



开放科学
(资源服务)
标识码
(OSID)

科学数据管理政策知识图谱构建与应用研究

支凤稳¹ 韩梦娇¹ 陈依萌¹ 郭诗悦¹ 郑彦宁²

1. 河北大学管理学院 保定 071002;
2. 中国科学技术信息研究所 北京 100038

摘要: [目的/意义] 汇总和梳理我国科学数据管理政策内容, 研究政策重点, 探究科学数据管理政策的现状。[方法/过程] 以我国科学数据管理政策为研究对象, 对政策进行可视化梳理, 基于 Neo4j 工具, 以可视化形式分析政策内容结构, 识别政策实施重点。[结论/结果] 研究构建的政策知识图谱, 共包含 844 个节点与 1254 条边。研究发现当前科学数据管理政策注重于数据管理规范、数据安全与隐私保障、数据开放共享, 通过对政策的可视化与聚类分析, 结合上述内容提出管理建议。

关键词: 科学数据管理; 政策内容; 知识图谱

中图分类号: G353.1

Research on the Construction and Application of Scientific Data Management Policy Knowledge Graph

ZHI Fengwen¹ HAN Mengjiao¹ CHEN Yimeng¹ GUO Shiyue¹ ZHENG Yanning²

1. School of Management, Hebei University, Baoding 071002, China
2. Institute of Scientific and Technical Information of China, Beijing 100038, China

Abstract: [Objective/Significance] Summarize and sort out the content of China's scientific data management policies, study policy priorities, and explore the current status of scientific data management policies. [Methods/Processes] Taking China's scientific data management policies as the research object, this study visualizes and sorts out the policies. Based on the Neo4j tool, the content structure of the policies is analyzed in a visual form to identify the key implementation points. [Results/Conclusions] The policy knowledge graph constructed in this study includes 844 nodes and 1254 edges. Currently, scientific data management policies focus on data management standards, data security and privacy protection, and data openness and sharing. Through visualization and clustering analysis of policies, suggestions are proposed based on the above content.

Keywords: Scientific Data Management; Policy Content; Knowledge Graph

基金项目 国家社会科学基金一般项目“基于科学数据关联的科研合作网络研究”(24BTQ041)。

作者简介 支凤稳(1987-), 博士, 教授, 主要研究方向为科学数据、竞争情报; 韩梦娇(1999-), 硕士研究生, 主要研究方向为科学数据; 陈依萌(2003-), 本科生, 主要研究方向为科学数据; 郭诗悦(2004-), 本科生, 主要研究方向为档案大数据; 郑彦宁(1965-), 通信作者, 博士, 研究馆员, 主要研究方向为情报学理论与方法, E-mail: ynzheng@istic.an.cn。

引用格式 支凤稳, 韩梦娇, 陈依萌, 等. 科学数据管理政策知识图谱构建与应用研究[J]. 情报工程, 2025, 11(6): 14-27.

引言

科学数据主要包括在自然科学、工程技术科学等领域通过基础研究、应用研究、试验开发等产生的数据,以及通过观测监测、考察调查、检验检测等方式取得并用于科学研究活动的原始数据及其衍生数据^[1]。随着科学数据的快速增长和多样化,科学数据管理政策的制定和执行变得日益重要。2018年,国务院办公厅出台《科学数据管理办法》,明确科学数据管理工作的原则,为科学数据管理与共享提供了宏观指导^[2]。科学合理的政策体系是科学数据管理工作不断推进与深化的重要保障^[3]。

知识图谱技术作为一种结构化知识表示方法,为政策文本分析提供了创新路径,其核心价值体现在三个方面:(1)通过多维度语义解析深度挖掘政策内容;(2)运用本体建模技术实现政策要素的细粒度拆解与标准化表达^[4];(3)依托图谱化的组织方式,对数据要素之间的关联结构进行整体刻画,来支撑内容关联分析与结构性理解^[5]。利用知识图谱构建政策要素的语义关联图谱,既能够清晰展现政策内部的逻辑链条,又能反映跨政策文件的要素映射关系,有助于把握科学数据管理的整体方向和战略部署,了解科学数据管理活动的目标和重心,揭示政策之间的关联;也能更好识别政策重点和内容结构,提升政策分析能力。

1 研究综述

随着科学数据管理政策的发展与完善,国内学者关于科学数据管理政策研究逐渐增多,主要从两个方向展开。一是对我国科学数据管

理政策文本的分析研究。有学者从政策演进视角出发,对我国科学数据管理政策的发展过程进行梳理,并将其划分为起步、发展、转型与创新四个阶段^[6]。部分学者对我国科学数据管理政策的实施效果与结构特征进行了评价,认为我国相关政策整体呈现出积极演化趋势,但在政策目标与政策工具协同性、政策目标连续性与创新性以及政策工具配置均衡性等方面仍有改进空间^[3,7]。也有研究从开放科学视角出发,指出为保障科学数据的健康有序发展,有必要构建系统化的开放科学政策体系,并进一步明确政府、科研资助机构及教育科研机构在科学数据管理中的责任分工^[8]。有学者发现我国相关政策制定进程不断加快,但在多主体协同机制、政策细化程度及协同实施等方面仍存在不足,并据此提出了相应的完善建议^[2]。二是对国外科学数据管理政策的研究。学者们对欧盟^[9-10]、美国^[10-11]、英国^[11-12]、澳大利亚^[11,13]等国家的科学数据管理政策进行研究,调查政策不同的实施机构^[14-15];或结合政策工具等研究方法分析政策文本内容^[9,15];或通过对比其他国家,对政策进行分析,如对中美科学数据管理政策进行量化评价^[15]和对比分析^[16]。这些研究都在分析各国数据开放管理政策特点,为我国政策完善提出建议。

知识图谱作为一种融合符号逻辑、语义网络及人工智能等多学科优势的知识表示与推理框架^[17],已被广泛应用于政策领域,用于对政策元数据及文本内容进行结构化表示与语义化组织。在隐私政策^[18]、大数据政策^[19]、科技政策^[20]、产业政策^[21]等领域中,知识图谱技术以强大的语义解析、实体识别与关系抽取能力,

构建政策多维要素的语义关联网络模型，通过可视化手段展示非结构化政策文本数据。随着自然语言处理技术与深度学习模型等的迅速发展，知识图谱的政策研究偏向细粒度知识层面，以此挖掘并提取政策中的核心要素与关键信息^[4]，通过统计机器学习或深度学习方法^[22]将政策文本解构，实现政策情感与政策工具分类，以新技术为政策分析、评估与决策提供结构化、语义化的知识支持。

目前，科学数据管理政策研究多集中于文本量化分析，虽能统计与解读政策文本，但缺乏多维知识体系的构建，难以全面揭示政策之间的逻辑关联。而知识图谱在政策领域的应用广泛，却没有形成针对科学数据管理政策特点的顶层设计框架。鉴于科学数据管理政策具有多范围、多学科、实施动态调整等特点，构建适配的知识图谱框架需结合政策特性，不能只借鉴其他政策框架。因此本文利用 Neo4j 图数据库，构建适合科学数据管理政策的多层级知识图谱模型，实现科学数据管理政策的可视化，清晰展现政策内容结构和实施重点，识别科学数据管理政策之间的关联。

2 研究设计

通过文献研究、政策解读，总结科学数据管理与政策文本分析的相关理论，构建科学数据政策知识图谱本体概念模型，利用 Neo4j 进行可视化和深层次语义分析，本文的研究设计主要包括政策文本选择、研究方法和研究思路。

本文研究思路如图 1 所示，首先进行政策选择，以政策文件为依据和支撑，对我国科学数据管理政策进行可视化和深层次语义分析，通过检索国务院政府文件库、科学技术部科技政策文件、“北大法宝”等专业政策数据库，最大限度获取相关研究样本。

本文遵循以下原则筛选整合数据：确保数据的权威性和准确性，选择由中央与地方立法机关、行政机构制定的全国性与地方性政策法规；对关联度较低的样本进行过滤，以确保数据的完整性和针对性^[2]。选取与科学数据相关的现行有效的部门规范性文件、行政法规、部门工作文件中附件等政策文件。通过筛选整合最终确定 75 份科学数据管理政策作为分析样本（样本收集截至 2025 年 5 月 1 日）。

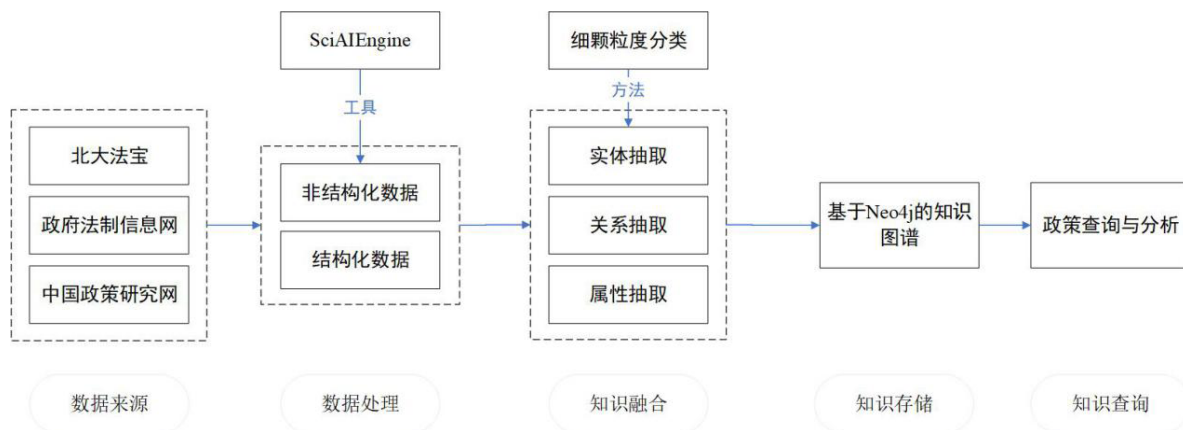


图 1 科学数据政策知识图谱构建整体流程

确定知识图谱构建方法：采用自底向上的方式构建科学数据政策知识图谱。借助人工处理、科技文献知识 AI 引擎 (SciAEngine) 对所获取的政策文本进行初步的数据结构化，再对所获得数据进行人工修订，之后进行实体抽取、关系抽取和属性抽取，利用 Neo4j 对结构化数据进行可视化展示，从而开展政策内容的分析与研究。

3 知识图谱构建

3.1 本体建构

本体是对一切真实存在的事物进行分析研究，是一个用于对客观世界实际存在的事实进行分析描述或者分类的分类体系^[23]。本体作为知识图谱的模式层，在约束概念关系、明确概念属性上起到了重要作用，其设计在很大程度上影响着知识图谱构建的质量，因此在本体建模过程中，有必要结合科学数据管理政策文本自身的结构特征与表达方式，对政策知识进行系统化抽象与规范化表达。通过剖析采集到的政策文本知识本体的构成及特征，确定科学数据政策知识实体的概念、关系及属性，利用本体理论对政策文本数据进行逻辑层的抽象化、概括化描述，以此实现政策文本的细粒度本体建模，搭建政策细粒度知识图谱模型，实现政策可视化。

根据政策数据在处理过程中的结构化程度，本文将科学数据管理政策知识本体划分为显性知识本体和隐性知识本体。显性知识本体指政策文件直接呈现、已经整理成型的结构化数据

信息，如实行日期、法规类别等，此类本体便于快速检索和分析政策发布的相关信息。隐性知识本体则侧重于对政策内部非结构化文本内容的识别，即通过对政策文本内容的深入解析与重构，将其转化为结构化的知识形式，如政策目标、政策措施等，此类本体通过结构化处理非结构化文本，挖掘深层次的政策意图和逻辑关系，能够识别政策重点，提高政策信息利用深度和广度。显性知识本体和隐性知识本体的区分有助于政策制定者、执行者及研究者快速识别并获取所需信息。

3.2 显性知识本体建构

政策文件层面的显性知识主要来源于“北大法宝”等权威政策法规数据库，包括政策标题、发文部门、政策执行时间、法规类别、发文字号、效力位阶、引用情况等政策信息。由于这类信息为结构化数据，因此可以直接用于知识图谱构建。显性知识涉及的实体包括政策标题、发文部门、发文字号、法规类别、政策执行时间及效力位阶等，实体之间的关系包括政策制定→发文部门、政策执行→时间、政策引用→其他政策等。具体实体与关系的定义及更多细节如表 1 所示。

这些实体及其相互关系是科学数据管理政策的重要内容，反映政策的制定背景、执行时效、法律地位及影响力，引用与被引用的关系揭示政策间的逻辑联系与演进脉络。将内容结构化并通过知识图谱的可视化展示，能够更直观了解政策信息，分析政策间复杂关系。

表1 政策显性实体—关系—实体表

实体	实体关系	实体	说明
政策	制定机关	政府部门	表示政策的制定者和来源
	实行于	实行日期	说明政策开始执行的时间
	属于	法规类别	表明政策所属的法规或政策类别
	具有	发文字号	表示政策文件的编号或标识
	被赋予	效力位阶	说明政策在法律体系中的层级和效力
	引用	其他政策	表示政策之间的引用关系,体现关联性和影响

3.3 隐性知识本体建构

政策文件层面的政策隐性知识通过对政策文本进行分析与处理获得。本文利用文本爬取、SciAEngine工具、人工阅读筛选的方法,对政策文本中隐含的知识实体进行提炼与抽象,最终获得隐性知识实体,如表2所示。

在提炼政策文件层面的政策隐性知识时,本文选择聚焦于表格内容,主要是考虑以下几个方面:(1)政策目标是政策制定的核心所在,指明政策方向和预期效果,能够理解政策背后的动机和目的从而把握政策的整体框架和核心理念。(2)政策工具是达成目标的方式,列举其类型与特性,能识别政策实施路径及潜在影响。(3)政策措施是工具的细化呈现,解析其内容与关联能够精准把握政策着力点与推进节奏。(4)政策主体是政策制定与执行的关键角色,为政策执行提供了明确的组织保障和人员支持,能够清晰了解政策执行的责任体系,促进各部门之间的协同配合和相互监督。(5)数据要素作为科学数据管理政策的核心,是政策执行中必须明确界定的要素,决定着哪些对象或领域

将受到政策影响,明确政策实施范围。这五类内容共同构成了政策文件层面的隐性知识核心框架,是理解和分析政策文本的重要依据。

表2 政策隐性实体—特征—关系表

实体	实体特征	实体提炼	实体关系
政策	政策目标	愿景	(政策—目的—政策目的)
		项目计划	
		协同合作	
	政策工具	信息公开	(政策—利用—政策工具)
		创造	
		检验检测认证	
	政策措施	财政支持	(政策—资源—政策工具)
		效益评价	
		资源与环境技术	
	政策主体	信息公开	(政策—实施—政策措施)
协同合作			
使用管理			
数据要素	公共服务	(政策—主体—政策主体)	
	政		
	部门		
	平台	(政策—包含—政策要素)	
	科研院所		
	工程与材料科学		
	数理科学		
	生命科学		

3.4 政策主体建构

在政策主体构建过程中,本文以75篇科学数据管理政策文本为研究对象,对政策内容进行结构化处理,并再次进行人工复核,对政策的显性与隐性知识进行归纳和分类,得到实体—关系—实体层面的知识表示。需要说明的是,政策主体不是孤立存在的,而是存在于政策整体知识结构中,需要通过对政策显隐性知识系统抽取并识别。

以《国务院办公厅关于印发科学数据管理办法的通知》为例,如表3所示,列举该政策在显性与隐性知识层面的实体—关系—实体抽

取的示例，用以说明政策主体类实体是利用 SciAIEngine 网站和人工复核从整体政策知识中

识别与抽取，而不是直接呈现完整的政策主体构建结果。

表 3 《国务院办公厅关于印发科学数据管理暂行办法的通知》显隐性知识关系表

实体	实体关系	实体	说明 / 层级含义
	制定机关	国务院办公厅	政策发布与统筹主体
	实行于	2018.03.17	政策生效时间
	属于	机关工作综合规定	政策类型
	具有	国办发〔2018〕17号	政策编号
	被赋予	国务院规范性文件	法律效力层级
《科学数据管理暂行办法》	目的	经济社会发展，数据管理，科学数据安全，开放共享，科技创新，国家安全，科技成果，公共安全，环境保护	政策目标导向
	利用	检验检测，政府预算资金，数据共享，科学数据管理	政策应用方式
	实施	数据管理，开放共享，绩效收入，职称评定，增值服务，定制服务，网络安全保障，考核制	政策实施措施
		政府	宏观治理主体
	主体	科学技术行政部门	政府职能部门
		数据中心	数据管理与存储主体
		科研院所	数据生产与使用主体
	包含	自然科学，工程技术科学	数据覆盖领域
		知识产权，科学数据活动，单位和个人数据	数据对象与活动范围

对多项政策样本的实体系统抽取与归纳，通过区分宏观治理主体、职能部门与执行主体，为后续政策主体建构与分类分析提供基础。本文实现了科学数据管理政策主体的细粒度分

类，其分类内容来源于本文研究数据，并在图示呈现方式上借鉴了《科技报告政策细粒度知识图谱构建及应用研究》^[4]的表达形式，如图 2 所示。

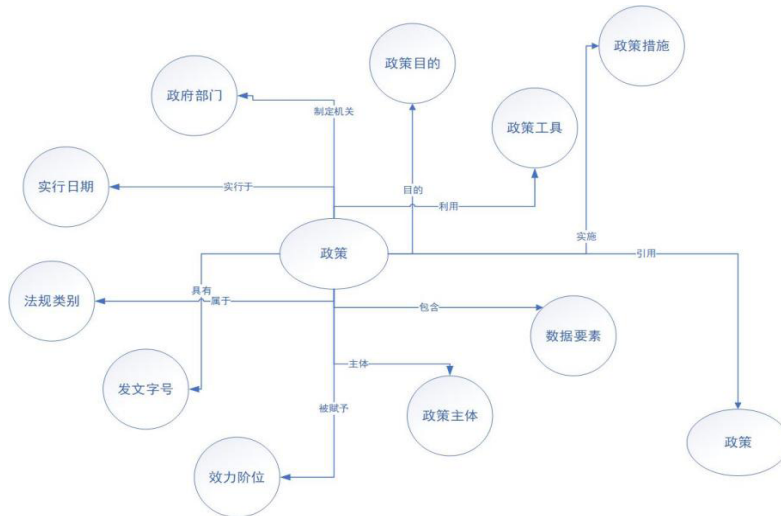


图 2 知识关系图

4 知识图谱实践应用

通过对政策的显性知识和隐性知识的抽取和解析，对政策所包含的关键要素进行了细粒

度拆解与结构化表达，最终形成了科学数据政策的知识图谱。一共包括 844 个主体和 1254 条关系。具体如图 3、图 4 所示。

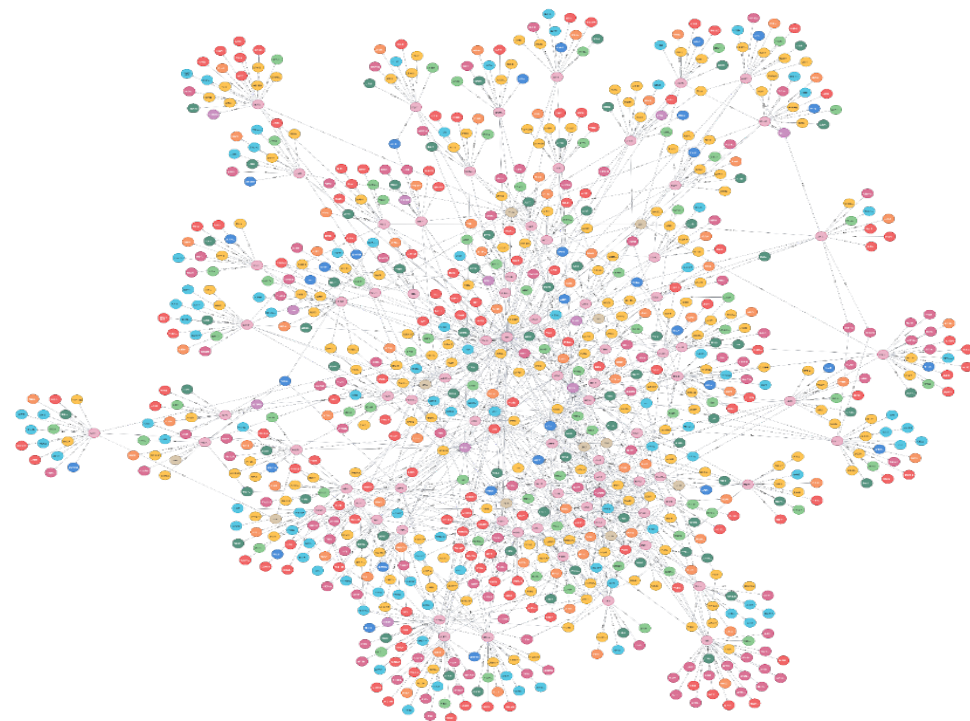


图 3 科学数据政策细粒度知识图谱



图 4 《国务院办公厅关于印发科学数据管理办法的通知》知识图谱

以颜色区分实体内容,粉色对应政策实体,棕色对应实行日期,深绿色对应制定机关,浅绿色对应发文字号,深蓝色对应法规类别,紫色对应效力层级,黄色对应政策目标与政策措施,红色对应政策工具,橘色对应政策措施,浅蓝色对应政策主体,深粉色对应数据要素。以颜色区分能够直观呈现各类要素的对应关系,提升政策文本的读取效率与信息识别精度。

4.1 政策内容查询

Neo4j 作为信息管理与数据分析的强有力工具,其独特的图数据库结构辅助处理复杂关系

数据,在构建和解析政策知识库方面也能提供帮助。通过 Neo4j 构建政策知识图谱,能够实现高效、精确的相关知识查取^[24]。如图 5 所示,以《国务院办公厅关于印发科学数据管理办法的通知》为例,知识图谱能够快速检索政策名称、精确追溯政策发布时间、清晰了解政策颁布部门,以及准确掌握政策实行日期。快速获取基础信息为政策分析奠定了坚实的基础。更进一步,知识图谱的深度挖掘能力使得用户能够深入探究政策内容,如解析政策发布目的,明确政策执行的主要责任方及其相互关系,以及政策利用的措施及工具。

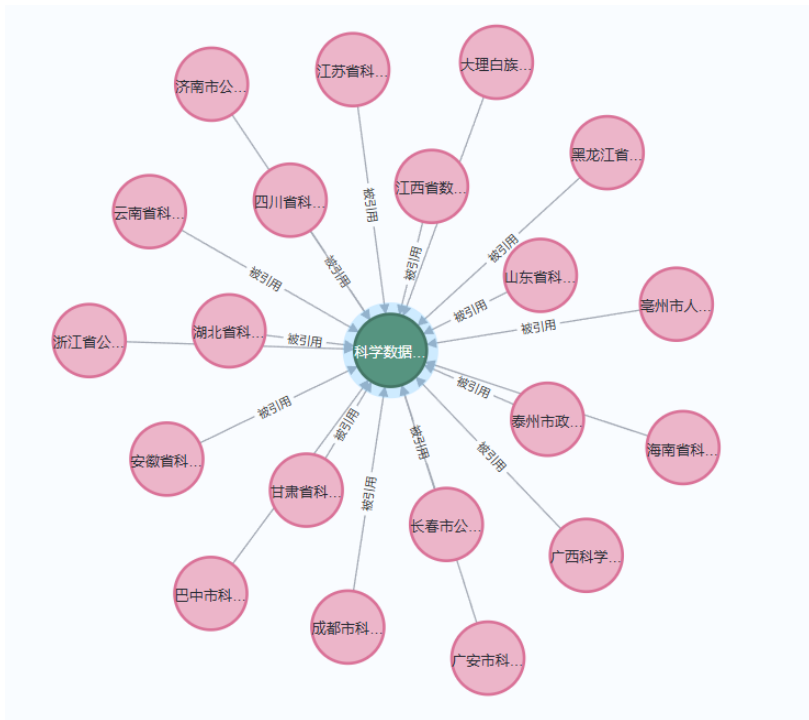


图 5 《国务院办公厅关于印发科学数据管理办法的通知》被引用关系

在此基础上,知识图谱支持单一政策文本中关键信息与核心内容的深度挖掘,还能够对不同政策之间的关联进行研究分析,比如梳理

某一政府部门发布的科学数据政策体系,或者识别引用《国务院办公厅关于印发科学数据管理办法的通知》的地方政策及其关联关系。

一些政府部门颁布一个科学数据类型政策，也有政府部门颁布多个科学数据政策，如图6所示，国家航天局颁布《澳门科学一号卫星科学数据管理办法》《“羲和号”卫

星科学数据管理办法》《月球样品及科学数据国际合作管理细则》《月球与深空探测工程科学数据管理办法》《月球样品管理办法》五个相关政策。

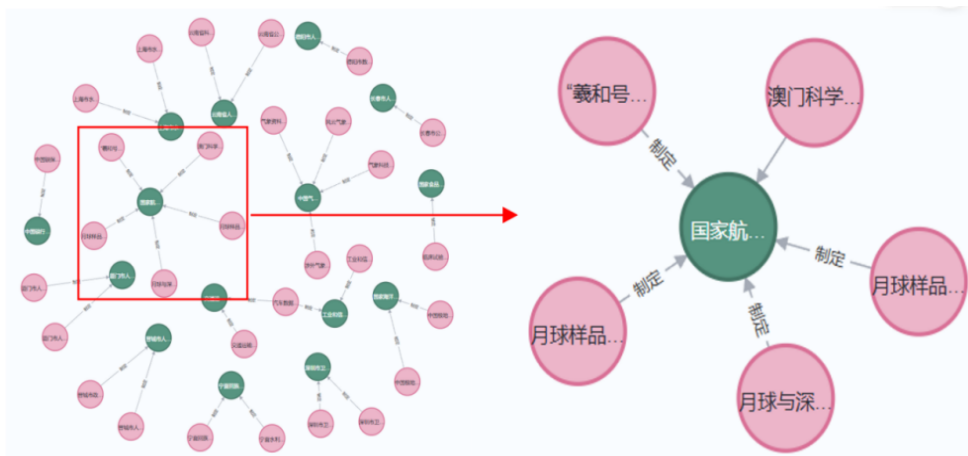


图6 国家航天局制定政策图谱

知识图谱使得政策内容的查询更加高效与直观。用户无需面对冗长的政策文本，而是能够通过可视化的界面，直接点击或搜索关键词快速定位到相关政策条目，并获取其详细信息。这种方式降低了信息检索的时间成本，提高了政策研究效率。

知识图谱能够揭示政策之间的内在联系与逻辑结构，将孤立的政策文件通过关系链接起来，形成一个有机整体，使用户能清晰地看到政策之间的引用、层级或冲突关系。这种全局视角的呈现，有助于用户更全面地理解政策背景、目的与影响，为政策制定、执行与评估提供有力的支持。知识图谱还具备强大的可视化能力，它能够将复杂的政策信息转化为直观的图形或图表，如政策网络图、时间线、关系树等，使得用户能够一目了然地掌握政策的核心要点

与关键路径。这种直观性的呈现方式，不仅降低了政策理解的门槛，也使政策间的关系网络更加清晰。

4.2 政策关系分析

4.2.1 普及性应用

科学数据作为现代科研活动不可或缺的核心资源，覆盖到各个领域，范围广泛，影响深远。国家积极推动科学研究发展，鼓励科学数据的开放共享，促进科学数据管理政策的制定，政策的应用范围不断拓展。这些政策覆盖基础科学研究各个领域，如物理学、化学、生物学等，还深入到了应用科学的诸多方面，包括农业、林业、海洋、气象、地震、地球系统科学以及人口与健康等。相关领域的政策如图7所示。

Node properties		Node properties	
政策		政策	
<id>	2473	<id>	3076
name	国家农业科学观测工作管理办法(试行)	name	江西省林业信用数据管理办法
Node properties		Node properties	
政策		政策	
<id>	3196	<id>	3075
name	中国极地考察数据管理办法	name	风云气象卫星数据管理办法(试行)

图7 各领域政策实例

科学数据管理政策的广泛应用具有社会和经济价值。在农业、林业等领域，科学数据的开放共享有助于精准农业、林业的发展，提高资源利用效率，保障粮食安全和生态安全等。在气象、地震等自然灾害领域，科学数据的及时共享则为防灾减灾提供了有力支持，有利于保障人民生命财产安全。在人口与健康领域，科学数据的开放共享则有助于公共卫生事件的应对和疾病的预防控制，提升全民健康水平。在数据运用方面，通过政策引导，科研机构和个人能够更加便捷地获取和利用所需数据，减少数据收集、整理和分析的重复劳动，从而有

更多的时间和精力投入到科研创新。此外，数据的开放共享还能够激发新的科研思路 and 方向，推动科学研究的不断深入和发展。政策完善也促进不同学科的交流与融合，打破学科壁垒，为跨学科研究提供了坚实的数据基础。在大数据和人工智能等新兴技术的推动下，跨学科合作已成为科学研究的重要趋势，而科学数据政策规范为这一趋势的加速发展提供了支撑。

4.2.2 层级网络

科学数据管理政策构成了一个复杂而有序的层级网络，如图8所示，其中不同层级的政

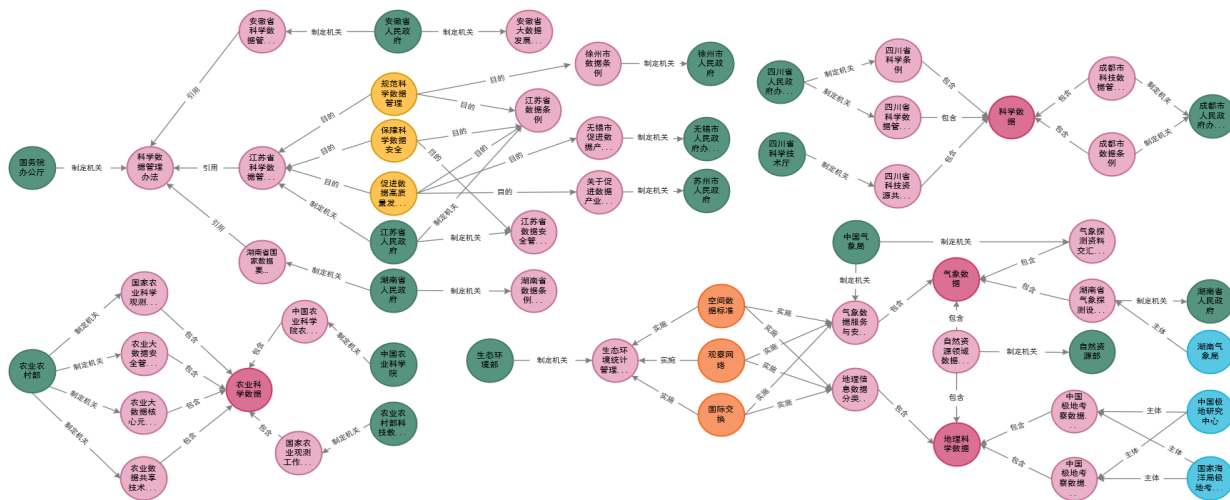


图8 各政策层级网络

策文件相互关联、相互影响，共同构成了科学数据管理的政策框架。地方政府基于本地科学研究、产业发展以及公共服务等多方面对科学数据的实际需求，通过制定规范性文件，因地制宜地构建适合本地的科学数据管理细则，来满足不同地区的科学数据管理需求。各部门依据自身职能，针对特定领域的科学数据管理工作，以保障部门业务范围内科学数据的有效利用与管理。地方在科学数据管理中同时关注实际操作和制度建设。为落实具体工作任务和应对阶段性管理需求，地方政府需要对相关政策进行动态调整。同时，将实践中形成的成熟经验转化为制度规范。所以地方层面科学数据管理政策呈现出数量较多、更新频繁的特点。

4.2.3 多维目标导向

政策目标深刻影响着科学数据管理的每一个环节，从数据的生成、存储、处理到最终的共享与利用。研究目标宗旨，能把握政策整体方向和预期效果。通过提炼深入理解政策背后的动机和目的，来更好架构政策整体框架和核心理念。

输入 cypher 语句：

```
MATCH (gm: 政策目标 )-[r]-(p: 政策 )
RETURN gm.name AS 政策目标 ,
COUNT(DISTINCT r) AS relationshipCount
ORDER BY relationshipCount DESC
LIMIT 50
```

内容深刻揭示了科学数据政策的多维目标导向，对文章目的进行统计并聚类，如表 4 所示。

表 4 政策目标与宗旨(部分)

目的与宗旨	出处政策	知识聚类
保障科学数据安全	《国务院办公厅关于印发科学数据管理暂行办法的通知》《甘肃省科学数据管理实施细则》《成都市科技数据管理实施办法》……	数据安全 保障
保障数据安全	《甘肃省科学数据管理实施细则》《宁夏回族自治区科学数据管理实施细则》《云南省科学数据管理实施细则》……	
加强和规范科学数据管理	《嘉峪关市科学数据管理实施细则》《江苏省科学数据管理实施细则》《昆明市科学数据管理实施办法》……	数据管理 规范
数据基础制度建设	《山西省数据工作管理办法》《数据知识产权登记管理办法（试行）》	
开放共享	《国务院办公厅关于印发科学数据管理暂行办法的通知》《中国极地考察数据管理办法》《甘肃省科学数据管理实施细则》……	数据开放 共享
促进数据开发利用	《汽车数据安全若干规定（试行）》	

科学数据管理政策目标具有多元化与系统性。通过知识聚类分析，能够明确地辨析出几个核心的政策目标领域：数据安全保障、数据管理规范、数据开放共享。基于这些核心领域，提出一系列针对性建议，旨在进一步完善科学数据管理政策体系，提升数据管理效能，促进数据资源的合理利用与共享，为科研创新和社

会进步提供数据支撑。

5 科学数据管理策略与建议

5.1 数据管理规范针对性

在政策制定层面，提高针对性与精细化成

为迫切需求。鉴于地域、科研机构及研究领域的多样性，科学数据的特性与管理需求呈现显著差异。由图 5 也能看出，我国各省市及科研机构普遍以《科学数据管理办法》为构建其数据管理政策框架的基础，在这一普遍性指导原则下，需进一步细化为适应各地具体情境的定制化方案。因此，在制定科学数据管理政策时，既应遵循中央层面的总体战略导向，又紧密贴合地方科研生态、数据特征、基础设施发展状态、科研活动结构以及人员技术能力等实际情况，精准施策，提升政策有效性和适应性，促进科学数据管理的精细化与定制化发展^[24]。

5.2 数据开放共享

科学数据作为数据密集型科学研究范式成果，其开放共享具有不可估量的价值。我国科学数据开放整体水平并不平均，地区与地区之间数据开放不均衡，在获取资金、工具、技术上都存在一定的差异。通过政策能够看出，有的省市颁布政策积极响应号召，有的市暂时搜不到相关的科学数据共享政策^[25]。而各省份的数据开放平台程度的差异也体现出不同地区的数据开放程度，部分省市积极响应国家号召，已率先出台了一系列旨在促进科学数据开放共享的政策措施，具有较高的行动力。这些政策不仅为本地科学数据的开放提供了制度保障，也为其他地区树立了典范。但是也有城市在科学数据共享政策的制定与实施上相对滞后，没有找到相关政策出台的信息，这也在一定程度上限制着科学数据挖掘与利用。进一步观察各省份的数据开放平台，其建设程度与开放水平的差异更是直观反映了地区间数据开放的不均

衡状态。我国科学数据开放共享工作已取得一定进展，但仍面临区域不均衡、资源配置不均等挑战，需要通过加强政策引导、优化资源配置、提升技术能力等多方面措施，推动科学数据开放共享向更加均衡、高效的方向发展，以充分释放科学数据的巨大价值，为科技创新和社会发展注入强劲动力。

5.3 数据安全与隐私保护

在科学数据开放共享过程中，发展与安全是两个相互依存的重要目标。当前在推进可信、可靠与合规的数据开放共享方面仍面临诸多困难和挑战。数据安全不能只从宏观上进行保障，从研究中也能看出科学数据管理范围广泛，笼统的数据安全并不全面覆盖所有范围。当前数据安全的重点是关注内容，但也面临执行、技术层面的挑战。数据交叉融合安全存在高隐藏风险，多源、多学科、多时空的科学数据不断汇集带来数据交叉融合方面的问题。数据的规模、精度呈现动态的变化，原有的安全边界不断被打破，导致隐私信息、隐私属性及隐私关联的泄露，失去对数据安全的控制^[26]。为克服这些问题，更要加强政策的细化与落地，确保每一条款都能转化为具体的操作指南，并辅以有效的监督与激励机制，采取奖惩并重的措施，激发数据管理者与用户的安全意识，共同构建安全可信的数据环境。此外，建立定期的政策效果评估机制，通过实证分析反馈政策执行成效，及时总结经验，动态调整优化数据安全政策，以应对科研实践中不断变化的安全威胁与隐私保护需求，确保数据安全政策的持续有效性与前瞻性^[27]。

6 总结

本文汇总并深入梳理了我国科学数据管理政策的核心内容,通过采用 Neo4j 工具对政策进行可视化分析,清晰呈现政策内容的结构框架,识别政策实施的关键领域,我国科学数据管理政策在不断发展与完善中,已构建起较为完善的政策框架和实施机制,从数据的产生、收集、存储、共享到利用等各个环节,都有相应的政策规范与引导,跨部门、跨领域的协调机制也加强了政策在各主体间的衔接,为科学数据管理工作的顺利开展提供了制度保障。未来,随着技术的不断进步和需求的日益多样化,科学数据管理将面临更多新的挑战与机遇,期待政府、科研机构、高校及企业等各方共同努力,持续推动科学数据管理的创新与发展,为构建开放共享、安全可信的科研数据生态贡献力量。

参考文献

- [1] 国务院办公厅. 国务院办公厅关于印发科学数据管理的通知 [EB/OL]. (2018-03-17) [2023-06-01]. https://www.most.gov.cn/xxgk/xinxifenlei/fdzdgknr/fgzc/gfxwj/gfxwj2018/201804/t20180404_139023.html.
- [2] 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.
- [3] 刘莉, 岳丽欣, 刘自强. 我国科学数据管理政策内容演化态势研究 [J]. 情报科学, 2024, 42(1): 162-171.
- [4] 支凤稳, 孟佳琪, 郑彦宁. 科技报告政策细粒度知识图谱构建及应用研究 [J]. 情报理论与实践, 2024, 47(7): 170-180.
- [5] 张维冲, 王芳, 赵洪. 基于全要素网络构建的大规模政策知识关联聚合研究 [J]. 情报学报, 2023, 42(3): 289-303.
- [6] 阳巧英, 夏义堃. 我国科学数据政策结构特征与演进逻辑分析 (2001—2022 年) [J]. 情报资料工作, 2023, 44(6): 6-14.
- [7] 凡庆涛, 张士运, 李博, 等. 从政策工具视角审视我国开放科学政策文本特征 [J]. 大学图书馆情报学刊, 2024, 42(5): 90-99.
- [8] 叶钰铭. 我国开放科学政策体系构建研究 [D]. 北京: 中国科学院大学 (中国科学院文献情报中心), 2022.
- [9] 姜鑫, 侯裕馨. 欧洲国家开放科学政策文本多维量化评价分析与启示 [J]. 现代情报, 2025, 45(9): 150-164, 176.
- [10] 马合, 黄小平. 欧美科学数据政策概览及启示 [J]. 图书与情报, 2021(4): 84-91.
- [11] 刘冰, 王晋明, 晁世育. 英、美、澳高校科研数据管理政策实证研究 [J]. 情报理论与实践, 2021, 44(8): 59-67.
- [12] 张耀南, 任泽瑶, 康建芳, 等. 英国科学数据发展政策与规划 [J]. 中国科学数据 (中英文网络版), 2024, 9(1): 21-35.
- [13] 张耀南, 张彩荷, 康建芳, 等. 澳大利亚科学数据发展政策与规划 [J]. 中国科学数据 (中英文网络版), 2024, 9(1): 36-51.
- [14] 王译晗. 国外科研机构开放科学政策及其启示 [J]. 数字图书馆论坛, 2024, 20(5): 1-12.
- [15] 姜天海, 贾萍萍, 张增一. 政策工具视角下欧美国家与国际组织开放科学政策文本分析及其启示 [J]. 图书情报工作, 2022, 66(22): 119-133.
- [16] 姜鑫. 世界一流大学开放科学数据政策体系框架及内容要素研究 [J]. 现代情报, 2024, 44(5): 153-165.
- [17] 张吉祥, 张祥森, 武长旭, 等. 知识图谱构建技术综述 [J]. 计算机工程, 2022, 48(3): 23-37.
- [18] 张西珩, 李昕, 唐鹏, 等. 基于知识图谱的隐私政策合规性检测与分析 [J]. 网络与信息安全学报, 2024, 10(6): 151-163.

- [19] 张维冲,王芳,黄毅.基于图数据库的贵州省大数据政策知识建模研究[J].数字图书馆论坛,2020(4):30-38.
- [20] 张雨,吴俊.科技政策知识图谱构建研究[J].数字图书馆论坛,2021(8):31-38.
- [21] 于小涵.我国产业政策知识图谱构建与应用[D].济南:济南大学,2023.
- [22] 黄佳妮,陆伟,于丰畅.基于实体知识网络的政策文本关联分析——以国家创业政策为例[J].图书馆论坛,2024,44(6):71-80.
- [23] 冯一展.基于深度学习的长城知识图谱构建[D].北京:华北电力大学(北京),2022.
- [24] 王瑞萍,刘峰,杨媛琦,等.审计知识图谱的构建与研究——基于 Neo4j 的图谱技术[J].中国注册会计师,2020(9):109-113.
- [25] 黄如花,吴应强,李白杨.中国省级政府数据开放平台利用水平的组态效应研究——基于 NCA 与 fsQCA 的实证分析[J].图书情报工作,2024,68(10):66-80.
- [26] 廖方宇,李婧,龙春,等.开放科学背景下科学数据开放共享安全挑战及我国对策思考[J].农业大数据学报,2024,6(2):146-155.
- [27] 刘春年,佟雪.制度文本理论下的中外大数据安全政策比较[J].情报理论与实践,2022,45(11):68-77.

(责任编辑:李琳娜)