

基于CiteSpace的网络舆情研究现状可视化分析

施璇

上海工程技术大学, 上海

收稿日期: 2022年8月19日; 录用日期: 2022年9月12日; 发布日期: 2022年9月21日

摘要

网络舆情是社会舆情在互联网空间的映射, 是社会舆情的直接反映。近年来, 网络舆情对社会稳定和生活秩序的影响与日俱增。对网络舆情研究领域的相关文献进行可视化分析, 有助于把握我国网络舆情研究现状, 对进一步的研究提供依据和参考。本文以CNKI数据库2012~2022年收录的2033篇CSSCI文献为基础, 利用CiteSpace可视化分析工具对网络舆情相关文献的年代、作者、研究机构、关键词进行统计分析。研究发现, 近年来网络舆情领域的发文量经历快速上升后发展趋于平稳; 出现了一些有影响力的核心作者、科研机构, 但作者间、机构间的合作强度低; 近几年的研究热点主要集中在网络舆情的识别与传播、监管与治理、网络舆情治理的参与主体及其相互关系上。因此, 未来应加强网络舆情研究领域作者间、机构间的合作, 更多关注信息生态、政府回应等热点问题。

关键词

网络舆情, CiteSpace, 可视化分析, 研究热点

Visual Analysis of the Research Status of Online Public Opinion Based on CiteSpace

Xuan Shi

Shanghai University of Engineering Science, Shanghai

Received: Aug. 19th, 2022; accepted: Sep. 12th, 2022; published: Sep. 21st, 2022

Abstract

Internet public opinion is the mapping of social public opinion in Internet space and the direct reflection of social public opinion. In recent years, the influence of Internet public opinion on so-

cial stability and living order is increasing day by day. Visual analysis of the relevant literature in the field of online public opinion research will help to grasp the current situation of online public opinion research in China and provide basis and reference for further research. Based on 2033 CSSCI documents collected in CNKI database from 2012 to 2022, this paper uses CiteSpace visual analysis tool to make statistical analysis on the age, author, research institution and keywords of online public opinion related documents. The research finds that the number of articles published in the field of online public opinion has experienced a rapid rise in recent years, and the development tends to be stable. There are some influential core authors and scientific research institutions, but the cooperation intensity between authors and institutions is low. In recent years, the research focuses on the identification and dissemination of network public opinion, supervision and governance, the participants of network public opinion governance and their relationships. Therefore, in the future, we should strengthen cooperation among authors and institutions in the field of online public opinion research, and pay more attention to hot issues such as information ecology and government response.

Keywords

Network Public Opinion, CiteSpace, Visual Analysis, Research Hotspots

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

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



Open Access

1. 引言

网络舆情社会舆情在互联网空间的映射,是社会舆情的直接反映。网络舆情以网络为载体,以事件为核心,是社会心态最直观的反应,也是社会治理水平的表现[1]。一般来说,网络舆情是互联网上人们关于某类事件的认知、态度、情感和行为倾向的集合[2]。目前,网络是信息传播的重要载体,在此背景下,及时、准确地对网络舆情的发展进行识别、监管和治理显得尤为重要。因此,了解网络舆情研究领域的研究现状和发展趋势具有重要意义。本文运用信息可视化分析软件 CiteSpace,以中国知网(CNKI)收录的 CSSCI 文献为数据来源,对网络舆情领域的研究现状和研究热点进行分析,以便对后续的研究提供参考。

2. 数据来源与研究方法

2.1. 数据来源

本文在中国知网(CNKI)的高级检索界面中,以“网络舆情”为检索词,选择时间为 2012~2022 年,由于 CNKI 收录期刊众多,期刊差距较大,因此选择中文核心期刊和 CSSCI 来源期刊,以更好地反映中国网络舆情研究进展,共得检索结果 2054 条。对初次检索的结果进行人工筛选,去除期刊会议征稿、卷首语、个人学术成果介绍、科研机构介绍、书评以及署名为课题组及无作者等文献,最终得到 2033 篇相关文献。

2.2. 研究方法

在学术研究中,主要有 Bibexcel、ECharts、VanagePoint 以及 Citespace 四种可视化分析工具[3]。其中,CiteSpace 工具是由美国费城德雷克塞尔大学陈超美教授所开发的一种基于引文分析理论的科技文本挖掘

及可视化分析软件。其是应用 Java 语言开发的一款信息可视化软件,它主要基于共引分析理论(co-ciation)和寻径网络算法(path Finder)等,对文献信息进行细致的归类分析,其独有的时区视图和突变词检测功能更是能帮助学者快速发现课题的热点和发展前沿[4]。本文借助 CiteSpace 软件,使用 5.8.R3 版本,以网络舆情研究领域的中文核心期刊和 CSSCI 来源文献为对象,对文献的作者、机构以及关键词做可视化分析。

3. 网络舆情研究的基本情况

3.1. 文献数量特征

由图 1 来看,网络舆情研究在 2012 年至 2014 年处于快速上升期,2014 年至 2017 年处于发展期,并在 2017 年达到顶峰,17 年仅收录在 CSSCI 来源期刊中的网络舆情文献就已达 232 篇。2017 年至 2020 年舆情领域研究趋于平稳,文献发表数量随当年有无热点事件发生而上下波动,但总体维持在一个较高水平[5]。值得关注的是,2020 年至今,有关网络舆情的发文量持续降低,该研究领域的发文量是否会保持高位波动态势,进入研究的相对成熟期,还是从此持续走低,进入低迷期,有待进一步的数据挖掘与分析[6]。

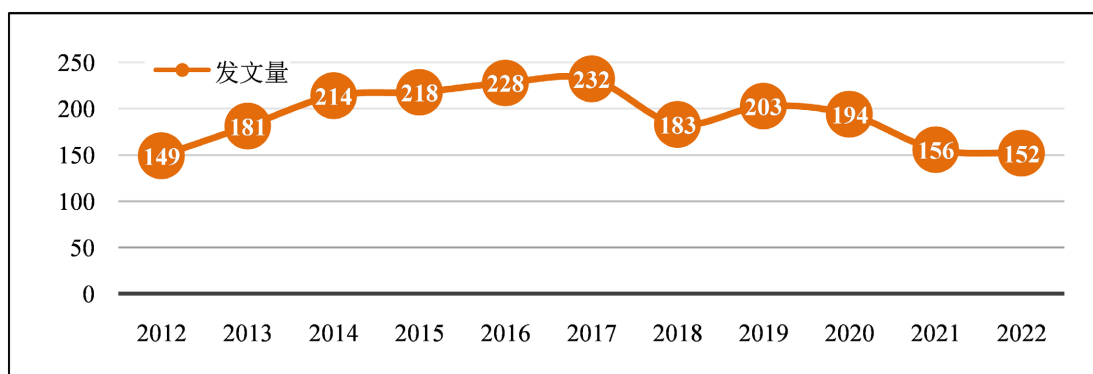


Figure 1. Annual trend of literature publication

图 1. 文献发表年度趋势

3.2. 文献作者分析

通过分析文献作者之间的合作网络以及在该研究领域的发文量,可以知道该领域中的重要研究团队关键研究人员。借助 CiteSpace 软件,对文献作者进行可视化分析。将 CiteSpace 软件中的时间跨度选择为 2012~2022 年,时间单位设置为 1 年,Node Types 设置为 Author,运行结果如图 2。

由图 2 可知,文献作者合作网络中共有文献作者合作网络中共有 448 个节点,每一个节点代表一位作者,即 2033 篇文献涉及 448 位作者。节点大小反映作者发文量,在 2033 篇文献中,发表论文最多的作者是兰月新,为 58 篇。以作者发文量为依据,兰月新、陈福集、黄微等学者是网络舆情研究领域的核心作者。节点之间的连线表示作者间存在合作关系,连线的粗细代表作者间合作关系的紧密程度。图谱的网络密度为 0.0033,作者间的合作强度较低。从连线情况看,存在几个成果丰硕的研究团队(分别以兰月新教授、陈福集教授、黄微教授、曾润喜教授为首的团队),但图谱整体比较松散,作者间的连线数量较少,没有形成联系紧密的合作网络,说明该领域大多数作者仍处于独立研究状态。学者之间的这种分散研究的状态,显然不利于网络舆情研究领域部分研究问题达成共识,因此,学者之间的交流与合作需要加强。

CiteSpace, v. 5.8.R3 (64-bit)
 August 25, 2022 10:18:38 PM CST
 WoS: /Users/zmyddst/data for citespace/网络舆情/data
 Timespan: 2012-2022 (Slice Length=1)
 Selection Criteria: g-index (k=25), LRF=3.0, L/N=10, LBY=5, e=1.0
 Network: N=448, E=335 (Density=0.0033)
 Largest CC: 40 (8%)
 Nodes Labeled: 1.0%
 Pruning: None

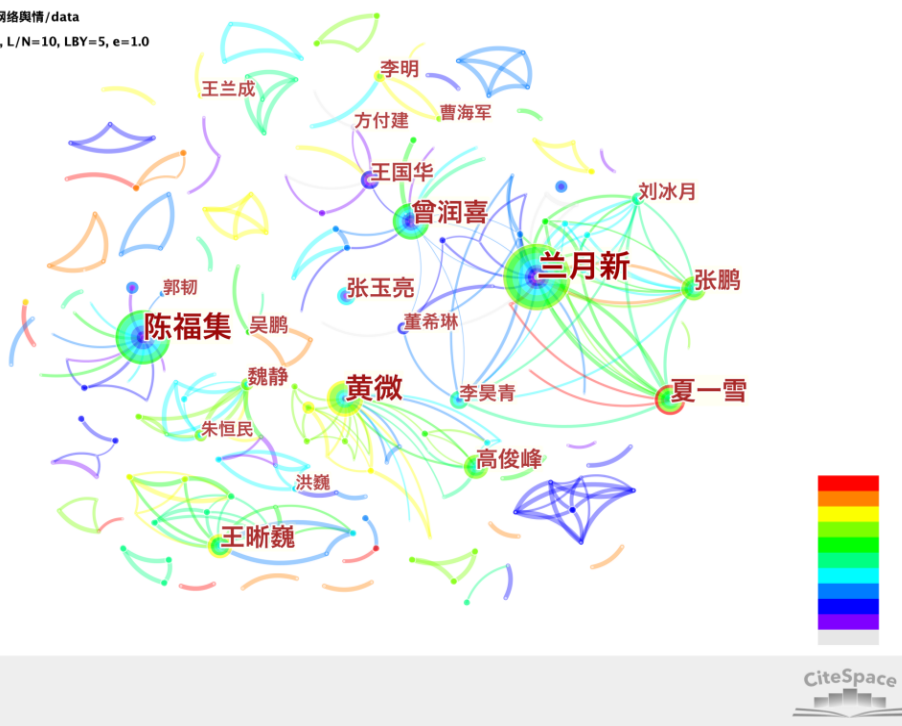


Figure 2. Literature author cooperation network
 图 2. 文献作者合作网络

CiteSpace, v. 5.8.R3 (64-bit)
 August 25, 2022 11:19:03 PM CST
 WoS: /Users/zmyddst/data for citespace/网络舆情/data
 Timespan: 2012-2022 (Slice Length=1)
 Selection Criteria: g-index (k=25), LRF=3.0, L/N=10, LBY=5, e=1.0
 Network: N=398, E=126 (Density=0.0016)
 Largest CC: 27 (6%)
 Nodes Labeled: 1.0%
 Pruning: None

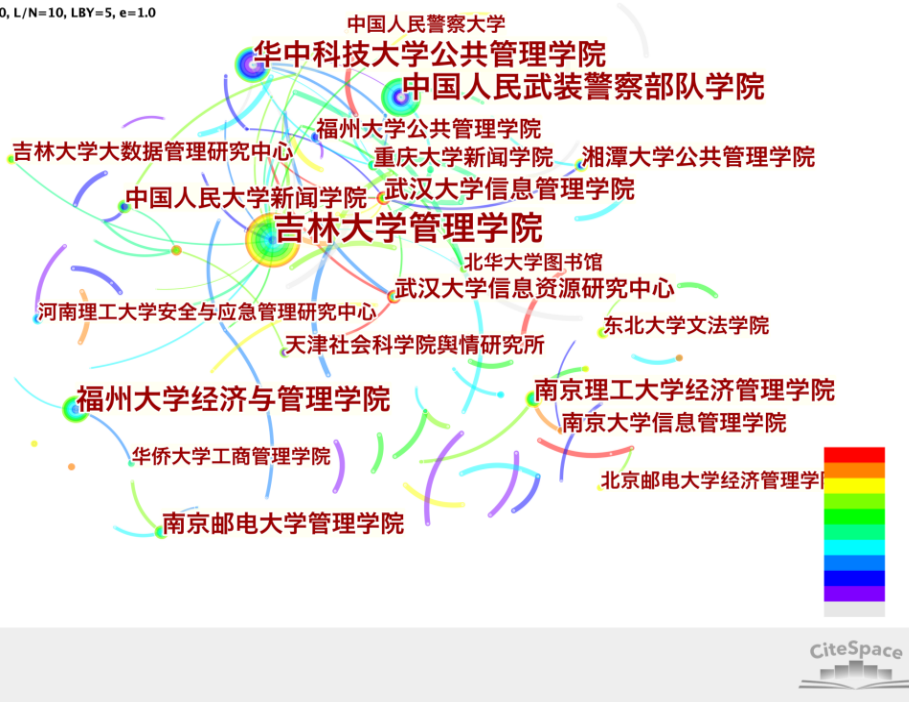


Figure 3. Analysis of cooperation between research institutions
 图 3. 研究机构合作分析

3.3. 研究机构分析

发文机构合作网络反映机构间的科研实力对比和知识的流动状况[7]。将 CiteSpace 软件中 Node Types 设置为 Institution，运行结果如图 3。

由图 3，图谱共有 398 个节点，126 条连线，网络密度为 0.0016，机构间的合作强度较低。节点越大，说明机构的发文量越多。图中可以看出，发文量最多的是吉林大学管理学院。表 1 为发文量前 10 的机构，如表 1 所示，网络舆情研究领域发文机构主要集中于高等院校。发文量最多的机构是吉林大学管理学院，其次是中国人民武装警察部队学院和华中科技大学公共管理学院。发文量较多的机构间连线数量少，合作性不强。未来实力较强的机构间应加强合作，推动研究的发展。

Table 1. Top 10 institutions by the number of publications
表 1. 发文量前 10 机构

序号	研究机构	年份	发文量
1	吉林大学管理学院	2014	89
2	中国人民武装警察部队学院	2012	52
3	华中科技大学公共管理学院	2012	48
4	福州大学经济与管理学院	2015	38
5	南京理工大学经济管理学院	2015	25
6	武汉大学信息管理学院	2013	23
7	南京邮电大学管理学院	2015	22
8	中国人民大学新闻学院	2012	21
9	湘潭大学公共管理学院	2013	20
10	武汉大学信息资源研究中心	2015	19

4. 网络舆情研究的关键词分析

4.1. 关键词共现分析

关键词共现分析是知识图谱分析的基础，通过识别高频关键词和高中介中心性关键词可以揭示某一研究领域的重要知识基础[8]。将 CiteSpace 软件中的 Node Type 设置为 Keyword，得到网络舆情研究领域文献的关键词共现图(图 4)。

如图 4，关键词共现图中共有节点 528 个，连线 1044 条，关键词间具有密切的联系。节点越大说明该关键词出现的频率越高，从图中可以看出“网络舆情”、“突发事件”和“大数据”出现的频次较高。即大多数文献都是围绕“网络舆情”、“突发事件”、“大数据”展开的。

为了更清晰地表现关键词的地位和关系，对关键词的频次和中心度进行统计，综合考虑后，分别选取频次排名前 10 和中心度排名前 10 的关键词。表 2 为网络舆情研究领域的高频关键词和高中心度关键词。



Figure 4. Keyword co-occurrence diagram
 图 4. 关键词共现图

Table 2. High-frequency keywords and high centrality keywords
 表 2. 高频关键词和高中心度关键词

排序	高频关键词		排序	高中心度关键词	
	关键词	频次/次		关键词	中心度
1	网络舆情	1266	1	网络舆情	0.94
2	突发事件	134	2	突发事件	0.24
3	大数据	70	3	大数据	0.16
4	新媒体	64	4	舆情	0.16
5	微博	61	5	舆情传播	0.16
6	舆情	54	6	新媒体	0.13
7	舆情传播	48	7	微博	0.13
8	高校	43	8	舆情治理	0.10
9	网络治理	38	9	情感分析	0.08
10	电子政务	36	10	高校	0.07

分析表 2, 从关键词频次可以看出研究主题较多, 其中网络舆情最高(1266 次), 其次是突发事件(134 次)、大数据(70 次)、新媒体(64 次)等。从中心度来看, 网络舆情、突发事件、大数据、舆情、舆情传播、新媒体、微博、舆情治理的中介中心性都达到或超过了 0.1, 在网络舆情研究领域中较为重要。从出现频次来看, 网络舆情的研究对于突发事件和大数据方面较为关注, 其次是新媒体、舆情及舆情传播、网络治理等方面。

4.2. 关键词聚类分析

关键词聚类分析是将研究领域中的相似度较高的研究点进行归类, 以得到更精炼的主题[9]。在关键词共现操作的基础上, 进行关键词聚类分析, 最终生成关键词聚类图(图 5)。

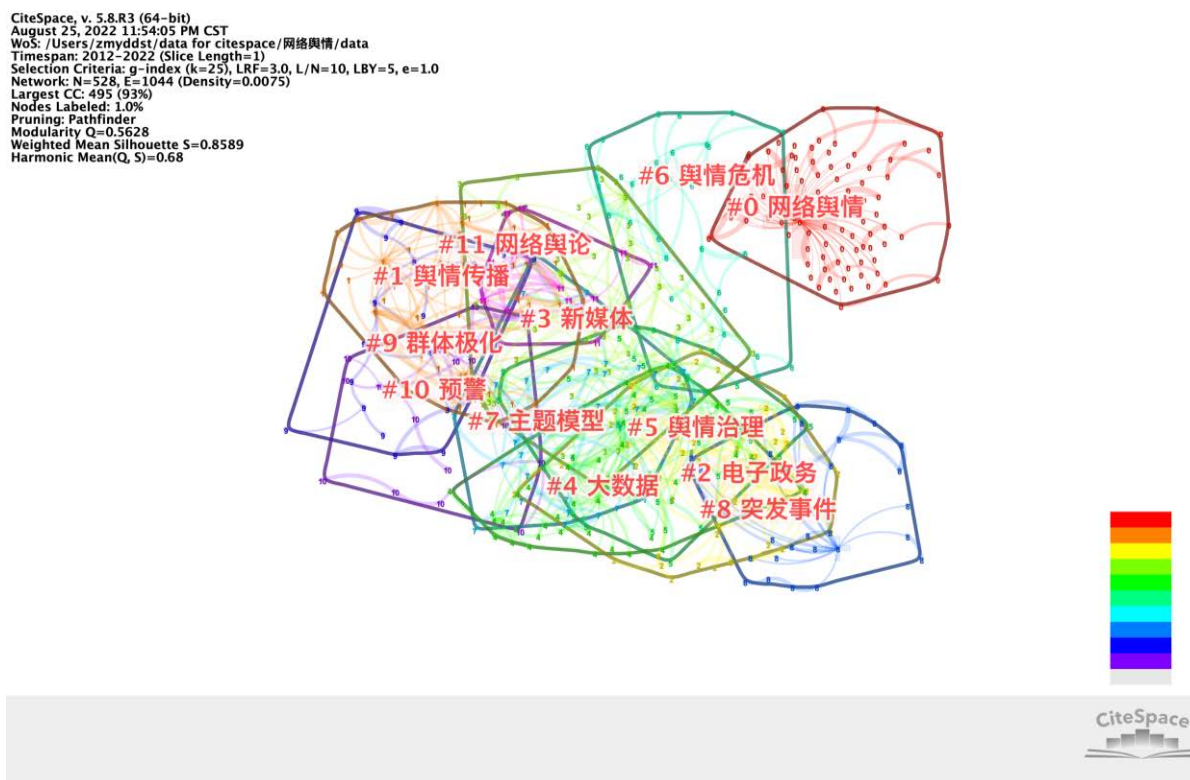


Figure 5. Keyword clustering diagram

图 5. 关键词聚类图

一般来说, 模块值 $Q > 0.3$ 意味着划分出来的社团结构是显著的, 当平均轮廓值 $S > 0.7$ 时, 聚类是高效率令人信服的。由图 5 可知, 模块值 $Q = 0.5628$, 说明聚类结果结构显著, 平均轮廓值 $S = 0.8589$, 聚类高效且结果令人信服。综上所述, 该图谱有效。分析图谱共得到 12 个聚类, 分别为: #0 网络舆情、#1 舆情传播、#2 电子政务、#3 新媒体、#4 大数据、#5 舆情治理、#6 舆情危机、#7 主题模型、#8 突发事件、#9 群体极化、#10 预警、#11 网络舆论。通过对前 10 个聚类标签下包含的重点关键词进行梳理, 得到聚类标签及主题关键词(如表 3)。

从聚类结果看, 学者对网络舆情的研究主要集中在以下几个方面:

(一) 网络舆情的识别与传播

对网络舆情的发展进行及时、有效的识别可以从初期对舆情态势进行监测、了解、干预以及疏导,

有助于为相关部门工作决策提供理论和数据支持[10]。主要为聚类#0、#1、#7，其中包括网络舆情、舆情传播、情景模型等关键词。学者们利用挖掘算法等先进手段，对网络舆情的发展进行及时的、有效的识别，对舆情的传播规律和传播机制进行研究。

(二) 网络舆情监管与治理

主要为聚类#2、#4、#5、#6、#8，其中包括网络治理、网络谣言、舆情检测、舆论引导、舆情危机、治理、政府干预等关键词。

(三) 网络舆情治理的参与主体及其相互关系

从网络舆情治理的参与主体及其相互关系入手、对新媒体、高校、政府等参与主体进行研究，如聚类#3、#9，代表关键词如新媒体、自媒体、政府官媒、高校。

Table 3. Clustering tags and keywords

表 3. 聚类标签及关键词

聚类序号	聚类名称	聚类规模	轮廓值	平均年份	关键词
0	网络舆情	95	0.965	2016	网络舆情、多媒体、思政教育
1	舆情传播	55	0.843	2015	微博、舆情传播、舆情演化
2	电子政务	53	0.867	2014	电子政务、网络治理、公共管理
3	新媒体	50	0.849	2015	新媒体、舆情、高校
4	大数据	48	0.759	2016	大数据、情感分析、网络谣言
5	舆情治理	44	0.809	2017	舆情监测、知识图谱、舆论引导
6	舆情危机	34	0.83	2014	政府、舆情危机、治理
7	主题模型	34	0.786	2016	社交网络、社交媒体、舆情事件
8	突发事件	30	0.784	2017	突发事件、事理图谱、政府干预
9	群体极化	19	0.92	2015	群体极化、社会心态、政府决策

4.3. 关键词突现分析

CiteSpace 中的关键词突现图谱用以展示关键词出现频率突然变化的情况，可考察突然出现、频率较高的关键词，这类词被称为“突现词”(Burst Terms)，通常是某领域的研究热点与前沿[11]。在控制面板 Burstness 中发现共有 32 个突现词，根据突变强度(Strength)选择排名前 25 的突现词(如图 6)。

在 25 个突现词中已经有 13 个(电子政务、信息社会、政务微博、虚拟社会、预警、微博、电子治理、公共管理、新浪微博、政府治理、意识形态、预测、多媒体)停止突变，表明这些关键词虽然曾经受到高度关注，但近年来有所弱化，出现的频率降低。有 12 个突现词(舆情事件、社交网络、社交媒体、知识图谱、新时代、政府回应、舆情治理、情感分析、信息生态、新冠疫情、社会治理、影响因素)出现较晚，突变强度较大且至今仍未停止突变，是网络舆情研究领域当前以及未来一段时间内的热点，是该研究领域需要大力挖掘的前沿方向。

Top 25 Keywords with the Strongest Citation Bursts

Keywords	Year	Strength	Begin	End	2012-2022
电子政务	2012	13.94	2012	2013	
信息社会	2012	7.44	2012	2013	
政务微博	2012	4.87	2012	2014	
虚拟社会	2012	3.63	2012	2013	
预警	2012	3.27	2012	2013	
微博	2012	3.16	2012	2013	
电子治理	2012	5.84	2014	2015	
公共管理	2012	5.08	2015	2016	
新浪微博	2012	3.85	2015	2016	
政府治理	2012	4.07	2016	2018	
舆情事件	2012	4.52	2018	2022	
意识形态	2012	3.29	2018	2019	
预测	2012	3.12	2018	2019	
社交网络	2012	4.34	2019	2022	
社交媒体	2012	3.55	2019	2022	
知识图谱	2012	3.42	2019	2022	
多媒体	2012	3.07	2019	2020	
新时代	2012	3.06	2019	2022	
政府回应	2012	2.99	2019	2022	
舆情治理	2012	6.94	2020	2022	
情感分析	2012	5.91	2020	2022	
信息生态	2012	4.07	2020	2022	
新冠疫情	2012	3.81	2020	2022	
社会治理	2012	3.17	2020	2022	
影响因素	2012	2.99	2020	2022	

Figure 6. Top 25 keywords with the strongest citation bursts

图 6. 突变强度排名前 25 的突现词

5. 结论与展望

5.1. 研究结论

本文在中国知网(CNKI)网络舆情研究领域的 CSSCI 文献数据的基础上,运用 CiteSpace 软件对数据进行知识图谱可视化分析。从文献发表的时间、发文量、作者以及研究机构四个方面对网络舆情的研究现状进行描述性分析,运用关键词共现分析、聚类分析及突现分析的方法呈现网络舆情研究的热点研究主题及前沿趋势,得出以下结论:

1) 从研究时间来看,网络舆情研究在 2012 年至 2014 年处于快速上升期,2014 年至 2017 年处于发展期,并在 2017 年达到顶峰,17 年仅收录在 CSSCI 来源期刊中的网络舆情文献就已达 232 篇。2017 年至 2020 年舆情领域研究趋于平稳。2020 年至今,有关网络舆情的发文量持续降低,研究进入相对成熟期,还是从此进入低迷期,有待进一步的数据挖掘与分析。

2) 在研究作者方面, 2033 篇文献共涉及 448 位作者, 其中, 兰月新、陈福集和黄微发文的引用频次较高。作者间存在一些小型合作网络, 同时, 该研究领域内存在几个成果丰硕的研究团队(分别以兰月新教授、陈福集教授、黄微教授、曾润喜教授为首的团队)。在研究机构方面, 网络舆情研究领域发文机构主要集中于高等院校, 共有 398 家研究机构。发文量最多的机构是吉林大学管理学院, 为 89 篇。发文量较多的机构间连线数量少, 合作性不强。目前大多数研究处于独立研究状态, 学者与机构间的合作需要加强。

3) 由网络舆情研究的关键词分析得出, 学者对该领域的研究主要集中在以下几个方面: 网络舆情的识别与传播; 网络舆情监管与治理; 网络舆情治理的参与主体及其相互关系。从关键词突现图可以看出, 新时代、政府回应、情感分析、信息生态、新冠疫情将是网络舆情前沿研究中需要大力挖掘的。

5.2. 展望

在未来针对网络舆情的研究里, 学者们可以多开展主题会议或者论坛活动, 拓展研究思路, 汇集研究知识成果[12], 进一步加强与其他学者的合作。同时, 机构与机构间也应加强合作, 以促进本领域的研究发展。

根据关键词突现图, 未来网络舆情研究应重点挖掘新时代、政府回应、情感分析、信息生态、新冠疫情等方面, 除此之外, 应注意推进网络舆情研究的多学科交叉研究, 促进网络舆情研究体系的形成。

参考文献

- [1] 殷轶. 网络舆论生态特点、问题及治理路径——以 2021 年河南省网络舆情事件为例[J]. 郑州航空工业管理学院学报(社会科学版), 2022, 41(2): 101-106. <https://doi.org/10.19327/j.cnki.zuaxb.1009-1750.2022.02.016>
- [2] 张筱荣, 郭圳凝. 突发公共卫生事件网络舆论的生成演化及治理研究[J]. 江苏海洋大学学报(人文社会科学版), 2022, 20(4): 23-35.
- [3] 王星星, 唐可歆, 戴熙月. 我国港口物流研究现状——基于 CiteSpace 的知识图谱分析[J]. 中国储运, 2022(8): 63-65. <https://doi.org/10.16301/j.cnki.cn12-1204/f.2022.08.055>
- [4] 邵际珍, 陈卓武, 陈嘉锐, 胡凤. 基于 CiteSpace 的大学生创新创业研究可视化分析[J]. 科技管理研究, 2022, 42(10): 98-104.
- [5] 张福俊, 李玉华, 庄晓, 张海艳. 1998-2019 年国内舆情研究热点及主题演化可视化分析[J]. 图书情报研究, 2021, 14(2): 75-82+120.
- [6] 赵蕾. 我国网络舆情研究的知识图谱——基于 CiteSpace 软件的可视化分析[J]. 秘书, 2020(4): 40-56.
- [7] 曹丽英, 邸玉琦, 陈帅. 基于知识图谱的玉米病虫害研究可视化分析[J]. 东北农业科学, 2022, 47(1): 145-150. <https://doi.org/10.16423/j.cnki.1003-8701.2022.01.032>
- [8] 韦长传, 仝德, 袁玉玺, 储君. 城中村研究热点及区域差异——基于 CiteSpace 的文献计量分析[J]. 地域研究与开发, 2022, 41(3): 68-74.
- [9] 樊自甫, 肖文睿. 数字经济和实体经济融合发展趋势及对策研究——基于 CiteSpace 知识图谱的分析[J]. 重庆邮电大学学报(社会科学版), 2022, 34(3): 87-96.
- [10] 张鑫, 刘鑫雅, 张雅文. 我国网络舆情识别的研究脉络、核心主题与发展趋势分析[J]. 河北科技图苑, 2021, 34(3): 89-96. <https://doi.org/10.13897/j.cnki.hbkjty.2021.0054>
- [11] 秦智聘, 林强, 李保东. 基于 CiteSpace 的供应链研究演进动态、热点与前沿分析[J]. 重庆第二师范学院学报, 2022, 35(3): 21-29+127.
- [12] 张颖雯, 檀秀侠. 后疫情时代中国大学生就业知识图谱研究——基于 CiteSpace 的计量可视化分析[J]. 中国大学生就业, 2022(15): 21-28. <https://doi.org/10.20017/j.cnki.1009-0576.2022.15.003>