

Visual Analysis of Literature on Domestic Anti-Corruption Research: Based on CiteSpace

Xingchang Yan

Science and Technology of China, Chengdu Sichuan
Email: 13069392948@163.com

Received: Nov. 12th, 2019; accepted: Nov. 25th, 2019; published: Dec. 2nd, 2019

Abstract

In order to fully understand the evolution and process of domestic anti-corruption research, and seek breakthroughs in future theoretical innovation, based on CiteSpace analysis software, this paper comprehensively uses big data crawling, cleaning and visual analysis techniques, takes the literature on anti-corruption research included in the CNKI database from 1992 to 2019 as the analysis object, and demonstrates the macroscopic picture of anti-corruption research with visual knowledge map as the presentation method. We found out that domestic anti-corruption research shows that external representations which are the scopes of participation are "wide", but the cooperative relationship is "distant" and the field is "wide", but the degree of innovation is "low". In general, anti-corruption is still a hot topic in academic research in the future, however, it must make theoretical breakthroughs in research ideas, viewpoints, perspectives, and methods.

Keywords

CiteSpace, Anti-Corruption, Research Literature, Visual Analysis

基于CiteSpace的国内反腐倡廉研究文献可视化分析

闫兴昌

电子科技大学马克思主义学院, 四川 成都
Email: 13069392948@163.com

收稿日期: 2019年11月12日; 录用日期: 2019年11月25日; 发布日期: 2019年12月2日

摘要

为全面把握国内反腐倡廉研究的演进脉络与发展过程,寻求未来理论创新的突破口,本文综合运用了大数据爬取、清洗、可视化分析技术,通过CiteSpace分析软件,以1992~2019年中国知网(CNKI)数据库收录的反腐倡廉研究文献为分析对象,以可视化知识图谱为呈现方式,生动呈现了反腐倡廉研究的宏观图景。分析结果显示,国内反腐倡廉研究突显出参与范围“广”,但合作关系“疏”、涉及领域“广”,但创新程度“低”的外在表征。总体而言,反腐倡廉在未来仍旧是学界研究的一大热点,但必须要在研究思路、观点、视角、方法等方面实现理论突破。

关键词

CiteSpace, 反腐倡廉, 研究文献, 可视化分析

Copyright © 2019 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

近些年来,运用可视化知识图谱科学呈现研究现状,已然成为学界的一大热点。究其缘由,主要在于以下三点:其一,可视化知识图谱一改基于片面与局部文献的传统文献分析方式,极大地提升了文献梳理的科学性和标准化;其二,可视化知识图谱一改基于经验与常识的传统热点探测模式,将研究热点的定位放置到更为宏大的研究全景当中;其三,可视化知识图谱为研究者把握研究脉络与热点前沿提供了“利器”,迈出了社会科学研究走向现代化、数据化与量化的重要一步。现有可视化分析软件众多,如VOSviewer、Ucinet、Gephi等,但各个软件具备不同的优势点,在数据处理与数据可视化方面各有侧重。考虑到可视化分析软件对中国知网(CNKI)数据处理的精准度以及知识图谱的美观度,本研究选择借助陈超美教授研发的CiteSpace 5.0版本进行后续的可视化分析,力求科学全面地呈现出国内反腐倡廉文献的研究图景与演化路径。

2. 研究方法 with 数据来源

(一) 研究方法

CiteSpace是Citation Space的简称,译为“引文空间”,是美国德雷克赛尔大学陈超美博士与大连理工大学WISE实验室联合开发的科学文献分析工具——基于Java编程语言的文献计量学软件,其旨在绘制某研究领域的可视化知识图谱,以挖掘文献背后所蕴含的潜在知识。本文首先采用中国知网(CNKI)数据库自带的文献分析功能,对2019年之前的反腐倡廉研究文献情况进行初步的汇总。然后,使用陈超美教授基于Java平台研发的CiteSpace分析软件对核心作者群、核心研究机构群、关键词共现情况、高频与中心性关键词、关键词时序演进以及突现关键词等进行深度的计量学分析与可视化分析。

(二) 数据来源

本研究以中国知网(CNKI)期刊数据库中的“核心期刊”与“CSSCI”为数据来源,整个研究的数据准备主要分为两个步骤:第一,数据爬取。在中国知网(CNKI)中启动高级检索功能后,以“反腐倡廉”为主题词,以“不限~2019年”为时间维度,共检索出2266篇文献材料。第二,数据清洗。由于用作

CiteSpace 软件可视化分析的文献必须包含作者、关键词等字段，因而要得到准确的可视化数据源，就必须对检索文献进行“过滤和清洗”。在 2266 篇文献中剔除会议通知、征稿通知、期刊简介等无效文献后，共得到 2073 条有限文献用以后续定量研究。通过数据爬取和数据清洗，有限文献得以“净化”和“提纯”，为可视化分析奠定了良好的数据基础。

3. 反腐倡廉研究的宏观动态分析

(一) 反腐倡廉研究文献年际分布

研究文献的年际分布客观上呈现了研究进程的宏观趋势，在一定程度上为研究者了解和把握学科领域的“昨天、今天与明天”提供了可靠遵循。本文将 CNKI 数据库中清洗后的数据以时间为维度进行划分，并将梳理结果转化为折线图，如图 1 所示。

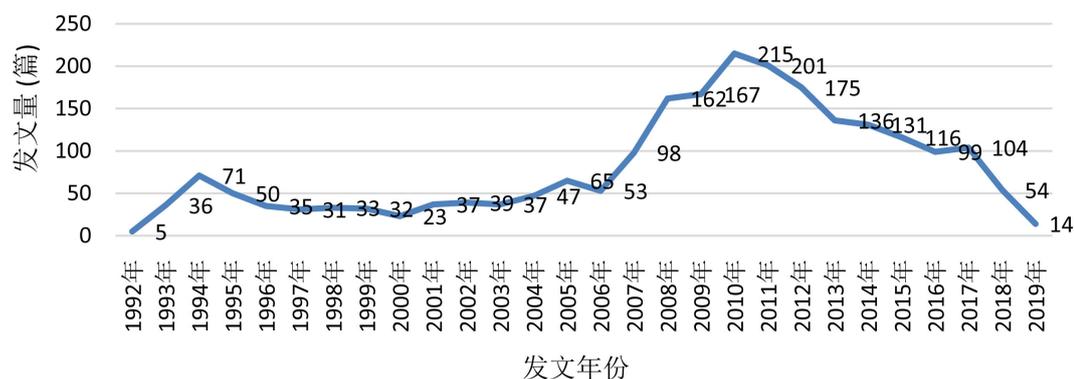


Figure 1. Inter-annual distribution of anti-corruption research literature

图 1. 反腐倡廉研究文献年际分布

由图 1 可知，CNKI 数据库中关于反腐倡廉的研究文献从 1992 年开始出现，在 2010 年到达峰值点。根据曲线图的升降趋势，可将整个研究过程划分为四个阶段：第一，萌芽时期(1992~1994 年)；该阶段发文总量为 112 篇，占总发文量的 4.94%。第二，探索阶段(1995~2006 年)；该阶段发文总量为 482 篇，占总发文量的 21.27%，较萌芽阶段有大幅增长。第三，发展阶段(2007~2012 年)；该阶段发文总量为 1018 篇，占总发文量的 44.92%，是探索阶段内发文量的两倍。第三，成熟阶段(2013~2019 年)；该阶段发文总量为 654 篇，占总发文量的 28.86%，与发展阶段相比，发文总量有所下降。

(二) 反腐倡廉研究文献作者分布

核心作者是指在研究领域内发文量达到一定标准，并产生了一定影响力和引领力的作者群体，对核心作者的分析不仅能够认清研究领域的领军人和先锋者，亦能够为后继研究者把握研究领域的前沿动态提供参照的目标。为准确呈现反腐倡廉研究的核心作者，本文依据普赖斯定律计算公式确定该研究领域的核心作者，即“ $M = 0.749 \times \sqrt{N_{\max}}$ (其中 M 为核心作者最低发文篇数， N_{\max} 为发文最多作者的文篇数)” [1]。

本文将 CNKI 数据库中爬取的数据进行格式转换后，导入 CiteSpace 可视化分析工具，并将运行结果输出。分析结果显示，1992~2019 年间参与反腐倡廉研究的作者共有 324 位，其中发文量最多的学者是邵景均，发文量为 21 篇，即 $N_{\max} = 21$ ，代入公式后计算得出 $M \approx 3.43$ ，取整数为 4，可得出以下结论：发文量在 4 篇及以上的作者可确认为反腐倡廉研究领域的核心作者，如表 1 所示。由表 1 可知，共有邵景均、李斌雄、杜治洲、胡杨、杨绍华、徐喜林、曹建明、张金煌、郭兴全、刘占虎、陈泽伟、李一帆、刘建华、赵喜泉、田旭明、夏云强等 46 位核心作者，总发文量为 258 篇，占有研究文献的 12.46%。

Table 1. Core author statistics (issued volume ≥ 4)
表 1. 核心作者统计表(发文量 ≥ 4)

发文量(篇)	核心作者	发文量(篇)	核心作者	发文量(篇)	核心作者
21	邵景均	6	许云昭	4	丁俊萍
9	李斌雄	6	黄红平	4	唐晓清
9	杜治洲	6	任建明	4	商志晓
8	胡杨	5	张增田	4	吴官正
7	杨绍华	5	蒯正明	4	范春梅
7	徐喜林	5	谢旭人	4	肖芸
7	曹建明	5	吴海红	4	王真
7	张金煌	5	贺国强	4	王平一
6	郭兴全	5	毛昭晖	4	孙载夫
6	刘占虎	5	勾清明	4	毛蔚兰
6	陈泽伟	5	李茜	4	唐贤秋
6	李一帆	5	齐卫平	4	虞崇胜
6	刘建华	4	臧献甫	4	王洪彬
6	赵喜泉	4	于滨	4	李法泉
6	田旭明	4	董瑛		
6	夏云强	4	秦正为		

(三) 反腐倡廉研究文献机构分布

研究机构是研究工作的重要依托平台, 给予学者开展学术研究的物质基础。对核心机构的分析, 不仅可以呈现出研究领域内的权威机构, 更能为研究者开展跨地域的学术交流明确方向。为了精准锁定反腐倡廉研究领域的权威机构与核心单位, 本文根据普赖斯定律统计发文的核心机构, 即 $N \approx 0.749 \times \sqrt{N_{max}}$, (N 为核心作者单位发表最低论文数, N_{max} 为最高产作者单位发表论文数) [2]。通过 CiteSpace 运行结果显示, 发文量最多的机构是中央纪委, 其发文量为 14 篇, 即 $N_{max} = 14$, 代入公式后计算可得, $N \approx 2.8$, 取整数为 3。根据计算结果可得出以下结论: 发文量在 3 篇及以上的研究机构为核心研究机构, 如表 2 所示。

Table 2. Statistics of core institutions (issued volume ≥ 3)
表 2. 核心机构统计表(发文量 ≥ 3)

发文量	核心机构	发文量	核心机构	发文量	核心机构	发文量	核心机构
14	中央纪委	5	大连海事大学 马克思主义学院	4	西南交通大学 马克思主义学院	3	四川省社会科学院
12	武汉大学政治与 公共管理学院	5	甘肃省财政厅党组	4	中共黑龙江省委党校	3	中共山东省委党校
12	武汉大学 马克思主义学院	5	清华大学 马克思主义学院	3	广西大学政治学院	3	北京航空航天大学 大学廉政研究所
9	郑州大学 公共管理学院	5	华东师范大学 政治学系	3	燕山大学经济管理学院	3	河南财政税务 高等专科学校

Continued

8	中共中央党校	4	中共湖北省委党校 党史党建教研部	3	天津师范大学 政治与行政学院	3	南开大学 马克思主义学院
8	西安交通大学 人文学院	4	中共中央党校 党建教研部	3	中共山西省委党校	3	河南教育学院 法律与经济系
8	中共吉林省委	4	福建省委党校	3	中共中央政治局	3	河南省廉政 评价研究中心
7	中共中央纪律 检查委员会研究室	4	南通大学政治学院	3	东北师范大学政法学院	3	东北师范大学 马克思主义学部
7	河南省社会科学院	4	海军大连 舰艇学院政治系	3	中共中央	3	中共天津市委 党校党建教研部
7	中国人民大学 马克思主义学院	4	最高人民检察院党组	3	中共河南省委党校	3	中共中央 党校党建部
6	中国科学技术 大学管理学院	4	江苏师范大学马克思 主义学院	3	电子科技大学	3	清华大学廉政 与治理研究中心
6	福建农林大学 廉政研究中心	4	中共山西省委	3	中共上海市委 党校党史党建教研部	3	湖南师范大学 公共管理学院
5	中央纪委驻财 政部纪检组	4	中共中央文献研究室	3	北京大学医学部纪委	3	吉林大学行政学院
5	武汉大学	4	山东大学	3	福建农林大学 经济与管理学院	3	河南大学 廉政研究中心
5	中共湖南省委	4	西安交通大学 马克思主义学院	3	扬州大学社会发展学院	3	中央纪委研究室
5	武汉大学 廉政研究中心	4	中共天津市委	3	清华大学公共管理学院	3	中共上海市委党校
5	中共湖北省委		中共天津市纪委	3	南通廉政文化研究所		
5	中共北京市委		株洲师范 高等专科学校	3	华中师范大学 马克思主义学院		

从统计数据可知,参与反腐倡廉的研究机构共有 309 个,其中核心机构共有 72 个,其中发文量在 5 篇以上的机构有:中央纪委、武汉大学政治与公共管理学院、武汉大学马克思主义学院、郑州大学公共管理学院、中共中央党校、西安交通大学人文学院、中共吉林省委、中共中央纪律检查委员会研究室、河南省社会科学院、中国人民大学马克思主义学院、中国科学技术大学管理学院、福建农林大学廉政研究中心。普赖斯定律还认为只有核心作者单位发文量约占总发文量的 50%时,学科的高产作者单位群才可以形成[2]。但据研究机构发文量的统计结果显示,核心机构发文总量为 314 篇,仅占文献总量的 15.15%,远没达到 50%的核心认定标准。因此,通过对核心机构的分析可得出以下结论:1992~2019 年,国内反腐倡廉研究尚未形成核心作者单位群。

4. 反腐倡廉研究的可视化分析

(一) 反腐倡廉研究文献的热点剖析

关键词是一篇研究文献主旨大意的高度浓缩,能够以小见大,映射出研究文献的精髓与特色。高频关键词是指在所有研究文献中出现频次较高的关键词,能够体现出一定时间范围内研究领域的热点话题,而高中心性关键词是指在整个文献结构网络中发挥重要衔接作用的节点。本研究借助 CiteSpace 分析软件

对 1992~2019 年反腐倡廉研究文献进行可视化分析。一方面,对软件进行参数设定,确定文献分析的“入围圈”。在分析对象(Node Types)栏勾选关键词(Key Words)选择框,将时区分割(Time Slicing)设为“1992~2019 年”,时间切片(Years Per Slice)设置为 4,数据筛选的阈值设置为 TopN = 50,即提取每 4 年中出现频次排名在前 50 的关键词。另一方面,对成像进行美化调整,选择适当的图谱外形。将图像剪裁方式(Pruning)设置为 Pathfinder 与 Pruning the merged network,可视化(Visualization)设置为静态的聚类视图(Cluster View-Static)与显示整体的分析网络(Show Merged network)。点击运行后得到 672 个节点、2610 条连接线、网络密度为 0.0116 的知识图谱,如图 2 所示。图 2 中的每个节点代表一个关键词,节点越大则关键词出现的频率越高,节点之间的连线表示关键词之间的共现关系,连线愈粗则共现关系越明显,即两个关键词在同一篇或多篇文献中出现的频次越高。由图 2 可知,知识图谱中节点较大的有反腐倡廉、反腐倡廉建设、反腐败斗争、党风廉政建设、反腐倡廉工作、反腐败、中国共产党、预防腐败、党的建设、廉政文化、反腐败工作、制度建设、习近平、反腐倡廉教育、全面从严治党、网络反腐、科学化。值得注意的是,定量分析要求任何的结论都必须依赖于客观的数据,但是可视化知识图谱只能从感性直观的视角观察研究热点,并不能精准反映出关键词的频次与中心性排序。因此,为了进一步提升研究的科学性和可信度,本文依据多诺霍提出的公式,用精准的计算梳理出高频关键词,即 $T = (-1 + \sqrt{1 + 8I}) / 2$,其中 I 为关键词个数, T 为高频关键词出现的最低频次[3]。通过 CiteSpace 导出的数据可知,关键词共有 672 个,即 $I = 672$,代入计算公式后得出 $T \approx 36.164$,取整数为 37。由此可得出以下结论:反腐倡廉研究文献中出现频次在 37 次及以上的关键词可认定为核心关键词。依据关键词出现的频率可知,符合核心关键词标准的关键词仅有 17 个,分别为反腐倡廉(423)、反腐倡廉建设(269)、反腐败斗争(255)、党风廉政建设(160)、反腐倡廉工作(129)、反腐败(104)、中国共产党(102)、预防腐败(90)、党的建设(72)、反腐败工作(70)、廉政文化(62)、制度建设(56)、反腐倡廉教育(50)、习近平(43)、党风廉政建设责任制(42)、廉政文化建设(41)、科学化(40)。与高频关键词相比,高中心性关键词并没有公认的计算公式,但学界普遍认定中心值在 0.1 以上的关键词可认定为高中心性关键词,依据现有的处理数据可知,中心值在 0.1 以上的关键词共有 6 个,分别为反腐败斗争(0.37)、反腐倡廉(0.33)、党风廉政建设(0.18)、反腐败(0.13)、反腐倡廉工作(0.12)、反腐倡廉建设(0.11)。

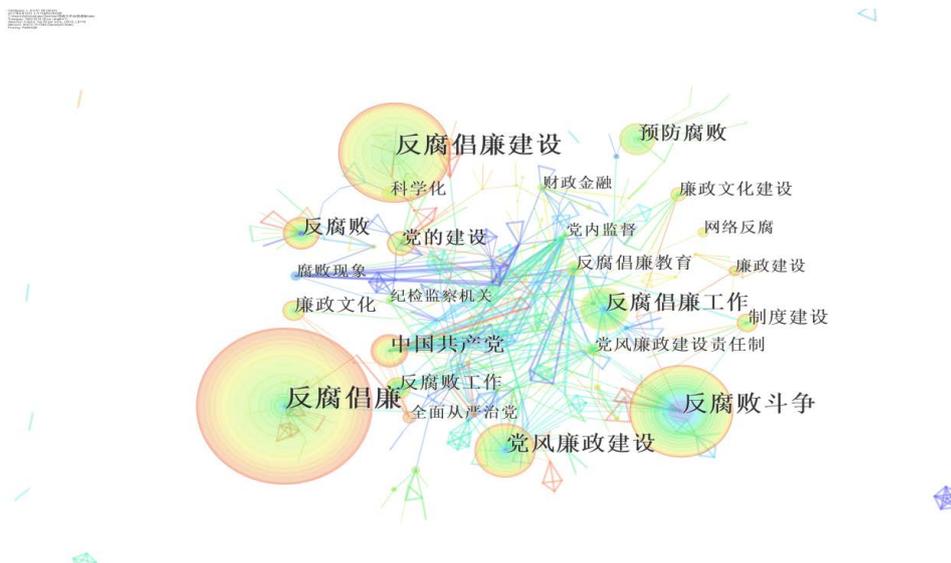


Figure 2. Anti-corruption keyword collaboration map
图 2. 反腐倡廉关键词共现图谱

(二) 反腐倡廉研究文献的主题演化

主题路径演化图谱是研究主题变迁的全景视图,能够呈现某一领域研究主题的传承关系和发展趋势,为后继研究者寻求视角、观点、方法的三维创新提供了方向指引。本文在 CiteSpace 关键词共现知识图谱的基础上,从时间维度对关键词进行划分,形成了关键词 TimeZone View 视图,如图 3 所示。图 3 中横坐标为关键词首次出现的时间,每一个节点都代表一个关键词。值得注意的是,图中节点的大小并不是指关键词在某个年份出现的频次,而是指关键词在整个研究进程中出现的频次,节点之间的连线则同样代表关键词之间的共现关系。

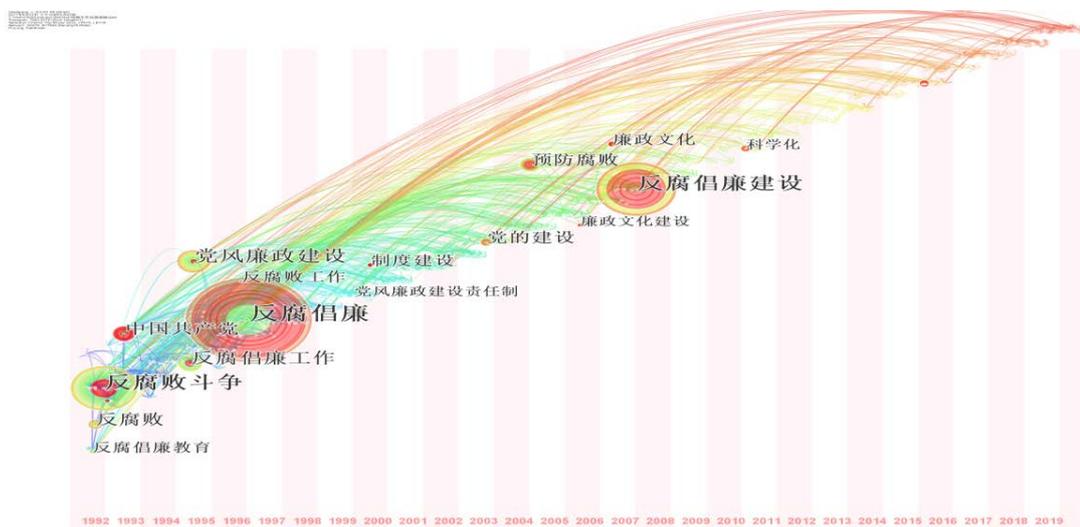


Figure 3. Anti-corruption keyword time zone map

图 3. 反腐倡廉关键词时区图谱

由图 3 可知,1992~2019 年反腐倡廉研究呈现出从聚焦走向多元的趋势。依据主题路径演化,可将研究分为三个阶段:第一,起步阶段(1992~2002 年)。该阶段的研究主题有反腐败、反腐倡廉、中国共产党、反腐倡廉工作、财政金融、制度建设、党风廉政建设责任制、党内监督、腐败现象,集中于反腐败领域的研究,视角相对而言较为聚焦,这与当时党中央的政策方针有密切关系,表明 1992~2001 年党中央高度重视反腐倡廉工作。第二,探索阶段(2003~2012 年)。该阶段出现的研究主题词有反腐倡廉建设、党的建设、预防腐败、廉政文化建设、制度反腐、网络反腐、科学化、科学发展观,表明这一阶段在继承上一阶段研究成果的基础上继续深化,主题词逐渐多元化和新颖化,不仅加入了网络新元素,还上升到了文化层面,研究深度和研究广度都有质的提升。第三,突破阶段(2013~2019 年)。该阶段出现的关键词有全面从严治党、习近平,集中于习近平总书记关于全面从严治党重要论述的研究,这表明十八大以来学界关于反腐倡廉的研究与党中央全面从严治党战略思想具有密切的关系。纵观 1992~2019 年学界研究主题的变迁过程可得出以下结论:其一,反腐倡廉在时间上具有持续性,需要各代共产党人齐心协力;其二,反腐倡廉在内容上具有复杂性,需要研究者透过腐败的外在表征找寻腐败行为滋生的内在根源;其三,反腐倡廉在成效上具有滞后性,需要党和人民在实践中不断检验理论的真理性与实效性。

(三) 反腐倡廉研究文献的前沿动态

突现关键词是指某一研究主题在特定时间内呈现出了急速增长状态,反映了一定时期内研究的前沿动态,对研究者预测未来研究趋势走向提供了客观标准。本文根据 CiteSpace 的突现算法计算出突现关键词,如图 4 所示。图 4 中的“Year”指的是突现词首次在研究文献中出现的时间,“Strength”是指突现

词突现的强度，“Begin”与“End”则是指关键词呈现出突现状态的初始年份与结束年份。由图4可知，在2019年突现的关键词共有6个，分别为中国共产党(11.4689)、习近平(22.2364)、全面从严治党(19.3587)、从严治党(8.613)、党内法规(3.6916)，这一组突现词代表着未来一段时间学者研究的重点领域，引领者反腐倡廉研究走上新平台。

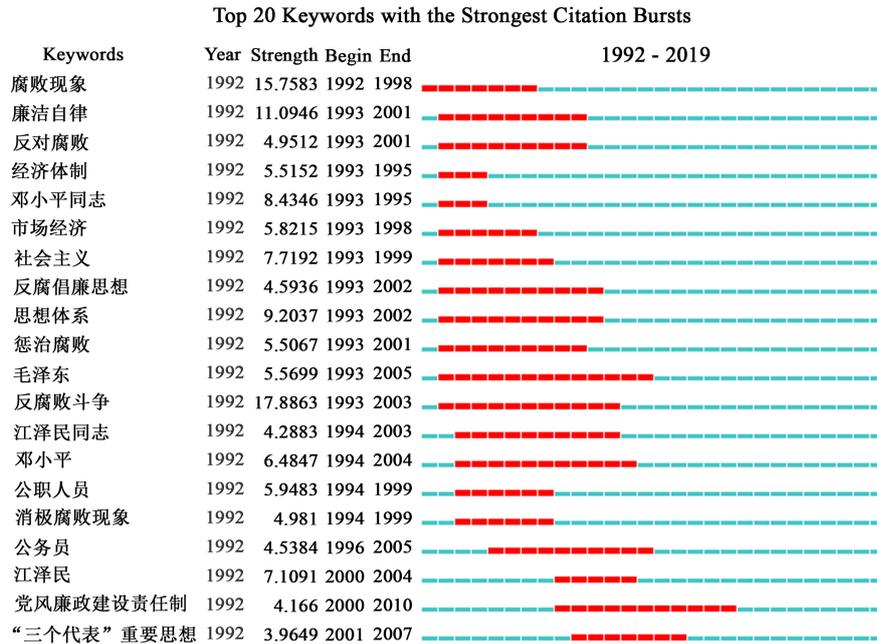


Figure 4. Anti-corruption research literature highlights keywords
图 4. 反腐倡廉研究文献突现关键词

5. 总结

本文借助 CiteSpace 可视化分析软件对 1992~2019 年中国知网(CNKI)中的文献资料进行了量化分析，通过数据的分布和图谱的呈现，可得出以下三点研究结论：

从研究主体上看，参与范围“广”，但合作关系“疏”。已有反腐倡廉研究领域作者与机构众多，凝结成了一股强大的理论生产力量，夯实了反腐倡廉工作的持续推进理论驱动。但是，研究者之间，研究机构之间的脱离关系与孤立状态无疑给未来的研究主体的跨领域交流拉起了无形的障碍，不利于未来反腐倡廉理论的创新和发展，更不利于营造和谐共进的研究氛围。

从研究内容上看，涉及领域“广”，但创新程度“低”。理论生产活动本质上也是一种实践活动，且实践活动的展开离不开其所处的“市民社会”，但由于生产力的不断发展，整个理论生产活动如人类历史一样在经济基础与上层建筑的矛盾运动中不断升级深化。反观当前的研究成果涉及政治、公共管理、教育、法学、工商管理、马克思主义、新闻传播、文学、金融、影视、美容等多个学科视角，形成了跨学科的新型研究范式。但是，通过深入的可视化分析发现，当前研究文献中虽涉及领域广，但研究内容多是对反腐倡廉政策的宏观解读与践行路径的探索，对基层反腐倡廉工作实况与新时代以来反腐倡廉新局面的研究较为匮乏。

从研究走向上看，研究热度“久”，且研究质量“升”。从严治党是中国共产党一直以来的优良传统，反腐倡廉作为从严治党的重要环节一直备受关注。新时代以来，习近平总书记高度重视党的建设，将全面从严治党上升到了国家的战略高度，激发了学界对反腐倡廉的研究热忱。随着中国特色社会主义

事业迈入新时代，我国面临的国际形势更严峻，对党的执政能力要求愈发严格，这也预示着未来反腐倡廉必定成为学界研究的持续性热点。但是，我们在维系研究热度的同时必须追求实质性的理论创新，在继承既有研究成果的基础上实现研究方法的突破与研究视域的扩展。

基金项目

本文系 2019 年度四川警察思想政治研究中心资助项目“新时代公安机关全面从严治党创新路径研究”（项目编号：CJS19A02）；2019 年度成都市哲学社会科学规划重点资助项目“新时代弘扬‘五四精神’与培育‘时代新人’的协同路径研究”（项目编号：2019A08）阶段性成果。

参考文献

- [1] 洪波. 我国高等职业教育研究的知识图谱分析——基于 1992~2016 年核心期刊文献[J]. 职业技术教育, 2017, 38(6): 45-50.
- [2] 纪谦玉. 新媒体时代大学生思想政治教育研究文献的计量分析[J]. 黑龙江高教研究, 2016(4): 116-120.
- [3] 魏宏君, 唐保庆. 近十年国际教师教育研究的热点与知识基础分析[J]. 黑龙江高教研究, 2016(2): 101-107.