

·信息分析与科学评价·

大数据研究空间格局的演变及研究力量布局分析*

李丽梅^{1,2} 吴新年¹

(1. 中国科学院兰州文献情报中心 甘肃兰州 730000)

(2. 中国科学院大学 北京 100049)

摘要 文章以1996–2015年Web of Science数据库收录的大数据领域文献为研究样本,从总体态势、空间格局演变特征和当前研究力量布局三个角度进行了分析。结果显示,大数据领域研究正处于中前期加速发展期,学科交叉性强;研究中心正在向中、美、英、德等多个核心区演变,呈现“核心–边缘”结构,并且各国研发实力相差悬殊,核心区主要发达国家研究力量和影响力稳步增长,而中国及亚洲地区的研究成果产出量增长较快,但研究成果学术质量亟待提升。

关键词 大数据;空间格局演变;多中心;核心–边缘;研究力量布局

中图分类号: G250.252 文献标识码: A DOI: 10.11968/tsyqb.1003-6938.2016061

Space Pattern Evolvement and Distribution of Research Strengths of the Study of Big Data

Abstract The paper is to scan the study characteristics of big data in detail, thereby to provide reference for an in-depth development of big data. The author collects literatures on the subject of big data of Web of Science from 1996 to 2015. The overall trend, space pattern evolvement and distribution of research strengths are analyzed. The results show that the study of big data is stepping into the golden development period, and has a strong interdisciplinary nature. Global research has evolved from America to multiple centers—China, America, England and Germany, and demonstrates a “core-peripheral” structure. And there are wide disparities in countries around the world. The strength and impact of R&D about major developed countries in core areas steadily increase over time, but China still has a way to go.

Key words big data; space pattern evolvement; multi-center; core-peripheral; distribution of research strengths

大数据是信息技术与经济社会交汇融合的产物,强调采集和存储数量巨大、来源分散、格式多样的数据,并从中挖掘新知识、创造新价值。大数据“病毒”式地快速蔓延,正日益深刻变革各行各业的商业模式、经济运行机制、国家治理方式以及人类生产生活的方方面面。当前,各国政府、企业、高校和研究机构都开始积极探索大数据发展和应用。

本文基于Web of Science核心集合数据库,采用文献计量方法,对大数据研究的相关文献数据进行分析,以期全方位了解大数据领域空间格局演变特征及当前研究力量布局情况。

1 数据来源

任何科学计量分析和知识图谱绘制的科学性都

根源于数据基础,即保证精确全面地检索到拟定研究主题的全部文献是一项关键问题。尽管学术界正式提出“大数据”始于2008年9月《自然》杂志发表的“Big Data: Science in the petabyte era”系列专题文章,但它并不是一个新概念,此前早已存在“大数据量”“海量数据”“大数据集”等多种与之相近的表达方式,因此有必要根据“大数据”概念演化的过程,构建完整而精确的检索式,以保证获取全面且可靠的数据集。

本文基于大量文献确定出“big data”的相近概念,并经过多次检索实验和抽查验证,制订了相对完善的检索策略,并进行了检索式构建和数据获取。

数据选自于Web of Science平台。检索策略为:主题=((big data or big-data) or “mega-data” or “enor-

* 本文系国家科技图书文献中心(NSTL)科研支撑项目“新疆跨越发展的联合知识服务推进”(项目编号 2014XM081)研究成果之一。
收稿日期 2016-04-27,责任编辑 魏志鹏

《中国学术期刊(光盘版)》电子杂志社编者注:本文中涉及台湾的“国家”均应为“国家(地区)”,“国”均应为“国(地区)”。

mous data” or “huge data” or “magnanim* data” or “mass* data” or “large data” or “tremendous data” or (massive -scale -data or data -massive -scale) or (large -scale -data or data -large -scale)), 出版年=1996-2015; 设定检索范围为: 数据库=SCI-EXPANDED、SSCI 和 CPCI-SSH, 时间跨度=ALL-YEAR (检索时间为 2016 年 4 月 16 日), 文献类型=(ARTICLE OR PROCEEDINGS PAPER)。执行检索, 选择“全记录与引用的参考文献”, 以纯文本格式下载, 共获得 21771 条文献数据。

2 总体态势分析

2.1 文献增长趋势

科学文献数量是衡量科学知识量的重要尺度之一, 某一时期文献数量的增长速度在一定程度上揭示出该学科领域研究的理论水平和发展速度^[1]。本文利用专业绘图软件 OriginPro 8.0, 将 WOS 中大数据领域 1996-2015 年期间文献产出量和累积量随时间变化曲线绘制成图(见图 1)。

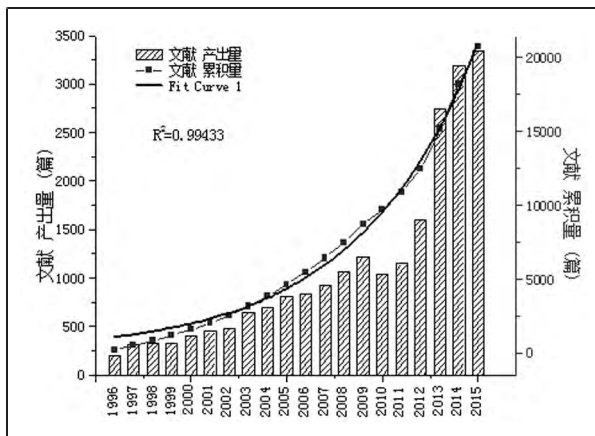


图 1 大数据研究领域文献增长曲线(1996-2015)

可以发现 1996-2015 年, 大数据研究一直处于良好发展态势, 文献数量呈逐年波动增长趋势, 大体可以分为三个阶段: (1)1996-2002 年, 文献增长缓慢, 年发文量不足 500 篇; (2)2002-2009 年, 文献开始呈现快速增长状态, 2008 年突破千篇; (3)2010-2015 年, 文献持续猛增(2010 年虽较 2009 年有所下降, 但仍有千余篇研究成果产出), 尤其是 2013 年增幅很大。

经计算, 拟合曲线为指数型函数, 曲线拟合度, 表明大数据研究领域文献累积增长趋势符合普赖斯

增长规律。按照科学文献增长的四阶段理论^[2], 大数据领域的研究目前正处在中前期加速发展期, 文献数量进入一个相对稳定的指数增长阶段, 呈现“知识爆炸”态势。

2.2 学科类别分布

通过分析 WOS 数据库中 1996-2015 年文献的学科类别构成, 有利于把握研究内容的侧重点, 发现驱动该领域发展的核心学科。依据 Web of Science 数据库对文献的学科分类标准, 发现样本数据分属于 232 种不同类型的学科。(排名前 20 位的学科类别占比情况见图 2)。

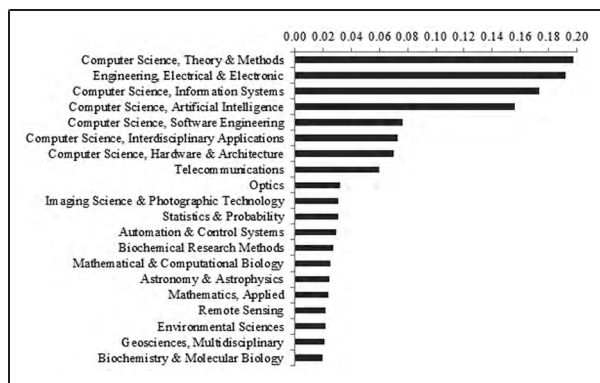


图 2 大数据研究领域文献所属学科分布情况(TOP20)

不难看出, 计算机科学(Computer Science, Theory & Methods; Computer Science, Information Systems; Computer Science, Artificial Intelligence; Computer Science, Software Engineering) 和电子与电气工程(Engineering, Electrical & Electronic) 一直是推动大数据发展的主要学科领域。此外, 大数据领域学科交叉性强, 来自不同学科的研究人员各有侧重, 针对大数据的产出与应用, 不断产生重大创新成果, 造就了大数据领域研究内容宽、应用范围广的现状, 通常涉及天文学、生物化学、环境科学、地球科学、数学与计算生物学、遥感、通信等。

3 研究中心的空间格局演变特征分析

从大数据研究领域文献增长情况可以看出, 2002 年和 2009 年分别是一个拐点年。2002 年以前大数据领域一直处于缓慢发展态势, 2002 年以后出现快速增长, 2009 年达到一次小的峰值, 之后发生短暂下滑, 转而呈现猛增势头。因此, 选择 1996、2002、

2009、2015年四个横断面数据,以世界各国作为空间观测单元,以文献数量作为科研产出的衡量指标,对WOS中大数据领域科研产出的分布情况进行逐一刻画,旨在揭示本领域研究中心在全球的空间格局演变特征(见图3)。

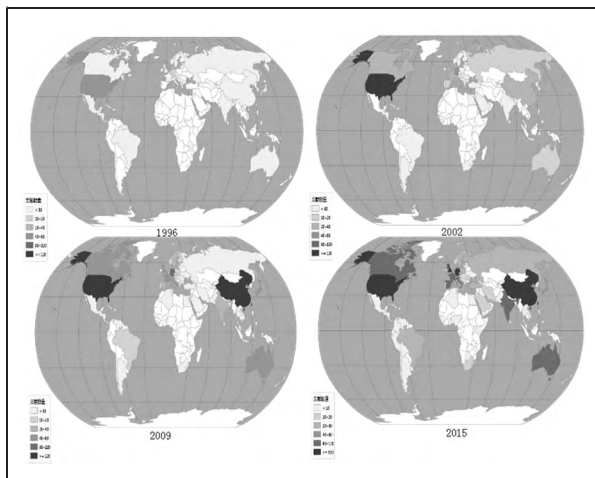


图3 大数据领域研究中心的空间分布格局演变情况

注:颜色越深,则该国家或地区的科研产出越多,反之亦然。

纵观1996-2015年大数据研究领域科研产出排名前10的国家历年排名变动情况(见图4),可以发现美国一直是大数据研究的主要推动者和重要贡献者。1996-2000年,仅次于美国的英、德两国科研产出稳步增长,法国小幅度下滑,加拿大则大幅度上升,日本上下波动较大;2000-2004年,中国研发实力大增,成为除美、英、德三国之外重要的研究力量;2004-2009年,美、中、英、德四国稳居前四,意大利、法国、加拿大、日本等国家发文量不相上下;2010-

2015年,各国大数据研究势头强劲,竞争激烈,逐渐涌现一批如西班牙、澳大利亚、韩国、印度等新生研究力量,而原来的强国日本2015年则跌出了前十位。另外,一些国家只是在个别年份进入了世界前十,如瑞典在1996-1999年间排名由第八下滑至第十,之后再未进入过前十;瑞士和韩国分别出现过两次,且排位均靠后。

综上所述,大数据领域研究中心在全球的总体格局和基本走向呈现以下特征:

(1)空间格局呈现多中心发展。20世纪90年代末,美国在大数据研究领域的发文量占全球总量的1/3以上,远远超出德国(稳居第二名)5倍之多;如今,美国发文比例降至1/5左右,逐渐失去霸主地位。全球大数据研究中心空间分布核心区的国家或者地区已经从20世纪90年代末占据绝对优势的美国逐步转变为当前由北美的美国、西欧的英国和德国以及亚洲的中国等组成的多个核心区。

(2)各地区呈现“核心—边缘”结构。除上述核心区外,各地区也呈现明显的“核心—边缘”结构。在研究期内,北美的加拿大,欧洲的法国、西班牙和意大利,亚洲的日本、韩国和印度,以及大洋洲的澳大利亚等多个国家或地区逐渐成为除核心区之外的一批新兴大数据研究力量,并且在地域上与美国、英国、德国和中国等主要研究中心相近。

4 当前大数据领域研究力量布局分析

4.1 基于国家层面的产出及影响力分析

被引率是美国学者Markusov^[3]和Smart等^[4]分别

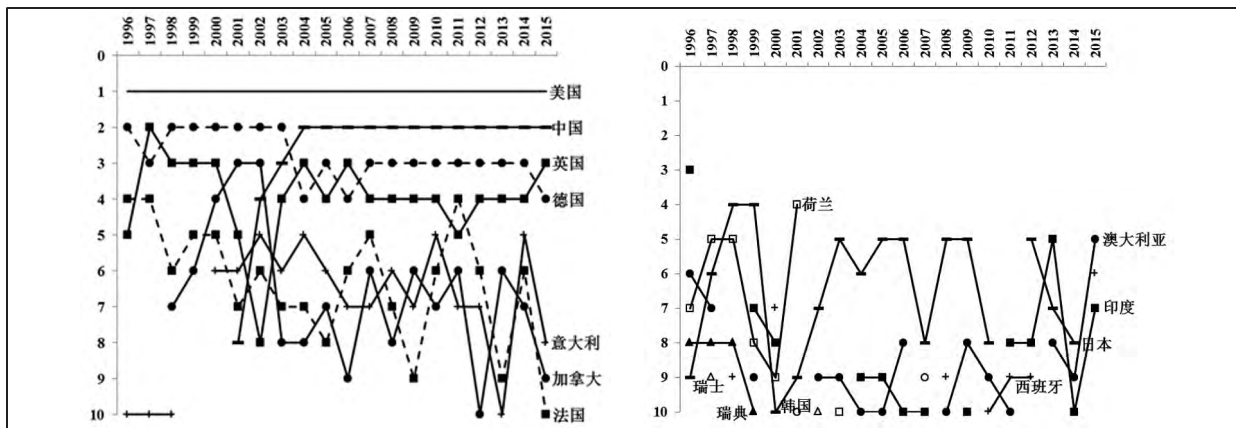


图4 大数据研究领域科研产出排名前10的国家历年排名变动情况(1996-2015)

于1973年和1981年提出的概念,其高低反映着科学文献的学术价值和适用价值,因此逐渐成为评估学术机构文献质量高低的重要参考标志。被引频次是衡量文献质量的另一重要指标,被引用次数越高说明该文献影响力越大^[5]。一般情况下,高被引文献可能极具创新性和前瞻性,提出了本领域亟待解决的实际问题,具有极大的指导意义。需要指出的是,本文所提到的被引率、篇均被引频次等多个文献计量指标均指某时期数据,如被引率是指某观测单位在某时期文献的被引篇数/某时期文献总篇数,篇均被引频次是指某观测单位在某时期文献的总被引频次/某时期文献总篇数。

H指数是2005年由美国加州大学圣地亚哥分校物理系J.E.Hirsch教授设计的一种混合量化指标^[6],其计算兼顾了文献数量和质量^[7],弥补了单纯以文献数量测度和评价研究机构等学术共同体成员影响力和贡献度的不足,能够较为全面客观地反映科研生产力和影响力。

本文借助WOS强大的数据分析功能和文本挖

掘软件TDA,对文献的国家分布情况展开分析。通过统计和计算1996-2015年世界上发文总量居前20位的国家在大数据研究领域的出产及影响力情况,可以发现1996-2015年全球共有143个国家发表了大数据领域相关研究文献,其中:(1)发文量方面:美国稳居第一,占世界总量的33.774%,是位居第二的中国的1.863倍;发文量达1000篇以上的国家还包括英、德两国;(2)文献被引方面:比利时、瑞典、荷兰、瑞士、英国、德国和加拿大被引率达70%以上,远远超出二十国的平均被引率(63.526%);而中国在大数据研究领域的文献被引率仅为35.445%,即每100篇文献中约有35篇被引用,研究成果的学术影响力偏低;美国总被引频次最高,而德、英、法和加拿大四国被引频次和篇均被引频次均在20国平均水平之上(20国平均值分别为20498.2次和18.32次),研究成果具有一定影响力;而中国篇均被引频次排名倒数第二;(3)H指数方面:北美地区的美国稳居第一,西欧的英、德两国稍次之,加拿大、法国、荷兰、意大利、中国和澳大利亚是除上述3个国家之外H指数

表1 大数据研究领域主要国家科研产出及影响力情况(1996-2015)(TOP20)

排序	国家	地区	文献数量	全球占比(%)	被引率(%)	被引频次	篇均被引频次	H指数
1	美国	北美洲	7353	33.774	65.334	153422	20.87	143
2	中国	亚洲	3947	18.130	35.445	19634	4.97	47
3	德国	欧洲	1606	7.377	70.672	36672	22.83	81
4	英国	欧洲	1348	6.192	72.255	48803	36.20	81
5	法国	欧洲	899	4.129	69.633	24039	26.74	57
6	加拿大	北美洲	882	4.051	70.181	21618	24.51	60
7	意大利	欧洲	875	4.019	67.314	11479	13.12	51
8	日本	亚洲	850	3.904	51.294	11727	13.80	40
9	澳大利亚	大洋洲	758	3.482	69.525	9345	12.33	46
10	印度	亚洲	737	3.385	43.555	3393	4.60	23
11	西班牙	欧洲	603	2.770	64.510	9864	16.36	36
12	韩国	亚洲	529	2.430	45.180	2954	5.58	22
13	荷兰	欧洲	524	2.407	75.382	16410	31.32	57
14	瑞士	欧洲	423	1.943	75.177	9030	21.35	44
15	中国台湾	亚洲	383	1.759	51.958	2017	5.27	22
16	瑞典	欧洲	309	1.419	77.023	8530	27.61	37
17	比利时	欧洲	277	1.272	80.144	13665	49.33	43
18	奥地利	欧洲	254	1.167	65.748	3315	13.05	28
19	巴西	南美洲	247	1.135	58.300	1993	8.07	26
20	波兰	欧洲	244	1.121	61.885	2054	8.42	20

注:不同国家/机构/作者之间合作发表的文献,在统计时存在重复计算。

相对较好的国家,但与英、德相比还存在明显差距,与美国的差距更大。

可见,以美国和加拿大为主的北美洲和以德国、英国和法国为主的欧洲地区科研实力相对较强,而以中国为主的亚洲和以澳大利亚为主的大洋洲地区也正在积极开展大数据研究工作。虽然中国发文量较高,但研究成果的学术影响力较低,需要引起重视。

本文利用 TDA 绘制大数据研究领域发文国家之间的合作网络(见图 5),旨在了解各国之间科研合作联系的紧密度,以及各国在合作研究中所处位置的情况。结果显示:大数据研究领域发文国家主要存在独立研究和小团体合作两种模式,且跨国合作具有明显的地域特征,尤其是欧洲地区。诸如美国、中国、日本、韩国、印度、新加坡、马来西亚、泰国、墨西哥等国家虽然都是大数据研究领域的主要国家,但对外的合作强度普遍较低;相反地,诸如丹麦、挪威、瑞典、芬兰等北欧地区,英国、法国、荷兰、比利时、爱尔兰等西欧地区,德国、瑞士、波兰、奥地利等中欧地区,俄罗斯、爱沙尼亚等东欧地区,意大利、西班牙、希腊、斯洛文尼亚、塞尔维亚等南欧地区国家间的合

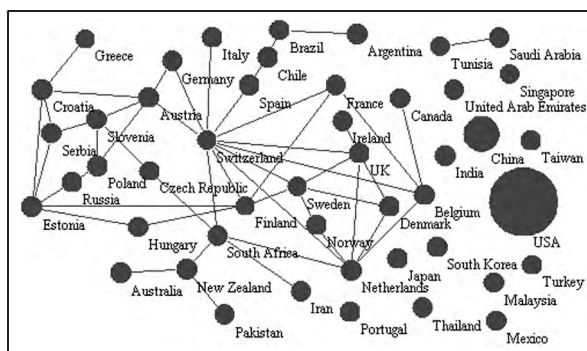


图 5 大数据研究领域国家合作网络

注:本图仅选取合作紧密度排名前 50 的连线,众多次要连线不做展示,下同;圆形节点代表发文国家,节点越大,科研产出越高,连线粗细代表国家之间的合作紧密度,合作越多,连线越粗。

作相对活跃。此外,南非在大数据研究领域的国际合作也较多。

4.2 基于机构层面的产出及影响力分析

本文以发文机构作为空间观测单元,以文献数量作为科研产出的衡量指标,得出 1996-2015 年世界上发表大数据相关研究文献总量居前 20 位的科研机构科研产出及影响力情况(见表 2)。

表 2 大数据研究领域主要研究机构科研产出及影响力情况(1996-2015)

排序	研究机构	国家	文献数量	全球占比(%)	被引率(%)	被引频次	篇均被引频次	H 指数
1	中国科学院	中国	406	1.865	53.941	2153	5.30	19
2	哈佛大学	美国	191	0.877	86.911	14197	74.33	41
3	加州大学伯克利分校	美国	170	0.781	70.000	3662	21.54	31
4	清华大学	中国	169	0.776	45.562	526	3.11	9
5	麻省理工学院	美国	169	0.776	72.189	8939	52.89	26
6	伊利诺伊大学	美国	159	0.730	67.296	1682	10.58	22
7	加州大学洛杉矶分校	美国	152	0.698	72.368	4258	28.01	23
8	斯坦福大学	美国	50	0.689	78.000	11911	79.41	33
9	华盛顿大学	美国	142	0.652	84.507	11654	82.07	32
10	明尼苏达大学	美国	134	0.615	73.881	2482	18.52	24
11	马里兰大学	美国	132	0.606	75.000	2314	17.53	25
12	牛津大学	英国	130	0.597	80.000	5676	43.66	26
13	威斯康辛大学	美国	130	0.597	74.615	2889	22.22	26
14	密歇根大学	美国	124	0.570	70.161	4345	35.04	25
15	加州大学圣地亚哥分校	美国	117	0.537	75.214	3889	33.24	27
16	俄亥俄州立大学	美国	115	0.528	74.783	2015	17.52	22
17	宾夕法尼亚州立大学	美国	114	0.524	66.667	5578	48.93	20
18	佐治亚理工学院	美国	114	0.524	59.649	860	7.54	15
19	加州大学戴维斯分校	美国	113	0.519	78.761	1836	16.25	22
20	多伦多大学	加拿大	113	0.519	76.106	3317	29.35	24

对 1996–2015 年发文机构进行查重和归并,结果显示共有 10618 个研究机构有成果产出,其中:(1)从发文量看,排名前 20 的机构中,除有中国的 2 个、加拿大的 1 个机构外,其余全部是美国的高校,其中排名第一的中国科学院发文量是排名第二的哈佛大学的两倍多;(2)从文献被引情况看,哈佛大学、华盛顿大学和牛津大学被引率高达 80%以上(20 个研究机构的平均被引率为 71.781%);哈佛大学、斯坦福大学和华盛顿大学的总被引频次则突破万次,篇均被引频次超过 70 次;(3)从 H 指数看,即使在大数据这样的新兴研究领域,美国高校的整体学术水平依然是很高的,处于大数据研究领域的核心地位,诸如哈佛大学、加州大学伯克利分校、斯坦福大学、华盛顿大学等院校的 H 指数均在平均值(24)之上,加拿大的多伦多大学对本领域的贡献度处于中等水平,而中国的两家研究机构发文量位居前列,但 H 指数不高,表明在本领域的学术影响力不高,研究实力还需提升。

本文利用 TDA 绘制大数据研究领域发文机构合作网络(见图 6),旨在发现领域内相对活跃的研究机构,并快速了解其合作关系概貌。结果表明:(1)大数据研究领域发文机构中,英美地区的一些著名高校之间存在较为密切的合作交流,但中美著名高校之间的合作关系薄弱。如美国的哈佛大学、耶鲁大学、宾西法尼亚大学、约翰-霍普金斯大学、密歇根大学、华盛顿大学、哥伦比亚大学、北卡罗莱那大学等与英国的牛津大学、剑桥大学、帝国理工学院等形成了较为紧密的合作网络,中国的清华大学、北京大学、北京理工大学、中国科学技术大学和中国科学院之间有合作产出,但合作度仍然偏低。更多的研究机

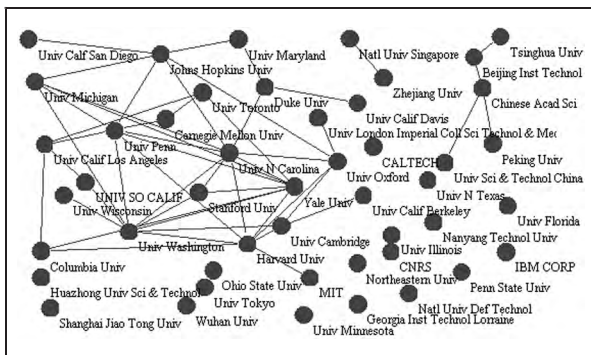


图 6 大数据研究领域机构合作网络

构对外的合作研究则很少或几乎没有,尤其是图片右下方分布着数量庞大但孤立的节点,如明尼苏达大学、伊利诺伊大学、南洋理工大学、法国国家科学研究院、IBM 等,都是大数据研究领域的重要力量,但对外合作似乎很少,说明大数据研究领域机构之间合作还不是很普遍;(2)尽管一些研究机构拥有跨机构的合作研究,但合作对象相对集中,说明这些主要机构的对外合作广泛度并不太高;(3)即使是主要研究机构也尚未在合作网络中形成权威地位和影响力。

4.3 基于作者层面的产出及影响力分析

本文以发文作者作为空间观测单元,以文献数量作为科研产出的衡量指标,得出 1996–2015 年世界上发表大数据相关研究文献总量居前 20 位的研究人员科研产出及影响力情况(见表 3)。

对 1996–2015 年发文作者名称进行清洗和合并,结果显示,共有 60855 位研究人员有成果产出。其中:(1)从发文量看,排名前 20 位的作者中中国的作者占 30%、美国的作者占 25%、澳大利亚的作者占 15%、韩国的作者占 10%,但人均发文不足 24 篇,发文量排名前 5 位的作者是韩国生物科学和生物技术研究室的 Kim J(45 篇)、德国波鸿鲁尔大学的 Lee J(42 篇)、韩国高丽大学的 Lee S(39 篇)、美国弗吉尼亚理工大学的 Kim S(38 篇)和美国南佛罗里达大学的 Hall Lawrence O.(30 篇);(2)文献被引方面,比利时布鲁塞尔大学的 Chamel N.教授、加拿大蒙特利尔大学的 Pearson J.M. 教授和澳大利亚墨尔本大学的 Bezdek James C.教授在被引率和被引频次方面统计结果最佳,表明他们在本领域具有较高的学术影响力;(3)从 H 指数看,前 20 位作者的平均值仅为 5.5,其中加拿大蒙特利尔大学的 Pearson J.M. 教授排名第一,比利时布鲁塞尔大学的 Chamel N.教授排名第二,中国及其他亚洲国家的作者进入发文量前 20 的不少,但篇均被引率和 H 指数都不高。

另外在大数据研究领域,除科研院所和大学外,IBM 等跨国公司在研发产品的同时也积极开展大数据基础和应用技术研究,并将可共享的成果以论文形式进行发表。

本文利用 TDA 绘制大数据研究领域发文作者合作网络(见图 7),旨在了解该领域合作密切的作者

表 3 大数据研究领域主要研究人员科研产出及影响力情况(1996-2015)

排序	研究人员	所属机构	国家	文献数量	被引率(%)	被引频次	篇均被引频次	H 指数
1	Kim J	韩国生物科学和生物技术研究所	韩国	45	37.778	120	2.67	5
2	Lee J	波鸿鲁尔大学	德国	42	40.476	147	3.50	7
3	Lee S	高丽大学	韩国	39	48.718	127	3.26	5
4	Kim S	弗吉尼亚理工大学	美国	38	52.632	131	3.45	6
5	Hall Lawrence O.	南佛罗里达大学	美国	30	56.667	212	7.07	6
6	Kim H	莱顿大学	荷兰	26	65.385	208	8.00	7
7	Wang Lizhe	中国科学院	中国	23	78.261	170	7.39	7
8	Zhang Wei	中国药科大学	中国	23	47.826	48	2.09	5
9	Pearson J.M.	蒙特利尔大学	加拿大	20	85.00	724	36.20	12
10	Liu Yang	康涅狄格大学	美国	20	35.000	24	1.20	3
11	Ranjan Rajiv	联邦科学与工业研究组织	澳大利亚	19	68.421	120	6.32	5
12	Liu Jun	田纳西大学	美国	18	38.889	10	0.56	2
13	Liu Chang	华中科技大学	中国	18	55.556	51	2.83	4
14	Chen Jinjun	悉尼科技大学	澳大利亚	18	72.222	66	3.67	5
15	Zhang Xuyun	华中科技大学	中国	18	66.667	55	3.06	4
16	Chamel N.	布鲁塞尔大学	比利时	17	88.235	659	38.76	10
17	Blasch Erik	美国空军研究实验室	美国	15	26.667	9	0.60	2
18	Li Yan	IBM	中国	15	53.333	49	3.27	3
19	Bezdek James C.	墨尔本大学	澳大利亚	15	80.000	196	13.07	9
20	Guo Lei	东北大学	中国	15	66.667	56	3.73	3

群、每位作者在合作中所处的地位以及活跃度。结果显示:(1) 大数据研究领域发文作者目前仍然表现为独立发文和小团队合作为主,合作范围仍然较窄。其中,合作网络中的大部分作者只与少数作者之间存在合作关系,尤其以同一机构内的同事之间合作为主。如,合作者群 #1 中的 Li Jian 教授来自 IBM 奥斯汀研究院,而与其合作的 Li Yan 教授和 Wang Kun 教授均来自 IBM 中国研究院;合作者群 #2 中的 Goriely S. 教授和 Chamel N. 教授均来自布鲁塞尔大学;合作者群 #3 中的 Leckie Christopher 教授和 Bezdek James C.

教授均来自墨尔本大学;合作者群 #4 中的 Yang Chi 教授、Liu Chang 教授、Chen Jinjun 教授等均来自悉尼科技大学,而与其合作的 Ranjan Rajiv 教授和 Nepal Surya 教授均来自联邦科学与工业研究组织(简称 CSIRO)。另外,Ranjan Rajiv 教授是连通合作者群 #1 中两个合作簇的关键节点,其拥有丰富的社会资本,有助于科研交流的发生;(2) 大数据领域研究人员组成的合作网络存在明显的同地域特征。如合作者群 #1 中,IBM 中国研究院的 Li Yan 教授、Wang Kun 教授和东北大学的 Guo Lei 教授均来自中国;合作者群 #2 中,Goriely S. 教授和 Chamel N. 教授均来自比利时;合作者群 #3 中,Leckie Christopher 教授和 Bezdek James C. 教授均来自澳大利亚,而与其合作的 Hall Lawrence O. 教授和 Havens Timothy C. 教授分别来自美国的南佛罗里达大学和密歇根大学;合作者群 #4 中,CSIRO 的 Ranjan Rajiv 教授和 Nepal Surya 教授、悉尼科技大学的 Chen Jinjun 教授和 Yang Chi 教授均来自澳大利亚,南京大学的 Dou Wanchun 教授和中国科学院的 Wang Lizhe 教授均来自中国地区;(3) 一些作者依托不同的机构

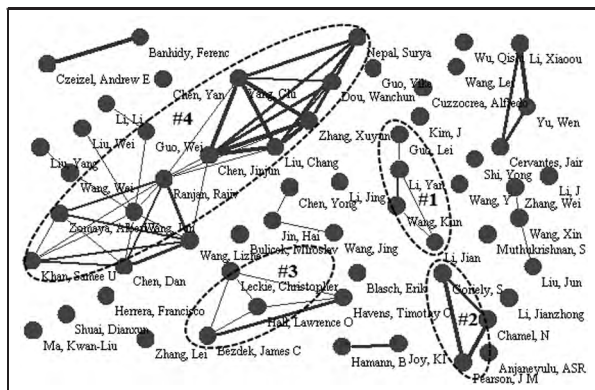


图 7 大数据研究领域作者合作网络

开展研究,为机构间的合作交流起到推动作用。如合作者群 #4 中的 Zhang Xuyun 教授和 Liu Chang 教授同时依托于华中科技大学和悉尼科技大学,而 Khan Samee U 教授同时依托于清华大学和美国的北达科他州立大学,这有利于带动中澳和中美地区实现资源共享及研究交流;(4)还可以发现,绝大部分发文作者不具备相对权威的学术影响力。

从前文分析可得,当前大数据领域的研究力量布局庞大且分散,包括美国、中国、英国在内的世界各国纷纷制定了相应的大数据发展战略或计划,但该领域的科研合作紧密度和广泛度普遍偏低,合作对象主要还集中在同一地区、同一机构内部,且科研团队规模较小,跨地域、跨机构的对外合作仍然偏少,另外大数据研究领域也还尚未出现核心机构和

核心领袖人物。

5 结语

本文通过分析大数据研究领域空间格局基本走向和当前研究力量布局形势,发现全球大数据研究领域正在形成多个核心区和“核心—边缘”结构,其中以北美的美国、西欧的英国和德国以及亚洲的中国为主。这一空间格局的演变与大数据正日益成为国家重要的基础性战略资源、正引领新一轮的科技创新,以及各国或各地区的政策倾斜和扶持力度大幅上升具有密不可分的关系。另外,各国在大数据领域研发实力相差悬殊,合作项目较少,今后应该鼓励加强合作,尤其是加强跨机构甚至跨国家的合作研究,以提高全球大数据领域的整体研究水平,推动其快速发展。

参考文献:

- [1] 邱均平.信息计量学[M].武汉:武汉大学出版社,2007:36-66.
- [2] 邱均平.信息计量学(二)第二讲文献信息增长规律与应用[J].情报理论与实践,2000,23(2):153-157.
- [3] MARKUSOV.VA.Comparative study of citation rates of Soviet scientific and technical publications[J].NAUCHNO-TEKHNICHESKAYA INFORMATSIYA SERIYA 1-ORGANIZATSIYA I METODIKA INFORMATSIONNOI RABOTY,1973(1):27-31.
- [4] Smart J C,Elton C F.Structural characteristics and citation rates of education journals[J].American Educational Research Journal,1981,18(4):399-413.
- [5] 苏君华.中国档案学核心期刊影响力分析——以 2000-2009 年所载论文为研究对象[J].档案学通讯,2010(5):15-20.
- [6] 赵基明,邱均平,黄凯,等.一种新的科学计量指标——h 指数及其应用述评[J].中国科学基金,2008(1):23-32.
- [7] 张凌云,齐飞,吴平.近十年我国旅游学术共同体成果的 h 指数测度与评价[J].旅游学刊,2014,29(6):14-23.

作者简介:李丽梅(1991-),女,中国科学院兰州文献情报中心、中国科学院大学硕士研究生,研究方向:情报理论与方法;吴新年(1968-),男,中国科学院兰州文献情报中心研究员,研究方向:情报理论与方法、产业竞争情报分析、知识管理与知识服务、区域发展研究。

(上接第 144 页)

动及这些耳熟能详的英雄人物早就成为民间艺人创作的优秀素材,并且通过木版年画使其教育意义在广大乡村更得以强化。正是因为东昌府木版年画包

含了如此丰富的爱国主义题材,其已经成为对青少年进行爱国主义教育、民族传统教育、革命体会教育不可多得的生动教材。

参考文献:

- [1] 张士闪.中国传统木版年画的民俗特性与人文精神[J].山东社会科学,2006(2):54.
- [2] 冯骥才.中国木版年画的价值及普查意义[J].民间文化论坛,2005(1):57.

作者简介:张兆林(1980-),男,山东大学儒学高等研究院博士研究生,聊城大学美术学院副教授,研究方向:文化遗产;束华娜(1979-),女,聊城职业技术学院讲师,研究方向:传统文化。