类分级管理,不仅能够显著提升数据资源的利用效率,还能激发数据创新的活力,为数字经济的发展注入强劲动力。

(二)将政府责任具体化规范化

在敏捷治理的体系之中,生成式人工智能肩负着为政务服务注入活力的关键使命,然而,这并不意味着我们可以忽视治理过程中非技术逻辑的重要性。尽管生成式人工智能能够提供极富价值的参考意见和建议,但它绝无法取代人类智能的核心地位:因为最终的判断与决策仍需仰赖人类治理主体的经验积淀和实践智慧。生成式人工智能在敏捷治理中担当的辅助角色,恰如其分地展现了技术价值与人文价值相融合的理念。虽然生成式人工智能能够自主、高效地处理海量复杂信息,但其运作依然离不开人类提供的训练数据以及支持其自主学习的环境。在生成式人工智能参与数字政府建设责任体系的构建中,必须严格遵守人类的价值观和伦理规范,以确保其不会被滥用,从而坚定地维护人类的主体性地位不受任何侵犯。具体而言,这一责任体系可分为前置责任与后置责任两个方面。

一方面,要确立政府的算法监管责任与快速反应机制。在技术赋能的过程中,政府有责任对算法的研发、部署、运行等环节进行监管,以确保算法的合理、合法应用,防止生成式人工智能滥用带来的风险。前置责任,即算法应用监管责任。随着数字技术的飞速发展,尤其是生成式人工智能技术的广泛应用,算法应用监管责任日益凸显。在技术赋能社会、经济、文化等各个领域的进程中,算法作为核心驱动力,其公正性、透明度及合法性直接关系到技术应用的成效以及社会的和谐稳定。作为公共利益的守护者,政府承担着不可推卸的前置监管责任。具体而言,政府对算法应用的监管应贯穿算法的研发、部署、运行乃至评估的全过程。在研发阶段,政府需引导企业遵循伦理规范,确保算法在设计之初就融入了公平、正义的原则,避免出现算法偏见和歧视性设计。同时,要鼓励技术创新和开放合作,促进算法技术的健康发展。在部署及运行阶段,政府应建立严格的审查监督机制,对算法的应用场景、数据处理方式、输出结果,等等进行全面评估,确保算法运行符合法律法规的要求,不侵犯个人隐私,不损害公共利益。此外,政府还需建立快速反应机制,对发现的算法滥用行为,能够迅速介入、有效处置,防止风险扩大。在评估阶段,政府应定期组织专家团队对算法应用的效果进行独立评估,包括其对社会、经济、文化等方面

的影响，以及是否存在潜在的风险和问题，评估结果将作为调整监管政策、优化算法应用的重要依据，推动后续的算法技术持续向善、健康发展。例如，广州市中级人民法院曾审理过一起案例，涉及某信用信息查询平台的算法错误导致个人信息被错误关联。本着平台监管责任的理念，法院最终判决该平台的开发方某科技公司对算法技术的使用负责，并要求其承担相应的经济赔偿责任。总之，今后政府有责任处理好技术赋能和政府主体资格侵蚀的风险，与之相应，必须精确界定政府在数据管理、信息安全等关键领域的职责范围，构建一套高效且适应性强的监管责任体系。

另一方面，在权责一体原则下落实政府的行政责任。后置责任中的算法侵权追责，涉及当生成式人工智能在数字政府建设过程中造成侵权行为时相关行政机关应承担的行政责任。在这种情况下，行政机关作为数字政府建设的主导者和监管者，应对管理失误造成的后果负责。后置责任深刻体现了权责一致的原则，要求行政机关在享受技术赋能带来的治理效率提升的同时，也必须对可能出现的监管漏洞和失误负责。当生成式人工智能在数字政府建设的过程中，因算法设计缺陷、监管不力或管理疏忽等原因导致侵犯公民权益、损害公共利益时，相关行政机关必须承担起相应的行政责任。这种追责机制不仅是对受害者的一种公正补偿，更是对行政机关行使公权力的一种有效监督。它要求行政机关在推进数字政府建设时，必须建立健全的算法监管体系，包括但不限于算法审查、风险评估、应急响应等机制，确保算法应用的合法合规性和安全性。同时，行政机关还需加强内部管理和人员培训，提升工作人员的法律意识和技术素养，以便更好地履行监管职责。在追责过程中，应坚持公开、公正、透明的原则，依法依规进行调查取证、责任认定和处罚决定。对确属行政机关监管失误造成的侵权行为，应依法追究相关责任人的行政责任，甚至刑事责任，以儆效尤。同时，还应建立健全赔偿机制，确保受害者的合法权益得到及时有效的救济。总之，后置责任中的算法监管失误追责机制是维护数字政府建设秩序、保障公民权益的重要保障。它要求行政机关在享受技术红利的同时，必须时刻保持警醒，切实履行好监管职责，确保技术应用的合法合规性和安全性。

（三）有效激活个人隐私信息的私法规范

从敏捷治理的视角审视数字政府建设，其核心目的在于迅速响应并精准满足人民群众日益多样化的公共服务需求，同时确保这一过程中的风险防控与公众利益保护并行不悖。在此背景下，“知情—同意”原则框架体系的建立显得尤为重要。它不仅是保障公民知情权自主权的关键，也是构建生成式人工智能应用事前风险预防模式的基础。作为数字治理领域的一项基本原则，“知情—同意”原则的核心在于确保公众在充分了解算法运作机制、潜在风险及数据使用方式的前提下，能够自主决定是否接受特定的服务或技术处理。这一原则在数字政府建设中尤为重要，因为它直接关系到人民群众对公共服务的信任度和满意度。以《中华人民共和国个人信息保护法》为核心，辅以《中华人民共和国数据安全法》《中华人民共和国网络安全法》《互联网信息服务算法推荐管理规定》等法律法规，共同构成了“知情—同意”原则的规范体系，为数字政府建设提供了坚实的法律支撑。

在生成式人工智能嵌入数字政府建设的过程中，私主体无法完全消解生成式人工智能的算法风险及应用风险，需要借助行政机关的公权力来有效排除对私主体数据主权和数据安全的侵害。首先，行政机关应当制定并执行一系列更加严格的监管政策，为生成式人工智能在数字政府中的应用划定清晰的边界。这些政策不仅应当明确技术应用的合法合规性要求，还应当建立严格的监管体系与问责机制，确保科技企业能够在法律的框架内有序开展技术创新活动。其次，行政机关应当加强跨部门协作，打破信息孤岛，实现资源共享和优势互补。通过建立跨部门的工作机制和联动机制，行政机关能够更加全面地掌握生成式人工智能在数字政府建设中的应用情况，及时发现并应对潜在的风险与挑战。在技术审查方面，行政机关也应当展现高度的责任感并具备专业水平。通过建立算法审查与评估机制，行政机关要对生成式人工智能的算法设计、数据处理、结果输出等关键环节进行全面而深入的审查。这不仅可以确保算法应用的合法合规性与透明度，还可以提高政府服务的公信力与民众的满意度。同时，行政机关还

应当高度重视数据安全防护工作。通过建立完善的数据加密、访问控制、应急响应等机制,行政机关应为数字政府建设筑起一道坚不可摧的安全防线。这些措施能够有效防止数据泄露与滥用现象的发生,保护公民的隐私权益。此外,行政机关还应当积极构建公众参与和反馈机制,鼓励公众对算法应用提出宝贵的意见和建议。通过搭建多元化的沟通平台与参与渠道,行政机关让公众能够更加直接地参与到数字政府建设的进程中来,形成政府、企业、公众三方共治的良好局面。这种共同参与、共同治理的模式,不仅能够增强公众对数字政府建设的信任与支持,还可以促进技术创新与社会发展的深度融合。

在日新月异的技术浪潮之下,无论是庞大的组织机构还是渺小的个体,都无法置身其外,免受技术发展的影响。技术的迅猛发展如同一把双刃剑,既带来了前所未有的机遇,也潜藏着难以预料的挑战。科林格里奇困境这一深刻揭示技术与社会关系的理论,便是对此现象的精准描绘。面对新兴技术的涌现,若因过度担忧其可能带来的不良后果而采取过早且过度的控制措施,无疑会束缚技术的自由发展,使技术潜力无法得以充分释放,技术爆发和革新将难以实现。这种保守态度,往往会错失推动社会进步的重要契机。然而,当某项技术一旦成功渗透并融入整个经济社会,成为社会不可或缺的一部分时,想要再回头处理和解决其带来的不良效应,却会发现难度极大,需付出的代价也异常高昂。此时,技术的负面影响往往已经根深蒂固,难以轻易消除。因此,如何在技术发展的初期既保持必要的审慎态度,又不妨碍其正常发展,如何在技术广泛应用后有效应对和解决其带来的问题,成为一个亟待解决的难题。

数字政府建设是我国2022年政府工作报告提出的目标任务,离不开新兴生成式人工智能技术的运用,其中功用与风险并存。在风险社会理论的帮助下,准确识别生成式人工智能技术在数字政府建设中可能产生的风险,是有效、精准预防该技术风险的前提,同时,也能够为后续风险控制措施设计提供参考,避免因此不合理地阻碍技术效用的发挥。社会与技术交互视角下的数智治理手段涵盖“技术—社会—主体”三个维度,旨在通过多要素联动,打造技术、社会、主体三位一体的数智治理体系,以技术进步与社会治理创新的良性互动提升数字政府建设的整体效能[35](P47-56)。面对国家、政府和个体三个层面的风险,敏捷治理思路下的综合规制路径或有益处,关键在于这些规制措施如何落地。同时,在实践中其如何发挥何种作用以及如何及时修正风险,也是敏捷治理的重要组成。

参考文献

[1] 习近平.高举中国特色社会主义伟大旗帜 为全面建设社会主义现代化国家而团结奋斗——在中国共产党第二十次全国代表大会上的报告.党建,2022,(11).

[2] 白文静.人工智能参与数字政府建设的范式革新——基于行政领域人工智能通用大模型(GovGPT)的交互性敏捷治理.西北民族大学学报(哲学社会科学版),2024,(3).

[3] 朱嫣丹.类 ChatGPT 模型辅助数字政府建设的可能、风险与模式.企业经济,2024,(4).

[4] 刘玮.ChatGPT类生成式人工智能嵌入数字政府建设:可供、限制与优化——基于技术可供性视角.情报理论与实践,2023,(10).

[5] 路丝·利维塔斯.风险与乌托邦的话语//芭芭拉·亚当,乌尔里希·贝克,约斯特·范·卢恩.风险社会及其超越:社会理论的关键议题.赵延东、马缨译.北京:北京出版社,2005.

[6] 曾宪才.风险、个体化与亚政治:贝克风险社会理论视域下的社会状态与风险应对.社会政策研究,2021,(3).

[7] 郭群英,夏雪.西方风险社会理论的主要流派及其审思.湖北行政学院学报,2023,(5).

[8] 乌尔里希·贝克.风险社会.何博闻译.南京:译林出版社,2004.

[9] 张广利,黄成亮.风险社会理论本土化:理论、经验及限度.华东理工大学学报(社会科学版),2018,(2).

[10] 郑作彧,吴晓光.卢曼的风险理论及其风险.吉林大学社会科学学报,2021,(6).

[11] 杨东,黄尹旭.人工智能发展面临的风险挑战及应对策略.秘书工作,2023,(7).

[12] 刘益东.底线思维与科技审度:高风险社会治理的要义与进路.哲学分析,2024,(1).

[13] 易承志,龙翠红.风险社会、韧性治理与国家治理能力现代化.人文杂志,2022,(12)

[14] 黄尹旭.和合共生:公共数据治理的“传统——现代”互融式建构.法治社会,2024,(3).

[15] 崔中良.生成式人工智能作为叙事主体的社会风险及治理.云南社会科学,2024,(2).

[16] 孟天广.政府数字化转型的要素、机制与路径——兼论“技术赋能”与“技术赋权”的双向驱动.治理研究,2021,(1).
[17] 孙全胜.人工智能赋能数字法治政府治理的制度构建.河南社会科学,2024,(4).
[18] 李盛竹.跨国公司国际竞争背景中的技术霸权现象——理论回顾与展望.社会科学家,2011,(9).
[19] 薄贵利.论优化政府组织结构.中国行政管理,2007,(5).
[20] 王文君.论数据主权的管辖权和控制权.南京理工大学学报(社会科学版),2023,(3).
[21] 余凌云.数字政府的法治建构.中国社会科学院大学学报,2022,(1).
[22] 蔡翠红.大变局时代的技术霸权与“超级权力”悖论.人民论坛·学术前沿,2019,(14).
[23] 李立丰,王俊松.人工智能嵌入刑法体系的障碍与定位——兼论刑法教义学体系下风险社会理论的反思.法治研究,2023,(1).
[24] 郑冬芳,秦婷.数字帝国主义技术霸权的政治经济学批判.理论学刊,2022,(3).
[25] 邓经超.数字政府技术资本侵蚀的生成机理与法律规制.东北师大学报(哲学社会科学版),2023,(2).
[26] 许开轶,谢程远.数字政府的技术资本侵蚀问题论析.政治学研究,2022,(2).
[27] 郑智航.数字技术对政府权力的侵蚀及其法律规制.行政法学研究,2024,(5).
[28] 刘艳红.生成式人工智能的三大安全风险及法律规制——以ChatGPT为例.东方法学,2023,(4).
[29] 于文轩.奔跑的大象:超特大城市的敏捷治理.学海,2022,(1).
[30] 何小勇,张艳娥.风险社会视域下科技理性的悖论与超越.科技进步与对策,2009,(4).
[31] 王学忠,张宇润.技术社会风险的法律控制.科技与法律,2008,(4).
[32] 陈祥玲.政府数据分类分级保护的理论逻辑、现实困境与实践路径.征信,2023,(4).
[33] 商希雪,韩海庭.数据分类分级治理规范的体系化建构.电子政务,2022,(10).
[34] 高磊,赵章界,林野丽等.基于数据安全法的数据分类分级方法研究.信息安全研究,2021,(10).
[35] 徐明月,安小米.数智治理概念体系构建:要素、特征及关系.中国科技术语,2024,(3).

On the Risks in Applying GAI in Digital Government Construction and Their Mitigation

Fu Jianping（Central South University）

Abstract In the process of digital government construction, generative artificial intelligence (GAI) technology can play beneficial roles in enhancing information organization and information-driven capabilities, as well as optimizing governmental operational structures. At the same time, embedding GAI technology into digital government construction can also pose risks. The risk society theory, which aims to systematically address the anxieties and hazards induced by modernization, can provide assistance in identifying and overcoming such risks. Regarding risk identification, the application of GAI in digital government construction may generate risks related to data security, blurred responsibility, and privacy infringement. While in risk governance, adopting an agile governance approach, a comprehensive regulatory path should be taken to establish a framework for classifying and regulating data of different levels and categories, clarify the primary responsibility of the government, effectively activate private laws and norms for privacy information protection and pay attention to timely assessing effectiveness and making corrections.

Key words risk society; generative AI technology; digital government; privacy

■ **作者简介** 傅建平,中南大学马克思主义学院副教授,湖南 长沙 410083。
■ **责任编辑** 涂文迁