

# 基于文献计量学方法的我国知识图谱研究的回顾与展望<sup>\*</sup>

颜鑫<sup>1</sup> 孙悦<sup>2</sup> 孙芸<sup>1</sup> 王倩飞<sup>1△</sup>

(<sup>1</sup> 济宁医学院医学信息工程学院,日照 276826; <sup>2</sup> 南开大学商学院,天津 300071)

**摘要** 以中国知网(CNKI)《中国学术期刊(网络版)》(CAJD)为数据来源,以知识图谱相关论文为研究对象,利用文献计量学方法从年代、作者、机构、研究热点等方面进行统计分析,以揭示我国知识图谱研究的现状及趋势。

**关键词** 知识图谱;CNKI;文献计量

中图分类号:G250.252 文献标识码:A 文章编号:1000-9760(2017)12-453-04

## Review and prospect of knowledge mapping research based on bibliometrics in China

YAN Xin<sup>1</sup>; SUN Yue<sup>2</sup>; SUN Yun<sup>1</sup>; WANG Qianfei<sup>1</sup>

(<sup>1</sup> School of Medical Information Engineering, Jining Medical University, Rizhao 276826, China;

<sup>2</sup> Business School, Nankai University, Tianjin 300071, China)

**Abstract:** The paper takes China Academic Journal Network Publishing Database (CAJD) as the data source, the research papers related to the knowledge mapping as the research object, and uses the bibliometrics method to analyze from some aspects, such as the time, authors, institutions and research hotspots to reveal the status quo and trend of knowledge mapping in China.

**Keywords:** Knowledge mapping; CNKI; Bibliometrics

知识图谱是以科学学为基础,涉及应用数学、信息科学等多门学科的交叉领域,是科学计量学和信息计量学的新发展<sup>[1]</sup>。它以知识域为对象,显示科学知识的发展进程与结构关系,旨在揭示科学发展态势,对科学知识的结构、关系与演化过程进行可视化呈现,具有可视化和形象化等特征<sup>[2]</sup>。自 2005 年知识图谱概念被引入国内以来<sup>[3]</sup>,广泛应用于多个领域并且取得了丰硕的成果。本文旨在利用文献计量学方法对我国知识图谱研究论文的年代、作者、机构、研究热点进行定量研究,从而揭示我国知识图谱研究的发展历史、现状,为知识图谱未来的发展提供有益借鉴。

## 1 数据来源

本文以中国知网(CNKI)《中国学术期刊(网

络版)》(CAJD)为数据来源,检索策略为:(篇名知识图谱 或者 关键词 知识图谱,时间为不限-2016 年)。数据收集时间为 2017 年 6 月 20 日。经过数据清洗后,将 1898 篇研究论文作为统计资料,按照 Notefirst 格式套录论文题目、作者、作者单位、年代、关键词等数据,并采用 Bicomb(书目共现分析系统)、Excel 对数据进行统计分析。

## 2 结果分析

### 2.1 年代分布

科学文献作为科学信息的主要载体形式,其数量变化情况是反映科学发展的重要标志<sup>[4]</sup>。2005-2016 年 12 年间共计发文 1898 篇。根据论文数量特征,可将知识图谱研究分为三个阶段。2005-2008 年是知识图谱研究的初期阶段,年发文量不足 20 篇,共计发文 36 篇,占 1.90%(36/1898)。2009 年起知识图谱研究进入新的发展时期,2009~2012 年论文年增长速度快,共计发文 394 篇,占 20.76%(394/1898)。2013 年起,研究进入蓬勃发

<sup>\*</sup> [基金项目] 济宁医学院 2015 年度大学生创新训练计划立项项目(cx2015066); 济宁医学院青年基金项目(JYQ2011IKZ004)

<sup>△</sup> [通信作者] 王倩飞, qianfei wang@126.com

展阶段,2013 - 2016 年共计发文 1468 篇,占 77.34% (1468/1898)。这一时期知识图谱应用研究

得到学术界的高度重视及认可。见表 1。

表 1 我国知识图谱研究论文的年代分布

年代	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	N
n	2	6	11	17	41	68	107	178	246	293	371	558	1898

### 2.2 作者分析

本文对全部作者的发文量进行统计,其中发文 12 篇及以上的作者共 15 人,是我国知识图谱研究的高产作者。发文量最多的是刘则渊教授,于 2005 年首次将知识图谱这一概念引入中国并在科学学领域开展了应用性研究,并牵头创建了网络 - 信息 - 科学 - 经济计量 (WISE) 实验室<sup>[5]</sup>,带领 WISE 实验室团队开拓了知识计量学与知识图谱研究的新方向。累计发表论文 37 篇,其中 12 篇为第一作者,不仅高产而且研究连续性好。其次是大连理工大学侯海燕教授,发表论文 25 篇,其中 12 篇为第一作者,研究持续稳定。研究主要集中在基于知识图谱原理及可视化方法研究国际科学学,特别是科学计量学研究热点及演进趋势<sup>[6-7]</sup>,界定学术群体及学科代表性人物,为相关研究提供有益的启示。发文量第三位的是南京大学宗乾进教授,发表论文 25 篇,其中以第一作者发文 10 篇,采用知识图谱绘制工具对我国图书馆学、情报学、金融学等领域文献进行可视化分析<sup>[8-13]</sup>,揭示各领域的研究热点和知识来源谱系。见表 2。

表 2 我国知识图谱研究的高产作者信息 (发文量≥12 篇)

序号	作者	机构	发文量	第一作者 发文量	初始 年份
1	刘则渊	大连理工大学	37	12	2005
2	侯海燕	大连理工大学	25	12	2006
3	宗乾进	南京大学	25	10	2011
4	许振亮	大连理工大学	22	13	2007
5	侯剑华	大连大学	20	9	2007
6	袁勤俭	南京大学	20	1	2011
7	姜春林	大连理工大学	18	14	2006
8	汤建民	浙江树人大学	18	10	2009
9	邱均平	武汉大学	16	13	2010
10	赵蓉英	武汉大学	16	10	2010
11	王琪	天津师范大学	15	14	2010
12	陈悦	大连理工大学	14	6	2005
13	王晴	山西大学	13	8	2013
14	陈瑜林	广东嘉应学院	12	11	2011
15	蔡建东	河南大学	12	6	2011

### 2.3 机构分析

结果显示共涉及 505 个不同的机构,发文量≥12 篇即年均发文一篇以上的有 31 个,合计发文 725 篇,占论文总量的 38.20%。除中国科学院国家科学图书馆外,全部为高等学校,其中综合性大学最多,其次是师范类大学。发文量≥50 篇的机构有 3 所,分别为大连理工大学、武汉大学、南京大学,发文量 20 ~ 50 篇的机构有 10 个,发文量 < 20 篇的有 18 个。

大连理工大学发文量居首,共发表论文 93 篇。主要来自于大连理工大学 21 世纪发展研究中心暨 WISE 实验室,该研究团队主要包括刘则渊、侯海燕、姜春林、陈悦、陈超美等。此外,大连理工大学人文社会科学学院、公共管理与法学学院等部门也借助本校丰富的研究资源开展知识图谱相关研究。发文量第二位的是武汉大学,发表论文 80 篇。主要来自武汉大学信息资源管理学院、中国科学评价研究中心等机构。研究团队是由邱均平、赵蓉英带领的团队。发文量居第三的是南京大学,发表论文 56 篇。作者主要来自南京大学信息管理学院,团队成员以宗乾进、袁勤俭、沈洪洲为核心。见表 3。

表 3 我国知识图谱研究的机构分布 (发文量≥12 篇)

序号	机构	n	序号	机构	n
1	大连理工大学	93	17	黑龙江大学	15
2	武汉大学	80	18	华东师范大学	15
3	南京大学	56	19	山西大学	15
4	温州大学	34	20	天津师范大学	15
5	华中师范大学	33	21	安徽大学	14
6	北京师范大学	27	22	华中科技大学	14
7	江苏大学	26	23	北京大学	13
8	吉林大学	23	24	河北大学	13
9	上海体育学院	21	25	内蒙古大学	13
10	四川大学	21	26	浙江大学	13
11	浙江树人大学	21	27	国家科学图书馆	13
12	陕西师范大学	20	28	华南师范大学	12
13	中南大学	20	29	上海交通大学	12
14	辽宁师范大学	18	30	同济大学	12
15	武汉体育学院	16	31	湘潭大学	12
16	河南大学	15			

合计 725 篇 占论文总量的 38.20%

### 2.4 高频关键词分析

经统计,1898 篇知识图谱论文共标引关键词 2658 个,共计频次 8175 次。为提高分析的准确性,对同义词或同一概念的不同表达方式进行规范化处理,如“CiteSpace”、“citespace”、“Citespace III”统一为“CiteSpace”,“研究热点”、“热点”统一为“研究热点”,“文献计量”、“文献计量学”、“信息计量分析”统一为“计量分析”。然后选取频次 ≥ 10 次的关键词为高频关键词,共 53 个。除本研究的检索词“知识图谱”外,其余高频词总频次 3030 次,占总词频的 37.06%。见表 4。

表 4 我国知识图谱研究的高频关键词

序号	关键词	频次	序号	关键词	频次
1	知识图谱	1676	28	文献共被引	20
2	可视化	480	29	大数据	18
3	CiteSpace	469	30	竞争情报	16
4	研究热点	390	31	档案学	16
5	计量分析	261	32	VOSviewer	16
6	共词分析	186	33	中国	15
7	研究前沿	168	34	数字图书馆	14
8	科学计量	95	35	CNKI	14
9	社会网络分析	79	36	技术创新	13
10	CSSCI	62	37	Web of Science	13
11	共被引分析	52	38	体育科学	12
12	演进路径	42	39	Ucinet	12
13	词频分析	41	40	综述	11
14	共现分析	36	41	知识服务	11
15	图书馆学	35	42	翻转课堂	11
16	研究现状	34	43	多维尺度分析	11
17	引文分析	31	44	作者共引	10
18	共引分析	29	45	专利分析	10
19	聚类分析	28	46	知识地图	10
20	研究进展	26	47	云计算	10
21	知识基础	25	48	学术群体	10
22	关键词	25	49	信息素养	10
23	情报学	24	50	数据挖掘	10
24	图书情报学	23	51	合作网络	10
25	作者共被引	22	52	MOOC	10
26	知识管理	22	53	信息检索	10
27	研究主题	22			

论文高频关键词的统计分析显示,知识图谱研究主要涉及知识图谱构建的基本理论、绘制方法及应用研究等方面。

知识图谱是利用可视化方式形象地展示学科的发展历史、前沿领域、核心结构以及整体知识架构<sup>[14]</sup>,为学科研究提供切实的、有价值的参考。从应用领域来看,其已被广泛应用于各学科领域,主要包括科学学、图书情报学、档案学、知识管理、信息检索等领域。原因在于知识图谱属于图书情报

领域的研究范畴,因此研究总是从该领域学者最熟悉的领域开始,进而扩展延伸到其他领域,如体育科学,以及一些学术界的热点主题,如技术创新、云计算、信息素养、MOOC 等。

### 3 结论与展望

本文利用文献计量学方法对我国知识图谱研究的发展历程进行了回顾,得出以下结论:自 2005 年以来知识图谱研究蓬勃发展,文献量逐年递增;形成了以刘则渊、侯海燕、宗乾进等为代表的高产作者队伍;以大连理工大学、武汉大学、南京大学为代表的高产机构;知识图谱绘制以 CiteSpace 为主,研究内容已从科学学、图书情报学领域延伸到众多其他新兴领域。目前研究内容以知识图谱的绘制与应用研究为主,基础理论研究较少,且主要由大连理工大学刘泽渊团队完成。知识图谱的绘制工具以 CiteSpace 为主,多为利用软件直接构建图谱,对图谱的科学解读还不够深入。因此,未来知识图谱的研究应以基础理论研究为重点,在已有研究成果基础之上进行更深入的研究,充分利用知识图谱的可视化特征,直观、形象地揭示学科之间的联系,揭示知识领域的动态发展规律,为学科研究提供切实的、有价值的参考<sup>[15]</sup>。

### 参考文献:

[1] 郭颖涛. 21 世纪我国情报学研究知识图谱[D]. 湘潭市:湘潭大学,2014.

[2] 陈悦,陈超美,刘则渊,等. CiteSpace 知识图谱的方法论功能[J]. 科学学研究,2015,33(2): 242-253. DOI: 10. 3969/j. issn. 1003-2053. 2015. 02. 009.

[3] 陈悦,刘则渊. 悄然兴起的科学知识图谱[J]. 科学学研究,2005,23(2): 149-154. DOI: 10. 3969/j. issn. 1003-2053. 2005. 02. 002.

[4] 王伟. 信息计量学及其医学应用[M]. 北京:人民卫生出版社,2009:45.

[5] 孙晓宁,闫励,张强. 科学知识图谱在学科可视化研究中的应用[J]. 图书馆,2014(5): 87-91. DOI:10. 3969/j. issn. 1002-1558. 2014. 05. 025.

[6] 侯海燕,刘则渊,陈悦,等. 当代国际科学学主流学术群体及其代表人物[J]. 科学学研究,2006,24(2): 161-165. DOI: 10. 3969/j. issn. 1003-2053. 2006. 02. 001.

[7] 侯海燕,刘则渊,赫尔顿·克雷奇默,等. 中国科学计量学国际合作网络研究[J]. 科研管理,2009,30(3): 172-179. DOI: 10. 19571/j. cnki. 1000-2995. 2009. 03. 023.

- [8] 宗乾进,袁勤俭,沈洪洲. 2001-2010 年我国图书馆学研究知识图谱——基于知识图谱的当代学科发展动向研究[J]. 国家图书馆学刊, 2012, 21(2): 84-91. DOI: 10.3969/j. issn. 1009-3125. 2012. 02. 016.
- [9] 宗乾进,袁勤俭,沈洪洲,等. 2001-2010 年国内情报学研究回顾与展望——基于知识图谱的当代学科发展动向研究[J]. 情报资料工作, 2012(1): 10-15. DOI: 10.3969/j. issn. 1002-0314. 2012. 01. 003.
- [10] 童玲玉,宗乾进,袁勤俭. 中国 2009 年金融学研究的知识图谱分析——南京大学知识图谱研究组系列论文[J]. 现代情报, 2011, 31(5): 16-20. DOI: 10.3969/j. issn. 1008-0821. 2011. 05. 004.
- [11] 潘晨,宗乾进. 国内数量经济学 2009 年研究热点分析——南京大学知识图谱研究组系列论文[J]. 现代情报, 2011, 31(5): 21-24. DOI: 10.3969/j. issn. 1008-0821. 2011. 05. 005.
- [12] 章以金,宗乾进,袁勤俭. 我国开放存取研究主题的知识图谱分析——南京大学知识图谱研究组系列论文[J]. 现代情报, 2011, 31(5): 8-11. DOI: 10.3969/j. issn. 1008-0821. 2011. 05. 002.
- [13] 宗乾进. 中国极地研究知识图谱——南京大学知识图谱研究组系列论文[J]. 现代情报, 2011, 31(5): 12-15, 20. DOI: 10.3969/j. issn. 1008-0821. 2011. 05. 003.
- [14] 陈楚湘,郭晓峰,陈冰,等. 基于 CiteSpace III 的系统工程学科知识图谱初探[J]. 时代教育, 2015(3): 197, 200. DOI: 10.3969/j. issn. 1672-8181. 2015. 03. 150.
- [15] 姜阳阳. 基于共词分析的组织变革知识图谱研究[D]. 大连市: 东北财经大学, 2011.

(收稿日期 2017-10-09)

(本文编辑:石俊强)

(上接第 452 页)学生的调查资料(61.2%)<sup>[6]</sup>。在男女居民对药物知识知晓率对比发现,男性基本用药知识知晓率总体高于女性。其中,“购买非处方药时是否需要开具处方”、“长期应用抗生素是否可以引起新感染”男女知晓差异具有统计学意义( $P < 0.05$ ),其他 5 个问题男女知晓差异不具有统计学意义( $P > 0.05$ );男性居民对“不同服药时间是否影响药物疗效”知晓率最高,女性居民对“不同服药时间是否影响药物疗效”知晓率最高。通过本次用药知识的调查发现大部分居民对药物知识有一定了解。分析其原因可能与以下几个方面有关:1)调查对象为中青年人,自主学习能力和理解力较强,平时注重药物知识的学习;2)社区健康知识的宣传教育比较到位;3)家庭环境和亲朋好友的影响;4)科普教育和媒体宣传的影响。

用药行为调查情况:本次调查发现,居民在用药行为上存在的问题比较突出,其中漏服药物和擅自增减的现象比较常见,发生率最高;男女不良的用药行为比较差异不具有统计学意义( $P > 0.05$ )所以,居民在用药行为上存在的问题容易造成药物的不合理应用,造成药源性疾病<sup>[7-8]</sup>。这些用药的不良行为应引起我们高度重视。因此,要加强用药知识的宣传教育,如利用各种媒体,开设专题栏目,发放宣传册,向公众讲解用药知识等;有关部门需规范广告及网络等媒体,避免一些不合理的用药知识误导广大消费者<sup>[9-10]</sup>。总之,要多渠道、多部门、全方位共同营造浓厚合理用药的社会支持性环境,不断提高居民安全用药意识和水平,改变不良的用

药习惯,确保人民身体健康。

#### 参考文献:

- [1] 高任萍. 改善不合理用药现象及应对措施[J]. 中国现代药物应用, 2011, 5(11): 93-94.
- [2] 王璐. 临床不合理用药现象及促进合理用药措施分析[J]. 中国药业, 2016, 25(10): 74-76.
- [3] 刘倩. 不合理用药现象及促进合理用药措施分析[J]. 临床合理用药杂志, 2011, 4(8): 54-55.
- [4] 王凤潇,许卫芳,华婷婷,等. 全国合理用药现状抽样调查[J]. 中国药房, 2010, 21(17): 1541-1544.
- [5] 季鸿雁,李道兵,李萍,等. 2011 年陕西省咸阳市城区居民用药知识和行为调查[J]. 中国健康教育, 2014, 30(8): 694-697.
- [6] 卢永,李英华,程玉兰,等. 2011 年中国五省(市)居民用药知识与行为现状调查[J]. 中华预防医学杂志, 2012, 46(6): 495-499.
- [7] 苏睿. 某校大学生用药知识和行为的现状调查[J]. 河北医学, 2016, 22(2): 337-339.
- [8] 朱玉芬. 某高校大学生不合理用药知识行为的现状及对策[J]. 西南军医, 2015, 17(2): 174-175.
- [9] 赵加奎,林军,陆瑛,等. 上海市卢湾区居民安全用药知识与行为调查[J]. 中国初级卫生保健, 2012, 26(3): 28-31.
- [10] 季鸿雁,李道兵,李萍,等. 2011 年陕西省咸阳市城区居民用药知识和行为调查[J]. 中国健康教育, 2014, 30(8): 694-697.

(收稿日期 2017-10-09)

(本文编辑:石俊强)