基于情景分析的人工智能赋能智慧城市信息服务风险治理研究*

郭骅[1]　董怡彤[1]　李辉[2]
（1. 河海大学商学院，南京 211100；2. 南京市江北新区城市数字治理中心，南京 211899）

摘要：为识别信息服务中人工智能应用的多元风险，提出针对性的治理策略，助力智慧城市的高质量与可持续发展。运用情景分析法，首先构建人工智能赋能智慧城市信息服务应用场景，设定情景基线；其次建立风险框架和识别风险因素，通过环境分析识别关键驱动力和不确定性；最后提出面向未来情景的治理原则和对策。本研究描绘了应用场景构成和风险特征，构建共融善治、技术孤岛、包容探索、失序危机4种未来情景，提出以协同治理与门径管理为基石的风险治理框架，并结合差异化的领域重点与情景特征，统筹创新发展与风险防控，为安全、可靠的智慧城市信息服务体系建设提供理论参考。

关键词：智慧城市；信息服务；人工智能；情景分析；风险治理
中图分类号：G35；TP18　DOI：10.3772/j.issn.1673-2286.2026.02.007
引文格式：郭骅，董怡彤，李辉. 基于情景分析的人工智能赋能智慧城市信息服务风险治理研究[J]. 数字图书馆论坛，2026，22（2）：63-72.

人工智能的快速发展正深刻推动智慧城市转型升级。2024年5月，国家发展改革委等四部门联合发布《关于深化智慧城市发展 推进城市全域数字化转型的指导意见》[1]，明确提出要大力发展人工智能技术在智能分析、调度、监管、辅助决策等信息服务场景中的应用，全面支撑并赋能城市数字化转型。在政策引导与技术创新的双重驱动下，人工智能赋能的智慧城市信息服务日益成为城市高质量建设的核心引擎，实现了从基础数据加工到高层次信息增值的转化。

传统城市信息服务主要依赖结构化数据库和预设规则，风险多集中在数据录入错误、系统宕机和权限滥用等易识别、可量化的静态环节，通常可通过流程管理、日志审计等成熟机制管控。相比之下，人工智能技术的应用具有难以捕捉的复杂性、动态性和发展性，既推动了信息服务的数智化升级，又带来高度不确定的风险，给城市可持续发展带来挑战。明晰人工智能赋能智慧城市信息服务的应用场景，识别和应对潜在风险，对提升智慧城市智能化水平和风险防范能力具有重要意义。

1　相关研究

1.1　智慧城市信息服务发展

在相当长的时期内，信息服务主要指由专业机构提供的检索、复制和咨询等传统信息活动[2]。党的十八大以来，数智化创新实践为信息服务注入了新的场景和内涵[3-4]：人工智能的引入大幅提升了信息处理、分析和应用能力，驱动信息服务的主客体关系、增值要素、服务内容、服务形态和服务流程发生转变[5-6]，为智慧城市转型升级提供了底层逻辑[7-8]。智慧城市建设破除信息孤

收稿日期：2025-11-17
*本研究得到国家社会科学基金项目“人工智能赋能数智城市的信息服务与治理研究”（编号：24BTQ056）资助。

岛和数据分割，使信息服务具备融合与跨界的特性，引发学者对众多应用场景的广泛关注。现有研究聚焦于医疗、交通、教育、政务和商务5个领域[9]。在医疗领域，智能信息服务覆盖“检、诊、疗、护”全过程，显著提升医疗服务效率与智能化水平[10-11]。交通领域聚焦交通管理与规划[12]和出行服务体验[13]，旨在通过智能信息服务提升出行体验[14]。教育领域依托人工智能赋能图书馆科研、咨询服务等场景[15]，推动教育资源配置与知识服务智能化。政务领域的智能信息服务涵盖智能决策、智能管理、智能服务与智能监管，逐步构建多元协同、智能联动的一体化治理体系[16-17]。商务领域通过基于用户数据的智能营销、个性化推荐、动态定价与市场趋势预测等场景，实现更高水平的商业智能与经济效益提升[18]。

1.2 人工智能的风险和治理

在技术进步带来社会繁荣的同时，人类必须正视人工智能风险与治理挑战。一般认为，人工智能风险分为技术和制度两个维度：技术风险指新技术产生的负面影响，如对人类认知的冲击、系统安全隐患和隐私泄露；制度风险指既有制度和规范难以有效协调各方利益所引发的潜在问题[19]。麻省理工学院研究团队构建的人工智能风险数据库（MIT AI Risk Repository）将风险划分为歧视偏见、隐私与安全、错误信息、恶意和滥用、人机交互风险、社会经济与环境风险、人工智能系统安全7个领域[20]，为风险识别提供通用参考框架。在智慧城市领域，学者多通过文献综述构建风险图谱，例如：Ullah等[21]基于TOE（Technology-Organization-Environment）框架识别56个风险要素，Kirytopoulos等[22]梳理65种风险并以分解结构图呈现，Wolniak等[23]回顾了在智能治理、智能出行等领域中实施人工智能的障碍，Al Sharif等[24]指出现有研究缺乏跨领域、跨风险类型的整体性治理方案。

针对这些风险，学界提出多种治理方法和路径。在通用层面，适应性治理[25]、预见性治理[26]、试探性治理[27]、敏捷性治理[28]、自反性治理[29]等模式被广泛用于解释各类新兴技术的治理实践。在具体应用场景下，学者结合不同领域特征提出针对性策略。吴培熠等[30]主张通过明确场景范围、保障数据安全等策略，推动智能政务服务高质量发展；李少帅[31]提出在商业应用、公商混合以及公共服务3种场景下的差异化治理策略，并将此转化为适配的治理模式；顾东晓等[32]围绕医疗知识服务场景提出全过程风险数字治理策略。

1.3 研究述评

总体来看，相关研究多聚焦特定场景的应用实践与路径优化，系统性的场景构建、风险识别及治理策略研究相对薄弱。首先，多数学者关注人工智能赋能智慧城市的应用场景，对其风险治理少有研究；其次，风险研究多依赖文献综述，缺乏更具前瞻性的分析视角；最后，在从治理理论到实践操作的转化过程中，缺乏全面的应用场景支撑，因而难以提出有针对性的治理对策。

随着智慧城市的转型升级不断推进，人工智能赋能智慧城市信息服务的风险体系难以完全沿用现有经验，风险识别与应对更具前瞻性和不确定性。因此，有必要采用情景分析方法，系统探讨在高度不确定条件下可能出现的风险情景及相应的治理策略。

2 研究设计

2.1 研究方法

场景是指在特定时间和地点发生的一系列相关事件或活动，反映现实[33-34]。本研究中的应用场景指人工智能嵌入城市信息服务系统后，在特定功能领域与业务流程中形成的结构化运行形态及其要素组合。参照国家标准提取相关系统要素，用作构建情景基线的基础。情景基线是场景在系统要素层面的结构化归纳，是风险识别和情景分析的前置环节。情景是对事物未来多种可能发展趋势的描述[35]，用来规划未来可能呈现的多种发展状态，从而分析风险结构的演化差异。

情景分析法通过整合不同的未来状态组合，在保留不确定性的同时，为风险管理和战略决策提供多元化参考[36]。情景分析有多种学术流派，常见情景类型包括预测性情景、探索性情景、规范性情景[37]。本研究主要采用直觉逻辑定性方法[38]和GBN（Global Business Network）矩阵技术[39]，识别关键及不确定的驱动因素，为探索性情景的构建奠定基础。

2.2 研究思路及框架

传统风险研究侧重于量化风险概率及其后果严重

性，但人工智能的未来发展充满不确定性，难以依赖概率与历史经验进行有效评估[40]。因此，有必要以更加前瞻的视角，探索多种通向未来的可能路径。

具体而言，人工智能赋能智慧城市信息服务主要面临三大不确定性。①应用场景模糊性。创新应用不断涌现，但尚无结构化体系，场景描述零散且不全面。②风险因素伴随性。风险不仅源自技术本身，还会随着应用场景的变迁被重新定义、放大或转化。③社会和技术发展不确定性。人工智能发展历程起伏多变，未来社会对新技术的适应接纳存在争议。

针对第一点，通过构建应用场景模型与设定情景基线，弥补场景描述的不足；针对第二点，从风险与场景互动的根源出发，构建适配的风险框架；针对第三点，通过识别人工智能赋能智慧城市信息服务发展的核心驱动力，描绘未来情景并提出治理对策。具体研究框架如图1所示。

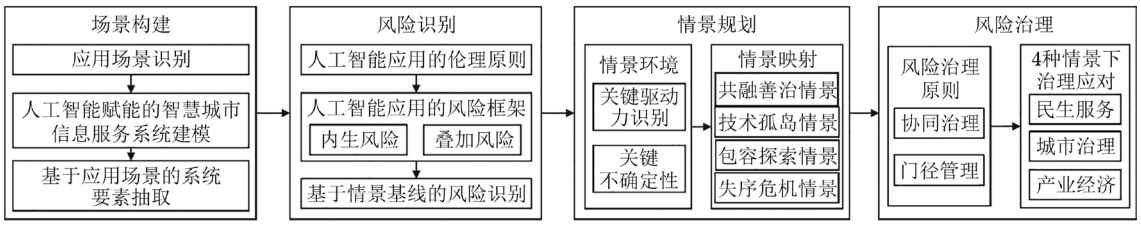


图1　研究框架

3　人工智能赋能智慧城市信息服务的应用场景构建

3.1　应用场景识别

本研究选用国家标准《智慧城市 人工智能技术应用场景分类指南》[41]（以下简称《分类指南》）作为数据来源。该标准由全国信息技术标准化技术委员会提出，汇聚多方力量对智慧城市中人工智能的应用进行了结构化梳理与前瞻性展望，构建了包含4个层次的场景体系：大类场景依据城市功能划分为民生服务、城市治理、产业经济和生态宜居4个领域；中类场景根据行业属性进一步细分为智慧政务、智慧交通等类型；小类场景以用户对象或业务职能为分解依据；子类场景按行业功能需求进一步细分，如智能预约、智能审批。

总体来看，应用场景可分为两种：一种是直接面向实体操作的智能行动场景，如在自动驾驶场景中，机器替代人类执行驾驶任务；另一种则以信息增值为核心，提供智能信息服务。本研究以《分类指南》中的283个子类场景为分析单元，基于信息链[42]和情报流程[43]理论识别信息服务场景。识别标准为信息在服务过程中经历增值转化、存在明确的信息服务主体与目标以及应用人工智能技术，信息服务数据抽取详情具体如表1所示。反复阅读每个分析单元的描述，对相关语义内容进行编码标注。分析过程由3名研究人员独立完成，并对判定结果交叉比对，讨论达成一致后，最终共筛选出140个信息服务应用场景。

表1　信息服务数据抽取

抽取信息	文本对应内容
信息服务目标	子类应用场景
信息服务主体	提供和参与信息服务的用户或系统
技术	人工智能技术
情报流程	①需求感知与获取：用户交互、需求感知、数据采集；②加工与处理：数据处理、数据存储；③分析与挖掘：信息分析、发现新知识/新规律；④转化与传播：情报输出、情报对接、情报传播
信息链	①数据：对事实的客观记录，是通过传感器、设备或系统采集的原始数值或文本，如采集、获取、输入、传感器等；②信息：通过加工处理数据生成的，能够揭示特定上下文中数据的意义，如描述、展示、显示、解释等；③知识：通过对信息的总结、分析和模式识别形成的更高级层次，是对信息的系统理解和归纳，如分析、识别、总结、理解、规则等；④情报：针对特定需求进行深入挖掘和预测，帮助决策者在复杂环境中进行决策，如预测、决策、风险评估、建议等

3.2　应用场景建模

从系统角度看，信息服务是嵌入智慧城市业务的技术赋能系统，人工智能技术则为信息服务二次赋能（见图2）。因此，人工智能赋能智慧城市信息服务可视为一个由城市业务子系统、信息服务子系统和人工智能赋能子系统共同构成的整体系统。城市业务子系统决定最终目标和需求；信息服务子系统通过信息加

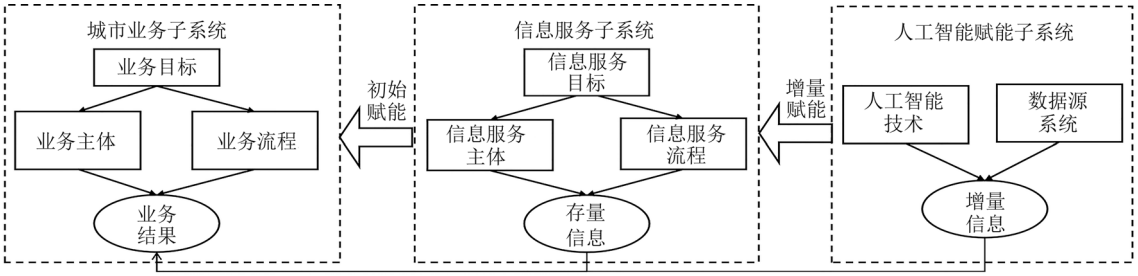


图2　系统模型

工与处理为业务提供技术支撑；人工智能赋能子系统则利用人工智能生成增量信息，进一步赋能城市业务。

3.3　场景-模型映射

从已经识别的应用场景中提取系统要素，重点揭示人工智能技术赋能前后，系统在主体、目标与流程等方面的变化，映射方法如表2所示。完成映射后，整合具有相同要素的应用场景，最终提炼出覆盖民生服务、城市治理和产业经济3个大类场景的73个差异化子类场景，作为本研究的情景基线，为后续的风险识别提供结构化支撑，部分映射结果如表3所示。

映射结果显示，机器学习、人工智能生成内容及自然语言处理等共同构成了人工智能赋能子系统的技术支柱，对各场景深度赋能。其中：所有场景均依赖机器学习技术；人工智能生成内容技术主要服务于方案、建议、报告等自动生成场景，生成式对抗网络与大规模预训练模型被广泛采用；自然语言处理技术在咨询、客服等场景中发挥关键作用。信息服务子系统的场景映射则显现出应用领域差异，民生服务涵盖政务、医疗、交通、教育、社区、旅游、体育7个中类场景，主要服务对象是城市居民和访客，服务流程注重效率和便捷，服务成效以用户体验为评价标准。城市治理以政府管理和监督部门为服务对象，覆盖安防、应急、能源、环保、城市管理、市场监管6个中类场景，主要依靠感知数据构建“智能识别-预警评估-闭环处置-决策支撑”的信息增值链，强调服务流程的规范性和结果的可靠性。产业经济服务于企业及产业链上下游，包括金融、农业、零售、园区、物流、制造、环保7个中类场景，服务侧重于分析辅助决策和反馈优化，聚焦流程自动化，以提升企业竞争力为核心目标。

表2　映射方法

系统	要素	映射方法
人工智能赋能子系统	人工智能技术	对应“技术”文本，并与《关键数字技术专利分类体系（2023）》知识对齐
	数据源系统	基础数据：对城市运行管理有全局性、标准化要求的数据
		业务数据：在信息服务或业务流程中产生并积累的数据
		感知数据：反映环境、用户行为与事件动态的数据
	增量信息	对应“场景描述”中由基础、业务或感知数据经过信息过程增值转化而来的信息
信息服务子系统	信息服务目标	对应“子类应用场景”
	信息服务主体	对应“场景描述”，自行归纳或推理得到人工智能赋能前后的差异
	信息服务流程	
	信息服务结果	

表3　应用场景-系统模型映射结果（部分）

大类场景	中类场景	子类场景	人工智能赋能子系统		信息服务子系统
			人工智能技术	数据源	
民生服务	智慧政务	政务事项预约导办	计算机视觉、机器学习、自然语言处理、语义理解	基础数据、业务数据	服务目标：从排队引导转变为自动预约、智能导办；服务主体：从引导员、排队机、信息机转变为人工智能机器系统；服务流程：从分散流程变为动态交互的信息服务集成
城市治理	市场监管	企业信用监管评估	知识图谱、机器学习、人工智能生成内容	基础数据、业务数据	服务目标：从静态档案式信用查询转变为动态信用评估与预警；服务主体：从分析师转变为人工智能机器系统；服务流程：从多部门业务协同转变为报告自动生成
产业经济	智慧零售	零售商品智能定价	机器学习	业务数据、感知数据	服务目标：从经验定价转变为基于市场和供应链数据的动态定价；服务主体：从营销人员转变为人工智能机器系统；服务流程：从竞争情报分析程序转变为算法驱动的自动化程序

4　基于场景的风险识别

4.1　风险框架

人工智能风险源自其对人类尊严和社会繁荣核心价值的挑战，可能导致人类能力被削弱、责任被转移、控制力下降和自决权受损。AI4People项目提出5项伦理原则：受益原则、非恶意原则、自主原则、公正原则和可解释性原则[44]，为人工智能风险的规范性识别提供了重要伦理参照。在确认风险因素方面，麻省理工学院发布的人工智能风险数据库[20]通过系统梳理65篇既有研究成果，基于领域分类法将人工智能风险划分为24个子领域，为不同应用场景下的风险识别提供共同参照，具有较强的时效性与代表性。既有研究进一步表明，人工智能风险本质上通常被概括为两类[45-46]：一是人工智能技术自身的内生风险，包括数据、模型、系统框架等技术环节所固有的风险问题；二是人工智能技术应用后衍生的风险，即系统赋能风险，是技术赋能行为所衍生的结构性与社会性风险，这类风险体现为技术在运行过程中对信息服务目标、信息服务主体与社会关系产生的连锁影响。

本研究以系统内生风险和系统赋能风险为基础构建风险框架，将伦理原则与人工智能风险数据库结合，置于系统模型，以其具体表现形式和作用方式为依据，对照原有框架风险内涵，对风险因素进行筛选与内涵界定。具体而言，将与智慧城市信息服务关联度较低或在当前技术条件下尚不具备现实适用性的5个风险因素剔除，并对内涵重叠度高的风险因素进行合并与重组，最终形成含有16个风险因素的风险框架（见表4）。

4.2　风险识别

将风险因素映射至结构化的情景基线中，构建如表5所示的风险矩阵。剔除各子类场景中存在的共性风险R7~R10后，梳理各子类场景的风险共现关系，构建

表4　风险框架

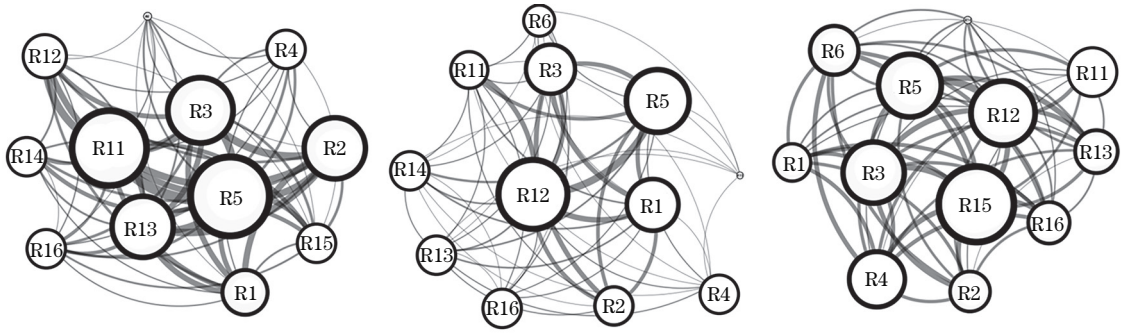
风险因素		伦理原则	风险因素的内涵
系统内生风险	R1 算法偏见与不公平歧视	公正原则	因数据采集、模型训练及算法设计偏差，系统对不同社会群体产生不公平处理，形成基于特征数据的不平等待遇
	R2 群体表现不一致	受益原则	训练数据多样性不足，导致系统理解、适配能力下降，模型识别与决策精度降低
	R3 隐私泄露与敏感信息推断	非恶意原则	系统未经授权记忆、存储或推断用户敏感数据，造成信息泄露，侵害用户隐私、网络身份及知识产权
	R4 欺诈操控与有针对性的欺骗	非恶意原则	恶意行为者利用智能问答、推荐等系统设计漏洞，实施诈骗或定向操纵，谋取不当利益
	R5 虚假或误导性信息	受益原则	系统因数据或算法缺陷，在生成政务信息等内容时产生错误或欺骗性信息，误导公众及管理者决策
	R6 不正当内容生成	非恶意原则	若信息服务对外公开交互并允许生成式对话，可能存在生成或展示有害、侮辱等不当内容的风险，危害公共安全和社会治理
	R7 缺乏系统稳健性	受益原则	系统无法在某些条件下可靠或有效运行，导致错误和故障，带来严重后果
	R8 系统安全漏洞与攻击	非恶意原则	系统、开发工具及运行设备存在安全漏洞，被恶意攻击者非法访问、操纵，威胁城市管理与关键基础设施安全
	R9 环境危害	受益原则	系统开发和运行消耗大量数据，可能对环境造成危害
	R10 缺乏透明度和可解释性	可解释性原则	信息过程不透明，使理解或解释系统决策面临挑战，导致业务或信息服务主体对结果缺乏信任，无法识别和纠正错误
系统赋能风险	R11 过度依赖与自主性削弱	自主原则	业务或信息服务主体信任和依赖系统提供的决策建议，甚至将决策和行动委托给人工智能，导致批判性思考和独立解决问题的能力下降
	R12 责任认定模糊	可解释性原则	当人类和系统同时完成某项信息服务时，出现错误决策导致损失，其责任主体难以明确界定，给监管和治理带来难度
	R13 就业质量下降	受益原则	人工智能替代大量城市业务岗位，导致相关从业者失业，或被迫接受低薪酬、低福利的工作条件
	R14 劳动与文化价值贬低	受益原则	复制人类创新与创造，降低社会对人类技能的认可度，引发知识型产业中断、文化体验同质化问题
	R15 技术竞争态势	非恶意原则	人工智能开发和部署者优先考虑短期内部目标（例如提升利润或影响力）以确保市场竞争地位，可能增加发布不安全且易出错的系统风险
	R16 信息生态污染与共识缺失	公正原则	系统基于主体偏好推送个性化内容，形成信息茧房，削弱社会凝聚力与群体共识，阻碍城市可持续治理

表5　风险矩阵（部分）

应用场景			风险因素															
大类场景	中类场景	子类场景	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14	R15	R16
民生服务	智慧政务	政务事项预约导办	1	1	1	0	0	0	1	1	1	1	1	0	1	0	0	0
城市治理	市场监管	企业信用监管评估	1	1	1	0	1	0	1	1	1	1	0	1	0	1	0	0
产业经济	智慧零售	零售商品智能定价	0	0	1	0	0	0	1	1	1	1	1	1	1	0	1	0

注：0表示该应用场景不存在该风险因素，1表示存在该风险因素。

3个领域的风险因素共现网络（见图3），用以分析不同城市功能领域的风险格局。在共现网络中，节点大小由加权度决定，用以表征风险因素的中心性水平，节点越大表示其在网络中的中心性越强。节点之间的连边用于刻画风险因素之间的关联关系，其粗细反映关联强度，连边越粗表示风险因素之间的共现关系越紧密。



(a)民生服务　(b)城市治理　(c)产业经济

图3　3个领域的风险因素共现网络

总体来看，在受益原则与非恶意原则的伦理指导下，虚假或误导性信息、隐私泄露与敏感信息推断已成为普遍面临的系统内生风险挑战。除共性风险外，各领域还呈现出由系统赋能带来的差异化风险特征：民生服务领域，过度依赖与自主性削弱风险、就业质量下降风险较为显著，表明该领域的系统赋能可能损害人类的发展利益和自我决策权；在城市治理领域，责任认定模糊、算法偏见与不公平歧视问题突出，说明人工智能促进的社会公正必须以技术开发和部署中人和组织的公正为前提，不能任由人工智能承担信息过程却无法有效追责；在产业经济领域，技术竞争态势风险与责任认定模糊风险更为明显，反映出企业在人工智能系统研发和行业应用部署中的激烈竞争，可能导致产业发展中出现难以溯源与修复的严重后果。

5　人工智能赋能智慧城市信息服务的情景规划

5.1　环境分析

（1）驱动因素分析。在人工智能视域下，智慧城市信息服务的发展受到多元驱动力影响，驱动力的确定基于前文风险分析结果，通过对高影响风险因素的归纳与抽象提炼关键驱动力，用以构建情景。在民生服务领域，技术的应用有助于解决多方面的难点问题，但也反映出公众对人工智能技术介入的顾虑。唯有在技术持续优化和公众充分信任的基础上，人工智能才能在民生服务中稳健落地。在城市治理领域，政府主导、责任主体多元且数据来源复杂，这对信息安全和责任追溯提出了更高要求。提升数据体系成熟度以及完善机制体系是保障城市治理信息服务高效运行的关键。产业经济领域则重视人工智能与传统行业的深度融合，推动数字化转型升级。技术投入强度和行业对技术的适应能力，成为推动人工智能广泛应用的重要因素。因此，发展驱动力可归纳为两大类：一是技术系统成熟度，包括人工智能的技术进步、数据体系的健全以及行业适用性；二是社会系统成熟度，涵盖社会公众的接受度、治理和监管机制的完善程度，以及经济投入的持续性和保障能力。

（2）不确定性分析。情景分析中，有必要排除发展趋势明确、对未来情景分化影响有限的确定性因素，聚焦关键不确定性因素。技术系统方面，生成式人工智能

等技术快速发展，算法模型创新加快、性能稳步提升，其发展趋势具有确定性，核心不确定性集中于技术公正性、行业融合和信息安全；社会系统方面，人工智能训练成本攀升、国际竞争格局基本确立是确定性事实，而社会公众的技术接受度、监管体系的完善程度成为未来情景分化的关键影响因素。模型性能提升、训练成本增加以及国际竞争格局等因素均具有确定性趋势。在情景分析中，真正影响未来分化的关键不确定性主要集中于技术系统的公正性、可解释性、行业融合与安全，以及社会系统的接受度与监管机制的完善程度。这些因素将直接影响人工智能在智慧城市信息服务体系中的应用路径与治理复杂性。

5.2　情景叙述

结合GBN矩阵，以技术系统成熟度和社会系统成熟度为两条关键不确定性轴，分别设定高、低发展状态，构建2×2情景矩阵，规划4种未来情景。

（1）共融善治情景：技术与社会系统高水平协同发展。人工智能具备可解释、公正、安全的核心技术，社会公众高度认可，监管与伦理体系健全，智能信息服务在推动社会治理创新和经济高质量发展中发挥重要作用，显著提升民生福祉与公共价值。民生服务的智能应用可靠、可信且负责任，公众对技术和数据安全充满信心，人工智能应用广泛普及，技术创新和岗位升级助力个人价值实现；城市治理的大数据体系完善，可融合、可追溯，政府信息共享和监管机制成熟，决策过程透明且可问责；政府和企业重视研发、基础设施和人才培养，推动行业应用实现规模化商用，有效满足适配产业应用多样化需求。

（2）技术孤岛情景：技术系统领先但社会系统适应滞后。尽管技术迅猛发展，但社会公众对人工智能缺乏信任，数据安全和伦理问题引发广泛担忧，监管机制和社会认知滞后于技术发展。民生服务中公众对信息准确性和安全性的疑虑阻碍了智能信息服务的普及，原有信息服务岗位转型升级空间有限，就业保障进一步制约了技术推广；城市治理方面，部门间信息壁垒依然存在，决策难以实现充分透明与追责，监管措施相对滞后；产业经济领域，由于行业壁垒和利益冲突，投资回报低于预期，核心技术瓶颈难以突破，行业应用规模十分有限，加剧了技术与社会的脱节。

（3）包容探索情景：社会系统包容引领技术系统发展。尽管人工智能在可解释性和安全性方面尚有不足，但社会整体仍持开放包容态度，监管与伦理机制相对完善，为技术试错和优化创造了良性环境。在民生服务领域，智能信息服务快速落地并广泛试点，通过迭代优化用户体验和应用模式，提升服务质量和个人获得感；城市治理领域，政府开展包容审慎监管，完善城市大数据体系和智能信息服务平台建设，通过反馈机制提升治理能力；产业经济领域，政府政策与企业资源持续投入，构建产业平台和行业场景，形成示范带动效应，助力经济社会转型升级。

（4）失序危机情景：技术系统与社会系统低水平错配。技术缺乏可靠性、可解释性和安全保障，社会接受度低，信息服务生态薄弱，智慧城市信息服务的发展停滞。在民生服务领域，数据安全和用户权益难以保障，公众信任度低，就业结构失衡，社会不满加剧；城市治理领域，数据体系不健全、信息孤岛严重，监督监管缺失，决策不透明且责任不清，公共管理能力下降；产业经济方面，盲目创新导致风险积聚，规范和协同不足，政府和企业投入低效，技术难以形成规模化效益，严重威胁社会稳定与可持续发展。

6　人工智能赋能智慧城市信息服务的风险治理策略

6.1　治理原则

随着人工智能深度赋能智慧城市信息服务，风险呈现出多元化和复杂化趋势，且不同领域风险类型、表现各异，风险间高度关联。情景分析进一步显示，未来信息服务的发展路径高度依赖技术系统与社会系统的协同成熟，单一主体或单一治理手段难以应对多样化、动态演化的风险挑战。因此，亟须构建以协同治理[47]与门径管理[48]为基石的风险治理框架，统筹创新发展与风险防控，提升信息服务系统的整体韧性与可持续性。

协同治理强调政府、行业组织、社会公众等多元主体在风险治理中的互动协作，通过资源整合、信息共享和利益协调，提升风险识别的全面性和风险应对的灵活性，适用于系统复杂、广泛部署的情景。门径管理则对高风险技术应用和敏感场景实施分级准入、动态审查和阶段性放行，通过门径机制实现前置防控和动态纠偏。通过小范围试点、监管沙箱等方式，门径管理为

技术完善和规制演进提供缓冲空间，保障智慧城市信息服务的可持续发展，适用于新兴技术早期、敏感环节或高风险不确定情景。在信息服务实践中，应结合不同功能场景与风险状态，灵活运用两类治理原则，实现高质量发展与有效风险防控的有机统一。

6.2 领域重点

民生服务领域风险主要集中在信息真实性、隐私保护以及对社会结构与信任体系的影响。治理应重点强化公众参与和伦理嵌入，完善数据安全与内容监管机制，提升公众数字素养和信息辨识能力，防范信息失真和数据滥用。同时，应健全就业支持和转型保障机制，缓解技术替代带来的就业冲击，维护社会公平，增强公众信任。

城市治理领域核心风险包括责任认定不清、算法偏见与不公平歧视、缺乏透明度等。因此，应坚持明晰责任追溯、提升算法透明度和公正性，完善政务信息审核和应急响应机制，防止决策失序和算法歧视。同时，应加强城市大数据体系的建设、治理和应用，通过多元主体协作，打造高效开放的治理格局。

产业经济领域风险主要表现为恶性竞争、责任模糊和敏感信息推断。人工智能推动数字经济转型的同时，也带来了行业规范缺失、企业责任不落实和劳动者权益受损等问题。治理应聚焦健全行业标准和企业责任体系，加强合规管理和行业自律，防范市场无序扩张，通过动态门径机制防止风险外溢和系统性失控。

6.3 情景对策

将智慧城市信息服务未来发展的4种典型情景统一在协同治理与门径管理两大治理原则下，结合不同情景下风险结构的差异性，因情施策，动态调整治理重心与手段。

（1）共融善治情景。在技术与社会高度协同的最优状态下，治理应由“被动应对”转向“主动塑造”，不仅关注风险防控，更强调生态优化和系统韧性。民生服务领域强化人机协同与数字理性教育，推动“人机共决策”，避免过度依赖人工智能。政府和企业协同创造新的就业机会，将技术红利转化为就业增长。城市治理领域建设城市数据治理中台，统一身份认证和数据标准，实现数据同步、业务流转和风险全程留痕，支撑跨部门的信息过程和责任追溯。产业经济领域倡导企业在透明、安全、伦理等方面展开良性竞争，防止技术逐利损害生态平衡。

（2）技术孤岛情景。在技术系统领先但社会系统适应滞后的情景下，治理应以“风险前置、门控优先”为主，防止技术与社会脱节带来的系统性风险。在民生服务领域，强化对公众的人工智能素养教育，推动就业转型与技能培训，缓解技术替代冲击；对高风险智能信息服务实施严格的分级准入和动态审查机制，防止风险外溢。在城市治理领域，完善伦理规范和监管机制，建立跨部门风险预警和应急响应体系，确保技术创新与社会承接能力同步提升。在产业经济领域，推动企业自律和运营透明化，防止技术滥用和责任转嫁，鼓励行业协会和第三方机构参与风险评估和合规审查。

（3）包容探索情景。在社会系统包容开放、监管与伦理机制较为完善但技术系统尚存短板的情景下，治理应突出“包容审慎、协同迭代”。在民生服务领域，建立多方参与的评议与反馈机制，定期发布系统性能与安全评估报告，培养公众对智能信息服务的理性认知。在城市治理领域，依托现有城市数据共享交换平台，完善政府数据开放机制，推动信息互联互通，明确人工智能代理的责任分配和问责流程。在产业经济领域，制定差异化扶持政策，引导企业稳步推广智能应用，鼓励合规创新。

（4）失序危机情景。在技术与社会系统双重低谷、风险高度聚集的危机情景下，治理应以“底线防控、能力补短”为核心。在民生服务领域，坚持数据最小化和透明授权，确保以最少必要数据提供技术应用，通过量化指标定期评估模型性能，及时纠错，逐步恢复公众信任。在城市治理领域，加强数据治理和质量监管，引入第三方独立审计与仲裁，多主体协同保障服务安全可靠。在产业经济领域，加大基础研发与平台建设投入力度，建设跨行业共享算力与测试平台，降低试点门槛，鼓励可持续发展。

7 结语

人工智能技术为智慧城市信息服务的智能化升级带来了重要机遇，但其快速发展同样伴随着诸多风险与挑战。本研究以场景抽取、风险映射、情景规划与治理应对为研究主线，系统构建了人工智能赋能智慧城市信息服务系统模型，识别了各类应用场景中的关键风险

因素，并进一步刻画了典型发展情景，提出治理原则和应对策略，为未来的智慧城市建设提供了风险分析和治理框架。需要指出的是，当前相关研究仍缺乏充分的实证支持，未来可加强案例分析和实地调研，完善理论与实践框架，推动构建更加安全、可靠的智慧城市信息服务生态。

参考文献

[1] 关于深化智慧城市发展 推进城市全域数字化转型的指导意见[EB/OL]. [2025-11-12]. https://www.gov.cn/zhengce/zhengceku/202405/content 6952353.htm.

[2] 张潇月，刘畅. 面向用户信息需求表达的信息问题空间构建研究[J]. 情报学报，2023，42（3）：365-379.

[3] 张海涛，张春龙，张鑫蕊，等. 新文科背景下“情报学+”人才的时代意蕴[J]. 中国图书馆学报，2022，48（4）：67-78.

[4] 何劲，王曰芬，傅柱. 场景式情报服务及其发展路径研究[J]. 情报理论与实践，2024，47（3）：67-72.

[5] 胡昌平，邓胜利. 数智背景下的信息交互与服务发展[J]. 情报理论与实践，2024，47（4）：1-9.

[6] 孙建军，马亚雪，李阳. 面向前沿交叉领域的信息服务路径思考[J]. 图书与情报，2024（6）：1-9.

[7] 张艳丰，邹凯，彭丽徽，等. 数字空间视角下智慧城市全景数据画像实证研究[J]. 情报学报，2020，39（12）：1330-1339.

[8] BEN RJAB A，MELLOULI S，CORBETT J. Barriers to artificial intelligence adoption in smart cities: a systematic literature review and research agenda[J]. Government Information Quarterly，2023，40（3）：101814.

[9] 吴丹，刘子君. 大数据视角下的智慧信息服务：应用实践与未来趋势[J]. 信息资源管理学报，2018，8（2）：28-39，75.

[10] 张传洋，郭宇，庞宇飞，等. 数智化医疗信息利用与服务模式框架构建[J]. 图书情报工作，2023，67（13）：49-58.

[11] 翟兴，肖源，王若佳，等. 数智环境下智慧健康信息服务体系构建研究[J]. 情报科学，2022，40（10）：43-50.

[12] ENGLUND C，AKSOY E E，ALONSO-FERNANDEZ F，et al. AI perspectives in smart cities and communities to enable road vehicle automation and smart traffic control[J]. Smart Cities，2021，4（2）：783-802.

[13] TURO K，TÓTH J. Innovations in shared mobility: review of scientific works[J]. Smart Cities，2023，6（3）：1545-1559.

[14] 赵光辉. 人工智能驱动交通运输高质量发展逻辑机理及发展路径[J]. 当代经济管理，2025，47（7）：18-30.

[15] 郭亚军，郭一若，李帅，等. ChatGPT赋能图书馆智慧服务：特征、场景与路径[J]. 图书馆建设，2023（2）：30-39，78.

[16] 陈潭，刘璇. 智能政务ChatGPT化的前景与隐忧[J]. 电子政务，2023（4）：36-44.

[17] 张亨明，章皓月. 城市治理智慧化的理论分析与实践探索[J]. 求索，2021（6）：156-164.

[18] KHAMOUSHI E. AI in food marketing from personalized recommendations to predictive analytics: comparing traditional advertising techniques with AI-driven strategies[EB/OL]. [2025-11-12]. https://arxiv.org/abs/2410.01815.

[19] LIU Z Y，ZHENG Y J. AI ethics and governance: black mirror and order[M]. Singapore：Springer Nature Singapore，2022.

[20] SLATTERY P，SAERI A K，GRUNDY E A C，et al. The AI risk repository: a comprehensive meta-review，database，and taxonomy of risks from artificial intelligence[EB/OL]. [2025-11-12]. https://arxiv.org/abs/2408.12622.

[21] ULLAH F，QAYYUM S，THAHEEM M J，et al. Risk management in sustainable smart cities governance: a TOE framework[J]. Technological Forecasting and Social Change，2021，167：120743.

[22] KIRYTOPOULOS K，CHRISTOPOULOS T，DERMITZAKIS E. Smart cities: emerging risks and mitigation strategies[M]// FITSILIS P. Building on Smart Cities Skills and Competences: Human Factors Affecting Smart Cities Development. Cham：Springer International Publishing，2022：123-139.

[23] WOLNIAK R，STECUŁA K. Artificial intelligence in smart cities: applications，barriers，and future directions: a review[J]. Smart Cities，2024，7（3）：1346-1389.

[24] AL SHARIF R，POKHAREL S. Smart city dimensions and associated risks: review of literature[J]. Sustainable Cities and Society，2022，77：103542.

[25] CHAFFIN B C，GOSNELL H，COSENS B A. A decade of adaptive governance scholarship: synthesis and future directions[J]. Ecology and society，2014，19（3）：53.

[26] GUSTON D H. Understanding “anticipatory governance”[J]. Social Studies of Science，2014，44（2）：218-242.

[27] 李冲，李霞. 人工智能风险的试探性治理：概念框架与案例解析[J]. 中国软科学，2024（4）：91-101.

[28] MERGEL I，GONG Y W，BERTOT J. Agile government: systematic literature review and future research[J]. Government Information Quarterly，2018，35（2）：291-298.

[29] DRYZEK J S，PICKERING J. Deliberation as a catalyst for

reflexive environmental governance[J]. Ecological Economics，2017，131：353-360.

[30] 吴培熠，申红梅，梁正. 政务大模型的敏捷治理研究：基于“数据-场景”的多案例分析[J]. 电子政务，2025（8）：2-12.

[31] 李少帅. 场景驱动视角下的大模型治理策略与模式选择[J]. 当代经济管理，2025，47（5）：27-33.

[32] 顾东晓，黄智勇，朱凯旋，等. 医疗健康大模型知识体系构建、服务应用与风险协同治理[J/OL]. 情报科学：1-29[2026-01-11]. https://link.cnki.net/urlid/22.1264.G2.20240925.0948.002.

[33] 秦春秀，马茜茜，马续补，等. 信息资源管理视角下的场景概念及场景分析模型[J]. 图书情报工作，2024，68（13）：78-86.

[34] 范炜. 情报学视角的情境概念及情境观认识[J]. 图书情报工作，2020，64（12）：3-10.

[35] 宗蓓华. 战略预测中的情景分析法[J]. 预测，1994，13（2）：50-51，55.

[36] 孙建军，柯青. 不完全信息环境下的情报分析方法：情景分析法及其在情报研究中的应用[J]. 图书情报工作，2007，51（2）：63-66，120.

[37] AMER M，DAIM T U，JETTER A. A review of scenario planning[J]. Futures，2013，46：23-40.

[38] BRADFIELD R，WRIGHT G，BURT G，et al. The origins and evolution of scenario techniques in long range business planning[J]. Futures，2005，37（8）：795-812.

[39] OLIVEIRA A S，DE BARROS M D，DE CARVALHO PEREIRA F，et al. Prospective scenarios: a literature review on the Scopus database[J]. Futures，2018，100：20-33.

[40] CORDOVA-POZO K，ROUWETTE E A J A. Types of scenario planning and their effectiveness: a review of reviews[J]. Futures，2023，149：103153.

[41] 国家市场监督管理总局，国家标准化管理委员会. 智慧城市 人工智能技术应用场景分类指南：GB/Z 42759—2023[S]. 北京：中国标准出版社，2023.

[42] 梁战平. 情报学若干问题辨析[J]. 情报理论与实践，2003，26（3）：193-198.

[43] 谭晓，靳晓宏. 数智时代情报流程模型构建研究[J]. 情报理论与实践，2023，46（10）：32-39.

[44] FLORIDI L. Ethics, governance, and policies in artificial intelligence[M]. Cham：Springer International Publishing，2021.

[45] 蔚海燕，朱苇琦. 人工智能风险识别与治理路径构建[J]. 情报理论与实践，2026，49（1）：30-40.

[46] 人工智能治理蓝皮书（2024年）[EB/OL]. [2025-11-12]. http://www.caict.ac.cn/kxyj/qwfb/bps/202412/P020241227660032159191.pdf.

[47] 李汉卿. 协同治理理论探析[J]. 理论月刊，2014（1）：138-142.

[48] COOPER R G. Stage-gate systems: a new tool for managing new products[J]. Business horizons，1990，33（3）：44-55.

作者简介

郭骅，男，博士，副教授，硕士生导师，研究方向：智慧城市、人工智能治理，E-mail：guohua@hhu.edu.cn。
董怡彤，女，硕士研究生，研究方向：人工智能治理。
李辉，男，高级工程师，研究方向：大数据治理。

Risk Governance of AI-Enabled Smart City Information Services Based on Scenario Analysis

GUO Hua[1] DONG YiTong[1] LI Hui[2]
(1. Business School of Hohai University, Nanjing, 211100, P. R. China;
2. Nanjing Jiangbei New District Urban Digital Governance Center, Nanjing 211899, P. R. China)

Abstract: In order to identify the multiple risks of artificial intelligence application in information services, targeted governance strategies are proposed to help the high quality and sustainable development of smart cities. First, we use scenario analysis to construct application scenarios of artificial intelligence empowering smart city information services, and set scenario baselines. Then, we establish a risk framework and identify risk factors. Key drivers and uncertainties are also identified through environmental analysis. Finally, we analyze and propose governance principles and countermeasures for future scenarios. This paper depicts the composition and risk characteristics of application scenarios, constructs four future scenarios (inclusive good governance, technological islands, inclusive exploration, and disorderly crisis), and proposes a risk governance framework based on collaborative governance and stage-gate. This framework integrates differentiated domain priorities and scenario characteristics, balancing innovation development with risk prevention, thereby providing a theoretical reference for the construction of a safe and reliable smart city information service system.

Keywords: Smart City; Information Service; Artificial Intelligence; Scenario Analysis; Risk Governance

（责任编辑：常春）