

定量化研究方法在学科情报研究中的应用研究

张晋辉 刘 清

(中国科学院国家科学图书馆武汉分馆, 武汉 430071)

[摘要] 针对学科情报研究中人们对定量化研究方法应用的一些困惑, 文章首先阐释了这两个概念并解释了学科情报研究对于定量研究方法的特殊需求, 然后通过具体实例说明定量化研究在学科情报研究工作中的应用分类, 并介绍了常用的定量研究方法以及指标、数据库等定量研究资源, 最后针对情报工作者在研究工作中如何更好地应用定量化研究方法进行了一些思考。

[关键词] 定量化研究 学科情报研究 应用 方法 指标 数据库

[中图分类号] G350

The study of quantitative study's application in discipline information research

[Abstract] Towards to same confusion on applications of quantitative study in discipline information research, this paper firstly explains the two concepts and the special demand of discipline information research on quantitative study, and then through the specific examples the article illustrates the categories of quantitative study's application in discipline information research. In addition, the article introduces general quantitative methods, indicators and databases in discipline information research. Finally give some thoughts for information researchers to apply quantitative study in their works.

[Keyword] quantitative study discipline information research application method indicator database

随着科学研究定量化发展趋势的日益明显, 定量化研究被广泛应用于情报研究的各个领域, 定量研究方法得到不断丰富, 指标和数据库不断增加。学科情报研究作为情报研究的一个重要组成部分, 很早便将定量化研究方法引入到它的具体研究工作之中, 大大提高了研究质量和效率。然而对于定量化研究应用于学科情报研究的一些问题, 比如定量化研究可以为学科情报研究做什么? 到底哪些方法、指标、数据库等可以应用在学科情报研究工作中? 如何应用? 如何在学科情报研究工作中更好的应用定量化研究方法? 等等, 这实际上是比较复杂的问题。本文通过调研相关文献并结合具体工作, 探讨这些问题, 旨在帮助有关人员更好的了解定量化研究及其应用于学科情报研究中的常用方法、指标和数据库等, 更好地开展学科情报研究工作。

1 定量化研究与学科情报研究

定量化研究是针对可计量研究对象，利用某些定量化研究资源，采用一定的定量研究方法和手段（数学方法、统计方法和引文分析法等），寻求将数据定量表示以揭示数据之间的关系和规律的一种研究方法。情报科学的定量化研究，就是在情报学研究中采用定量方法，进行定量分析，揭示数据间的关系和规律，其最初是从文献计量学开始的。早在 30 年代中期，当情报学尚处在形成初期时，定量化研究就在文献领域内展开，并取得了突出成果^[1]。随着情报研究活动的深入开展，定量化研究逐渐延伸到网络信息等其它领域，派生出信息计量学、网络计量学等其它的分支学科。

作为情报研究的重要组成部分，学科情报研究是针对特定学科，运用科学的研究分析方法，通过分析、对比、推理判断综合等逻辑思维过程，收集、整理、加工和分析有用的社会信息、科学知识和新的科研成果，有计划的、有目的、准确、及时地为学科科研活动提供服务的一种情报活动。其最终目的就是总结过去（如学科发展脉络）、把握现在（如学科发展态势研究）和展望未来（如学科发展趋势预测、前瞻等），从而把握学科发展方向，确定学科发展重点，为学科决策提供参考方案，换句话讲，学科情报研究可以揭示学科的内变化规律及其与其它相关学科的联系，弄清其历史和现状，并预测学科未来发展的趋势^[2]。学科情报研究工作的顺利开展必然要涉及大量理论和方法的应用，其定量化研究在学科情报研究方法体系中发挥了越来越重要的作用。

2 学科情报研究对定量化研究的特殊需求

由于现代科技、经济和社会的不断发展，信息资源的网络化及科学研究的信息化和社会化，使得情报学处于急剧变革和发展之中，也更直接地影响到了情报研究工作。其中定量研究一直是情报学乃至整个科学发展的重要方向和必然趋势。这是因为从学科发展的一般规律来看，情报学必然从定性阶段向定量阶段发展，只有加强情报学的定量研究，才能不断提高它在整个学科体系中的学科地位（王宏鑫，2003^[3]）。正如英国著名情报学家布鲁克斯所指出的：“情报学如果不实现定量化，它将是一堆支离破碎的技艺，而不会成为科学。”^[4]因此实现情报研究定量化，才能提高情报学的科学性和精确性，从而有助于确立和提高情报学在整个科学体系中的学科地位。

学科情报研究在情报收集、整理和分析的过程中对于理论和方法的需求是多样的。在情报研究的过程中，对于量化信息的需求变得日益突出。情报研究所分析对象的特点决定了分析过程中存在许多可计算的项目，比如科技文献、人力资源、经费投入等等。对于这些可计算单元的整理、分析和展示，就是定量化研究的专长。一般认为，基于定性方法的情报研究，其结果的精确度和可信度往往有限。特别是，如果面对的是极其复杂的系统，定性分析往往无能为力（人力投入和资金投入的需求巨大，而且往往还很耗费时间）。而量化方法相对而言具有精

确的特点，面对复杂系统的可操作性比较好。一般而言，结合定量方法的情报研究成果直观、可靠，易于接受，说服力强。因此，学科情报研究过程中便产生了对量化研究的特殊需求，量化研究方法随之应用到情报研究工作中，为情报学的发展提供了一种操作性强而又卓为有效的方法，弥补了定性研究的不足。

3 量化研究方法在学科情报研究中的应用分类

一般认为，在学科情报研究领域，量化研究方法主要可用于绩效评价和科学映射研究。其中科学评价功能是以科学产出和引用为基础，用以评价国家、大学、机构、个人等科研绩效水平；而科学映射的功能，目的在于揭示科学研究的结构和变化趋势。两者在情报研究中的地位和比重正在提升，下面笔者对量化研究方法在学科情报研究中的应用概况进行了总结分类，并分别列举案例加以说明（主要借鉴了我国学者的研究成果）：

（1）学科群规律分析。包括：学科群中各学科相互引用情况、间接预测学科交叉趋势，选择当采学科，确定科学中心，计算科学中心科学兴隆期等。例如程莹，刘念才（2005）统计分析了属于两个或多个学科类别的期刊跨学科类别和跨学科群，通过并研究交叉期刊定量地揭示了学科之间的联系程度^[5]。

（2）学科产出与影响力基本特征分析。包括：学科国际论文产出对国际论文产出的贡献，被评价国家对学科国际论文产出的贡献，被评价国家学科论文产出对本国论文产出的贡献，学科国际论文质量，学科被评价国家论文质量，学科国际论文影响力，学科国内论文影响力，学科被评价国家在国际发表论文的影响力等等（“被评价国家”可根据需要换成被评价机构、个人等科研实体）。如储宗元，王凤英（1998）依据《中国天文学文摘》，统计分析北京天文台等九个单位（部门）在该期间发表在各类期刊上的论文产出情况，简单分析了各单位（部门）的研究状况并展望中国天文学的发展情景^[6]。

（3）学科结构分析。包括：学科子学科产出规模及分布等。例如冯光卉等1996年以《高原医学杂志》著文与引文统计为基础口，研究了高原医学文献学科结构，作者及其情报吸收能力和高原医学的学科优势等特点^[7]。

（4）学科与相关学科分析。包括：学科独立性、稳定性，识别学科的相关学科及其联系程度。比如邢珏珏等人06年按照旅游学的十个分支学科，对1973-2003年国际旅游学界权威学术刊物《旅游研究纪事》的《主题词索引》进行了多元相关分析和聚类分析，阐述了这十个分支学科与整个旅游学以及相互之间的相关性，并分别探讨了这十个分支学科各自的进展态势^[8]。

（5）学科研究热点及未来发展趋势分析。包括：识别学科研究领域，学科活跃研究领域，学科子学科活跃研究领域，学科国际关注热门论文，预测学科各子学科交叉发展方向，追溯学科发展历史、预测发展趋势，学科国际研究热点子学科，学科中国研究热点子学科，学科与其它学科分化、组合趋势。例如曹小玉，

唐允华等采用布拉德福“文献离散律”的情报分析方法，对中药的权威检索工具《中文科技期刊资料目录·中草药》1995年的文献主题词、说明词进行统计，从不同角度分析了1995年中药学学科研究热点，中药治疗疾病的研究热门、热点，以及薄弱环节，并将年中药学学科研究热点与前5年中药学学科研究热点进行对比，从而提示了中药的发展趋势^[9]。

(6) 学科内国家及地区分析。包括：识别学科主要研究国家，识别子学科主要研究国家，学科被评价国家国际论文产出，学科被评价国家地区布局及论文产出，学科国际有影响力国家，学科被评价国家的学科影响力。如叶立平，凌啟柏(2003)利用MEDLINE数据库，运用文献计量学方法对1995-2001年循征肿瘤学相关文献进行统计，分析其国家分布等情况，以为广大肿瘤临床工作者提供参考^[10]。

(7) 学科内机构分析。包括：识别学科发文机构类型，识别学科核心机构，学科研究机构的科学生产力、机构地位，学科机构主要研究领域、及在该领域活跃水平，学科核心机构的地区布局，学科核心影响力的机构，学科研究机构对学科的影响力分析，学科研究机构对各子学科影响力，机构对学科的影响力、机构论文质量，学科核心研究机构地区布局。谭宗颖，刘佩华(2003)采用科学计量学方法对1991-2002年我国及世界玻色·爱因斯坦凝聚研究论文进行了横向比较，分析了国际合作模式和主体研究机构群^[11]。

(8) 学科科研人员分析。包括：学科科研人员整体规模，学科主要科研人员，学科子学科主要科研人员，学科科研人员组织结构，学科核心影响力科研人员，学科被评价科研人员特征参数。例如张子平，霍媛桃通过对2002-2004年《林业科学》引文的数量、语种分布及引文类型、论文作者的机构分布及合作度、论文获得的基金资助数量等统计分析，来了解我国林业科研人员的科研现状等^[12]。

(9) 学科期刊分析。包括：学科主要研究期刊，学科期刊对学科的影响力等。早有学者利用引文分析等方法对专门学科的核心期刊进行了识别，刘竞，李立恒1993年就对《中国农业科学》等20种期刊1990-1991年两年各期所有论文的被引文献进行统计分析，通过对期刊的借读量和对期刊论文的引文量、引文语种、引文类型、引文年代分布和自引率的统计定量研究，找出草原生态学科的核心期刊，初探草原生态学科文献情报源的分布情况和利用规律^[13]。

(10) 学科基金分析。包括：学科可使用基金及其产出对学科的贡献，学科基金产出对学科的贡献，基金在各学科产出，学科基金论文产出的年度变化等。如岳洪江(2004)利用受各种基金资助的国内论文统计数据，通过相对指标和变异指标，考察了20世纪90年代我国科学基金组织资助完成的论文学科与机构分布特征^[14]。

(11) 学科合作情况分析。包括：学科合作率，国家的国际合作水平，机构合作水平，学科个人合作水平，国家合作伙伴分析，学科机构合作伙伴，学科个人合作伙伴。如公晓晖，赵公正(2003)根据对中国人文社会学科11194篇论文

的作者合作数目的统计数据,对该领域存在的合作研究情况进行了统计分析,并对比分析了国际人文学科和国内自然学科的合作情况^[15]。

其中(2)、(6)、(7)、(8)、(10)部分评价功能属于学科绩效评价,(6)、(7)、(8)部分功能及(1)、(3)、(4)、(5)、(9)、(11)属于学科映射。

4 学科情报研究中常用的定量研究方法

量化研究方法的应用是以定量方法的使用为核心的,量化研究方法是一个由众多方法如数学方法、统计方法、引文分析法、系统分析法、矩阵分析法、网络分析法、层次分析法等等构成的方法体系,下面笔者针对学科情报研究中的常用几种定量方法进行简单介绍:

3.1 统计分析法,就是指对一些计量单元的直接统计分析,即对一些科学指标的计量研究。目前常用的科学指标主要有出版物数量、著者数量、引证频次、词频数量、科学发现(成果数)、期刊数、科学机构数、获奖人数和科学家年龄等等。通过这些指标的计量研究可以统计分析学科成果数或期刊数的时间分布、科学家年龄及结构分布和科学家人数按语言、领域、国籍等分布。

3.2 引文分析法,是根据文献间存在的相互引证的关系和特点,利用图论、模糊集合、数理统计数学方法及比较、归纳、抽象、概括等逻辑方法对科技期刊、论文、著者等各种分析对象的引用或被引用现象进行分析,以揭示其数量特征和内在规律的一种文献计量方法,通过引文分析,能理清学科的发展态势、核心著作以及整体的知识架构等。在学科情报研究中,引文分析主要用于对学科科研人员分析、学科期刊分析、学科产出与影响力基本特征分析等。目前引文分析中的引文网络、共引、自引现象以及科学地图等都得到了相当学者的高度重视,它也将会在很长一段时间持续成为研究热点。

3.3 共词分析法,是通过分析在同一个文本主题中的款目对(单词或名词短语对)共同出现的形式,以此为基础对这些词进行聚类分析,从而反应出这些词之间的亲疏关系,进而分析这些词所代表的概念间的相互关系,确认文本所代表的学科领域中相关主题的关系,进而探索学科的发展,利用共词分析方法可以帮助科研人员了解学科结构及其变化,加之时间维度上的比较,揭示学科领域的发展演进过程,帮助科研管理者进行学科布局和相关的决策,共词分析在学科情报研究中主要用来对学科研究热点及未来发展趋势进行分析。

3.4 数学模型分析法^[16]。数学模型是用数学语言表达出来的系统特征和其内部联系与外界联系的等式、不等式和数理逻辑等。数学模型分析法是模拟对象实际情况的同态映象,是为了简化研究而构造出来的研究对象的假定性的方法,学科情报研究工作中常常利用这种能如实反映情况的数学模型法,例如我们可以通过它分析科学期刊的被引分布实际情况,其优点在于简化研究对象,运用数学语言,输出定量化,从而使结果更为精确可靠。

5 学科情报工作中常用的定量化研究资源

定量化研究方法在学科情报研究工作中的应用往往会涉及一些定量化研究资源的使用,比如指标、数据库等,而要更好的应用定量研究方法就要对这些资源有所了解和掌握,下面笔者就对这些资源做下简单介绍。

4.1 指标。定量化研究的应用为情报研究提供了丰富的评价指标,其中学科情报研究中常用的分析指标有(参考 CWTS 和美国 ISI 公司的提出的文献计量指标):(1)反映学科产出能力的指标,主要有国际学科论文(会议报告及其它)产出数量、国际学科论文产出比重、产出速度、论文所属期刊影响力等级等。(2)反映学科影响力的指标,主要有引文总数、篇均引文、未被引论文比例、引文份额、学科相对强度、学科影响因子等。(3)反映学科间合作关系情况分析指标,主要有学科单一(非单一)作者论文数及比重、学科国际(机构或个人)合作论文数、学科国际(机构或个人)合作率等。(4)反映学科研究热点与发展趋势的评价指标,主要有学科论文(题目)关键词及频次、学科各二级子学科国际论文数时间分布、学科高被引论文、学科被引(施引)率等。(5)反映学科科研人员情况分析指标,主要有学科发文作者总数、学科发文数量前 50 位著者、学科科研人员耦合强度、学科被评价科研人员被引频次等等。此外还有一些对学科期刊、基金、机构的分析指标,这里就不再一一列举。这些指标的构建是情报研究定量化的反映和重要体现,同时其使用也极大提高了学科情报研究的定量化程度,方便了广大情报分析人员,但其中一些指标的设计存在着一些主观性和不合理的地方,引发了业界关于影响因子、h 指数等的争论,所以指标的科学性亟待实践的不断检验,数量也需要不断丰富。

4.2 数据库。学科情报研究工作中,使用最广泛的还是世界三大检索系统 SCI、EI、ISTP,它们被普遍应用于对某国家、机构和个人学科水平和影响力的评价,是情报研究人员进行引文研究和学术评价的主要依据。随着情报研究实践不断展开和定量化研究的不断应用,新的数据库层出不穷,比如美国 ISI 公司 2001 年推出科学计量指标数据库(ESI),分 22 个学科领域对国家、机构、期刊、科学家、论文进行了文献统计分析和排序,主要用于衡量学科内科学产出与影响力,是学科生产力评价的重要工具;Scopus^[17]是全球规模最大的文摘和引文数据库,功能强大、方便使用,特别还可以提供引文分析评价功能,提供针对作者的 h 指数计算,可以帮助我们对学科文献、出版物、学科领域、学科人才和科学机构进行评价;Google Scholar 是 Google 于 2004 年底推出的专门面向学术资源的免费搜索工具,能够帮助用户查找包括期刊论文、学位论文、书籍、文摘和技术报告在内的学术文献;美国科学基金会编发的《科学与工程指标》,从基础教育、高等教育、科技人力资源、研发资金等多方面介绍美国和全球的科技状况,也是非常重要的信息计量学参考工具和基础数据来源。

此外，我国也产生了很多数据库，如中国科学院国家科学图书馆研制的《中国科学计量指标：论文与引文统计》和中国科技信息研究所的《中国科技论文统计与分析（年度报告）》等在我国科技评估和学科评价中发挥了重要作用，但是学科情报研究工作中使用最多的仍是国外数据库，国内数据库的影响力和使用率远不及国外，说明我们与国外数据库相比还有很大差距，还需要在加强和国外数据库出版商合作的过程中不断完善。

6 学科情报工作中加强量化研究应用的一些思考

综上所述，正是学科情报研究工作对量化研究方法的特殊需求促进了量化研究方法在实际工作中的广泛应用，也促进了量化研究资源的不断丰富，学科情报研究工作应用量化研究手段也不断产出更为科学性和准确性的研究成果。情报研究者应在以后的工作中更多地运用量化研究手段，提高研究工作的质量。对于如何更好地应用量化研究，笔者结合情报实际工作和目前研究热点进行如下思考：

5.1 加强量化方法的研究，完善量化方法体系。从情报研究工作的特点出发，不断引入新的定量研究方法，笔者强调就数据挖掘和数据清洗等数据分析方法进行研究，并进行分析工具开发，弥补商业软件功能的不足，开展数据的准确性和完整性研究。特别是数据清洗的研究，目前国内外文献计量工具中的数据存在很多问题，而专门针对数据清洗的研究还很少，该问题的研究对于控制研究数据的可靠性和准确性，净化情报研究数据源和提高分析结果质量具有很大价值。

5.2 科研合作（个人、机构、地区与国际合作）的信息计量研究和科技评估指标体系的建立和标准化。通过对个人、机构、地区与国际合作的网络、重点的研究和分析，研究科技合作的网络和领域特点；加强对新的科技指标的研究和设立，针对评价指标不统一的现状，构建标准指标体系和分析流程，以更好的指导情报研究人员科学地进行科技评价工作。

5.3 做好理论和方法研究的同时，加强定量研究方法的学习和交流。针对目前除了杂志之外国内还没有其他有助于国内研究人员进行量化研究方法交流的平台，建议有关部门或有识之士牵头，搭建我国研究人员交流、合作的综合平台、定期组织情报研究人员的相关培训。通过这些举措普及量化研究知识，尤其是加强对量化研究技术和方法的学习，使量化研究成为情报研究人员工作中的重要工具。

参考文献:

- [1] 邱均平, 颜金莲. 中国情报学定量化研究的发展特点和趋势 [J]. 图书馆论丛, 1999, (1)
- [2] 郭吉安, 李学静. 情报研究与创新 [M]. 北京: 科学出版社, 2006.
- [3] 王宏鑫. 信息计量学的基础与发展研究 [J]. 图书情报工作, 2003, (2): 7-12.
- [4] 霍国庆. 百年沧桑三次高朝四代学人——20世纪中国大陆和台湾地区图书馆学总评 [J]. 图书馆, 1998, (3)
- [5] 程莹, 刘念才. SCIE、SSCI期刊跨学科现象的定量分析 [J]. 情报科学, 2005, 23(2): 237-263.
- [6] 储宗元, 王凤英. 1991—1995年间中国天文学科论文统计分析 [J]. 中国科学院上海天文台年刊, 1998, (19): 188-193.
- [7] 冯光卉, 马建得, 王旭萍. 高原医学文献及引文学科结构分析 [J]. 高原医学杂志, 1996, 26(2): 62-64.
- [8] 邢珏珏等. 旅游学分支学科相关性及其进展态势分析——《旅游研究纪事》30年 [J]. 人文地理, 2006, 21(1): 51-55.
- [9] 曹小玉, 唐允华. 1995中药学科研究热点分析 [J]. 成都中医药大学学报, 1996, 19(4): 4-7.
- [10] 叶立平, 凌啟柏. 循证肿瘤学文献计量分析 [J]. 中华医学图书情报杂志, 2003, 12(3): 59-60.
- [11] 谭宗颖, 刘佩华. 10年来玻色-爱因斯坦凝聚研究的科学计量分析 [J]. 中国基础科学, 2003(4): 32-36.
- [12] 张子平, 霍媛桃. 我国林业科研人员科研现状分析 [J]. 农业图书情报学刊, 2007, 19(1): 142-145.
- [13] 刘竞, 李立恒. 对草原生态学科期刊的引文分析 [J]. 甘肃科学学报, 1994, 6(4): 70-77.
- [14] 岳洪江. 我国科学基金资助成果学科与机构分布特征 [J]. 情报杂志, 2004, 23(12): 33-35.
- [15] 公晓晖, 赵公正. 人文社会学科合作研究的数量分析 [J]. 辽宁工程技术大学学报(社会科学版), 2003, 5(1): 110-112.
- [16] 王崇德. 文献计量学引论 [M]. 桂林: 广西师范大学出版社, 1997
- [17] <http://info.scopus.com/overview/what/> [EB]