

科技创新

“目标-工具-主题-治理域”四维分析视角 的科学数据管理政策文本量化研究

宋振豪, 田文迪, 汪源

(武汉纺织大学管理学院, 武汉 430200)

摘要: 科学数据管理政策在规范数据管理和促进科研创新方面具有重要意义。对现行政策进行量化分析有助于为未来相关政策的制定和改进提供参考。构建了一个涵盖政策目标、政策工具、政策主题和治理领域的四维分析框架,综合采用多种文本分析方法,对相关政策文本进行了单维度和多维度的交叉比较分析。结果表明,我国已建立了相对完善的科学数据管理政策体系,但仍有提升空间。未来可以进一步优化政策目标,细化政策内容,加强执行监督机制的设计,增强需求型政策工具的应用,并提升政策工具的协同性。

关键词: 科学数据管理政策; 政策优化; 四维分析视角; 政策工具

中图分类号: G322 **文献标志码:** A **文章编号:** 1671-1807(2025)04-0001-08

科学技术是第一生产力,在如今科学技术快速发展和科学研究如火如荼的情况下,科研活动成为推动经济高质量发展不可或缺的强大动力。科研活动中产生的科研数据是促进创新、支持决策、推动产业升级的重要资源。科学数据同时也包含国家敏感信息、科研机密和个人隐私等关键数据。科学数据应用场景和新兴技术更替,不同学科的特点和内容各不相同,会造成科学数据的性质、汇交保存、共享利用等具体管理方法有所差异,造成科学数据管理面临着严峻挑战。为提升科学数据管理的效能和整合度,中央及地方政府陆续出台《科学数据管理办法》《科学数据共享条例》《成都市科技数据管理实施办法》等相关政策,形成相对完善的政策体系,但目前尚缺乏系统、深入、全面的对我国科学数据管理政策的梳理和多视角的综合性分析,这不利于我国相关政策的进一步改善和优化,基于此聚焦“目标-工具-主题-治理域”四维分析视角的科学数据管理政策文本量化研究。研究以我国科学数据管理政策为样本,构建涵盖政策目标、政策工具、政策主题和治理域的四维分析框架,综合采用文本主题挖掘、科学数据生命周期模型和内容分析法对相关政策文本展开量化比较分析,探究相关

政策的优势和不足,并提出改进建议。

1 文献回顾

当前科学数据管理政策研究主要分为以下几个方面:①特定领域的政策调查,包括 Rousi 和 Laakso^[1]对期刊、出版社的科学数据管理政策进行了调查,薛秋红和徐慧芳^[2]探讨了科研机构的科学数据管理政策要素,付少雄等^[3]分析了高校的科学数据管理政策,Devriendt 等^[4]对科研资助机构的科学数据管理政策进行了调查研究。②政策内容分析,王芳和卜昊昊^[5]对不同国家科学数据管理政策进行对比研究;周文能等^[6]梳理欧盟、英、美等国家和地区科学数据管理政策,分析了科学数据管理与共享的政策演进及做法;Bjaalie 等^[7]讨论了欧盟科学数据管理和共享的变化细节,对比美国进行相关政策和方法的区分;都平等^[8]探讨了《科学数据管理办法》中数据域范畴与管理边界;温亮明等^[9]构建“词语-条款”共现矩阵,从《科学数据管理办法》的文本词语把握国家对于科学数据管理政策的政策导向。③政策量化评价,宋李叶^[10]构建三维分析框架对国内外科学数据管理政策进行量化评价;霍倩等^[11]对档案部门科学数据管理政策的相关政策开展了

收稿日期: 2024-08-19

基金项目: 国家社会科学基金项目一般项目(22BGL139);湖北省普通高等学校人文社会科学重点研究基地项目(DSS20200705);湖北省普通高等学校人文社会科学重点研究基地项目(DSS20170303);湖北省教育厅科学技术研究计划指导性项目(B2013203)

作者简介: 宋振豪(1999—),男,湖北黄陂人,硕士研究生,研究方向为科学数据管理;田文迪(1983—),女,湖北罗田人,博士,教授,研究方向为项目管理;汪源(1998—),女,湖北黄冈人,硕士研究生,研究方向为政策文本分析。

量化研究;司莉和刘莉^[12]基于政策工具视角,结合利益相关者和生命周期理论,构建三维分析框架,对科学数据管理政策的内容特征进行分析;王流芳等^[13]构建政策工具、主体、生命周期多维框架对国家层面发布的科学数据政策进行解构和量化分析。

总体来看,当前关于我国科学数据管理政策的研究往往局限于单一维度,缺乏对政策目标、工具、主题和数据生命周期视角等多维度的综合分析,导致无法全面揭示政策在科学数据生命周期各阶段的具体实施方式和治理焦点,这不利于各维度政策要素协同作用的发挥,也导致政策效果评估、监督和优化依据不充分。另外相关研究大多采用定性分析的方法,而针对科学数据管理政策的量化评价所用研究方法则相对单一,缺乏对文本挖掘、内容分析和计量分析等多种方法的融合应用。综上所述,有必要通过多视角跨维度的政策内容分析框架,全面评估我国的科学数据管理现行政策,并加强对多种量化分析方法的综合应用。鉴于此,本文综合应用主题挖掘、科学数据生命周期和内容分析法,基于政策工具、主题、目标、治理域、相关政策文本多个视角展开多维度量化比较分析,最终为政策优化提出针对性的建议。

2 “目标-工具-主题-治理域”四维分析框架构建

2.1 政策目标维度

在政策目标维度划分上,通过查阅相关政策资料,最终选择具有统领性的国家官方政策文件内容作为政策目标划分参考规范。《科学数据管理办法》中明确提及“进一步加强和规范科学数据管理,保障科学数据安全,提高开放共享水平,更好支撑国家科技创新、经济社会发展和国家安全”。据此,将科学数据管理政策目标总结为保障科学数据安全、提高共享开放水平、促进科技创新发展。

2.2 政策主题维度

为了进一步厘清科学数据管理政策的内容重点,进行 LDA(latent Dirichlet allocation)挖掘政策文本主题研究,分析可知在主题数量为 7 时,主题一

致性得分处于极大值拐点位置。如表 1 和图 1 所示,考察不同主题间关键词分布上的差异性,最终综合为 7 个政策核心主题,即“数据共享规范”“数据安全治理”“主体权责明晰”“数据出版引用”“数据汇交计划”“数据生产采集”和“数据中心建设”。

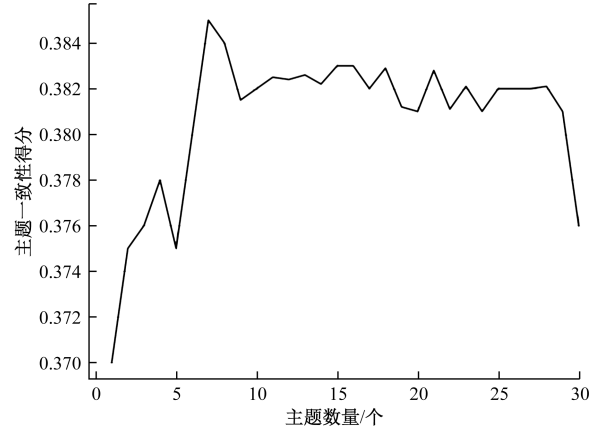


图 1 主题一致性得分

2.3 政策工具维度

根据政策工具分类思想,将科学数据管理政策工具分为供给型政策工具、环境型政策工具和需求型政策工具。

(1)供给型政策工具:是指政府为支持科学数据管理活动,围绕资源、技术、人才等要素,进行资金支持、平台建设、人才帮扶等一系列实质性活动,如表 2 所示。

(2)环境型政策工具:是指通过目标规划、法规管制、标准规范等方式为科学数据管理打造有力的政策环境,如表 3 所示。

(3)需求型政策工具:是指通过刺激科学数据管理各类需求的基本手段,为科学数据管理活动的持续发展注入强大动力,如表 4 所示。

2.4 治理域维度

治理域指科学数据生命周期各个阶段的管理活动,通过调研相关研究^[12],将关键治理域确定为规划与采集、汇交与保管、共享与利用、运行与维护、考核与监督 5 个阶段,如表 5 所示。

表 1 科学数据管理政策主题及核心关键词

主题编号	综合主题	核心关键词
综合主题(一)	数据共享规范	共享、开发、共享协议、交换、标准、平台
综合主题(二)	数据安全治理	保障、国防建设、公共安全、防灾减灾、安全审计、
综合主题(三)	主体权责划分	权利、义务、制度、明确、角色、分工、权限、责任
综合主题(四)	数据产业应用	知识产权、注明、出版、专利申请、使用者、应用、服务
综合主题(五)	数据汇交计划	汇交、审核、计划、申请、验收、采集、传输、汇总
综合主题(六)	数据生产调整	生产、采集、加工、整理、处理、转换、配置、整合
综合主题(七)	数据中心建设	数据中心、建立、主管部门、法人单位、存储、备份、设施

表 2 供给型政策工具说明

政策工具类型	子工具	工具定义描述	数据来源
供给型政策工具	基础设施	信息系统、科研平台等设施的建设	张霁等 ^[14] 、陈兰杰和赵元晨 ^[15]
	技术保障	信息处理技术、大数据和人工智能技术	卿倩和李桂华 ^[16]
	资金支持	直接投入资金、提供研发经费、纳入财政预算	陈兰杰和赵元晨 ^[15] 、Mons ^[17]
	人才队伍	政府通过完善培训体系、建设专业人才队伍	Bossaller 和 Million ^[18]
	公共服务	开展宣传、指导、咨询数据服务	Martin 等 ^[19] 、马续补等 ^[20]

表 3 环境型政策工具说明

政策工具类型	子工具	工具定义描述	数据来源
环境型政策工具	统筹规划	统筹协调、数据管理计划	陈兰杰和赵元晨 ^[15] 、马续补等 ^[20]
	法律法规	明确科学数据管理条例、实施细则	Marlina 等 ^[21]
	标准规范	编制数据目录、建立质量控制体系	Rafiq 和 Ameen ^[22]
	安全管理	数据保密管理、数据安全、容灾备份	李旭光等 ^[23]
	权责划分	政府明确不同主体的权利和义务	李旭光等 ^[23]
	策略性措施	激励、评价、考核、审查、监督、责任追究	张霁等 ^[14] 、卿倩和李桂华 ^[16]

表 4 需求型政策工具说明

政策工具类型	子工具	工具定义描述	数据来源
需求型政策工具	免费开放服务	编制数据资源共享目录、更新数据交换平台	邢文明等 ^[24] 、Devriendt 等 ^[25]
	增值开发与推广	通过分析挖掘和加工开发,引导鼓励开展创新应用和产业化推广	Mushi 等 ^[26]
	数据出版与引用	科学数据出版和传播工作、遵守知识产权规定	邢文明等 ^[24]
	交流合作	建立联盟、开展资源国际交流合作	司莉等 ^[12]
	政府采购	积极扶持龙头企业发展、优先采购国产数据服务	Bossaller 和 Million ^[18] 、司莉和刘莉 ^[12]

表 5 科学数据管理治理域

阶段编号	领域名称	关键活动
01	规划与采集	统筹规划、责任分工、收集范围、计划内容…
02	汇交与保存	汇交流程、质量控制、保存平台、保存期限…
03	共享与利用	共享范围、用户权限与责任、共享流程…
04	运行与维护	平台建设、培训宣传、应急管理、解释权…
05	考核与监督	违章处罚、表彰奖励、检查评估、资助扶持…

3 政策文本的选取和编码

3.1 政策文本的采集

我国系统形成有关科学数据管理的政策始于 2001 年的《气象资料共享管理办法》。以 2001 年为起点,首先对中国政府网、各部委门户网站、各地方政府及相关职能部门的官方网站开展政策检索工作;其次访问“北大法宝”政策信息专业检索平台;最后通过百度、谷歌学术等搜索引擎进行查漏补缺。检索表达式为“科学数据”OR“科研数据”OR“研究数据”OR[(“科研”OR“科学”OR“研究”)AND(“数据”OR“信息”)]。截至 2023 年 8 月,整理得到和研究主题高度相关的政策共计 63 份。语料具有较高权威性、可靠性、完整性和高相关性。

3.2 内容分析单元提取与编码

根据科学数据管理政策的文本特性,将分析单元确定为分割后的政策条款中单独的句。将遴选

出的分析单元依照“政策序号-章节条款-条款细分”格式进行编码,并对照构建的三维分析框架体系对政策工具进行两轮人工编码。通过二次编码检验进一步提高编码的可靠性,完成第二轮编码工作;最后复核二轮编码结果,确定最终编码内容。

3.3 政策文本的单维度分析

3.3.1 政策目标

根据前文目标维度划分结果,科学数据管理政策的政策目标总结为保障科学数据安全、提高共享开放水平和促进科技创新发展,提取出的政策分析单元共产生有效文本编码节点 2 376 个,如图 2 所示。

“提高共享开放水平”作为政策目标占据明显优势,它能增加数据的透明度和可验证性,从而提高数据质量和可信度,同时促进数据再利用,提高科学活动效率。“保障科学数据安全”是科学数据管理的根本目标,直接关系数据的可靠性和可信度,并构成数据合作与共享的基础。然而,对促进科技创新发展的关注明显偏低。尽管我国已开始重视科学数据的创新发展,但仍面临数据技术、标准及共享意识不足等问题。

3.3.2 政策工具

对收集的政策进行编码处理,最终整理得到 2 376 条信息,归类结果如表 6 所示。

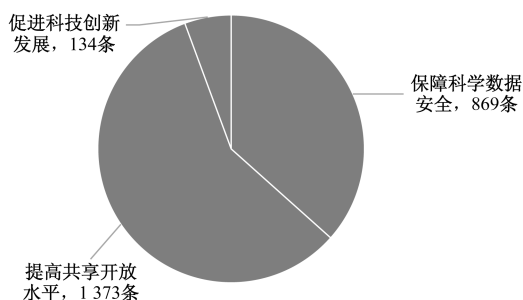


图 2 政策目标比例统计

从结果看出,环境型政策工具使用总数为 1 374 条(占比为 58%),供给型和需求型政策工具使用总数分别为 632 条(占比为 27%)和 370 条(占比为 15%),需求型政策工具使用最少。如图 3 所示,环境型政策工具占据支配地位,体现科学数据管理政策受到与外部环境相关的计划、法律、规范等因素影响较大。同时,这三种类型的政策工具使用差距较大,说明我国科学数据管理政策在工具选择上存在问题,政策工具单一使用为主,未形成政策工具间的有效协同配合。

3.3.3 政策主题

政策主题能表示具体政策内容和政策关注点,利用 LDA 主题模型构建了科学数据管理政策文本主题挖掘模型,得到了 7 个文本内容潜在主题。

在具体分布中,“主体权责划分”主题占比突出,共计 1 037 条(占比为 43.6%)。我国科学数据管理活动处在起步阶段,相应配套政策和措施正在不断完善,“主体责任划分”有助于提高政策执行效率,还可以防止权力滥用。科学数据中心建设和发展(总计为 215 条,占比为 9%)对政策主体的数据管理能力以及基本设施建设具有较高要求,因此真正能够承担数据中心建设任务的责任主体相对有限,反映在政策制定上相关内容也会相对较少。与此同时,我国科学数据管理政策关于科学数据产业的应用规范总计有 165 条(占比为 6.9%)。但是多数政策并未详细提及相关数据使用者应遵守何种出版和引用标准。

3.3.4 治理领域

我国科学数据管理政策的治理领域编码分析主要包括涵盖规划与采集、汇交与保管、共享与利用、运行与维护、考核与监督这五大环节。

如图 4 所示,整体来看,我国科学数据管理政策能够统筹兼顾多种治理领域。其中,规划与采集领域编码节点数最多,共计 897 条,占比为 37.7%。

表 6 科学数据管理政策工具编码汇总

政策工具	具体分类	小计/条	占比/%	总计/条	占比/%
供给型	基础设施	195	31	632	27
	技术保障	101	16		
	资金支持	60	10		
	人才队伍	109	17		
	公共服务	167	26%		
环境型	统筹规划	174	13	1 374	58
	法律法规	296	22		
	标准规范	153	11		
	安全管理	281	20		
	权责划分	200	15		
	策略性措施	270	19		
需求型	免费开放服务	135	36	370	15
	增值开发与推广	77	21		
	数据出版与引用	96	26		
	交流合作	36	10		
	政府采购	26	7		

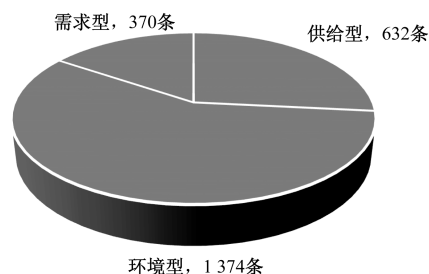


图 3 政策工具比例统计

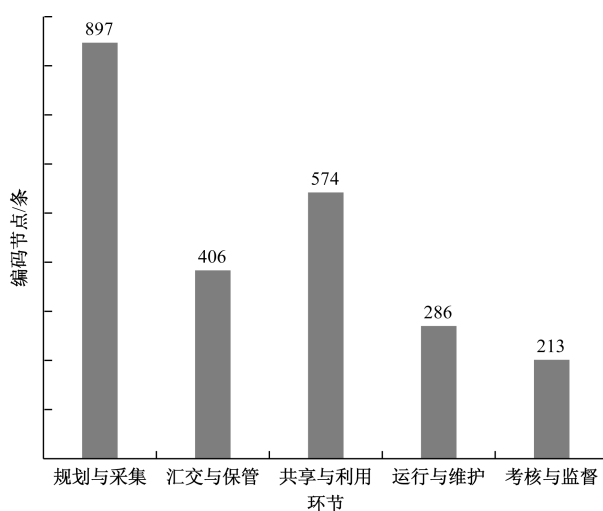


图 4 治理域分布统计

科学数据的规划和采集环节是推动后续活动顺利开展的基础,详细、全面地统筹规划、权责部署,能避免产生权责界定不明、采集条件受限等问题。运行与维护、考核与监督两个环节所存在的编码节点明显较少,仅占 12%和 9%。

基于科学数据全生命周期的视角,这些环节都

不应该被忽视和放弃,具有极大风险挑战的安全与管理问题,比如外部风险入侵、前端驱动力不足、运行效率低下,需要加大科学数据的运行维护、考核监督环节政策工具的使用力度以保证科学数据管理体系正常运行。

4 政策文本的多维度交叉分析

4.1 “政策工具-政策目标”二维交叉分析

为确定政策工具是否能够有效地实现政策目标,了解不同政策工具对不同政策目标的影响,评估选择政策工具来实现特定的政策目标,从而提高效果和效率。可进行政策工具与政策目标的二维交叉分析,其结果如图5所示。

政策工具与政策目标的交叉分析可知,不同政策目标下政策工具结构分布存在失衡。

在保障科学数据安全目标中,环境型政策工具较于其他工具使用频率差距较大;在促进科技创新发展目标中,供给型政策工具的使用频率最高,而环境型和需求型政策工具的使用相对较少。政策工具的平衡使用对政策目标的均衡发展至关重要,但目前主要偏向于使用环境型政策工具,整体呈现分布不均衡状态。需求型政策工具在促进科技创新发展目标中的使用最少,显示在促进科学数据创新发展方面社会参与和多元主体的引导力度有待进一步提升。

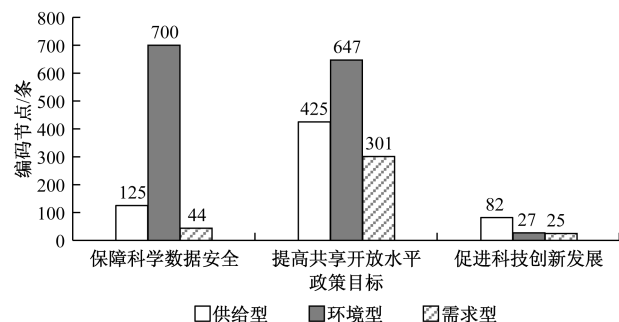


图5 “政策工具-政策目标”二维分布统计

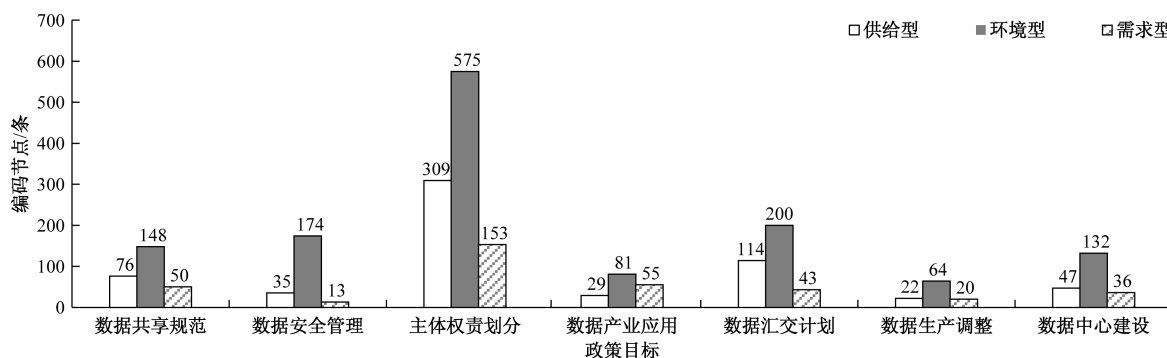


图6 “政策工具-政策主题”二维分布统计

4.2 “政策工具-政策主题”二维交叉分析

政策工具与政策主题两个维度交叉量化分析能够体现政策制定者的管理意图如图6所示。

政策主题上的政策工具存在分散不均问题,需完善政策体系内容,优化内容分布。主体权责划分使用政策工具占比最高,其政策工具使用占比高达43%。环境型政策工具在这两类主题中应用较多,可见相关部门希望通过外部作用的方式来保证汇交和共享管理的长期稳定性和综合协调性,加强对科学数据管理工作效果的监测评估。政策工具在各类主题中的应用还存在较大差异。需求型政策工具在数据安全主题中使用数量最低,仅占0.5%;环境型政策工具在主体权责划分主题中使用数量最多,占比高达24.2%,两者相差23.7%。供给型政策工具在数据产业应用主题和数据生产调整主题中使用极少,而环境型政策工具在工具与主题交叉维度中仍占相当大的优势,这也说明当前政策体系初步建成,工具分布结构还不太合理、应用差异较大。

4.3 “政策工具-政策治理域”二维交叉分析

政策工具与治理域的二维交叉分析可以更具针对性和科学性地剖析政策工具的应用情况,如图7所示,从整体上看,三类政策工具在科学数据生命周期不同阶段内的总体使用分布均呈现不均衡状态。供给型和环境型政策工具均主要运用在规划与采集阶段,而需求型政策工具则主要运用在共享与利用阶段。

环境政策工具在规划与采集领域内使用数量最多,共计541条,占比为23%。这主要是由于环境型政策工具相关的统筹规划、政策法规、标准规范多在科学数据规划和采集阶段部署和制定。值得注意的是,需求型政策工具在运行与维护领域和考核与监督领域内却极少被使用,二者之和仅占比为1.5%,使用分布上极度不均衡,缺乏必要的维护和监督手段。

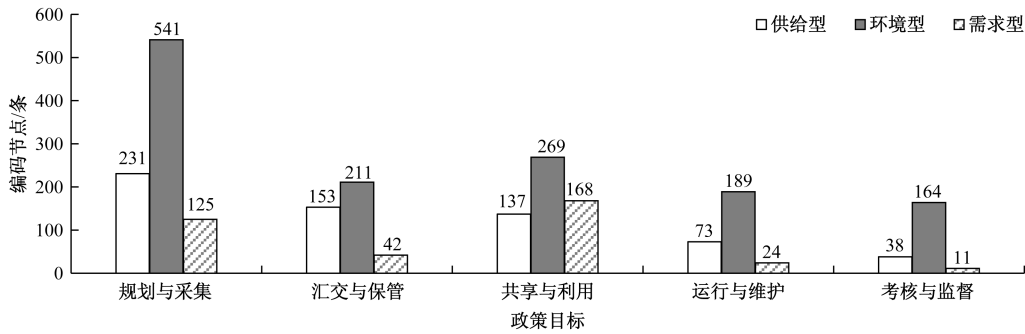


图 7 “政策工具-治理领域”二维分布统计

4.4 “目标-工具-主题-治理域”四维交叉分析

为整体把握科学数据管理政策的总体内容分布和使用偏好,并识别现有政策体系中可能存在的问题与不足之处,使用 MATLAB 软件综合进行四维交叉分析结果如图 8 所示。

环境型政策工具中,“政策目标 2(提高开放共享水平)-政策主题 3(主体权责划分)-治理领域 1(规划与采集)”交叉维度的工具最多,共 195 个;其次是“政策目标 1(保障科学数据安全)-政策主题 3(主体权责划分)-治理领域 1(规划和采集)”,有 146 个工具。环境型政策工具通常通过外部影响促进科学数据管理活动,与有效引导责任主体行为和治理规划阶段相契合,显示我国科学管理政策具有适配性。

供给型政策工具中,“政策目标 2(提高开放共享水平)-政策主题 3(主体权责划分)-治理领域 1(规划与采集)”交叉维度工具最多,共 110 个;“政策目标 2(提高开放共享水平)-政策主题 5(数据汇交计划)-治理领域 1(规划与采集)”交叉维度也有 76 个工具。供给型政策工具主要支持和保障科学数据管理活动,特别重视数据汇交,需要大量投入确保数据统一汇编和存储,以支持后续科研活动和数据再利用。

需求型政策工具中,“政策目标 2(提高开放共享水平)-政策主题 3(主体权责划分)-治理领域 1(规划与采集)”交叉维度的工具最多,共 71 个。然而,多个交叉维度存在使用空白,且使用力度较低,显示需求型政策工具在政策组合中的结构失衡,可能阻碍科学数据管理活动的有效进行。

总体而言,三类政策工具在“政策目标 2(提高开放共享水平)-政策主题 3(主体权责划分)-治理领域 1(规划与采集)”交叉维度使用最为频繁,突显当前我国科学数据管理政策的治理重心。未来应加强

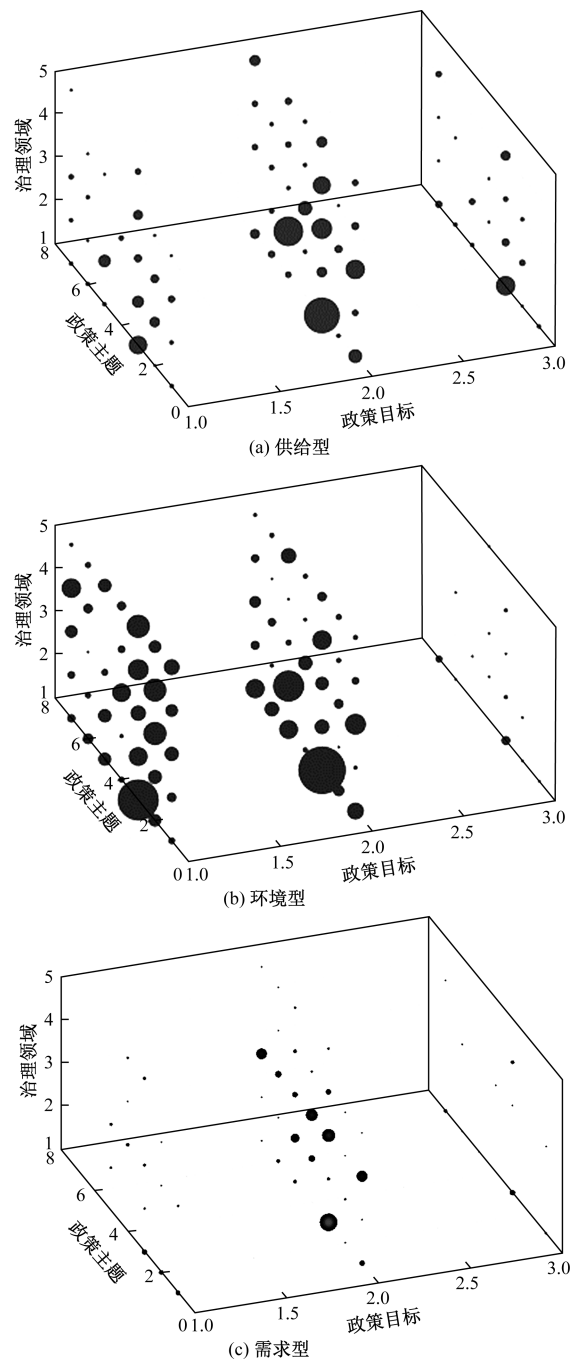


图 8 四维交叉分布

对需求型政策工具的关注,推动政策组合协同作用,优化科学数据管理的政策配置,以支持科技创新能力的提升和国家数字科学竞争力的增强。

5 结论与建议

5.1 因地制宜,细化和调整政策目标以增强灵活性与适应性

虽然我国科学数据管理政策体系的总体目标聚焦于数据安全、共享开放和创新发展,但不同地区和不同行业也具有其需求偏差。因此,为适应各地区和多行业的实际情况,根据其自身特点和特定需求制定方向具体、内容详细的政策目标,这样既可以发挥政策灵活性和适应性,又可以充分发挥其能动性,及时根据现实需求调整政策目标导向。

5.2 完善政策分类与细化,增强政策明确性与执行力

通过对政策内容进行分类细化,可以使政策更具体、清晰,减少歧义和漏洞,有利于政策的有效执行和落实。并且,将政策内容按照不同主题进行细化,也有助于发现政策之间的关联和差异,促进政策内容不断调整和优化,推动科学数据管理政策体系的持续完善和发展。

5.3 加强对运行维护与考核监督领域的政策关注

基于治理领域情境,科学数据管理政策较为注重规划采集和共享利用,而对于运行维护、考核监督领域关注较少。因此,注重科学数据全生命周期管理,未来加强有关运行与维护、考核与监督领域,重视科学数据建设的后期管理。

在运行与维护领域,加强投入基础设施建设、人才队伍支持、财政资金帮扶等措施的保障力度。例如,实现重要科学数据异地备份;各数据管理中心配备应急管理系统;分拨专项经费。在考核与监督领域,可以参考相关较为成熟的指标评估模型,并咨询领域内专业人员,建立标准、可行的评价考核机制。

5.4 加强需求型政策工具的应用,注重政策工具的协同效应

通过工具-目标交叉分析发现,当前不同目标下政策工具分布不均。例如,在保障科学数据安全方面,环境型政策工具使用频繁,而供给型和需求型工具较少。建议调整目标导向,以提升政策效果和效率。在工具-主题领域,需求型政策工具应用较少,表明社会参与和多元主体引导力度不足。在工具-治理领域,需求和供给工具主要用于共享与利用阶段,建议在数据生命周期后期加强需求型工具。

总体来看,未来应进一步优化科学数据管理政策配置,发挥政策工具的协同作用,加强对需求型政策工具的关注。

6 总结和展望

科学数据管理政策在规范数据管理和推动科研创新方面具有重要意义。本文通过文本内容量化分析,对我国相关政策进行了比较评估。研究结果表明,我国已建立了相对完善的科学数据管理政策体系,但仍有改进空间。为此,提出了相应的优化建议。未来将进一步研究政策的执行效果,以促进政策的持续改进。

参考文献

- [1] ROUSI A M, LAAKSO M. Journal research data sharing policies: a study of highly-cited journals in neuroscience, physics, and operations research [J]. *Scientometrics*, 2020, 124(1): 131-152.
- [2] 薛秋红, 徐慧芳. 西方国家科研机构科学数据管理政策要素研究[J]. *情报理论与实践*, 2021, 44(7): 191-196.
- [3] 付少雄, 滕雨彤, 韩悦, 等. 跨区域视角下高校科学数据管理政策的对比研究及启示[J]. *图书馆杂志*, 2023, 42(11): 48-61.
- [4] DEVRIENDT T, SHABANI M, BORRY P. Policies to regulate data sharing of cohorts via data infrastructures: an interview study with funding agencies[J]. *International Journal of Medical Informatics*, 2022, 168: 104900.
- [5] 王芳, 卜昊昊. 科学数据管理政策发展比较研究 [J]. *中国图书馆学报*, 2022, 48(6): 77-96.
- [6] 周文能, 刘云, 王刚波. 国内外科学数据管理与共享政策分析及对国家自然科学基金的启示[J]. *中国科学基金*, 2023, 37(1): 150-160.
- [7] BJAALIE J G, GOBLE C, SANSONE S A, et al. Perspectives on data sharing and the New NIH policy from the European Union[J]. *Havard Data Science Review: A Microscopic, Telescopic and Kaleidoscopic View of Data-science*, 2022, 4(3): 1-9.
- [8] 都平平, 李雨珂, 张雪媛. 我国《科学数据管理办法》中概念视角数据域范畴与管理边界研究[J]. *图书馆杂志*, 2022, 41(4): 96-105.
- [9] 温亮明, 李洋, 张丽丽, 等. 《科学数据管理办法》语词分析[J]. *新世纪图书馆*, 2020(5): 81-89.
- [10] 宋李叶. 政策工具视角下的国内外科学数据管理政策文本量化分析[D]. 南昌: 南昌大学, 2022.
- [11] 霍倩, 梁宵萌, 潘亚男. 科学数据管理政策对档案部门科学数据管理的启示——基于政策工具的文本量化分析[J]. *图书情报工作*, 2021, 65(20): 23-30.
- [12] 司莉, 刘莉. 我国科学数据管理政策: 内容结构及优化策略——基于政策文本内容的量化分析[J]. *情报理论与实践*, 2023, 46(9): 16-23.
- [13] 王流芳, 荣注瑶, 贾晓峰, 等. 基于三维框架的我国科

- 学数据政策文本内容分析[J]. 知识管理论坛, 2024, 9(2): 177-194.
- [14] 张霁, 周晓英, 裴俊良, 于小宁. 我国突发公共卫生事件应急信息管理政策工具的量化分析[J]. 情报资料工作, 2021, 42(3): 44-51.
- [15] 陈兰杰, 赵元晨. 政策工具视角下我国开放政府数据政策文本分析[J]. 情报资料工作, 2020, 41(6): 46-53.
- [16] 卿倩, 李桂华. 政策工具视野下我国省级全民阅读政策研究[J]. 中国图书馆学报, 2020, 46(3): 88-101.
- [17] MONS B. Invest 5% of research funds in ensuring data are reusable[J]. *Nature*, 2020, 578: 491-491.
- [18] BOSSALLER J, MILLION A J. The research data life cycle, legacy data, and dilemmas in research data management[J]. *Journal of the Association for Information Science and Technology*, 2023, 74(6): 701-706.
- [19] MARTIN M R, HERNÁNDEZ P T, MARTÍNEZ C S. Research data services (RDS) in Spanish academic libraries[J]. *The Journal of Academic Librarianship*, 2023, 49(4): 102732.
- [20] 马续补, 吕肖娟, 秦春秀, 等. 政策工具视角下我国公共信息资源开放政策量化分析[J]. 情报理论与实践, 2019, 42(5): 46-50.
- [21] MARLINA E, HIDAYANTO A N, PURWANDARI B. A qualitative exploration of challenges and factors in research data management across Indonesian research institutes[J]. *Journal of Librarianship and Information Science*, 2024, 56(3): 1-13.
- [22] RAFIQ M, AMEEN K. Research data management and sharing awareness, attitude and behavior of academic researchers[J]. *Information Development*, 2022, 38(3): 391-405.
- [23] 李旭光, 王丹, 肖斯佳, 等. 交叉学科科学数据管理政策内容分析及启示——以地理学为例[J]. 情报杂志, 2022, 41(1): 139-148.
- [24] 邢文明, 洪芳林, 李晓妍. 科学数据管理体系的二维视角——《科学数据管理办法》解读[J]. 图书情报工作, 2019, 63(23): 30-37.
- [25] DEVRIENDT T, SHABANI M, LEKADIRK, et al. Data sharing platforms: instruments to inform and shape science policy on data sharing? [J]. *Scientometrics*, 2022, 127(6): 3007-3019.
- [26] MUSHI G E, PIENAAR H, VAN DEVENTER M. Identifying and implementing relevant research data management services for the library at the University of Dodoma, Tanzania[J]. *Data Science Journal*, 2020, 19: 1-9.
- [27] 谈洁. 基于数据生命周期的科学数据管理高端交流平台调研与构建探究[J]. 图书馆, 2024(6): 89-97.

Quantitative Research on Scientific Data Management Policy Texts from the Perspective of Objective-Tool-Subject-Governance Domain

SONG Zhenhao, TIAN Wendi, WANG Yuan

(School of Management, Wuhan Textile University, Wuhan 430200, China)

Abstract: Scientific data management policy plays an important role in standardizing data management and promoting scientific research innovation. The quantitative analysis of the current policy is helpful to provide reference for the formulation and improvement of relevant policies in the future. A four-dimensional analysis framework covering policy objectives, policy tools, policy themes and governance fields was constructed, and a variety of text analysis methods were comprehensively adopted to carry out a cross-dimensional and multi-dimensional comparative analysis of relevant policy texts. The results show that China has established a relatively perfect scientific data management policy system, but there is still room for improvement. In the future, policy objectives can be further optimized, policy contents can be refined, the design of implementation supervision mechanism can be strengthened, the application of demand-oriented policy tools and the synergy of policy tools can be enhanced.

Keywords: scientific data management policy; policy optimization; four-dimensional analysis perspective; policy tools