

· 探索与创新 ·

2012—2018 年国家社会科学基金项目图书馆、情报与文献学学科热点可视化分析^{*}

张 蒙 刘春艳

(黑龙江大学信息管理学院 黑龙江哈尔滨 150080)

摘 要: 文章从国家社会科学基金项目数据库中对图书馆、情报与文献学领域的相关项目标题进行信息爬取,使用武汉大学研发的ROST CM 6 对文本进行分词处理,再使用Citespace生成可视化知识图谱,直观地了解该学科项目申报的主题演进与热点前沿趋势。研究表明,该学科在2012—2018年仍以图书馆、大数据、信息资源为核心主题开展研究,跨学科交叉研究、大数据与人工智能、阅读推广和国家安全会成为该学科领域未来重要的研究热点。

关键词: 国家社会科学基金;图情档学科;研究热点;发展趋势;可视化分析

中图分类号: G250.2

文献标识码: A

Visualization Analysis on Hot Spots of Library, Information and Literature Research Based on 2012–2018 National Social Science Fund

Abstract Based on the National Social Science Fund Project database, we can collect the information on project titles in the field of library, information and literature. Through the ROST CM 6 software developed by Wuhan University, we can segment the text, and then use CiteSpace to generate a visual knowledge graph, to understand the evolution of the project application and hot frontier trends of this discipline intuitively. As we can see in the result, between 2012 and 2018, library and information resources are still the core subjects, while interdisciplinary research, big data and artificial intelligence, reading promotion and national security will become increasingly important in this field.

Key words National Social Science Fund; library information and archives science; research hotspot; development trend; visual analysis

国家社会科学基金(以下简称“国家社科基金”)是于1986年由国务院批准设立的。能够成功申报国家社科基金的学科项目研究都代表了我国社会科学研究领域的最高层次和水平。1991年,我国国家社科基金开始将“图书馆、情报与文献学”学科设立为一个独立的学科,成为了现有的26类学科中的一门单独学科。1994年又将其列为国家社科基金资助学科,这充分说明了“图书情报与档案管理”(以下简称“图情档”)学科在整个学科领域研究中占据了重要

地位^[1]。本文将从可视化角度出发,利用生成知识图谱的方法来直观展示近10年图情领域的研究现状及热点,以期帮助学者更准确地把握和分析本学科的立项特点和资助重点。

1 文献综述

近几年,很多学者针对国家社科基金图书馆、情报与文献学立项信息进行现状剖析和热点研究。赵

^{*} 本文系国家社会科学基金一般项目“新型智库协同知识管理能力与智慧服务创新研究”(项目编号:17BTQ052)的研究成果之一。

蓉英等人对我国图情档学科2001—2012年间的国家社科基金与国家自然科学基金的申报项目进行对比分析,认为该学科正处于蓬勃发展的阶段,但是发展的过程中存在较为明显的地区、机构以及个人差异,基金项目高度集中于高等院校以及学科带头人当中;学科研究主题稳中求新,在重视传统研究的同时把握不断涌现的新热点^[2]。李英同样以国家社科基金与国家自然科学基金为研究对象,对其在2009—2013年间的内容特征等方面进行系统分析比较,认为近年来该学科研究趋势将围绕Web2.0、信息资源建设、用户信息行为等主题开展^[3]。柴欢、阮建海阐释了2013—2017年图情档学科科研立项的基本情况,发现该学科在这5年的整体研究上呈现以下几个特点:以图书馆为核心的研究始终是该学科重点关注的主题,学科国际化和图情机构的数字化创新服务是学者研究的热点问题;以大数据环境为背景的相关研究成为研究的重要领域;古籍整理与文献研究在研究层面上更为深刻和细化^[4]。

上述学者所作的分析大多采取文字描述和统计图表形式展示,未有人将国家社科基金项目的数据信息展示为可视化知识图谱,且少有人将分析的时间跨度放在最近几年间(2012—2018年)。本文研究的一大创新点就在于突破了Citespace只能做现成数据库中的文献计量的限制,对作为非文献数据的国家社科基金项目的标题进行可视化分析。

2 研究设计与研究方法

2.1 主要步骤

在国家社科基金项目数据库中检索出符合时间、学科类别要求的项目标题;爬取标题信息,生成文本;应用ROST CM 6软件对已爬取生成的文本信息进行分词处理;将分词处理后的文本转化为Citespace可识别的格式;使用Citespace软件绘制可视化知识图谱并对研究热点进行分析。

2.2 数据来源

本文在国家社科基金项目数据库中爬取2012—2018年立项学科为“图书馆、情报与文献学”的项目标题等相关信息,共得到1 000条项目信息,并获取了每个项目的负责人姓名、职称、所在地区等相关信息。

2.3 研究方法

(1)文献计量法。文献计量法是采用统计学和数学的定量分析方法研究文献的基本特征,如文献发表时间、来源期刊、论文作者职称与所在机构、项目基金等有关信息,从而了解文献的数量关系、分布结构及规律^[5]。

(2)Citespace可视化分析软件。使用Citespace工具的各个学科领域的学者一直以来进行的都是文献或专利计量,如对中国知网、WOS数据库中的文献研究热点、研究前沿和研究趋势进行探索。在本文研究中主要使用该软件关键词共现网络、聚类分析以及时区图等功能,在对生成的图谱的解读中主要针对高频节点、聚类知识群、时区图的演进特征、图谱的基本图例进行说明。

(3)ROST ContentMining知识处理与内容挖掘软件。ROST CM是由沈阳教授团队研发编码的辅助人文社科研究的数字化计算平台。该平台可以提供网站分析、字频与词频统计、情感分析、流量分析、共现分析等一系列功能。在本文的研究中主要使用该软件的分词辅助功能。

3 研究过程与分析

3.1 关键词共现网络

关键词是一个项目研究的核心概括,对项目标题关键词进行分析可以对项目内容和主题窥探一二。一个项目标题中所包含的几个关键词一定存在着某种关联,而这种关联可以用共现的频次来表示。文本数据所代表学科中各主题之间的关系通常是利用汇集关键词或对名词短语共同出现的情况来确定。本文统计1 000条标题名称的关键词两两出现在同一标题的频次,形成一个由这些词对关联所组成的共词网络。

将分析时间设定为2012—2018年;Note Type选择Keyword;统计每一年度共现频次最高50个关键词,连线强度选择Cosine。得到国家社科基金的图书馆、情报与文献学的关键词共词网络如图1所示。

从图1中可以看出有下面几组共现呈现较明显:图书馆与阅读推广、社会教育、开放数据;大数据与网络舆情、用户个性化推荐服务、竞争情报;信息资源与元数据、云计算、互联网、一体化等。图中圆圈大小代表的是关键词频次,频次越大,圆圈越大(注意并不是中心性大小)。显而易见,图书馆、大数据、信息

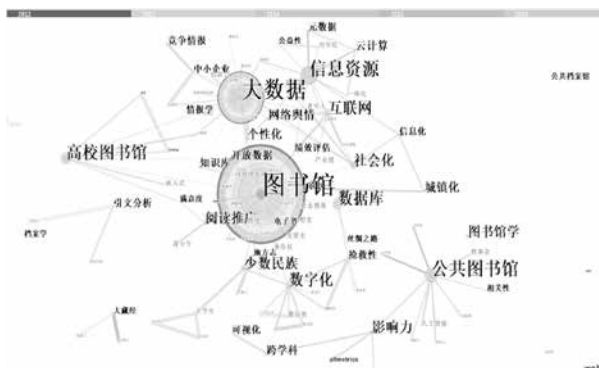


图1 关键词共现网络图

资源、公共图书馆、高校图书馆等关键词圆圈较大，即在国家社科基金项目标题中出现的频次较高。以上这些关键词处于网络的中心位置，其他关键词是围绕他们辐射开的。

这一点也可以通过频次中心性表格来进行验证，如表1所示。表1显示了国家社科基金立项课题的高频关键词与中心性，由此可以看出在国家社科基金项目中频次较前的关键词，如“图书馆”“大数据”出现频次分别达到129次和71次，两者的中心性分别为0.28和0.15；其次是“公共图书馆”“高校图书馆”“数据库”“互联网”出现的频次也比较高；再次是“阅读推广”“安全观”“网络舆情”“云计算”“开放数据”，最后频次较低的有“开放数据”“跨学科”“web”等。一般认为，频次和中心性的值大，说明该关键词在共现网络中处于中心位置，且在知识结构演变中占据重要地位^[6]。

表1 高频关键词词频中心性表

关键词	年份	频次	中心性	关键词	年份	频次	中心性
图书馆	2012	129	0.28	安全观	2018	8	0.02
大数据	2017	71	0.15	网络舆情	2012	8	0.01
信息资源	2012	31	0.07	云计算	2012	7	0.02
公共图书馆	2012	27	0.01	开放数据	2015	6	0.04
高校图书馆	2012	22	0.04	跨学科	2014	6	0.01
数据库	2012	18	0.02	web	2013	5	0.01
互联网	2016	17	0.03	竞争情报	2014	5	0.02
少数民族	2012	11	0.03	一带一路	2016	5	0.01
阅读推广	2014	10	0.01	数字图书馆	2016	5	0.01

3.2 聚类图

如图2所示，我们可以清晰地看到原本的关键词共现网络聚成了几个不规则区域，每个区域都对应一

个标签。其中，第一大聚类（#0）共包含21个立项标题，S值为0.879，LLR算法得到的聚类命名结果为“图书馆”；第二大聚类（#1）共包含19个立项标题，S值为0.914，LLR算法得到的聚类命名结果为“大数据”；第三大聚类（#2）共包含14个立项标题，S值为1，LLR算法得到的聚类命名结果为“公共图书馆”；第四大聚类（#3）共包含12个立项标题，S值为0.921，LLR算法得到的聚类命名结果为“信息资源”；第五大聚类（#4）共包含12个立项标题，S值为0.951，LLR算法得到的聚类命名结果为“数字化”；第六大聚类（#5）共包含7个立项标题，S值为0.914，LLR算法得到的聚类命名结果为“高校图书馆”；第七大聚类（#6）共包含6个立项标题，S值为0.867，LLR算法得到的聚类命名结果为“社会化”。

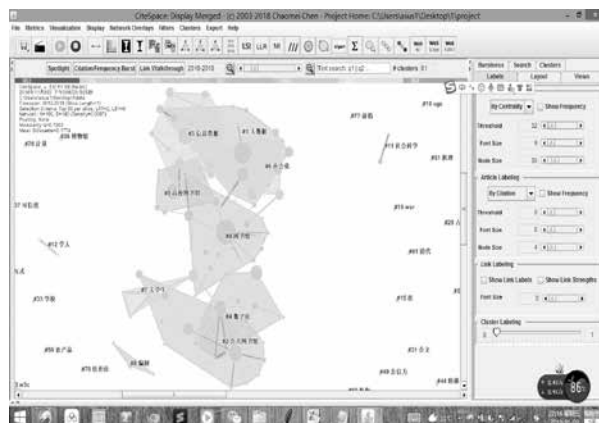


图2 主题聚类图

从图2中我们可以看到聚类较明显的有7类，分别是图书馆、大数据、公共图书馆、信息资源、数字化、高校图书馆、社会化。顺序编码是从0到6，编码数字越小，聚类中包含的关键词越多，每个聚类是由多个紧密相关的词组成的。

3.3 时区图

图3显示了国家社科基金项目标题关键词变化的时区图，其中节点大小代表关键词出现频次的高低，第一次出现的关键词会体现在具体的时间区间上，如果之后立项项目出现相同的关键词，只会在第一次出现的时间点上体现。从图3中可以看出，2012年出现的关键词较多，包含图书馆、公共图书馆、信息资源、互联网、网络舆情、高校图书馆、云计算等。从2013年开始，国家社科基金中新的标题关键词的出现呈现缓慢下降的趋势。但这并不代表图情档学科领域研究热



图3 时区图

点主题有所变少,也不代表研究方向趋向于单一化。2013年主要出现的关键词是大数据、数据库、个性化、社会化、引文分析、知识库等;2014年至2018年间出现了跨学科、阅读推广、人工智能、国家安全观等关键词。

通过关键词的时区图展现,可以了解不同时间节点上国家社科基金中图情档学科立项出现的新变化。通过对立项信息的把握,利于为该学科发展提供经验借鉴和立项参考。根据时区图显示可以看出:跨学科、人工智能、阅读推广和安全观这四个关键词在图中的圆圈较小,出现与发展的时间较短,但这恰恰是图情领域未来一段时间的学科研究关注点。

4 研究结论

4.1 高频热点主题分析

(1) 图书馆的热点主题研究。图书馆的相关研究一直都是图书馆、情报与文献学学科领域在国家社科基金中立项的选题热点。从上述图表可以看出“图书馆”几乎是每一年度的高频关键词,那么可以将其看作是该学科领域的核心关键词。该学科对图书馆的相关研究主要集中在图书馆社会教育职能的发挥、图书馆用户服务、馆藏资源建设、数字图书馆建设等内容,且又将研究范围细化为高校图书馆和公共图书馆两大类,将关注点延伸至基层、偏远地区的图书馆建设与开发以及完善社会弱势群体的阅读服务等。

(2) 大数据的热点主题研究。“大数据”频次仅次于“图书馆”,也是近几年国家社科基金的高频关键词。大数据时代的到来为图情档学科研究,尤其是

情报学领域研究带来了研究技术应用、研究思路与研究环境的变革,构建适用于大数据环境的情报工作流程以及方法体系,已经成为当前情报学研究的热点之一^[7]。大数据环境带来的改变不仅是情报学研究的对象由“信息”深入到“数据”,还有情报学和情报工作概念链“数据-情报”环节的增加。且在大数据环境下,“互联网”“云计算”“网络舆情”“数据挖掘”这些词汇的高频出现,说明情报学数据分析研究的技术方法也是紧跟时代发展脚步的。

(3) 信息资源的热点主题研究。“信息资源”的相关研究是国家社科基金重点资助的选题之一。信息资源作为一种提高国家竞争力和创新力的重要战略资源,对信息资源的研究是图情档学科最具活力和影响力的研究重点之一。从立项项目来看,对信息资源的研究包括信息资源的开发利用、信息资源管理、信息资源共享以及用来支撑信息资源存储的数据库的建设;从信息资源分类来看,不仅包括图书馆馆藏资源建设,还有网络信息资源挖掘、利用与存储。除此之外,政府信息资源的共建共享也是研究的热点。

4.2 学科未来发展趋势分析

(1) 大数据与人工智能。人工智能是一种新的技术科学,是吸收了大量数据由人工制造出来应用系统用于模拟和拓展人的智能。在大数据环境下,人工智能的发展使我们在海量数据中快速挖掘有效信息并形成知识成为可能。在未来,大数据技术与人工智能的持续发展会给情报学学科带来更为丰富完善的研究技术手段。

(2) 跨学科交叉融合。跨学科体系的构建是情报学未来应当着重思考的问题。各行业、领域之间技术、理论的相互联系、相互交流利于图情档学科在自有特点的基础上取长补短,发展构建军事、安全、科技、国防、医疗卫生、社会经济、生态环境、政府决策、历史文化为一体的学科体系。

(3) 阅读推广。为了响应国家倡导的“全民阅读”活动,对于图书馆阅读推广的服务质量管理的相关研究越来越多,如提供个性化阅读推荐服务、关注读者阅读行为与阅读心理等。尤其是在新媒体环境下,阅读推广策略与活动的变化研究在未来一段时间会受到越来越多的关注。

(4) 国家安全观。2017年6月,我国颁布了《中华人民共和国国家情报法》。这个法案的颁布为“总体

国家安全观”在图情领域发展提供了法律支撑。同年,在南京召开了“情报学与情报工作发展论坛”,会议达成“南京共识”。该共识提出情报工作应不遗余

力为国谋略,为建设情报强国而努力。基于此背景,学科立足于国家安全观与军民融合战略的相关研究会成为情报学学科未来发展的重要领域。

参考文献:

- [1] 杨宁,文奕.基于国家基金立项的图书情报学研究热点与趋势分析[J].情报科学,2017,35(2):138-141.
- [2] 赵蓉英,赵浚吟,陈必坤.我国“图书馆、情报与档案管理”学科的研究现状分析:以2001-2012年国家基金项目为视角[J].情报杂志,2013,32(7):106-112.
- [3] 李英.我国图书情报与档案管理学科研究现状剖析:基于2009-2013年国家自然科学基金和国家社会科学基金立项的分析[J].图书情报工作,2014,58(9):31-36.
- [4] 柴欢,阮建海.基于2013—2017年国家基金项目的“图书馆、情报与档案管理”学科研究现状分析[J].情报科学,2019,37(8):163-169.
- [5] 陈维军.文献计量法与内容分析法的比较研究[J].情报科学,2001(8):884-886.
- [6] 陈扬森,陈军.基于关键词共现与社会网络分析法的国内外社交媒体研究热点分析[J].信息资源管理学报,2017,7(3):45-55.
- [7] 李广建,祝振媛.基于硕博学位论文和国家社科基金项目的我国情报学研究热点分析[J].图书情报研究,2016,9(2):3-11.

作者简介: 张蒙,黑龙江大学信息管理学院2019级硕士研究生,研究方向为信息管理;刘春艳,黑龙江大学信息管理学院教授,博士,研究方向为知识管理与知识服务。

收稿日期: 2020-02-17