

科学知识图谱在科技期刊编辑工作中的应用

姜春林 王海龙

大连理工大学 21 世纪发展研究中心, 116024, 辽宁大连

摘要 科学知识图谱是科学学和科学计量学领域的重要研究成果。本文利用引文分析、共被引分析、多元统计分析、词频分析方法以及社会网络分析等方法,以可视化的图形展示科学知识的结构、热点和趋势。认为将科学知识图谱应用于科技期刊编辑工作,能对期刊选题策划、遴选审稿专家等提供重要参考。强调科技期刊编辑应加强科学学和科学计量学理论的学习,熟悉先进的信息处理软件,并正确认识科学知识图谱在编辑工作中的辅助作用。

关键词 科学知识图谱;科技期刊;编辑工作

Application of mapping knowledge domains in editing sci-tech journals editing work//JIANG Chunlin, WANG Hailong

Abstract Mapping knowledge domains is one of the most important fruits in science studies and scientometrics, and as a tool, it can visualize knowledge domain structure, hotpots and academic trends using citation, co-citation, multidimensional statistics, word frequency, social network analysis etc. methods. It can be applied for science and technology journals' editing process such as choosing topics, reviewers etc. It is emphasized that editors should study knowledge of science studies and scientometrics on purpose. Assisting role of mapping knowledge domains in editing process is also pointed out.

Key words mapping knowledge domains; science and technology journals; editing work

Author's address 21st Century Development Research Center, Dalian University of Technology, 116024, Dalian, Liaoning, China

1 科学知识图谱及其功能

近年来,科学计量学的研究在深度和广度上都得到极大的推进,科学知识图谱已成为科学计量学研究的热点之一。科学知识图谱是指对大量科学文献新信息,借助于统计学、图论、计算机技术等手段,以可视化的方式来展示科学学科体系的内在结构、学科特点、研究前沿等信息的一种计量学方法。其描述的对象主要包括科学技术活动中从事知识生产的人、作为知识载体的论文、期刊、显性或者可编码化的知识,以及科学研究过程。其基本原理是基于文献单元(科学家、引文、机构、关键词、期刊等)的相似度分析,根据各种数学和统计学的原理来绘制科学知识图谱。科学知识图谱已经历了从二维图表、三维构型(3DCN)、多维尺度图谱(MDSM)、社会网络分析图谱(SNAM)、自组织映射图谱(SOM)、寻径网络图谱(PFNET)等几个发展阶段^[1]。

科学知识图谱的绘制过程大体包括以下 4 个环

节:1)数据准备阶段。即确定和获取原始数据,形成原始数据空间。一般是从数据库中下载相关主题的文献信息。2)数据提取阶段。从原始数据中析取需要可视的数据,形成可视化数据空间,即通过软件形成共作者、期刊共引、共词等计量单元的共现矩阵。3)可视化映射。即采用一定的映射算法把可视化数据空间映射到可视化对象^[1-8]。4)借助于相关学科的背景知识,对形成的科学知识图谱进行深入解读。

科学知识图谱分析所使用的数据主要来源于美国汤姆逊公司的科学引文索引(Web of Science),该数据库整合了 SCI、SSCI、Conference Proceedings Citation Index-Science 等大型知名数据库,所收录的期刊都经过严格的遴选,来源期刊及数据库均具有国际性、权威性和前沿性;因此,利用该数据库提供的相关数据绘制科学知识图谱保证了数据的可靠性和权威性。当然,中国期刊全文数据库、中文社会科学引文索引数据库等中文数据库也是绘制知识图谱的主要数据源之一。

绘制科学知识图谱的主要软件有 SPSS、Ucinet (<http://www.analytictech.com/>)、Bibexcel、Citespace (<http://www.pages.drexel.edu/~cc345>)等,后三者均是仅用于学术研究的免费开放软件,都可以在 Google 中检索到并下载。这些软件的界面简捷友好,使用方法简单,便于学习掌握。

科学知识图谱广泛应用于科学学研究,如研究学科发展历史,确定学科前沿和热点、主流研究群体、代表性人物、主流期刊、热门论文等。科学知识图谱在科技期刊编辑工作中的应用主要体现在:帮助编辑了解与本期刊报道主题密切相关的国内外学术动态,识别主要的研究团队;为审稿专家提供判断待审论文的创新点的重要参考,便于把握撰写审稿意见的尺度;帮助期刊编辑辨识高影响力作者和研究群体,利于遴选和邀请研究方向一致或接近的审稿专家,以进一步优化审稿队伍。

2 科学知识图谱应用举例

以我国中长期科学和技术发展规划中所列示的生物前沿技术之一的干细胞组织工程技术为例,来介绍这种方法的使用。

利用 Web of Science 中的 SCI 数据库和 CPCI-S 数

数据库为数据源,以“stem cell”与“human tissue engineering”作为关键词,选择逻辑关系符“AND”,对数据库进行高级检索,检索时段选择 1998—2008 年,一共有 1 280 条文献记录。通过 Web of Science 自带的分析工具来分析,得到该领域发表论文数量随时间的变化趋势以及国别分布,得知该领域进展较快,美国以近 40% 的比例高居各国之首,中国的份额为 7%,表明中国已经在该研究领域占有一席之地,但是距离美日英德等科技强国仍有一定差距。再将数据导入 Citespace 软件,执行操作程序后,生成了该研究领域的共词关系知识图谱,见图 1。其中,图谱的颜色对应于关键词共现的时间,节点圈越大表明共现次数越多。组织工程(tissue engineering)、干细胞(stem cells)、体外培养(in-vitro culture)、分化(differentiation)、骨髓基质细胞(marrow stromal cells)、骨髓间充质干细胞(mesenchymal stem-cells)、移植(transplantation)、软骨(cartilage)等是国际学术界关注的热点领域。

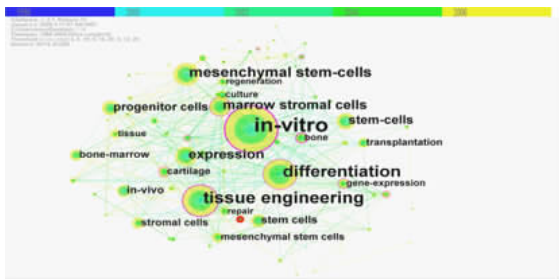


图 1 共词分析图(SCI, CPCI-S)

为比较中文文献的研究状况,我们还可以利用中国知网 CNKI 中的期刊全文数据库分别用“干细胞”“组织工程”“人体组织”进行主题搜索,时段范围为 1998—2008 年,共检索到 2 111 条数据,然后对文中关键词进行统计分析,对出现频次较高的前 50 个关键词利用自编软件作关键词共现分析。再用社会网络分析软件 Ucinet 绘制科学知识图谱,见图 2。

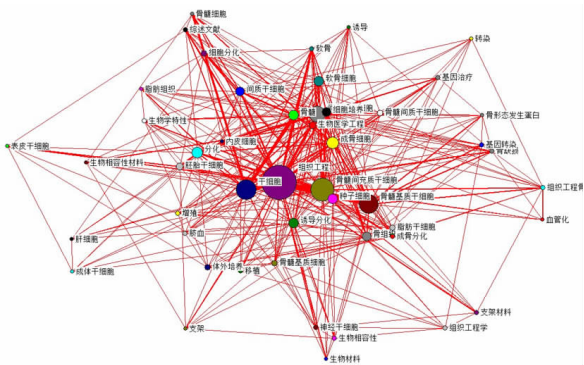


图 2 CNKI 50 个关键词共现网络图

经比较我们发现,中外文献在该领域的研究热点基本一致。

此外,还可以从文献共被引角度来探索在干细胞组织工程领域的主要研究方向。具体数据导入方法略。知识图谱中生成的引文聚类共有 4 个(见图 3),即:知识群 1——干细胞分化与移植;知识群 2——骨髓干细胞与组织工程;知识群 3——各类干细胞在人体组织中的作用;知识群 4——组织工程技术应用。

期刊的特色是期刊持续发展的基础,主要体现在报道内容上,因此,选题尤为重要。科技期刊只有围绕期刊定位和特色,并根据学科发展特点进行选题,不断扩大学术影响,才能保持长久的生命力^[9]。科学知识图谱展示的丰富信息,为编辑在期刊定位、栏目设计、选题重点等方面提供了重要参考。作为专业期刊,如果仅凭编辑部或者编委会主观的定性把握,有时难免会有失偏颇。

通过对图 3 中的关键节点进行分析,我们可以进一步了解哪些学者是该领域的权威科学家,哪些文献是重要文献。对这些引领学术发展方向的权威科学家,我们可以借助谷歌学术搜索(Google scholar)来进一步了解他们的工作机构和联系方式,从而为学术期刊寻找合适的中外审稿人提供可靠的保障,这就在一定程度上克服了以往期刊选定的审稿人要么不是该领域的权威,要么就是长期不在科研一线,最终影响审稿质量的弊端^[10-11]。同时,期刊编辑通过绘制科学知识图谱,可以在初审阶段对来稿的创新性、是否与国际前沿保持一致进行大致判断,进而决定下一步的审稿安排。

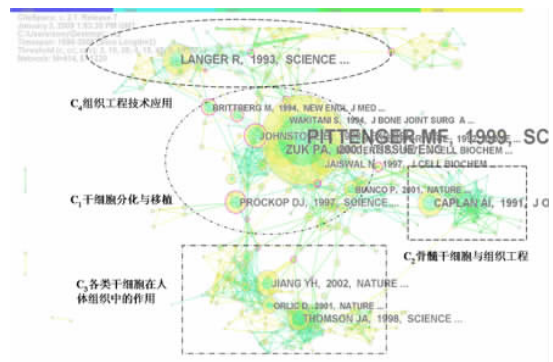


图 3 文献共引分析聚类图

此外,如果在外审阶段,把某一领域的科学知识图谱连同待审论文一并提供给审稿专家,也将极大地减少或避免某些评审专家囿于思维定势或由于精力有限而难以充分了解该领域大量文献,导致待审论文在科学性和创新性方面评判的失误。

3 科学知识图谱应用需注意的问题

科学知识图谱以大量文献为基础,基于词频分析、共词分析、共引分析、多元统计分析、社会网络分析等技术手段,为我们清晰地展示了科学技术领域的结构、

特征、热点、主流群体等信息,描绘了一幅幅生动、绚丽、直观的知识画卷,也为学术期刊提供了一把打开了解学科发展动态的钥匙。可见,发挥和拓展科学知识图谱的功用,必将对科技期刊编辑提出了更高的业务要求。当前,各类重要数据库的使用变得越来越方便,并且各种绘制科学知识图谱的工具简单易学,这为科学知识图谱在编辑工作中的普及应用提供了可能。当然,以下几个问题仍需要引起注意。

3.1 增强编辑工作创新意识 创新是科技期刊事业发展的源动力。随着计算机和网络技术的发展,大量高效的编辑工具和编辑平台出现,极大地提高了科技期刊的编辑效率。编辑工作中软硬件条件的改善并不能必然地保证科技期刊学术质量和影响力的提高。编辑人员在遵循基本编辑规律的同时,应不断强化科技期刊在科学交流系统和国家科技创新体系重要性的认识,增强编辑工作的历史使命感。积极创新编辑思路,开阔编辑视野,熟悉各种现代化的编辑软件,提高新技术新手段导入编辑流程的主动意识,从而进一步提高选题策划能力、对稿件创新性的识别能力和新知识传播的时效性。

3.2 加强科学学和科学计量学理论及最新研究成果的学习 要提高科技期刊的编辑质量,科技期刊编辑不仅仅要注意编辑学理论和编辑实务的学习,而且还要不断吸收与编辑工作密切相关的各学科知识。科技期刊编辑是不断与科学技术最新成果、学术专家和科学技术共同体近距离对话的特殊群体,如果不能深刻洞悉科学技术的内在特征、社会功能和发展趋势,就不可避免地影响编辑在科学技术交流系统中主体性功能的充分发挥,进而影响期刊的编辑质量。科学学是一门从总体上研究科学技术体系的结构、特征和发展规律的跨学科群,而科学计量学作为科学学重要的方法性学科,对推动科学学的发展具有根本性作用。科学知识图谱既是科学计量学研究的重要成果,也是科学计量学的研究前沿,它的萌生与发展对科学学及其相关学科的影响是深远的。科技期刊编辑应不断通过自学、培训、攻读学位等渠道,及时了解并掌握科学学和科学计量学领域的最新研究成果,提高科学学理论素养,深化对科学技术发展趋势以及编辑工作本质的认识,塑造学习型编辑的形象,锤炼学者型编辑的气质,真正扮演好科学技术成果“守门人”的角色。

3.3 正确认识科学知识图谱对编辑工作的辅助作用

科学知识图谱的功能无疑是强大的,应用于科技期刊编辑工作也给人以耳目一新之感;但是,我们仍然要清醒地认识到,任何先进的工具和技术手段,都不能完全取代科技期刊编辑在整个编辑流程中的主体地位,编辑也不应该完全摒弃独立的编辑思想,而完全依赖

某种新技术。“洞见”和“齐物”仍然是科技期刊编辑所追求的境界^[12]。科学知识图谱能否在期刊编辑中发挥更大的作用,一方面取决于编辑是否熟练掌握了科学知识图谱的有关理论和软件使用方法,另一方面还要求编辑具备很强的专业背景知识,否则很容易误读图谱、错读图谱,给编辑工作带来负面影响。对审稿专家,编辑也应力必说明为其提供的谱图仅供审稿时参考,以免干扰专家审稿。

4 结束语

科学知识图谱导入科技期刊编辑工作是一个新生事物,提倡科技期刊编辑们大胆尝试,扬科学知识图谱之长,弃其可能之短,使之能在科技期刊编辑工作中发挥应用的作用。

5 参考文献

- [1] 刘则渊,陈悦,侯海燕. 科学知识图谱:方法与应用[M]. 北京:人民出版社,2008:11-21
- [2] Schubert A. The web of Scientometrics: a statistical overview of the first 50 volumes of the journal [J]. *Scientometrics*, 2002, 53: 3-20
- [3] Braam R R, Moed H F, Van Raan A F J. Mapping of science by combined co-citation and word analysis: I. Structural aspects[J]. *JASIS*, 1991, 42(4): 233-251
- [4] Small H G, Griffith B C. The structure of scientific literatures I: Identifying and graphing specialties [J]. *Science Studies*, 1974, 4(1): 17-40
- [5] McCain K W. Mapping economics through the journal literature: an experiment in journal co-citation analysis[J]. *J Am Soc Inform Sci*, 1991, 42(4): 290-296
- [6] White H D, McCain K W. Visualizing a discipline: an author co-citation analysis of information science (1972-1995)[J]. *J Am Soc Inform Sci*, 1998, 49: 327-355
- [7] McCain K W. Mapping authors in intellectual space: a technical overview [J]. *J Am Soc Inform Sci*, 1990, 41 (6): 433-443
- [8] Chen C. CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature [J]. *J Am Soc Inform Sci*, 2006(3): 359-377
- [9] 王卫霞. 新时期科技期刊编辑的选题意识[J]. *编辑学报*, 2008, 20(3): 203-205
- [10] 李云霞. 加强审稿专家队伍的动态管理[J]. *编辑学报*, 2005, 17(1): 66-67
- [11] 陈斌. 专家审稿中存在的问题及其解决办法[J]. *编辑学报*, 2006, 18(6): 449-450
- [12] 李树华. 科技期刊编辑的高境界:洞见和齐物[J]. *编辑学报*, 2009, 21(1): 89-90

(2008-07-16 收稿;2008-08-29 修回)