

基于文本挖掘的我国科技创新政策量化评价研究

郝海拓, 张志娟

(中国科学技术信息研究所, 北京 100038)

摘要: 本研究基于2014—2019年国家科技创新政策进行文本挖掘, 以实现政策的量化评价研究, 对政策的高频词、网络性、中心性和小团体进行了直观的量化和列示。文本挖掘方法基于政策原文本进行分析, 避免了主观判断带来的偏差。研究表明, 2014—2019年国家科技创新政策中具有“创新”“科技”和“技术”等高频词, 文本挖掘结果具有较优的网络集中度, 并包含7个核心小团体。国家科技创新政策推动着国家创新事业的发展, 并为创新指明方向、提供支持。对政策深度挖掘有助于理解政策背后深层的意义, 以指导各主体的创新决策。同时, 对政策文本的挖掘有助于从深层次理解和认识政策, 为国家及区域、省、市等地政策制定者提供政策决策支持。

关键词: 科技创新政策; 文本挖掘; 语义分析

中图分类号: G322.0 **文献标识码:** A **DOI:** 10.3772/j.issn.1009-8623.2020.06.009

习总书记曾经提出“发展是第一要务, 人才是第一资源, 创新是第一动力”, 可见创新对于国家发展的重要程度。我国对创新工作的重视集中体现在国家创新政策的发布和落实上, 国家创新政策不仅引领着国家创新的方向、各主体的创新活动, 还为各区域、省域、城市等的科技创新政策的决策提供依据, 因此对国家科技创新政策进行深度的理解具有极为重要的意义。政策文本相对于数据而言, 其可读性、易读性和直观性都明显较低, 政策评价研究中往往有较强的主观性, 如专家打分和定性研究等。为解决这一问题, 本文选取2014—2019年的403项国家科技创新政策进行文本挖掘, 摒除了主观偏好对结果的影响, 增强了结果的客观性。基于文本本身进行挖掘, 得出高频词、网络型、中心性和小团体等结论, 以图、表和数据的形式展现政策重点, 增强了政策分析结果的直观性和易读性。

1 研究背景与研究意义

2016年, 中共中央、国务院印发了《国家创新驱动发展纲要》, 提出了三大战略部署, 第一条就是“实施创新驱动发展要促进科技创新和体制创新的双轮驱动”。创新驱动发展需要依靠科技创新, 而科技创新需要国家通过科技创新政策予以引导、支持和激励, 通过国家科技创新政策引导各创新主体的创新行为, 进而形成创新结果, 再经由创新成果转化形成能够助力国家发展的切实生产力。可以说, 国家科技创新政策是所有创新活动参与主体的行动依据和活动指南, 直接影响着国家创新活动的有序推进, 最终作用于国家的创新发展。

而国家科技创新政策众多, 内容涉及面广, 如何对其进行深入分析是学术界研究的重点。以往的政策评价研究往往难以实现量化, 采用专家打分的定量分析方法和其他定性分析方法很难规避主观性

第一作者简介: 郝海拓(1988—), 女, 助理研究员, 主要研究方向为国家创新型城市、科技创新政策和企业创新。

项目来源: 中国科学技术信息研究所创新研究基金青年项目“基于文本挖掘的国家创新型城市科技创新政策量化评价研究”(QN2020-06)。

收稿日期: 2020-05-03

导致的分析结果偏差, 因此实现客观准确的政策量化一直是政策研究的难点。本研究拟通过文本挖掘方法实现科技创新政策的量化评价, 完全以政策原文本作为分析的原始数据, 对政策进行深入的挖掘以实现科技创新政策的量化评价, 是“从文本中来到文本中去”的研究思路, 避免了主观性对量化评价的影响。

从理论意义角度来说, 基于文本挖掘的国家科技创新政策量化评价研究采用 ROST CM 和 Ucinet 软件实现了政策的深度挖掘, 对政策研究的深度有所加深, 立足政策原文本的客观研究也增加了研究结论的客观性和说服力。同时, 将文本挖掘用于国家科技创新政策研究, 拓展了语义文本挖掘的应用范围, 有重要的理论意义。

从现实意义角度来说, 基于文本挖掘的国家科技创新政策量化评价研究得到的高频词、网络性、中心性和小团体分析对于挖掘政策重点、侧重点、实现路径等都具有重要的现实价值, 有助于下级政策制定者加深对政策的解读和理解, 为政策制定者的政策立改废决策提供依据。同时, 对于政策受众, 特别是参与创新活动的主体来说, 本研究能够帮助受众理解政策意图, 为调整创新决策提供具体的参考, 因而本研究具有重要的现实价值。

无论从理论还是现实角度, 本研究都具有很强的研究价值, 这也是本文开展研究的重要立足点。

2 研究综述

2.1 科技创新政策研究综述

Amankwah-Amoah^[1]提出, 科技创新 (Science Technology and Innovation, STI) 政策是对科学、技术与创新三种政策类别的统称, 不是单一概念, 而应视为一种综合概念。科学、技术和创新是关于科技创新政策内涵的一致性观点, 科技创新政策融合了三个核心关键点的研究重心, 并以创新收益乃至经济增长为目标^[2, 3]。我国学者赵修卫^[4]从政策工具的角度解析科技创新政策, 包含人才激励、税收优惠、金融支持等诸多方面, 并认为不同政策工具的组合使用往往会带来更优的政策作用效果。

随着对科技创新政策内涵的逐渐深化, 政策分类的相关研究也逐渐发展并形成体系。基于国内外学者对于科技创新政策的定义, 本文认为科技创新

政策是依靠科学和技术实现创新目标并最终服务于国家科学、技术、创新以及经济发展的国家治理工具。

国外最为经典的科技创新政策分类方法是将其划分为供给型、需求型和环境型三类^[5], 这种分类方法是 1985 年由 Rothwell 等学者提出的, 后来许多学者认同, 并基于此对政策进行深化研究并扩展到政策工具角度, 基于国内外学者对于政策工具划分的相关研究对三类科技创新政策进行总结, 供给型科技创新政策包含政府研发补贴、信息服务、教育培训和人才激励等内容, 需求型科技创新政策则大致包含技术认定、政府采购、服务外包和示范工程等, 环境型科技创新政策则包含金融支持、税收优惠、知识产权、法律法规和服务支持等方面^[6-8]。我国学者范柏乃等^[9]在供给型、需求型和环境型的政策类型划分基础上提出了增加基础政策这一类型。1986 年, Rothwell^[10]在其研究的基础上将科技创新政策划分为科技政策和产业政策两大类型, 这种分类方式与学者们在研究科技创新政策内涵中提到的科学、技术和创新三点契合。国内其他学者研究科技创新政策分类时, 往往从政策工具角度出发, 如蒋栋等^[11]提出按税收优惠、政府采购、金融支持等政策工具对科技创新政策进行分类。葛丹明^[12]则提出按照政府补贴、专利保护、税收政策、政府采购四大分类准则对科技创新政策进行分类。另外还有学者从政策的特质层面来评价科技创新政策, 如刘会武等^[13]认为可以从评价目的、评价时间和创新活动三个方面对创新政策进行评价, 并从效果、效率和效益三方面评价政策支持的创新活动。2013 年程华等^[14]以 454 项创新政策为样本, 结合 C-D 函数及政策变量, 评价政策的演变、力度、稳定性和创新绩效。

2.2 科技创新政策量化研究综述

科技创新政策量化是该领域研究的难点, 但学者们利用各种工具探索政策量化的方法并取得了一定成果。专家打分法是早期政策量化评价的主要方法, 这种方法操作简单, 但不能有效避免专家的主观性所带来的结果偏差。我国学者彭纪生等^[15]2008 年提出了基于政策力度、政策措施和政策目标维度的三维度评价指标体系, 并设置了易对照操作的政策量化标准操作手册, 由于其具

备很强的可操作性，指标的选取也较科学合理，因而该政策量化标准操作手册被学者广泛采用。彭纪生的政策量化评价方式兼具主观性与客观性，是政策量化从主观性向客观性过渡的重要标志。张永安等^[16]参照彭纪生的政策量化评价方法评价创新政策，并提出以政策力度、目标和措施等为一二级指标下的11个二级指标实现国家创新政策的量化并研究其与国家创新能力之间的关系。Ruiz Estrada等^[17]于2008年提出了PMC指数（Policy Modeling Consistency Index）来结合主客观角度实现政策量化，且PMC指数模型方法更多地偏向客观量化。国内学者张永安等^[18]采用此方法将区域科技创新政策予以量化，此后，张永安等^[19]又对PMC指数模型进行了改进，将其中主观评价的部分转化为全部的文本挖掘，进一步增强了政策量化评价研究的客观性。

上述方法对科技创新政策都有极为重要的意义，但即使是PMC指数评价方法在选择指标时也很难避免主观性的部分影响，政策量化评价的趋势是从主观量化逐渐转向主客观结合的量化评价，最后实现客观的量化评价。因而本研究拟基于文本挖掘实现国家科技创新政策的量化，不选择指标，只分析和挖掘政策文本背后隐含的深层次信息，以完全规避主观性带来的量化评价偏差。

3 基于2014—2019年我国科技创新政策的文本挖掘研究

3.1 研究对象与数据来源

本研究选取2014—2019年国家科技创新政策，主要来源为人民政府官方网站、各关键部门官方网站（包含中华人民共和国科学技术部、中华人民共和国工业和信息化部、中华人民共和国人力资源和社会保障部、国家知识产权局、中华人民共和国国家发展和改革委员会、中华人民共和国财政部等）以及北大法宝数据库。在进行政策筛选的过程中，只选择有具体细则的精准政策，对于只对一般性事件进行简单通知或回复的政策文本则舍去，因而在政策筛选中需要进行手工整理和辨别，最终选取403项国家科技创新政策进行研究，研究对象来源与数量分布如表1所示。

表1 研究对象来源与数量分布

| 年份 | 官网 | 北大法宝 | 合计 |
|------|-----|------|-----|
| 2014 | 31 | 15 | 46 |
| 2015 | 37 | 27 | 64 |
| 2016 | 51 | 25 | 76 |
| 2017 | 51 | 40 | 91 |
| 2018 | 34 | 28 | 62 |
| 2019 | 40 | 24 | 64 |
| 合计 | 244 | 159 | 403 |

如表1所示，403项国家科技创新政策中，有244项来自于各官网，159项来自于北大法宝，对政策原文本进行整理后汇总成国家科技创新政策数据库，并对其进行文本挖掘。

3.2 研究方法及研究框架

本研究基于政策本身进行文本挖掘，将非结构化和半结构化的文本数据转化为结构化的关系和数据，增加了科技创新政策研究的深度，避免了人为政策评价所带来的主观性强、客观性无法预测等问题，为政策立改废提供客观可行的依据。研究主要依据政策科学理论和文本挖掘方法完成，符合挖掘科技创新政策背后的本质关联与学科发展新路径。语义文本挖掘方法凭借语义提取思想，能够将文本文件背后的重点、规律和关联关系进行深度挖掘，将文本形式的数据转化为可读、可计算、易识别的各类数据形式。相对于人工筛选等方法，文本挖掘避免了主观性对统计和分类的影响，将客观结果直观地呈现出来。凭借ROST CM与Ucinet软件对政策文本的高频词、关键词、网络图和中心性等信息进行高效的分析，实现政策文本量分析。

本文的研究框架可以分为以下部分。

第一步：样本收集。收集2014—2019年的国家科技创新政策并进行手动筛选，将政策原文件全部转为文本格式，形成2014—2019年国家科技创新政策数据库，以方便进行下一步研究。

第二步：政策文本挖掘。对政策文本进行处理，首先将所有政策文本汇总为一个文本文件并对该文本文件进行分词处理和词频分析，确定高频词后进

行高频词整理, 主要是筛除无实际意义的高频词和非研究重点的高频词, 完成后形成最终的高频词汇总表。此外还在此阶段形成共词矩阵, 这部分主要依靠 ROST CM 软件实现。

第三步: 进一步文本挖掘。对汇总的政策文本文件进行进一步挖掘, 基于文本文件与第二步形成的共词矩阵进行国家科技创新政策文本挖掘, 包含网络分析、中心性分析以及小团体识别与分析, 此部分主要依靠 Ucinet 软件实现。

第四步: 基于上述过程与结果提出研究创新点与本研究的研究展望。

具体如图 1 所示。

3.3 研究过程及结论

3.3.1 高频词分析

采用 ROST CM 软件对汇总后的国家科技创新政策进行高频词分析, 剔除其中无实际意义的词汇并对排序前 20 的高频词进行词语及频数汇总, 汇总结果如表 2 所示。

从高频词汇总表中可以看出, 排在第一位的高频词是“创新”, 频数为 11 236, 说明科技创新政策的核心在于创新。第二位和第三位分别是“科技”和“技术”, 这两个高频词是实现创新的主要方法和途径, 即依靠科技和先进技术实现创新。后面的关键词都是实现科技创新的路径和侧重点, 如在硬实力方面, 依靠企业、人才、资源、平台等并予以

相应的支持来实现科技创新, 在软实力方面, 也同样注重管理、相应的研究、体系的建立、机制的健全、能力的提升、成果的转化等, 双管齐下最终实现建设创新型国家。

高频词可以体现政策原文本所隐含的重点内容, 并据此对逻辑关系进行简单和直观的解读和判断, 而真正具体、客观和清晰的逻辑信息可以依靠 Ucinet 软件实现进一步的可视化。接下来就基于该

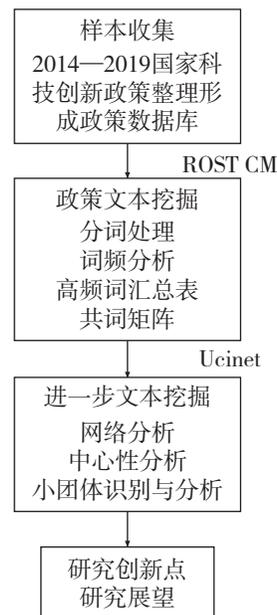


图 1 研究框架图

表 2 高频词结果汇总表

| 序号 | 高频词 | 频数 | 序号 | 高频词 | 频数 |
|----|-----|--------|----|-----|-------|
| 1 | 创新 | 11 236 | 11 | 人才 | 3 503 |
| 2 | 科技 | 9 240 | 12 | 加强 | 3 282 |
| 3 | 技术 | 9 068 | 13 | 机制 | 2 863 |
| 4 | 发展 | 7 735 | 14 | 成果 | 2 721 |
| 5 | 企业 | 6 648 | 15 | 资源 | 2 610 |
| 6 | 国家 | 5 699 | 16 | 体系 | 2 445 |
| 7 | 服务 | 5 465 | 17 | 平台 | 2 396 |
| 8 | 建设 | 4 873 | 18 | 能力 | 2 147 |
| 9 | 管理 | 3 871 | 19 | 提升 | 1 947 |
| 10 | 研究 | 3 724 | 20 | 转化 | 1 774 |

软件实现网络分析、中心性分析以及小团体识别与分析。

3.3.2 网络性分析

对文本进行网络分析，2014—2019年国家科技创新政策文本挖掘网络图如图2所示，可以看出，

“创新”“科技”“技术”“发展”“建设”在网络中重要性较强，处于相对关键的地位。

为了将网络图所体现的关系予以量化，接下来借助Ucinet软件展示网络节点上的相应数据，以实现中心性分析。

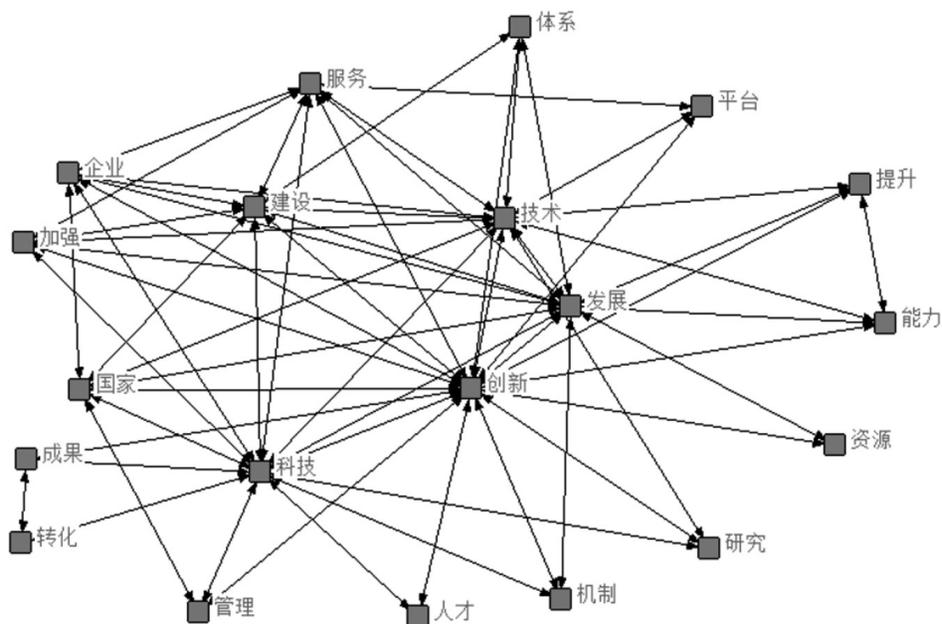


图2 2014—2019年国家科技创新政策文本挖掘网络图

3.3.3 中心性分析

以20个关键高频词为20个节点，对节点的点度中心度和网络中心势予以测算，并对节点代码、点度中心度及节点在网络中与其他节点连接的紧密程度予以列示。点度中心度体现的是节点在网络中的重要性，数值表示有多少个节点与之相连，结果显示，“创新”的点度中心度为22 149，是网络中最关键的核心节点，排在第二位的是“科

技”，其点度中心度为15 777，是网络中第二重要的节点，同时也是实现创新的最主要的途径。其他节点的中心度都在表3中予以列示，文中不予赘述，只需明确随着排序，越靠后的节点在网络中与其他节点相连的紧密程度越低。与其他节点连接的紧密程度与点度中心度对应，同样显示的是特定节点的重要性程度，具体数值见表3所示。

表3 网络中心势及节点点度中心度汇总表

| 序号 | 节点代码 | 节点 | 点度中心度 | 连接紧密程度 |
|----|------|----|--------|--------|
| 1 | 1 | 创新 | 22 149 | 0.163 |
| 2 | 2 | 科技 | 15 777 | 0.116 |
| 3 | 4 | 发展 | 15 710 | 0.116 |
| 4 | 3 | 技术 | 13 772 | 0.101 |

续表

| 序号 | 节点代码 | 节点 | 点度中心度 | 连接紧密程度 |
|----|------|----|--------|--------|
| 5 | 8 | 建设 | 10 262 | 0.076 |
| 6 | 6 | 国家 | 8 304 | 0.061 |
| 7 | 7 | 服务 | 8 078 | 0.059 |
| 8 | 5 | 企业 | 7 247 | 0.053 |
| 9 | 12 | 加强 | 5 871 | 0.043 |
| 10 | 16 | 体系 | 3 421 | 0.025 |
| 11 | 18 | 能力 | 3 358 | 0.025 |
| 12 | 19 | 提升 | 3 271 | 0.024 |
| 13 | 14 | 成果 | 2 851 | 0.021 |
| 14 | 10 | 研究 | 2 826 | 0.021 |
| 15 | 13 | 机制 | 2 791 | 0.021 |
| 16 | 9 | 管理 | 2 455 | 0.018 |
| 17 | 17 | 平台 | 2 388 | 0.018 |
| 18 | 15 | 资源 | 1 815 | 0.013 |
| 19 | 20 | 转化 | 1 759 | 0.013 |
| 20 | 11 | 人才 | 1 737 | 0.013 |

表 4 是对点度中心度和连接紧密程度的描述性统计结果。

表 4 节点的点度中心度与连接紧密程度的描述性统计

| 序号 | 类别 | 点度中心度 | 连接紧密程度 |
|----|----------|--------------------|--------|
| 1 | Mean | 6 792. 100 | 0. 050 |
| 2 | Std Dev | 5 738. 189 | 0. 042 |
| 3 | Sum | 135 842. 000 | 1. 000 |
| 4 | Variance | 32 926 820. 000 | 0. 002 |
| 5 | SSQ | 1 581 188 864. 000 | 0. 086 |
| 6 | MCSSQ | 658 536 384. 000 | 0. 036 |
| 7 | Euc Norm | 39 764. 164 | 0. 293 |
| 8 | Minimum | 1 737. 000 | 0. 013 |
| 9 | Maximum | 22 149. 000 | 0. 163 |

网络中心势 = 39. 750%

本研究中所有描述性统计保留到小数点后三位, 网络中心势保留到小数点后三位。

网络中心势体现的是在特定网络中, 各个节点的集中程度, 网络中心势取值介于 0 到 1 之间, 越趋近于 1 则表示网络越集中, 越趋近于 0 则代表网络越分散, 一般经验分析的界限是 10%, 即超过 10% 即代表该目标网络具有一定的集中度, 本研究中网络中心势达到 39. 750%, 网络集中度较优, 其可能原因有三方面: 第一是国家政策在规范程度、词语表达的精准性等方面较优; 第二是国家政策重点突出、核心明确; 第三是政策选取过程较精细, 选择的都是从题目到内容都与科技创新高度相关的政策。

3. 3. 4 小团体识别与分析

接下来进行小团体识别与分析, 小团体分析有助于将政策文本所体现的具体内容及其背后的逻辑进行团体划分。

样本共形成了7个小团体，将小团体汇总至表5，并对小团体命名。

表5 小团体情况汇总表

| 小团体序号 | 小团体名称 | 小团体包含关键词 |
|-------|-------|-------------------|
| 1 | 科技创新 | 创新、科技、建设、发展、加强、国家 |
| 2 | 企业服务 | 企业、服务 |
| 3 | 核心技术 | 技术 |
| 4 | 人才机制 | 人才、机制、研究 |
| 5 | 成果转化 | 管理、成果、转化 |
| 6 | 平台支持 | 能力、平台、资源 |
| 7 | 体系支持 | 提升、体系 |

从小团体结果可以看出，第1个小团体“科技创新”是整个网络中的核心团体，其中的“创新”“科技”“建设”“发展”“加强”以及“国家”，都是点度中心度排名靠前的高频词。

第2个小团体“企业服务”包含“企业”与“服务”两个高频词在科技创新乃至国家的经济发展中，企业都是极为重要的创新单元，为企业提供支持服务也是科技创新政策的重点之一，以此激发企业创新活力，带动区域乃至国家的创新能力提升。

第3个小团体是“核心技术”，其中只有一个关键词“技术”，“技术”的点度中心度排名第4位，之所以将其单独列为一个小团体也是因为技术的重

要性，核心技术的突破会带来全新的创新增长点和经济增长点。

第4个小团体是“人才机制”，包含“人才”“机制”和“研究”三个关键词，习近平总书记曾经提出“发展是第一要务，人才是第一资源，创新是第一动力”，足可见人才对于科技创新和国家发展的重要性，良好的人才机制能够激发研究人员的创新动力，提升创新效率。

第5个小团体是“成果转化”，包含“管理”“成果”和“转化”三个高频词。成果转化能够将最新的研究成果转化为切实的生产力，并反哺研究工作，是创新到发展的重要一环。

第6个小团体是“平台支持”，包含“能力”“平台”和“资源”三个高频词，不论是企业、人才还是其他创新资源，都需要平台才能实现资源与信息的交互，因而平台支持对于创新而言也是极为重要的。

第7个小团体是“体系支持”，包含“提升”与“体系”两个关键词，提高创新能力需要建立完善的国家创新体系，以确保各主体各司其职，创新绩效稳步提升。

总之，7个小团体代表了国家科技创新政策的重点内容，并共同构建了国家创新体系，落实创新战略，激发企业与创新人才的创新活力，疏通创新和成果转化通路，完善资源共享平台，最终提升创新绩效。

将小团体识别后2014—2019年国家科技创新政策网络图予以列示，如图3所示。国家创新能

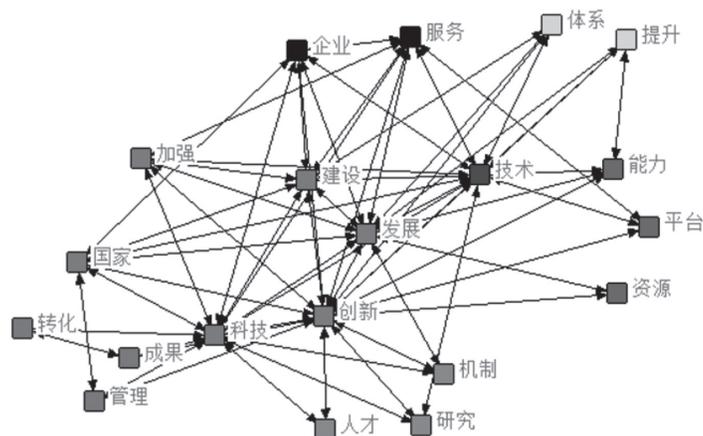


图3 小团体识别后2014—2019年国家科技创新政策网络图

力的提升、创新目标的实现需要各方的共同努力和政府及相关部门的全力支持。从量化结果看,我国科技创新政策已逐渐显现出强大的系统性,重点突出,保障有力。国家科技创新政策是国家创新活动的指南针,引导着创新的方向并对创新活动进行支持。

我国创新能力的稳步提升得益于国家对创新工作的高度重视,以及各创新主体的不懈努力。对政策进行挖掘有利于将隐含的信息以最直观的方式显示出来,加深对政策的理解,并为国家及区域、省、市等地政策制定者提供政策决策予以支持。

4 创新点与研究展望

本文的创新点主要体现在将 2014—2019 年国家科技创新政策文本后隐藏的深层次信息以高频词、网络图、中心性与小团体等方式全面体现出来,将非结构化的文字转化为结构化的直观易读的数据和关系。基于原文本的客观文本挖掘也摒除了研究人员的主观性带来的偏差,实现了客观的政策量化评价研究。

下一步本研究还希望对政策按照不同标准进行划分,以实现不同类型政策的对比研究。例如将国家科技创新政策按照供给型、需求型和环境型进行划分,以及按照强制型和引导型政策进行划分,对比不同类型的国家创新政策的不同特征以及在不同情境下的政策绩效,探索其深层次的原因,为政策制定者的政策决策提供借鉴和依据。■

参考文献:

- [1] Amankwah-Amoah J. The evolution of science, technology and innovation policies: A review of the Ghanaian experience[J]. *Technological Forecasting & Social Change*, 2016, 83(2): 134-142.
- [2] OECD. *Innovation Policy*[R]. Paris: OECD, 1982.
- [3] Dodgson M, Gann D, Phillips N. *The Oxford Handbook of Innovation Management*[M]. Oxford: Oxford University Press, 2013: 69-138.
- [4] 赵修卫. 现代科技创新政策发展的四个特点[J]. *科学学研究*, 2006, 24(6): 895-900.

- [5] Rothwell R, Zegveld W. *Reindustrialization and Technology*[M]. London: Longman, 1985: 2-27.
- [6] 袁健红, 龚天宇, 郭进芬. 中小企业创新政策效应评价[J]. *东南大学学报*, 2012, 14(4): 40-45.
- [7] Diaz-Garcia C, Gonzalez-Moreno A, Saez-Martinez F J. Eco-innovation: insights from a literature review[J]. *Innovation: Organization & Management*, 2015, 17(1): 6-23.
- [8] Mori D. Market distortions and optimal environmental policy instruments[J]. *Journal of Regulatory Economics*, 2017, 52(1): 1-13.
- [9] 范柏乃, 段忠贤, 江蕾. 中国自主创新政策的效应及其时空差异——基于省际面板数据的实证检验[J]. *经济地理*, 2013, 33(8): 31-36.
- [10] Rothwell R. Public innovation policy: To have or to have not?[J]. *R&D Management*, 1986, 16(1): 25-36.
- [11] 蒋栋, 李婷, 李志祥. 自主创新科技政策在河北省的实施效果评价[J]. *中国软科学*, 2009, 24(1): 88-92.
- [12] 葛丹明. 江西省农业龙头企业对技术创新政策满意度研究[D]. 南昌: 江西农业大学, 2013.
- [13] 刘会武, 卫刘江, 王胜光, 等. 面向创新政策评价的三维分析框架[J]. *中国科技论坛*, 2008, 24(5): 33-36.
- [14] 程华, 钱芬芬. 政策力度、政策稳定性、政策工具与创新绩效——基于 2000—2009 年产业面板数据的实证分析[J]. *科研管理*, 2013, 34(10): 103-108.
- [15] 彭纪生, 仲为国, 孙文祥. 政策测量, 政策协同演变与经济绩效: 基于创新政策的实证研究[J]. *管理世界*, 2008, 24(9): 25-36.
- [16] 张永安, 鄒海拓, 袁页. 科技创新政策与国家创新能力作用机理研究[J]. *华东经济管理*, 2019, 33(3): 104-110.
- [17] Ruiz Estrada M A, Yap S F, Nagaraj S. Beyond the ceteris paribus assumption: Modeling demand and supply assuming omnia mobilis[R]. Malaysia: Social Science Electronic Publishing, 2010.
- [18] 张永安, 耿喆. 我国区域科技创新政策的量化评价——基于 PMC 指数模型[J]. *科技管理研究*, 2015, 35(14): 26-31.
- [19] 张永安, 鄒海拓. 金融政策组合对企业技术创新影响的量化评价——基于 PMC 指数模型[J]. *科技进步与对策*, 2017, 34(2): 113-121. (下转第 76 页)

China's Industrial Innovation Ecosystem: A Visual Analysis of Research Hotspots and Frontiers

LIU Jing

(City College of Xi'an Jiaotong University, Xi'an 710018)

Abstract: Great attention has been paid to research on regarding innovation ecosystem as the novel mode of innovation, however, few scholars participate in its front focus in the middle-view of the industrial innovation ecosystem. Based on this point, relying on data from CNKI, visual analyses by Citespace V have been adopted in this paper to sort out the literature of industrial innovation ecosystem in China from 2000 to 2020. According to the study, there are five domestic research hot-spots: innovation, strategic emerging industry, industrial innovation ecosystem, high-tech industry, and industrial innovation. Therefore, a conclusion can be drawn that in China, scholars take “high-tech industry innovation ecosystem” as the front focus, but with rich fields and methods, they still give priority to basic theoretical research and qualitative methods in their studies. Therefore, this paper holds that in the forthcoming years, on the one hand, the academic circle can carry out further micro research in the middle-view; on the other hand, research methods and data types should be expanded.

Key words: industrial innovation ecosystem; research hotspot; research front; visual analysis

(上接第60页)

Research on Quantitative Evaluation of Science and Technology Innovation Policy in China Based on Text Mining

QIE Hai-tuo, ZHANG Zhi-juan

(Institute of Scientific and Technical Information of China, Beijing 100038)

Abstract: This research is based on the text mining of National Science and Technology Innovation (STI) Policy from 2014 to 2019 to achieve the quantitative evaluation of policy research. The high-frequency words, network, centrality and small groups results are listed and quantified directly. The text mining method is based on the analysis of the original policy text to avoid the bias caused by subjective judgment. The results show that there are high frequency words such as “innovation”, “science and technology” and “technology” in the national science and technology innovation policy from 2014 to 2019. The result of text mining has better network concentration and contains seven core groups. The National Science and Technology Innovation Policy promotes the development of the national innovation, points out the direction and provides the support for the innovation. In-depth policy mining helps to understand the deep meaning behind the policy to guide the main innovative decision-making. At the same time, the policy text mining helps to understand the policy from a deep level, and provides policy-making support for national and regional, provincial and municipal policy-makers.

Key words: science and technology innovation policy; text mining; semantic analysis