



情报杂志
Journal of Intelligence
ISSN 1002-1965, CN 61-1167/G3

《情报杂志》网络首发论文

题目: 从文献计量学到“五计学”:计量学方法的演化与发展
作者: 周子番, 邱均平, 魏开洋
网络首发日期: 2021-09-08
引用格式: 周子番, 邱均平, 魏开洋. 从文献计量学到“五计学”:计量学方法的演化与发展[J/OL]. 情报杂志.
<https://kns.cnki.net/kcms/detail/61.1167.G3.20210908.1004.022.html>



网络首发: 在编辑部工作流程中,稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定,且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式(包括网络呈现版式)排版后的稿件,可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定;学术研究成果具有创新性、科学性和先进性,符合编辑部对刊文的录用要求,不存在学术不端行为及其他侵权行为;稿件内容应基本符合国家有关书刊编辑、出版的技术标准,正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性,录用定稿一经发布,不得修改论文题目、作者、机构名称和学术内容,只可基于编辑规范进行少量文字的修改。

出版确认: 纸质期刊编辑部通过与《中国学术期刊(光盘版)》电子杂志社有限公司签约,在《中国学术期刊(网络版)》出版传播平台上创办与纸质期刊内容一致的网络版,以单篇或整期出版形式,在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊(网络版)》是国家新闻出版广电总局批准的网络连续型出版物(ISSN 2096-4188, CN 11-6037/Z),所以签约期刊的网络版上网络首发论文视为正式出版。

从文献计量学到“五计学”：计量学方法的演化与发展^{*}

周子番^{1,2} 邱均平¹ 魏开洋^{1,2}

(1. 杭州电子科技大学中国科教评价研究院 杭州 310018;

2. 杭州电子科技大学管理学院 杭州 310018)

摘要:[目的/意义]梳理我国从文献计量学到“五计学”方法的发展脉络能够明晰计量学方法的起源、演化与应用,为计量学方法的应用和创新整理思路。[方法/过程]利用余弦相似度方法、共现分析法、内容分析法等回顾与总结了“五计学”主题的演化发展过程以及“五计学”方法在不同发展阶段特征,并依据存在的问题提出了构建“五计学”方法体系的建议。[结果/结论]研究发现:“五计学”概念之间在不同研究主题中存在着融合使用的趋势;“五计学”方法之间关联程度较深,并存在着继承性与交叉性;应该秉承着问题导向的标准构建多层次、多维度 and 动态化的“五计学”方法体系。

关键词:“五计学”;方法体系;演化发展;文献计量

中图分类号:G353.1

From Bibliometrics to "Five-metrics": the Evolution and Development of Bibliometrics Methods

Zhou Zifan^{1,2} Qiu Junping¹ Wei Kaiyang^{1,2}

(1. Hangzhou Dianzi University Chinese Academy of Science and Education Evaluation Hangzhou 310018;

2. Hangzhou Dianzi University School of Management Hangzhou 310018)

Abstract:[Purpose/Significance] In recent years, the "five-metrics" has shown a trend of integration and development. Combing the development context of China from bibliometrics to "five-metrics" methods can clarify the origin, evolution and application of "five-metrics" methods, and sort out ideas for the application and innovation of "five-metrics" methods. [Method/Process] This article uses the cosine similarity method, co-occurrence analysis method, content analysis method, etc. to review and summarize the evolution and development process of the theme of "five-metrics" and the characteristics of the "five-metrics" method at different stages of development, and based on the existing problems, suggestions for constructing the "five-metrics" method system are put forward. [Result/Conclusion] The study found that there is a trend of fusion between the concepts of "five-metrics" in different research topics; The "five-metrics" methods are deeply related, and there are inheritance and intersectionality; We should build a multi-level, multi-dimensional and dynamic "five-metrics" methodology system based on problem-oriented standards.

Key words: five-metrics; method system; evolutionary development; bibliometrics

0 引言

所谓研究方法就是人们为了达到某种目的而采用

的方法、手段的总和。最初,文献计量学方法脱胎于图书情报学的交叉领域,因此在早期研究中也时常认为文献计量学就是图书情报学方法体系的一种专门方

基金项目:国家社会科学基金项目“我国‘五计学’融合和图书情报学的方法创新研究”(编号:18BTQ080)的研究成果之一。

作者简介:周子番(ORCID:<https://orcid.org/0000-0003-4445-9182>),女,1996年生,博士研究生,研究方向:信息计量与科教评价;邱均平(ORCID:<https://orcid.org/0000-0001-8660-3491>),男,1947年生,杭州电子科技大学中国科教评价研究院院长、资深教授,研究方向:信息计量与科教评价;魏开洋(ORCID:<https://orcid.org/0000-0003-0316-728X>),男,1997年生,硕士研究生,研究方向:信息计量与科教评价。

法。王崇德认为日益发展的文献计量学方法极有可能成为情报学的特殊方法,它们既可以解决情报学中的一些问题,也会被其他学科作为研究方法广泛移植^[1]。叶继元把图书馆学研究方法分为一般方法和专门方法,其中专门方法就包括文献计量方法、引文分析法等^[2]。近几年有关图书情报学研究方法使用情况调查研究的论文中,文献计量学方法也常常作为图书情报领域的方法来进行编码。但是方法的发展存在着泛化、分化和消亡的过程,引文分析由于对其研究地不断深入,最终也独立成为图书情报学的一类专门方法^[3],因此文献计量学方法已经成为一个方法集合^[4]。

另一方面,因为文献计量学、科学计量学与信息计量学在概念内涵、研究对象和应用方面既交叉重合又有区别经常被合称为“三计学”,而后学者发现文献计量学(Bibliometrics)、科学计量学(Scientometrics)、信息计量学(Informetrics)、网络计量学(Webmetrics)和知识计量学(Knowledgometrics)也在研究对象、研究内容、研究方法、数据来源、计量指标等方面也呈现出交叉关联、互为引用的发展特征和趋势^[5]。邱均平教授率先提出了“五计学”的概念,目前已经得到了中国学界的广泛认可,2018年还获批了国家级社会科学基金项目“中国‘五计学’融合和图书情报学的方法创新研究”。但目前直接研究“五计学”的文献还较少,主要分为两类:“五计学”的研究进展与“五计学”的演进过程。“五计学”进展研究从“五计学”个体的研究现状、学科分布、文献特征到作者合作情况全面分析了“五计学”的现状^[6-8]。从“五计学”的演进来看,关于“五计学”的研究主要集中于理论和应用两方面^[9]。随后,赵蓉英进一步从科研成果、科研基金、课程教育、人才与机构、科学评价、软件与工具等6个宏观角度发现“五计学”在理论基础、研究对象等方面相互区别又紧密联系,“五计学”的影响逐渐扩大,研究广度扩展、粒度变细,支持工具也在增加^[10]。

但上述“五计学”研究均集中于理论与应用的讨论,文献计量学作为图书情报学的一类专门方法,“五计学”方法研究也应是其重要的研究主题。虽然当前研究对图书情报学领域的研究方法梳理已经非常多,但是对“五计学”方法的分析仍然存在一定困难,一是国内并没有专门的计量学期刊,文献计量学方法一直作为图书情报学的专门方法存在于图书情报领域的研究中,在概念与来源上看,文献计量学、科学计量学、信息计量学、知识计量学和网络计量学在不同时期由不同学者提出,在概念上略有区别且研究对象不一致,然而纵观国内的研究,学者们对上述五个概念的使用存在着交叉混用现象,尤其是文献计量学、科学计量学和信息计量学这三个概念,最终文献中使用哪个概念通

常取决于学者的喜好,而且另一部分使用计量学方法的学者学科背景并非图书情报,对上述几个概念的了解和区分更不明确,因此要识别出使用了“五计学”方法的文献就存在一定困难;二是“五计学”存在微观与宏观两种性质的解释,以文献计量学为例,宏观上来看文献计量学本身就是一种方法,在文献中的使用方式是通常为采用文献计量学的方法分析某一领域的发展情况,但从微观上看,文献计量学是将数学和统计学的方法应用于文献分析中,因此文献计量学自身也是由其他方法组成并且不断丰富发展的,这也就意味着对“五计学”方法的梳理存在着两面性,不仅要在宏观上关注文献计量学、科学计量学、信息计量学、知识计量学、网络计量学在研究中的应用情况,还要在微观上关注文献计量学、科学计量学、信息计量学、知识计量学、网络计量学的支撑方法的发展变化。

因此,本文提出了以下研究问题:(1)宏观上,使用了“五计学”广义概念的研究如何演变发展?(2)微观上,“五计学”方法自身丰富与发展状况如何?(3)如何构建融合发展后的“五计学”方法体系?

1 数据与方法

1.1 数据说明

“五计学”方法还是较多应用于图书情报领域,因此本文以CNKI中“图书情报”类的论文作为筛选基础。又因为主题检索是在文章标题、关键词和摘要中检索,只要在这三个位置提到主题词的文献就能够被检索到,尽可能满足了查全率与查准率。第一组数据就使用了主题检索来筛选使用了“五计学”广义概念的研究,检索式为SU=‘文献计量’+‘信息计量’+‘科学计量’+‘知识计量’+‘网络计量’+‘Altmetrics’+‘替代计量’+‘选择性计量’,在筛除无关文献后最终得到了7067篇文献;第二组数据为了得到没有使用广义概念但是也做了计量研究的文献,选择以计量为主题词做主题检索,检索式为SU=‘计量’AND LY=(‘大学图书馆学报’+‘国家图书馆学刊’+‘情报科学’+‘情报杂志’+‘情报资料工作’+‘数据分析与知识发现’+‘图书馆建设’+‘图书馆论坛’+‘图书馆学研究’+‘图书馆杂志’+‘图书情报工作’+‘图书情报知识’+‘图书与情报’+‘现代情报’+‘信息资源管理学报’+‘中国图书馆学报’+‘情报理论与实践’),通过人工判读的方式筛选掉无关文献后得到了1657条数据,检索时间为2021.3.5。

如图1所示,两组数据的发文趋势呈现较为明显的阶段性特征,虽然两组数据并不完全一致,但是其发展阶段的切分仍然存在一定共性。从图中的分布特征来看,可以初步将研究阶段分为三部分:①萌芽发展阶段,论文长期增长平缓,出现了以文献计量学、科学计

量学为核心的研究,研究对象以文献信息为主,研究方法主要是文献统计方法;90 年代初,随着互联网的萌芽,对情报的认识加深,信息计量学也随之产生。②快速发展阶段,较前一时期论文增长率大幅度提升,互联网的快速扩张也促使网络计量学逐渐成为一支研究较多的分支,研究方法更推荐计算机辅助方法;在文献和信息的基础上,还出现了知识计量学;③发展瓶颈阶段,带有广义“五计学”概念的文献增长趋缓,出现下滑趋势,直接携带“计量”一词的文献数量更在逐年减少,但本阶段社交媒体的普及从网络计量学中催生出了 Altmetrics,更多学科的研究方法被吸收进来。

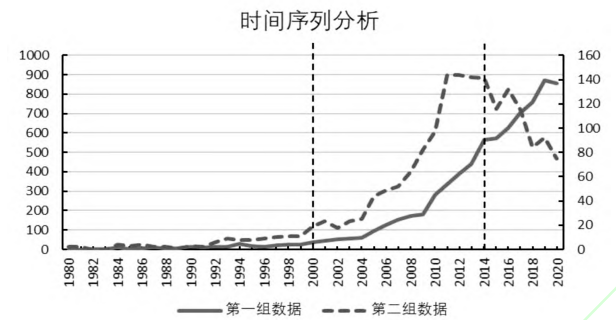


图1 两组数据时间序列分析

1.2 研究方法 (1)首先,在主题演化分析的研究中,Blei 把文献切分成几个时间片,并分别构建主题模型,再建立不同时间片之间的关联^[11],考虑到本文时间跨度较长,也将第二组数据平均分为 8 个时间切片,通过 SATI 对文献的标题、摘要和关键词进行分词处理,删除无意义通用词,合并近义词得到 8 个时间切片的词汇;再利用 Itginsight 对上述词汇进行聚类分析得到每个时间切片的研究主题,主题名称主要由聚类中出现频次较高的词语确定;其次,采用计算余弦相似度的方式计算不同时间切片中的研究主题的相似度,以此来构建“五计学”方法的演进过程。余弦相似度以向量空间中两个向量夹角的余弦值为衡量标准,以此来比较两个被向量化的个体之间的相似度。在二维向量空间中,两个二维向量:向量 $a(x_1, y_1)$ 和向量 $b(x_2, y_2)$ 在余弦定理中可表示为:

$$\cos(\theta) = \frac{ab}{\|a\| \times \|b\|} = \frac{x_1 x_2 + y_1 y_2}{\sqrt{+}}$$

同理,在 n 维向量空间中,以上公式仍然成立:

$$\cos(\theta) = \frac{\sum_{i=1}^n (x_i \times y_i)}{\sqrt{(x_i)^2} \times \sqrt{(y_i)^2}}$$

且余弦值越接近于 1,两个向量越相似。本文就将时间切片中描述研究主题的特征词汇向量化,构建词典并建立向量空间,通过两两计算比较余弦值来判断主题是否关联。

(2)利用词频分布特征发掘研究热点的研究已经非常成熟^[12]。首先,因文献的研究方法没有直接体现出来,本文参考了前人的研究方法编码形成了本文所使用的编码表,随后对第二组数据逐一做了研究方法编码,尽可能将每篇文献中涉及到的研究方法都列出来,能够明确的研究方法使用二级指标,不能确定具体研究方法的文献使用一级指标;其次,研究方法的使用往往无法脱离其研究对象,本文还将第二组数据中的关键词做了处理,去除掉方法类关键词,只保留依据词频、词量和累积词频算法计算得到的用以描述研究对象的中高频关键词;最后,通过对发文趋势的观察,本文将第二组数据分为三部分,再利用 access 建库得到研究方法和研究对象的二值矩阵并通过 ucinet 可视化。研究方法编码表如表 1 所示:

表 1 研究方法编码表示例

一级研究方法	二级研究方法举例
一般理论分析方法	抽象、归纳、演绎、推理、比较、分类、扎根理论、历史研究方法
统计分析方法	描述统计(集中、离散趋势)、推断统计(相关分析、方差分析、回归分析)、多元统计(主成分分析、因子分析、层次分析、多维尺度分析、数据包络分析、灰色关联分析)
调查分析方法	问卷调查、田野调查、访谈、德尔非法、焦点小组、面谈、专家会议
计算机分析方法	深度学习、支持向量机、遗传算法、最邻近分类、神经网络、决策树、贝叶斯网络、主题模型、聚类、关联规则、情感分析、文本分析
计量学方法	词频分析(共词分析、突变词分析)链接分析、引文分析(共引分析、共被引分析)
社会网络分析方法	中心性分析、核心-边缘结构分析
内容分析方法	主题分析、概念分析
专利分析	专利生命周期、技术关联度
案例分析	关键事件法
环境分析	PEST、SLEPT、SWOT、波特五力分析、情景分析、环境扫描、战争分析
其他	本体论、语义分析、句法分析、符号学、控制论

2 “五计学”主题演进过程

主题演进可以反映研究主题随时间的变化,包括主题内容或强度随时间推移的产生、增强、减弱、聚合、裂变、消亡过程,本文首先通过余弦相似度公式计算了 8 个时间切片的主题的相似度系数,再根据相似度系数大于 0.3 即为具有演化关系的原则^[13]。

“五计学”的研究主题主要是由文献计量和科学计量两个主题逐渐发展、裂变出信息计量、网络计量、知识计量等研究主题;在萌芽发展阶段,文献计量与科学计量始终是研究中极为重要的两类主题,直到 1990-1994 年出现了信息计量主题;而在快速发展阶段,

2000-2004 年网络计量研究主题最为突出,2005-2009 年又诞生了知识计量主题;在爆发增长阶段,聚类主题越来越多,主题间的演化关系也趋于复杂,引文分析、链接分析甚至可视化分析的主题出现意味着这些研究方法已经趋于成熟,2015-2020 年又在前五种计量学的基础上发展出了 Altmetrics 主题。然而,虽然将频次较高的词语作为聚类主题具有一定代表性,但是并不能进一步发现每个阶段具体研究方法的发展情况。因此,为了探索不同阶段的研究主题演化情况,本文分阶段分析了每个时间切片主题聚类内部高频词的演进关系。

2.1 萌芽发展阶段的主题演化 在本阶段,科学计量和文献计量是应用最多的方法之一,90 年代初产生的信息计量只在本阶段末期研究较多,信息计量的研究对象虽然扩展到任何活动中产生的信息,但信息计量的方法也强调数学统计方法的应用,因此也有研究确认文献统计方法、引文分析法等文献计量方法和洛特卡分布等规律在信息计量中仍然适用。总体上来看,作为萌芽阶段本阶段的主题演化还较不明显,因为更多研究主题立足于学科理论构建与研究。

从方法来源上来看,本阶段较多介绍和借鉴了国外的研究成果并在此基础上形成了国内的本土研究,介绍国外的著名学者如普赖斯、加菲尔德等的研究与贡献以及科学计量学杂志;理论与定律的探索为方法应用提供了基础,二八定律和文献计量学的三个定律是最多的研究如布拉德福定律的函数分布、词频分布规律,还有如普赖斯文献增长曲线、科学文献半衰期等;另一方面,学科之间的交叉特征在本阶段较为明显,学者对科学计量与科学学、科技管理的关系,文献计量与情报学、图书馆学等学科的关系做了较多的分析,论述了科学计量与文献计量的学科根基与方法意义。从主要方法上来看,文献统计方法是最先提出的方法,其次关于引文的方法如引文率、引文类型、自引率以及影响因子都有涉及,引文分析主要集中在引文数量的等级排序上;关于词频的分析,齐普夫第二定律确定了低频词的分布规律,Donohue J C 基于文章词频为 1 的词数量计算了低频词与高频词的临界值,孙清兰以不同词数为基础简化了高频、低频词分界临界值计算公式^[14];从工具上来看,本阶段计算机辅助研究的趋势在增强,前期工具以书目、索引、图书馆文献为主,在末期出现了数据库,利用计算机处理引文索引等,词频分析也在计算机普及之后重新焕发活力,它在揭示研究趋势上有较多应用^[15]。从方法应用上看,主要用于评价科学生产能力,构建综合评价指标。尤其是科学计量着眼于从投入产出评价科研成果,它还引入了专家加权法、量表打分法、排序法、关键事件法、目

标管理等方法^[16]来评价科研人员的绩效,而当某些因素无法直接计量时还可以采用兑换计量法,如转换计量法、互换计量法和成本计量法等将间接指标转化为可以计量的指标^[17]。

2.2 快速发展阶段的主题演化 在文献计量与科学计量两个主题之外,本阶段影响力最强的是网络计量方法,随着网络信息资源的极大丰富,网络计量学逐渐成为一个独立分支,知识计量萌生于 2005-2009 年,但影响力并不如网络计量广泛。

从方法来源上看,新生事物的蓬勃发展刺激了新方法的诞生,互联网的出现让网页、网站、网络用户等网络信息资源及其使用者成为新生力量,文献计量与科学计量并不完全适用于新生事物的计量,因此诞生了网络计量学。另一方面,对信息的认识深入到知识阶段使本阶段出现了知识计量、知识管理等研究主题,着重于“知识单元”层次上考察知识的投入和产出、流量与存量、生产与应用^[18],知识单元又通常表现为信息内容的概念、术语、词语等。然而在具体操作方法上,能够直接衡量文献选题材料方法与结论的方法主要为同行评议^[19]。尽管已经提出了“知识元”的概念,但是其概念还不能清楚界定,它在标引、组织、存储、检索、利用等方面具有不可操作性,无法脱离文献单元而计量^[20]。本阶段的主流方法分为以下几种:一是为了适应网络信息资源的存在形式而出现的新研究方法,如日志文件分析法、链接分析法、域名分析法、网络内容分析法、搜索引擎分析法、网络描述法等。又因网络计量学将网络之间的链接视为文献计量学中的引用关系,因此链接分析通过展示网页链接网络的密度、中心度、中心势,能够分析网站的网络影响力等。二是逐渐发展壮大数理统计分析方法,如聚类分析、因子分析等;三是对传统计量方法的继承发展,如本阶段不仅仅是引文统计和词频排序,而更多从共词、共引等角度观察文献中词频和引文的分布特征,从而提取研究主题,并利用多维标度分析将共现网络展示在二维坐标系中;四是内容分析的发展,上阶段对文献的计量主要集中在文献的外部特征的描述分析,而内容分析使研究内容深入到文献信息的内部;但知识计量方法的发展主要还是采取替代计算的方法,单篇论文知识流量可以从论文背景、期刊影响因子和被引计算^[21]。引入其他学科方法来辅助计算是另一种方式,知识的质量测量如知识普及度,知识价值的计量方法如投入法或重置成本法、市场价值法、边际生产法等^[22]。从方法的应用来看,在期刊评价、绩效评价之外,大学评价、学科评价出现,大学作为科研活动的主体更受关注;另一方面,门户网站的影响力逐渐增强,网站评价研究也在增加,新出现的网络部特征还可以作为计量指标,网页

数、链接数、流量、搜索量是描述网络载体的主要特征指标,网络影响因子是由此衍生的网络影响力评价指标,但国内外很多学者都对网络影响因子的计算方法进行过改进,而且计量结果各有差异。另一方面,传统文献计量指标的发展也有 h 指数、g 指数等指标。

2.3 发展瓶颈阶段的主题演化 总体上来看,网络计量和知识计量研究在本阶段影响力逐渐减弱,取而代之的是可视化分析和 Altmetrics 等新主题。

从方法来源上来看,大数据、云计算、人工智能等新技术为研究方法的进化提供了方向,开始有研究探索智能化的评价模型;同样的,社交媒体的发展也令网络计量发展到 Altmetrics 阶段,科学研究不仅仅局限于学术界,随着非正式科学交流活动的增加,科学活动的社会影响力也在增大,如何更好衡量社会影响力的 Altmetrics 指标的开发也就成为研究的重点。从方法本身来看,首先本阶段出现最多的研究方法是可视化分析,尽管从上一阶段就已经出现了知识图谱与可视化的研究,但是真正成为主流还是在本阶段,一系列可视化辅助工具如 Citespace、Ucinet、VOSviewer、Gephi 等的出现与应用也大大降低了可视化分析的进入门槛,使知识图谱、可视化等研究呈现出繁荣趋势。第二类方法是对 Altmetrics 指标的开发与应用,Altmetrics 出现之后,将 Altmetrics 指标与引文等传统指标结合成为主流,而且不仅只有外文指标,对中文 Altmetrics 指标的研究利用小木虫论坛中文期刊评价模块与中文期刊点评网的指标来评价中文期刊^[23]。针对 Altmetrics 数据多样化的特征,学者认为后续应该设计并研发相应的数据清洗、格式转换、词性标注和特征提取等方法来满足后续需求^[24]。从方法的应用来看,本阶段有关研究热点、研究现状、研究进展、发展态势、研究前沿和演进路径的研究急剧增多,主要还是源于研究方法的丰富。最初文献计量方法提供了最基本的文献统计分析,为了能够更加准确描述以及数据量逐渐增加

之后,数据收集方面,网络日志、搜索引擎等都提供了新的数据收集和预处理方法,在计算数据关系时可以使用数据共现分析、聚类分析、数据挖掘、推论统计、图论分析、关联分析、内容分析全面了解所分析的领域,在数据可视化展示方面,多维标度分析、社会网络分析和 Citespace 等可视化软件得出的各种图形结果都让研究热点与趋势的结果展示更直观。而且在前沿趋势研究中,对主题模型的探索让主题演进分析更为清晰,组织生态学中的 Lotka-volterra 模型能描述在学科融合背景下某一学科发展演化过程^[25]。另外一个得到发展的主题就是科学评价,包括机构、期刊、学者、论文、学科等评价,定性评价有同行评议,定量评价有构建评价指标体系;在确定指标时,不同的计量方法提供不同角度的分析指标,确定评价体系权重的方法有主观赋权法如专家咨询法、德尔菲法、层次分析法,客观赋权法如熵值法、变异系数法、相关系数法,筛选指标、发现指标间关系可以采取相关性分析、主成分分析、因子分析、聚类分析等,综合评价有层次分析法、结构方程模型、模糊综合评价法、Topsis 法、灰色关联法等。

3 “五计学”方法的发展阶段

3.1 萌芽发展阶段 如图 2 所示,萌芽阶段研究方法节点较少,节点网络密度也较小,节点之间联系并不紧密。研究主题节点中出现了文献计量、科学计量、情报计量学等概念,尽管情报计量学此概念在后期几乎不再使用,但本阶段对它们概念的界定和使用仍未形成统一标准。除此之外,本阶段也出现了一些较为明显的研究主题如学科评估、学术队伍建设、核心期刊确定和成果评价等;在研究方法的使用上,一般理论分析和描述性统计分析占据最主要的位置,因在发展初期理论梳理与概念构建是较为重要的课题,因此理论分析占据着重要位置,而且在此时期以前国内图书情报领域研究也以理论为主,较少使用定量方法。

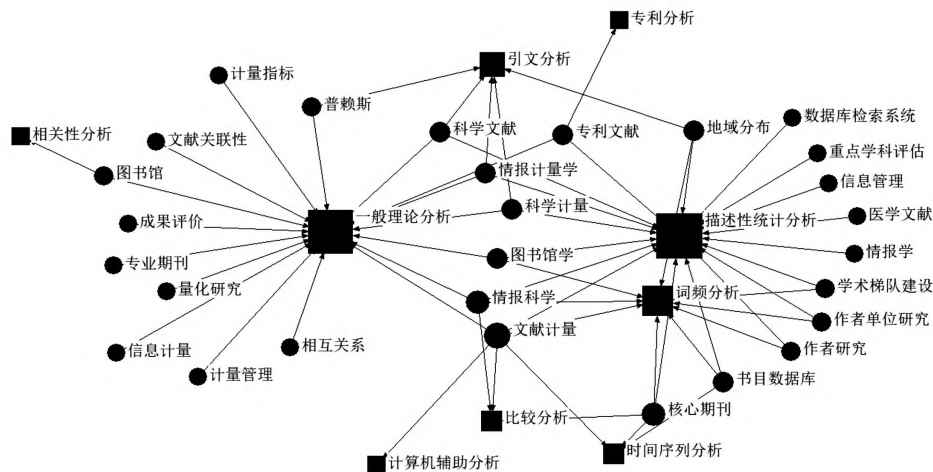


图2 萌芽阶段方法-关键词二模图谱

科学计量和情报计量关联的方法有一般理论分析、描述性统计分析和引文分析,文献计量关联的方法有一般理论分析、描述性统计分析、词频分析、时间序列分析、比较分析和计算机辅助分析等。因正处于诞生初期,文献计量、科学计量等作为方法集合时所包含的方法也是有限的,更多遵从于其定义将数学、统计学的方法应用于文献分析中,分析角度也多从文献外部特征如地域、作者、机构等出发。

3.2 快速发展阶段 虽然图3较前一阶段选用了同样数量的关键词,但是研究主题仍然发生了很大变

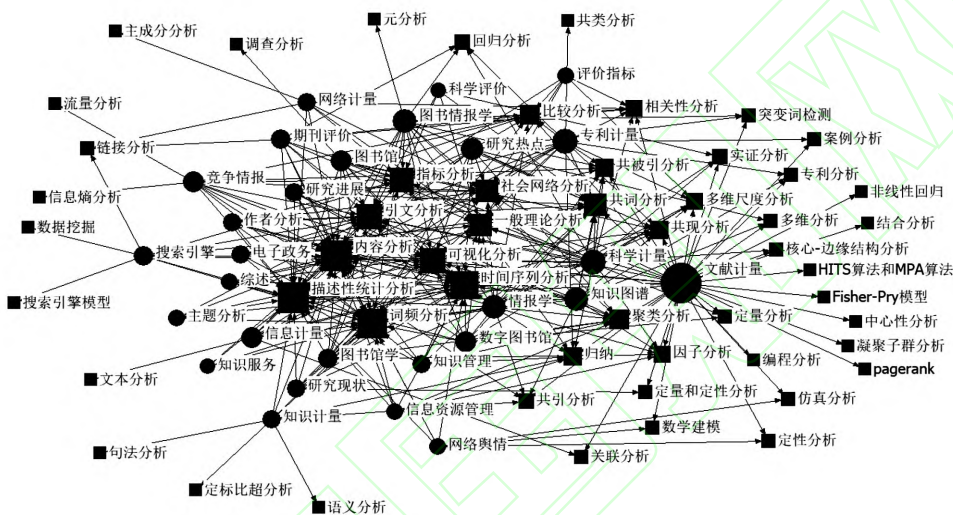


图3 快速发展阶段方法-关键词二模图谱

较之上阶段,文献计量发展极为迅速,它几乎囊括了图中三分之二以上的方法。不仅在传统描述性统计分析的基础上增加了相关分析、回归分析、聚类分析、因子分析等,引文分析和词频分析还扩展为分析引文之间的共现关系、突变关系检测,共引分析可从引文网状结构入手,通过聚类与多维标度方法将文章由共引形成的相互位置关系展示在图中^[26],而且文献计量方法从文献、作者、载文期刊等方面的计量深入到文献内容和主题的计量,新增了内容分析、社会网络分析、可视化分析等方法;因为科学计量方法围绕着科学家、科学成果、科学期刊、科学论文、科学交流活动、科学资源配置等展开,以全面展示科学活动水平和发展趋势,尤其是涉及到计量竞争力、绩效、效果等,科学计量方法借鉴其他学科理论的色彩更强,但是没有文献计量此概念使用广泛;在本阶段新增的信息计量本义是为了将对文献的计量深化到对信息的计量,虽然大部分研究方法继承自文献计量与科学计量,但是文本分析也能体现出其特色;知识计量则在继承前者的基础上增加了对句法、语义的分析;最后,网络用户和企业网站等新的研究对象出现,也促使以链接分析、流量分析、网络数据分析、搜索引擎法等研究方法为主的网络计量的诞生。

化,网络化的冲击使数字图书馆、电子政务等概念升温,网络舆情等数字化信息的研究爆发增长使信息计量、网络计量崭露头角,对知识的研究也让知识计量的概念出现,包括对知识管理与服务、知识图谱的研究;在研究方法数量上,计量学方法对其他方法的应用吸收逐渐明显,更多方法被吸收进来,尤其是文献计量所包含的方法发展迅速,计量学方法的应用逐渐划分为两个领域,一是对发展现状的分析,二是利用计量学指标开展科学评价。

3.3 发展瓶颈阶段

图 4 呈现出的研究主题较前一阶段发生的变化是:“五计学”主题基本稳定,不过网络时代交流活动的方式出现了巨大的变化,连带着科学交流也进入了新阶段,社交媒体对学界的冲击不容小觑,Altmetrics 指标备受关注;虽然也延续了上一阶段文献综述和科学评价的研究,但是本阶段较多讨论同行评议、学术影响力、学术评价,综述呈现模式也更多,可视化、知识图谱、演进路径等研究激增。从研究数量上来看,与前一阶段相比变化不大,尤其是核心研究方法并未有太多变化。

文献计量仍然是最大节点,几乎连接了所有出现的研究方法,科学计量与信息计量关联的方法差异不大,主要以聚类分析、可视化分析、共词分析、引文分析、内容分析、一般理论分析、描述性统计分析、词频分析等核心方法为主,但文献计量方法涉及面更广;知识计量在方法上并没有新的发展,仍以理论分析和继承其他方法为主;此阶段的网络计量连接的方法增加了计算指标权重的熵权法,不过与网络计量最相关的还是 Altmetrics 的发展,Altmetrics 的指标来源十分丰富,可以从多个角度计量,但这也是 Altmetrics 面临的问题之一,各个数据源中之间的标准不统一,Altmetrics 的相关研究也主要是对 Altmetrics 指标与传统指标相

关性的探索,Altmetrics 指标内部的结构与关系以及 Altmetrics 指标如何应用于科研活动评价中,因此它关

联了一系列诸如层次分析、主成分分析、因子分析等指标计算和构建的方法。

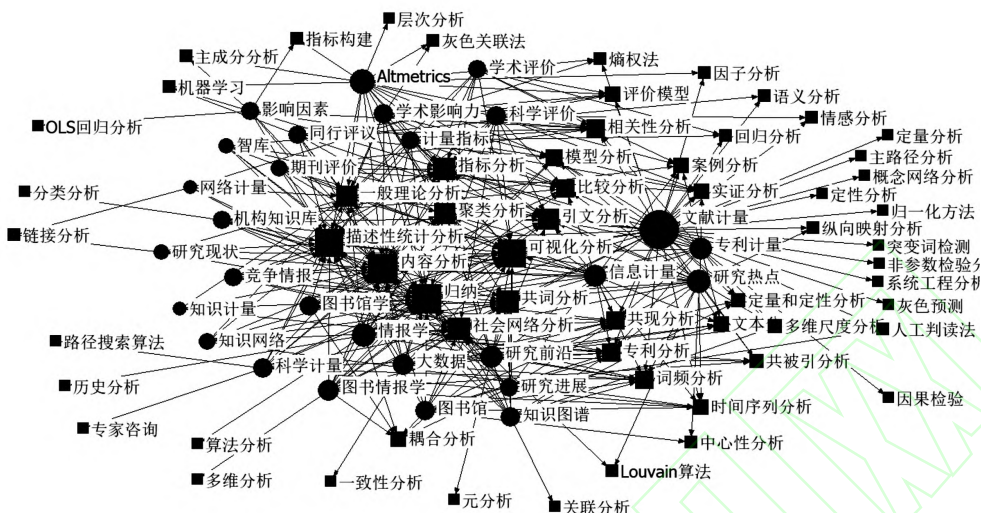


图4 发展瓶颈阶段方法-关键词二模图谱

4 “五计学”方法体系的构建展望

4.1 “五计学”方法的研究现状与问题 通过对“五计学”主题演进过程和“五计学”内部方法发展阶段的分析可以看出国内从文献计量学、科学计量学、信息计量学、知识计量学和网络计量学发展至今,尽管它们五者在提出时间、概念、研究对象和方法上存在着一定差异,但是已经在研究中出现了融合发展的现状,此种融合的原因可以归结为三点:其一,在主题演进图中“五计学”的研究对象是互有交叉的,或是随着研究主题的复杂化,它们都被用于同一研究之中,尤其是在评价体系构建研究多是混合使用了多种计量指标以达到客观评价的目的;其二,在具体分析“五计学”内部方法时也发现了在不同的发展阶段,即使研究方法总量上存在差异,但其所囊括的具体方法存在继承性和交叉性,尤其是文献计量、科学计量和信息计量三者的交叉融合程度更深,所关联的方法几乎一致,知识计量因为对知识单元的界定和研究还较困难,所以仍然以知识的载体即文献为研究对象,研究方法仍然沿用文献计量,网络计量也有将文献计量的方法迁移到网络信息资源上,因此它们在研究方法上的融合非常显著;其三,很多学者在使用这五个概念时,并没有仔细区分概念之间的区别,更多以自身的偏好为主,笔者在研究方法编码时发现了有较多文献虽然在标题或者关键词中使用了文献计量的概念,但实际上从概念上理解应该算作科学计量的范畴,而且对于非图书情报领域的学者来说,他们更倾向于使用那个被使用最频繁的概念即文献计量,这也导致在概念使用上出现了马太效应,给后来人造成一定误导。

4.2 构建“五计学”方法体系的建议

4.2.1 动态化 从发展过程来看,计量学方法随着认识加深以及环境变化而进化,从文献计量到信息计量到知识计量,从只关注图书馆期刊等文献情况发展到一切科学活动中的信息,另一方面又深入文献单元提出知识单元的概念,知识计量也成为发展方向。另外在互联网发展下,网络信息资源也立刻被纳入计量体系之中诞生了网络计量和 Altmetrics。上述都是环境的剧烈变化而催生出的新研究分支,另外新研究工具的出现也催生了可视化等方法成熟。因此,在构建“五计学”方法体系时首先要明确的一点就是要为未来将会新增的方法留有余地,尽管大多数方法体系都会尽可能追求本体系能够尽可能囊括当前研究中使用的所有的研究方法,但是固定化的方法体系随着时间发展会逐渐不适当当前研究现状,尤其是对教学有不利影响。

4.2.2 多维度 如果用网络图来描述“五计学”的方法体系,它应该是多层级、多维度的立体网络,“五计学”自身就存在着复杂的交叉融合衍生关系,并不能够简单地将它们分割成不同的部分,如引文分析在文献计量、科学计量、信息计量、知识计量中都存在,在划分归属关系难免无法抉择,在研究主题于方法关联图谱中也能够看到几者在核心方法都存在关联,而且随着某类方法的发展它也会逐渐扩展二级方法,如引文分析再细一级还可以分为共引分析、共被引分析等。另外,研究方法的划分标准不同,如按照研究对象或者分析角度划分时的方法体系也是不同的,但是不同的划分标准之间是不存在绝对的正确与错误的,因为研究方法本身就具有不同的属性,只有多维度的划分标准才能够更准确地描述研究方法。

4.2.3 问题导向 一方面,因为文献计量、科学

计量、信息计量、知识计量和网络计量等概念在使用上的双面性,其自身即使方法名称也是方法集合,给初学者或是非本领域学者使用造成一定障碍,即使文中提及使用了文献计量方法仍然不能知晓此方法在问题分析中起到了何种作用,又或者文献直接使用了较文献计量等概念次一级的概念如引文分析、链接分析等,这种概念使用混乱的状态不利于理清“五计学”的方法体系;另一方面研究方法诞生的使命就是为了解决实际问题,只要有新的需求出现就会促使学者寻找合适的方法来满足需求,如何归类研究方法必须要考虑到对研究人员来说如何使用研究方法,类似于情报收集、整理、分析、输出的过程,“五计学”方法在满足不同需求时也是不同的,如在评价体系中作为计量指标出现,在综述研究中提取研究主题可以采用共词法,分析作者合作可以使用社会网络分析。所以,在构建“五计学”方法体系时还应该分析“五计学”的研究对象,以及它们能够从哪些方面怎样解决问题。

参 考 文 献

- [1] 王崇德. 图书情报学方法论[M]. 科学技术文献出版社, 1988.
- [2] 叶继元. 图书馆学学术规范与方法论研究[M]. 北京: 科学出版社, 2014: 260.
- [3] 叶 鹰. 情报学基本教程[M]. 3 版. 北京: 科学出版社, 2018: 62-65.
- [4] 李 勤, 刘桂锋, 刘小文. 基于生命周期理论的图书情报学研究方法体系构建研究[J]. 图书馆学研究, 2020(15): 2-8.
- [5] 宋艳辉, 邱均平. 从“三计学”到“五计学”的演化发展[J]. 图书馆论坛, 2019, 39(4): 1-7.
- [6] 文庭孝. 近五年来国内外“五计学”研究进展[J]. 情报学进展, 2018, 12(0): 81-116.
- [7] 宋艳辉, 孙玉坤. 我国“五计学”作者合作研究[J]. 图书馆论坛, 2019, 39(4): 22-28.
- [8] 薛丽娜, 赵鸿玉. “五计学”研究的学科分布与文献特征分析[J]. 传媒论坛, 2020, 3(19): 94-97.
- [9] 赵蓉英, 魏明坤. “五计学”在我国的发展演进分析[J]. 现代情报, 2017, 37(6): 155-159.
- [10] 赵蓉英, 张心源, 张 扬, 等. 我国“五计学”演化过程及其进展研究[J]. 图书情报工作, 2018, 62(13): 127-138.
- [11] Blei D M, Lafferty J D. Dynamic topic models. 2005.
- [12] Kleinberg J. Bursty and hierarchical structure in streams[J]. Data mining&knowledge discovery, 2003, 4(7): 373-397.
- [13] 刘自强, 王效岳, 白如江. 多维度视角下学科主题演化可视化分析方法研究——以我国图书情报领域大数据研究为例[J]. 中国图书馆学报, 2016, 42(6): 67-84.
- [14] 孙清兰. 高频、低频词的界分及词频估计方法[J]. 情报科学, 1992(2): 28-32.
- [15] 邓路华. 词频分析——一种新的情报分析研究方法[J]. 大学图书馆通讯, 1988(2): 18-25.
- [16] 田志锋. 人力资源部门绩效指标体系及科学计量探析[J]. 科学管理研究, 2005(5): 97-101.
- [17] 刘振华, 于晓芳, 武春友. 科技人才绩效评估方法研究[J]. 科研管理, 2007(S1): 90-94.
- [18] 文庭孝, 邱均平. 科学评价中的计量学理论及其关系研究[J]. 情报理论与实践, 2006(6): 650-656.
- [19] 姜春林, 张立伟, 谷 丽, 等. 知识单元视角下学术论文评价研究[J]. 情报杂志, 2014, 33(4): 29-34.
- [20] 文庭孝. 知识计量单元的比较与评价研究[J]. 情报理论与实践, 2007(6): 731-736.
- [21] 姜春林, 刘则渊, 姜照华. 知识群的知识流量计量及其动力学模型[J]. 科学学与科学技术管理, 2010, 31(2): 82-85.
- [22] 余以胜, 张 洋. 知识的计量与评价研究[J]. 图书情报工作, 2008, 52(11): 18-21.
- [23] 李 勇, 邵钟钰, 赵 星. Altmetrics 背景下的期刊多维度测评指标研究[J]. 情报学报, 2017, 36(2): 190-196.
- [24] 赵蓉英, 郭凤娇. Altmetrics: 学术影响力评价的新视角[J]. 情报科学, 2017, 35(1): 14-18.
- [25] 邱均平, 何文静. 基于 Lotka-Volterra 模型的跨学科评价研究[J]. 大学图书馆学报, 2015, 33(5): 20-25.
- [26] 赵党志. 共引分析——研究学科及其文献结构和特点的一种有效方法[J]. 情报杂志, 1993(2): 36-42.