引文格式：张永平,鄢嘉昕. 面向智慧城市的数据要素政策文本时空挖掘分析[J]. 地理空间信息, 2026, 24(6):99-105. doi:10. 3969/j.issn. 1672-4623.2026.06.018

面向智慧城市的数据要素政策文本时空挖掘分析

张永平[1,2]，鄢嘉昕[1,2*]
（1. 浙江大学公共管理学院，浙江 杭州 310000；2. 浙江大学包容智治与城乡融合实验室，浙江 杭州 310058）

摘　要：在智慧城市建设深入推进的背景下，我国围绕数据要素出台了大量政策，但政策供给的主题结构、时空演变规律及区域扩散特征缺乏系统性的实证分析。本研究以212份数据要素政策文件为分析语料，在对比LDA与BERTopic 2种主题模型的适用性后，选择LDA模型对政策文本进行深度挖掘。研究发现：政策主题涵盖8个核心领域，形成了“价值导向-制度保障-实践落地”的三层递进结构；格局从中央向地方的空间扩散呈现总体适配与梯度分化；并经历了“基础性布局-市场化建构-纵深化推进”的三阶段嬗变。基于上述发现，本研究提出了完善差异化区域政策支持体系、推动政策工具向场景培育转型等对策建议。

关键词：数据要素；政策分析；LDA主题模型；时空演变；文本挖掘
中图分类号：P208　　**文献标志码：**B　　**文章编号：**1672-4623（2026）06-0099-07

Spatio-temporal Mining Analysis of Data Element Policy Texts for Smart Cities

ZHANG Yongping[1,2], YAN Jiaxin[1,2]
(1. School of Public Affairs, Zhejiang University, Hangzhou 310000, China;
2. Lab of Inclusive & Smart Governance for Urban-Rural Integration, Zhejiang University, Hangzhou 310058, China)

Abstract: Against the backdrop of the ongoing advancement in smart city development, China has introduced numerous policies surrounding data elements, yet systematic empirical analyses remain lacking regarding their thematic structure, spatio-temporal evolution patterns, and regional diffusion characteristics. Taking 212 data element policy documents as the analysis data, we selected LDA model to conduct in-depth mining of policy texts, through comparing the applicability of LDA and BERTopic topic models. The findings reveal that the policy themes cover eight core areas, forming a three-tiered progression from value orientation to institutional safeguards and practical implementation. The spatial diffusion from central to local governments exhibits both overall alignment and gradient differentiation. The policy themes undergo a three-stage evolution, “foundational planning, market-oriented construction, deepened advancement”. Based on these insights, we proposed recommendations including refining differentiated regional policy support systems and shifting policy tools toward scenario-specific cultivation.

Key words: data element, policy analysis, LDA topic model, spatio-temporal evolution, text mining

智慧城市作为新型城镇化与数字中国战略交汇的核心载体，已成为中国城市治理现代化的主流方向[1-2]。随着新一代信息技术与城市治理场景的深度融合，智慧城市的建设重心正从基础设施铺设和单点应用试点，转向以数据驱动的整体协同与价值释放阶段。在这一过程中，数据已成为继土地、劳动力、资本、技术之后的第五大核心生产要素[3]，数据要素的价值化过程已成为驱动经济高质量发展的关键引擎[4]。智慧城市的“智慧”在很大程度上取决于城市运行所产生与汇聚的海量数据能否被有效确权、安全流通并高效配置。然而，数据作为一种新型生产要素，具有非竞争性、非排他性等特殊经济属性，其在智慧城市场景下的市场化配置面临传统要素未曾遇到的挑战[5]。目前，学术界与实业界在数据确权、定价机制及隐私保护等方面仍存在显著分歧，导致数据要素在从“资源”向“资产”跨越时存在较大的制度摩擦[6]。如何通过政策的干预来矫正市场失控、加速数据价值化进程，已成为各国政府探索的方向。

面对全球数字竞争的新格局，我国政府展现出了极具前瞻性的顶层设计能力。2019年，党的十九届四中全会首次将数据列为生产要素[7]；2020年，《中共中央国务院关于构建更加完善的要素市场化配置体制机制的意见》将数据正式确认为新型生产要素；随后，“数据二十条”、《“数据要素×”三年行动计划（2024—2026年）》等政策陆续出台，形成了具有生态特征的“政策群”[8]。然而，政策数量的激增并不等同于治理效能的提升，现有政策体系的主题结构、时空演变规律及区域扩散特征仍需深入研究。

第一作者简介：张永平（1987—），博士生导师，研究方向为智慧城市、城市大数据分析与数字城市治理，E-mail：yongpingzhang@zju.edu.cn。
通信作者：鄢嘉昕（2004—），硕士研究生，研究方向为智慧城市与城市治理，E-mail：3220103769@zju.edu.cn。

现有研究在政策分析方面已有丰富积累。例如，孙倬等[9]运用BERTopic主题模型分析发现，我国数据要素政策经历了“基础构建—制度完善—价值释放”的递进路径。苏婉等[10]采用DTM模型和LDA模型，从时间和空间2个维度对597份数据要素政策文本进行了系统分析，揭示了政策主题强度的3种演化类型。金燕等[11]揭示了“中央主导布局—地方差异化响应”的政策扩散特征。金冬雪等[12-13]采用PMC指数模型对政策文本进行了量化评价。然而，多数研究或专注于政策文本的单向解读，或侧重于宏观经济指标的回归分析，在分析方法上仍大量依赖单一模型，却很少对多种模型进行系统对比后择优使用，也缺乏在智慧城市语境下对政策从中央到地方空间传导路径的量化刻画。

鉴于此，本研究面向智慧城市建设对数据要素配置的现实需求，以212份数据要素政策文件为分析语料，构建LDA与BERTopic 2种主题模型的对比框架，择优选取最适配的模型方案，从主题分布、空间扩散和时间嬗变3个维度系统揭示我国数据要素政策的演变规律，并提出优化政策体系、赋能智慧城市治理的对策建议。

1 研究数据与方法

1.1 数据来源

本研究基于北大法宝数据库、国家相关部门及省级人民政府官网等多源渠道，采用多轮迭代的检索方式系统收集与数据要素相关的政策文本。截至2026年2月5日，共收集到与数据要素相关的广义政策文件899份。

然后对收集到的政策文件进行严格筛选。筛选条件包括3个层面：第一，政策文本必须现行有效；第二，限定政策类型为法律法规、条例、规定、意见、通知和办法，排除批复、答复、报告等非正式决策文件；第三，进行人工筛选，筛除全文命中次数小于5的低相关性文件以及事务性文件，仅保留全文重点内容中含有“数据要素”的政策文件。经上述筛选，最终确定212份政策作为本研究的分析语料，其中中央和地方层面的文件数量分别为46份和166份。

1.2 研究方法

1.2.1 文本预处理

由于单份政策文件篇幅较长且涵盖多个议题，本研究以段落作为主题建模的基本单元。按换行符对212份政策文件进行段落切分，经格式清洗、关键词筛选（保留含“数据”的段落）和长度过滤（保留不少于30字符的段落）后，最终得到7 922条段落文本。在分词环节，采用jieba分词工具的精确模式，并构建了包含170个标准化术语的领域专用词典（来源包括国家数据局术语定义、政策文本高频搭配及学术文献关键词）和包含127个条目的自定义停用词表（涵盖行政套话、连接功能词和指代泛化词三类），以提升分词的领域适应性和主题识别的准确率。

1.2.2 主题模型选择

考虑到单一模型可能存在局限性，本研究构建了LDA模型与BERTopic模型的对比框架[14-15]。LDA（latent dirichlet allocation）模型是一种基于概率生成过程的主题模型，假定每篇文档由多个主题按一定比例混合而成，每个主题由词语的概率分布来表征。BERTopic模型则利用预训练语言模型（本研究采用paraphrase-multilingual-MiniLM-L12-v2）将文本编码为高维语义向量，再经UMAP降维和HDBSCAN聚类自动识别主题。模型评估将从主题一致性、主题多样性和主题可解释性3个维度展开，择优选取最适配本研究语料特征的模型方案。

1.2.3 语义相似度计算与空间分析

为量化不同阶段或不同区域主题之间的关联强度，本研究综合采用余弦相似度和Jaccard相似度2种度量方式。在政策主题扩散分析中，利用全量政策语料训练Word2Vec词向量模型，将各主题的Top 20主题词通过权重加权平均转化为主题向量，进而计算中央主题与地区主题之间的余弦相似度，以量化中央政策向各地区传导的路径和强度。在跨阶段主题演化分析中，根据各主题的Top 15主题词构建词权重向量，通过余弦相似度量化相邻阶段主题对之间的语义关联。在空间格局刻画方面，按照国家统计局提出的经济地区划分办法，将地方政策分别归入东部、中部、西部和东北四大经济区域[16]，通过分区域主题建模和区域间对比，揭示政策供给的空间分异特征[17-19]。

2 数据要素政策分布特征

2.1 地区分布与区域格局

从整体空间格局来看，地方数据要素政策呈现显著的“东密西疏”分布特征，与我国数字经济发展水平的东西梯度基本吻合。如表1所示，东部地区是政策供给的核心高地，山东省以17份政策位居全国首位，广东、福建、江苏、浙江、北京也均处于较高水平。中部地区的安徽在中部六省中遥遥领先，可能与

其依托合肥综合性国家科学中心积极布局数据要素市场密切相关。西部地区内贵州以15份政策高居全国第二，与其作为首个国家大数据综合实验区的先发优势相关；四川也处于较高水平，而其余西部省份普遍偏低。东北地区则整体处于政策洼地，辽宁、吉林和黑龙江三省政策数量均极低。

表1　政策数量按地区的分布

经济区域	省份	政策数量/条	总计/条
东部地区	山东省	17	74
	广东省	12	
	江苏省	11	
	福建省	11	
	浙江省	9	
	北京市	7	
	海南省	4	
	上海市	2	
	天津市	1	
	河北省	0	
中部地区	安徽省	12	35
	河南省	9	
	湖北省	4	
	湖南省	4	
	山西省	3	
	江西省	3	
西部地区	贵州省	15	53
	四川省	11	
	广西壮族自治区	8	
	甘肃省	6	
	云南省	5	
	内蒙古自治区	2	
	重庆市	2	
	陕西省	2	
	宁夏回族自治区	1	
	青海省	1	
	新疆维吾尔自治区	0	
	西藏自治区	0	
东北地区	辽宁省	2	4
	吉林省	1	
	黑龙江省	1	

空间分布中还可以发现两个值得关注的特征：一是政策密度与国家级试点高度相关，贵州、山东、广东等政策高密度省份大多承担了国家级数据相关试点任务；二是存在“有省无策”现象，河北、西藏、新疆在数据要素政策方面尚属空白，面临制度供给缺位的风险。这一宏观格局为后续的政策主题识别和扩散动向分析提供了背景信息。

2.2　时间分布与阶段划分

以212份政策文件的发布时间为依据，结合数据要素政策发展的关键节点，本研究将数据要素政策划分为3个阶段：2016—2020年为萌芽期、2021—2023年为起步期、2024—2025年为发展期。阶段划分的主要依据是两个关键的政策节点：第一，2020年发布的《关于构建更加完善的要素市场化配置体制机制的意见》首次将数据与土地、劳动力、资本、技术并列，正式确立了数据作为新型生产要素的地位，构成了萌芽期与起步期的分界线；第二，2023年12月制定的《“数据要素×”三年行动计划（2024—2026年）》标志着政策从制度框架设计进入行动落地实施阶段，构成了起步期与发展期的分界线。

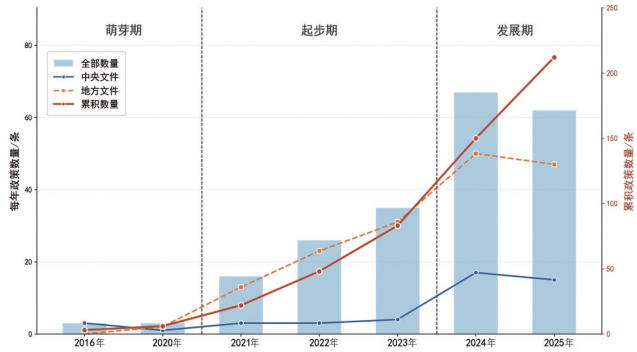


图1　各年度政策文件数量及累积趋势图

如图1所示，从年度数量来看，政策发布经历了从零星起步到稳步增长再到爆发式扩张的轨迹。最初，萌芽期年均仅约3份，反映出数据要素在这一阶段尚未成为政策议程的重点议题。起步期年均约25份，是萌芽期的约8倍，与《数据安全法》《个人信息保护法》等法律的密集出台以及“数据二十条”的发布有着较强的因果关联。发展期仅两年就产生约122份政策文件，年均超过60份，呈现爆发式增长。累积曲线呈现明显的指数型特征，表明政策体系正处于快速扩张的窗口期。从中央与地方的对比来看，地方政策在总量中的占比从萌芽期的约60%上升至发展期的约85%，反映出地方政府在数据要素市场化配置改革中的主动性正在增强。

3　数据要素政策主题识别

3.1　模型选择与对比

本研究构建了LDA与BERTopic 2个模型的对比框架，从量化指标和定性特征2个层面对2种模型进行了系统评估。首先，LDA模型在$K=5$至$K=20$的范围内搜索最优主题数，以U_Mass一致性为选优依据。U_Mass一致性曲线在$K=8$时达到峰值（−1.996 0），以全量7 917条语料训练最终模型，此时C_V一致性（Coherence Value）为0.512 3。而BERTopic模型经UMAP降维和HDBSCAN聚类后自动识别出了11个主题，同时产生33.4%的离群文档，C_V一致性为0.466 7，U_Mass一致性为−5.234 1。

从一致性指标来看，LDA在C_V和U_Mass两个维度上均优于BERTopic，表明LDA生成的主题内部语义连贯性更强。从主题多样性来看，BERTopic的得分（0.936 4）显著高于LDA（0.625 0），但其33.4%的离群率降低了模型对全部语料的覆盖能力。从主题可解释性来看，LDA的8个主题呈现出较为清晰的政策语义结构，各主题边界清晰、语义聚焦，便于人工归纳命名；BERTopic虽然行业细分粒度更细，但部分主题语义不够聚焦，增加了解读难度。综合评估后，本研究选择LDA模型（$K=8$）作为政策文本主题分析的主要工具。

3.2 主题识别与分布

基于LDA模型对7 917条政策段落进行主题建模后，共识别出8个核心主题，并进行人工命名。从整体词云来看，“数据要素”“数据安全”“公共数据”“数据资产”“数据交易”占据核心位置，如图2所示，表明政策体系已覆盖从数据资源开放到市场交易流通、从安全保障到产业落地的完整链条。

人工命名后的8个主题如表2所示，8个主题形成了三层递进结构，第一层次为战略导向与价值定位，“数据要素价值释放”是纲领性议题，集中体现了政策对数据要素战略地位的定位；第二层次为制度建设与治理保障，包含数据交易与资产管理、法规制度与授权运营、公共数据开放共享、数据安全与风险监管4个主题，共同构成数据要素市场化运行的制度框架；第三层次为实践落地与场景应用，包含行业场景应用、企业培育与试点推进、数字化应用与平台建设3个主题，体现了政策从宏观战略到微观实践的传导路径。



图2 数据要素政策主题词云图

表2 数据要素政策主题名称及主题词

主题	主题名称	Top15主题词
主题0	行业场景应用	改革, 农业, 健康, 发展, 贵阳, 医疗, 金融, 中国, 国家, 气象, 数据要素, 建设, 医保, 农村, 农产品
主题1	数据交易与资产管理	交易, 服务, 数据资产, 探索, 数据交易, 机制, 建立, 数据产品, 开展, 企业, 授权, 运营, 登记, 机构, 鼓励
主题2	法规制度与授权运营	应当, 运营, 机构, 登记, 授权, 进行, 规定, 服务, 部门, 提供, 使用, 或者, 管理, 依法, 其他
主题3	公共数据开放共享	公共数据, 政务, 部门, 公共数据资源, 开放, 管理, 共享, 工作, 目录, 公共, 政务数据, 建设, 单位, 建立, 机制
主题4	数据安全与风险监管	安全, 数据安全, 技术, 管理, 风险, 能力, 标准, 评估, 监管, 企业, 监测, 开展, 保护, 应急, 责任
主题5	数据要素价值释放	数据要素, 发展, 应用, 促进, 创新, 体系, 作用, 构建, 发挥, 提升, 融合, 价值, 产业, 治理, 建设
主题6	企业培育与试点推进	数据要素, 企业, 发展, 项目, 支持, 单位, 应用, 工作, 申报, 创新, 开展, 试点, 培育, 大数据产业, 组织
主题7	数字化应用与平台建设	应用, 建设, 服务, 企业, 平台, 领域, 创新, 支持, 数字, 打造, 产业, 提升, 发展, 智慧, 场景

4 政策主题的空间扩散

4.1 中央与各地区主题

按照中央、东部、中部、西部和东北5个维度分别进行LDA主题建模，确定中央、东部、中部和东北的最优主题数为$K=3$，西部为$K=4$。中央层面政策的主题词云和主题强度分布如图3所示。

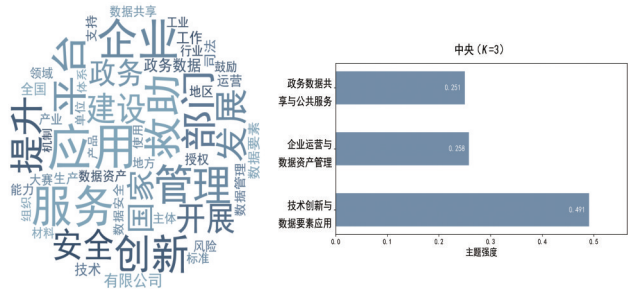


图3 中央政策主题词云和主题强度分布

如图4（地方政策主题强度分布）所示，各地区形成了各具特色的政策侧重。其中，东部地区体现出市场化先行的特征，“公共数据运营与资产管理”以0.503的强度占据主导地位，“数据资产”“运营”“登记”“评估”等市场化术语集中出现，反映了东部凭借成熟的市场经济基础在数据资产化方面走在前列。中部地区则呈现三足鼎立的均衡格局，“数据要素市场培育”“数据安全与政务开放”“公共数据授权运营”3个主题强度差距不大，发展与规范并重。西部地区是唯一产生4个主题的区域，议题分散度最高，“贵安”“贵阳”“大数据局”等地名和机构名称揭示了对个别先发省份的高度依赖，“算力”一词的涌现可能与东数西算工程相关。而东北地区呈现出了

鲜明的制度规范导向，“数据安全法规制度”下的主题词几乎全部是法律法规用语，与其他3个地区形成明显反差，反映出东北数据要素政策尚处于制度框架搭建阶段。

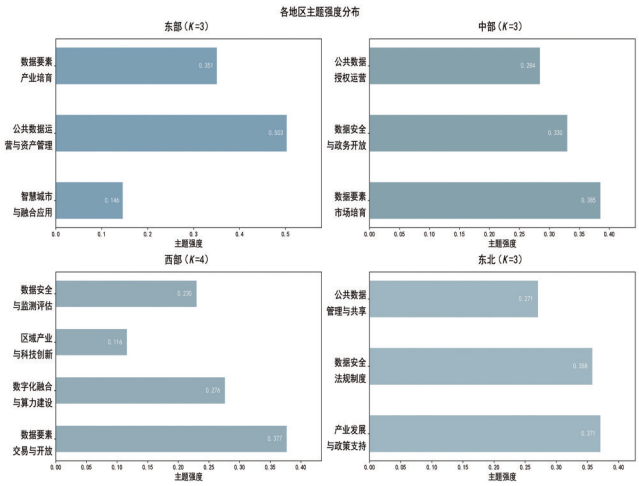


图4　地方政策主题强度分布

4.2　央地政策主题扩散

利用Word2Vec词向量模型量化中央与地区主题之间的关联强度，中央主题向各地区的扩散流向如图5所示。中央的“技术创新与数据要素应用”是扩散辐射面最广的主题，与4个地区的多数主题都形成了较高强度关联，其中，与西部“数字化融合与算力建设”相似度最高，与东部“智慧城市与融合应用”和中部“数据要素市场培育”也均在0.8以上。这表明技术应用和创新落地是中央政策向地方传导最为顺畅的议题领域。

中央的“企业运营与数据资产管理”是扩散选择性最强的主题。它与东部“公共数据运营与资产管理”和西部“数据要素交易与开放”形成了显著对应，但与东北3个主题的相似度普遍偏低。这种差异化扩散格局说明，市场化程度较高的政策议题更容易被经济基础和市场条件较好的地区所吸纳，而东北地区在这一议题上的政策承接明显不足。中央的“政务数据共享与公共服务”则与东北“公共数据管理与共享”形成最强对接，反映了东北政策重心仍集中在政务数据基础管理层面。

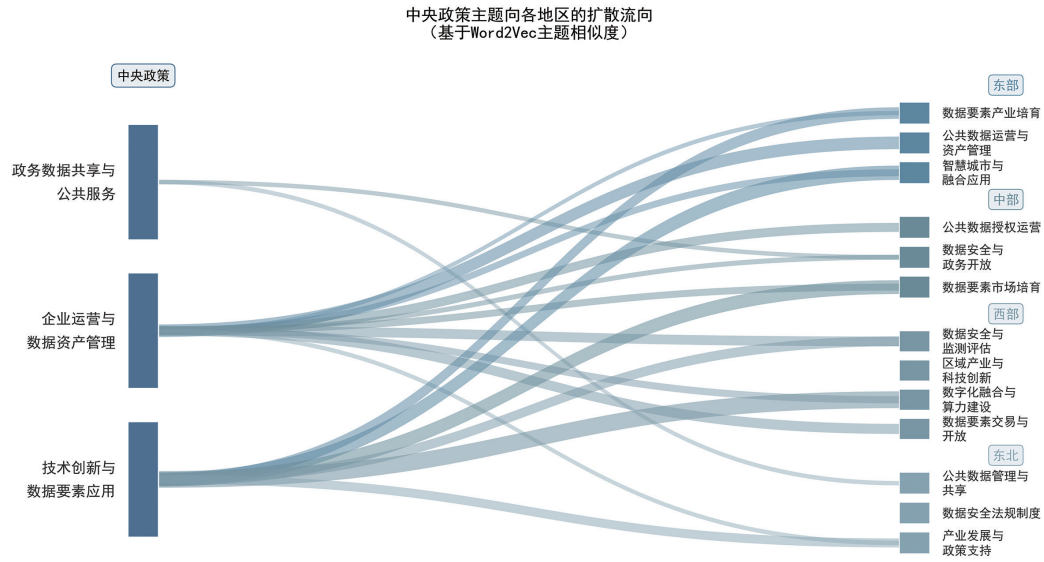


图5　中央政策主题向各地区的扩散流向图

综上所述，中央政策向地方的空间扩散呈现总体适配与梯度分化的动态格局，其中，东部地区作为先行承接者，在数据资产运营和市场化配置方面走在前列，其政策议题已从中央框架进一步深化为具有区域特色的实操方案。中部地区则在市场培育、安全治理和授权运营3个方向上同步推进，呈现均衡发展特征。西部地区的扩散路径带有资源领先锁定的特征，一方面依托“东数西算”在算力领域形成差异化优势，另一方面贵州等先发省份的实践为区域市场化进程提供了示范效应。东北地区处于制度规范导向型扩散阶段，与中央市场化导向主题的对接程度偏低。

5　政策主题的嬗变逻辑

5.1　各阶段主题内容

按照萌芽期、起步期和发展期3个阶段分别进行LDA主题建模，通过U_Mass一致性曲线确定各阶段最优主题数均为 $K=3$，即每个阶段可提炼出3个核心政策主题，各阶段的主题强度分布如图6所示。

萌芽期（2016—2020年）的3个主题为工业应用与市场探索、大数据产业与技术战略和数据中心与基础设施建设，围绕“工业应用”“产业战略”“基础设

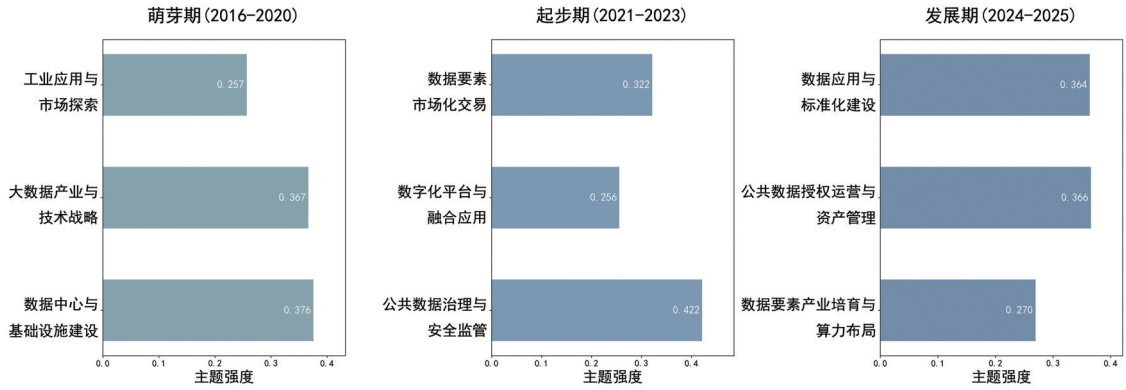


图6 各阶段主题强度分布

施”构成三角支撑格局。这一阶段，“应用”“发展”“技术”等偏宏观的表述出现频率最高，而“数据要素”一词尚未进入政策话语的核心位置，“大数据产业”“数据中心”才是反复被强调的专有名词。这意味着决策层对数据的认知基本停留在技术工具和产业方向层面，将数据从资源提升为要素的观念转变尚未完成。从强度分布来看，数据中心与基础设施建设的强度略高，说明硬件底座的铺设是这一时期政策投入最多精力的方向。

起步期（2021—2023年）的3个主题为数据要素市场化交易、数字化平台与融合应用和公共数据治理与安全监管。2020年《关于构建更加完善的要素市场化配置体制机制的意见》首次将数据列为新型生产要素后，“数据要素”“数据交易”迅速成为高频词汇。值得留意的是，公共数据治理与安全监管以0.422的强度占据3个主题之首，大幅领先于另外2个主题。这种治理导向主题占比大于市场导向主题的格局表明，政策制定者将制度规则的确立置于市场交易活跃度之前，选择了“先建规则、再育市场”的渐进式推进路径。

发展期（2024—2025年）的3个主题为数据应用与标准化建设、公共数据授权运营与资产管理和数据要素产业培育与算力布局，呈现出更为精细化的特征。“授权”“登记”“数据资产”等词汇的密集出现，标志着公共数据治理已从划定管理边界进入到授权运营和资产登记的实操层面。而“算力”“大数据局”“申报”等新词汇的出现，揭示了政策正向产业扶持和新型基础设施配套方向延伸，国家数据局的成立和东数西算工程的推进为这一趋势提供了制度支撑。

5.2 跨阶段主题演化

从萌芽期向起步期的过渡中，主题间流向呈现多源汇聚与新题涌现并存的特征。萌芽期的“工业应用与市场探索”同时向起步期的“数据要素市场化交易”和“数字化平台与融合应用”2个方向分流，早期较为笼统的工业数据应用实践由此被拆分为制度建设和技术落地两条并行路径。同时，萌芽期的“数据中心与基础设施建设”也与“数据要素市场化交易”保持了较高关联度，印证了基础设施的先期投入为后续市场化交易提供了物质条件。另一方面，萌芽期所有主题与起步期“公共数据治理与安全监管”之间均不存在显著关联，该主题伴随《数据安全法》《个人信息保护法》等法律法规密集出台而独立产生，体现了外部立法环境变化对政策议题设置的驱动作用。

从起步期向发展期的过渡中，最突出的演化路径是“数字化平台与融合应用”以最高的相似度向“数据应用与标准化建设”定向传导，表明政策重心从怎么建平台转向怎么立标准。而“数据要素市场化交易”向“数据要素产业培育与算力布局”的演化，揭示了市场交易制度的建立正在催生新的产业需求和基础设施配套需求。“公共数据治理与安全监管”向“公共数据授权运营与资产管理”的单向传导则描绘了公共数据领域从划定安全边界到推动资产价值变现的完整制度升级脉络。

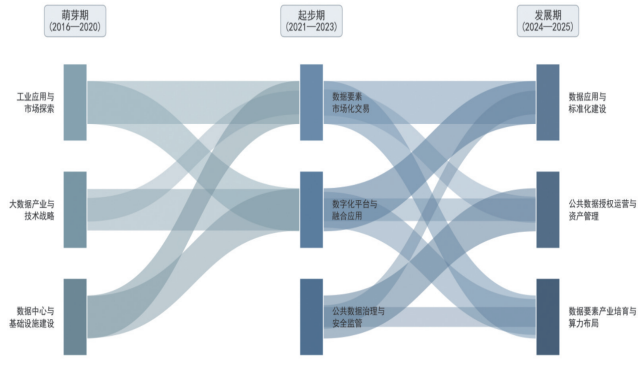


图7 政策主题演化流向图

总体而言，我国数据要素政策主题的嬗变遵循从

基础性布局到市场化建构再到纵深化推进的递进逻辑。萌芽期聚焦于大数据产业发展和数据中心等硬件设施铺设，对数据的定位尚停留在工具而非要素的认知层面。起步期转向市场化交易机制的构建、数字化融合应用的推进和公共数据治理框架的形成成为三大支柱，治理性主题强度显著高于市场性主题。发展期则聚焦于授权运营、资产登记、算力布局、标准化建设等精细化操作规程取代宏观方向性制度规范。公共数据从管好走向用好，市场交易从建规则走向育产业，应用场景从做示范走向定标准，构成了当前阶段政策供给的核心特征。

6 结　语

数据要素政策体系是支撑智慧城市从“建设阶段”迈向“治理阶段”的关键制度基础。本研究通过对212份政策文件的系统挖掘，回应了智慧城市语境下数据要素政策的基本问题。弥补了既有研究的不足。面向智慧城市数据要素治理的下一阶段优化，提出以下建议：第一，推动政策工具从“重规则制定”向“重场景培育”均衡转型。第二，构建差异化的区域政策支持体系。第三，建立政策效果的动态评估机制。第四，关注“有省无策”地区的制度供给缺位风险。

未来研究可进一步将政策文本挖掘结果与智慧城市建设的实际成效指标进行匹配分析，从政策到执行再到成效的全链条视角检验数据要素政策的治理效能，为构建更加敏捷、精准、协同的政策体系提供经验证据。

参考文献

[1] 张永平，张蔚文. 促进城市治理水平提升的智慧城市实践[J]. 人工智能，2021，5(5):94-1

[2] Zhang W, Zhang Y. Digital Urban Governance in China: Foundations and Applications[M]. Springer Nature，2026

[3] 中共中央 国务院关于构建更加完善的要素市场化配置体制机制的意见[EB/OL]. (2020-04-10)[2026-04-21]. 中华人民共和国国务院公报，https://www.gov.cn/gongbao/content/2020/content_5503537.htm

[4] 朱秀梅，林晓玥，王天东，等. 数据价值化：研究评述与展望[J]. 外国经济与管理，2023，45(12):3-17

[5] 王雪，夏义堃，裴雷. 国内外数据要素市场研究进展：系统性文献综述[J]. 图书情报知识，2023，40(6):117-128

[6] 鲍新中，李佳航，王子霖. 数据价值化：形态演进机制、主要障碍与推进策略[J]. 经济体制改革，2026，44(1):67-75

[7] 张庆国，薛德军. 人工智能、数据要素和人才队伍在新质生产力构建中的作用探析[J]. 江苏科技大学学报(社会科学版)，2025，25(2):33-41

[8] 颜赫隆，陈欣，王怡萱，等. 政策群视域下数据要素政策的群落结构建构与注意力差异演化研究[J]. 情报理论与实践，2025，39(12):83-93

[9] 孙倬，李嘉欣，曹玲静，等. 政策信息学视角下数据要素政策演变逻辑及发展趋向研究[J]. 情报理论与实践，2025，48(9):89-99

[10] 苏婉，于森，禚传阳. 我国数据要素政策主题演化与识别分析[J]. 图书情报工作，2024，68(16):90-103

[11] 金燕，陈登建，曹玲静. 中国数据要素政策主题的嬗变逻辑及扩散动向分析[J/OL]. (2025-06-17)[2026-04-21].现代情报，https://kns.cnki.net/KCMS/detail/detail.aspx?dbcode=CAPJ&dbname=CAPJLAST&filename=XDQB20250617002

[12] 金冬雪，赵建国. 基于PMC指数模型的我国省域数据要素发展政策量化与评价[J/OL]. (2025-09-08)[2026-04-12].情报科学，https://kns.cnki.net/KCMS/detail/detail.aspx?dbcode=CAPJ&dbname=CAPJLAST&filename=QBKX20250908001.

[13] Hu S, Wang X. A Text-mining-based Evaluation of Data Element Policies in China:Integrating the LDA and PMC Models in the Context of Green Development[J]. Sustainability，2025，17(15):6758

[14] Egger R, Yu J. A Topic Modeling Comparison Between LDA, NMF, Top2Vec, and BER Topic to Demystify Twitter Posts[J]. Frontiers in Sociology，2022，7:886498

[15] Nicolas C, Kim J, Chi S. Natural Language Processing-based Characterization of Top-down Communication in Smart Cities for Enhancing Citizen Alignment[J]. Sustainable Cities and Society，2021，66:102674

[16] 张兴华. 中华人民共和国2022年国民经济和社会发展统计公报_部门政务_中国政府网[EB/OL].(2023-02-28) [2026-04-21]. https://www.gov.cn/xinwen/2023-02/28/content_5743623.htm

[17] 梁文琼，王春晓，刘璐. 基于ArcGIS的人口重心和经济重心时空变化分析[J]. 地理空间信息，2022，20(2):67-70

[18] 胡璐嘉，郑新奇，郑劲颖. 基于空间句法的元明清北京城市功能时空演变分析[J]. 地理空间信息，2022，20(2):71-78

[19] 揭滢萤，陆妍玲，施钰琪，等. 融合轨迹特征的多要素时空大数据组织方法[J]. 地理空间信息，2026，24(3):7-11