

# 学科信息深度挖掘研究助力新刊出版的探索\*

李雪 董艺 李晓光<sup>†</sup> 崔晓健 王占坤

国家海洋信息中心《海洋通报》《海洋信息》编辑部,300171,天津

**摘要** 通过文献计量学与文本挖掘技术的研究,深入探索学科信息研究对期刊新创和全新改刊的科学价值,从主旨定位、热点发现、作者遴选、机构查找、地区判断、前沿跟踪、专题策划等方面进行分析,探讨信息深度挖掘研究助力新刊成长的方法和指导科学出版的重要意义。

**关键词** 数据挖掘;学科出版信息;信息反演;新刊创办

**Assisting new scientific journal publishing by deep mining subject data//LI Xue, DONG Yi, LI Xiaoguang, CUI Xiaojian, WANG Zhankun**

**Abstract** This paper explores the scientific value of the research of subject information on journals' launch and reform through the method of bibliometrics and text mining, discusses the method to promote the growth of new journals and the significance to guide scientific publishing by information deep mining research from the aspects of keynote location, hot spot discovery, author selection, organization search, regional judgment, frontier tracking and thematic planning, etc.

**Keywords** data deep mining; subject publishing information; information inversion; new journal establishment

**Authors' address** National Marine Data and Information Service, 300171, Tianjin, China

DOI:10.16811/j.cnki.1001-4314.2018.05.015

## 1 学科信息深度挖掘的意义和方法

新刊出版由于面市时间短,领域知名度和读者认可度还未全面建立起来,面临着前所未有的挑战,起步阶段通常难度较大。想要实现快速发展,必须在出版初期确立高起点的办刊方向,坚持高水平办刊道路,这是提高新刊学术质量、提升综合竞争力的有效途径<sup>[2]</sup>。一本好的期刊更需要全面掌握所属学科学术资源和行业信息,尤其是那些常识以外的以及常被忽略和遗漏的重要学术信息,对信息资源进行整合应用,按照出版特色进行内容的策划和执行,以优异的期刊作品面向学科读者。

文献计量学和文本挖掘技术可深度研究学科出版的潜在信息<sup>[3-4]</sup>,通过对动态的文献数据库的追踪、统计、计算、综合分析及可视化研究<sup>[5]</sup>,以定量的方法反演科技期刊出版所需的学科隐形情报和行业领域重要

资源,如科研热点、专家资源、研究机构、国家和地区学科优势、国际合作、专题资源等,以期刊大数据的研究理念将学科信息转换为出版实践,如定向约稿、选题策划、发掘学术领头人等,指导现代科学办刊。对于新刊来讲,学科信息挖掘研究更为适用和实用,特别是新刊面临的未知层面学术资源多,更有可待深度开发研究的意义和价值。

本研究主体思路是使用文献大数据技术分析获取出版潜在信息,形成各项指标所需的信息图谱并将重要文献情报通过专业分析生成办刊学术资源。其中,数据分析以2011—2015年海洋学为例,对中外引文数据库(Web of Science、中国学术期刊网络出版总库)和国家自然科学基金委资助项目检索库的相关论文题名、词频、引用、下载、作者、机构、地区、基金资助等信息进行分析,使用Navicat Premium和Microsoft Visual Foxpro工具对数据进行提取、处理和规范,建立模型并计算转换,再利用Citespace和词云图生成器进行可视化分析。

## 2 学科信息深度挖掘方向

从海量文献资源中分析数据辅助出版实践的研究雏形在《海洋通报》编辑部可追溯到2013年策划“渤海生态红线”“海平面与气候变化”“潮汐能资源”专题。当时,用数据支撑的选题使论文出版准确切入领域热点。这些稿件的被引和下载情况远远超过同行期刊。其后,编研人员不断完善学科信息挖掘的研究体系和方法,从主旨定位、热点发掘、作者遴选、机构查找、地区优势分析判断、科技前沿跟踪和专题策划7个维度(图1)进行深度挖掘,加强对出版信息的反演研究,强化对出版工作的主动谋略<sup>[6]</sup>,该刊影响因子近年来一跃成为海洋学期刊首位,具有较高的业界影响力。这种学科信息深度挖掘助力出版的模式在工作中为学术资源选择指点迷津提供了重要的出版策划指导。

**2.1 聚焦主旨定位** 新刊出版是学术生态格局的重新调整与读者布局的重新划分<sup>[7]</sup>。出版主旨是新刊的头等大事,在学科期刊中必须独树一帜、特色鲜明,找到最适合的专业方向定位和优势领域定位。

以《海洋信息》为例,原出版的期刊侧重于报道海

\* 中国科学技术期刊编辑学会基金项目面上基金(2017MSC4)

<sup>†</sup> 通信作者

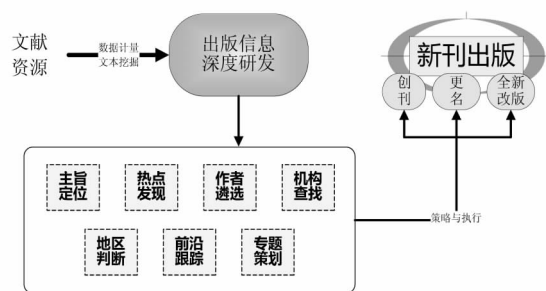


图1 学科信息深度挖掘助力新刊出版模式

洋管理与服务,外延大,缺乏专业显示度。出版单位采用学科数据信息挖掘研究,重新梳理并规划期刊出版工作,实行全面变革。据数据分析,海洋信息专业领域发展已由原先的“信息”概念发展成为“信息化”的概念。从该研究的“信息化”可视化关系图可知,海洋信息以数字海洋、物联网、云计算、大数据等作为支撑,应聚焦信息化技术进行专业出版;因此,新刊出版强调原创性成果报道,致力于快速推广新方法、新技术、新进展这一专业定位,同时将优势领域定位为学科国内前沿,这是我国“十三五”海洋信息化一盘棋工作的重要方向,具有较强学科前瞻性。根据读者反馈情况,新刊出版以来对学科领域科研人员、产业开发人员、涉海高校师生等受众群体具有较高的可读性,稿件下载量显著增长,仅《深度学习神经网络在海洋环境预测中的应用》一文在CNKI网络发布后10日内下载量达41次,这一数字对于海洋子学科小众期刊已属受关注程度高且迅速,说明主旨定位准确符合学科读者的学术需求。

**2.2 发现科研热点** 新刊出版要深入挖掘学科的科研热点,为学科研究提供切实、有价值的参考<sup>[8]</sup>。要明确学科热点就要全面客观掌握学科领域学术发展情况。大数据挖掘技术提供了词频跳变的特征、速率和相互关系,通过变化趋势可以分析学科热点和科研方向。从数据挖掘研究中生成的可视化关系的年轮节点内外圈的最早关注和最近年份关注情况,可帮助编辑获得学科知识点的演变情况,判断科研热点走势。由数据分析得知,平时办刊常识仅是一部分,更有意义之处在于科研热点在学科中的地位以及与其他知识点的共现关系,哪些处于核心层位置,哪些是从属,它们的远近关系如何。比如,海洋学科研热点可视化分析表明全球气候变化是核心热点,与其联系密切的有海水温度、生物多样性、海水酸化、海洋碳循环等次级层面的热点,各次级层面的热点还有与之相关的下一级热点,如北极、冰川、大堡礁等,这些为海洋学术期刊选题策划提供了重要的知识热点图谱,极大助益了科学指

导的高层次学术办刊。

新刊出版特别要紧抓科研热点,打造学术一流作品,抢占学科报道阵地。

**2.3 遴选优秀作者** 新刊稿件的作者科研水平直接关系到期刊的学术质量和读者对期刊的心理定位,而如何找到理想的学科作者是新刊常见的难题,毕竟办刊人的专家资源是有限的,能真正为新刊撰稿的知名专家也有限,且不一定能约到他们的稿件。大数据挖掘可帮助新刊寻找更多优秀作者,拓宽选择作者的范围渠道。

一方面,通过学科作者之间的文献合作与引用关系可发现学术流派情况<sup>[9]</sup>,形成作者关系图以帮助查找相应的专业人才。最理想的知名专家约稿未成,可分析寻找与其学术观点最密切的领域内优秀同行,与这些专家沟通,向他们定题约稿。

另一方面,大数据挖掘还可提供交叉学科和相近学科的领军人才、专家学者、潜在作者,这些是本学科办刊人所不熟知的,但却可以为期刊创作出更多的优秀稿件,比如关于红树林的研究在常识里是海洋滩涂生物学的课题,但研究发现也有很多研究成果出自从事植物生态学的科研人员之手,这2个关系并不密切的学科也可以在同一研究中紧密相关,这是数据深度挖掘研究的结果,能够另辟蹊径且科学有效地发现作者。

**2.4 查找科研机构** 科研机构是学术成果创造、孵化和产生的重要土壤,在一定程度上代表着相应的学科研究能力。以数据挖掘研究的科研机构信息可为新刊提供重要的学术资源。

一方面,挖掘科研机构信息可获取更多的一手行业情报,尤其是新刊更需要目标明确具体,精确锁定科研成果产生机构,这项行业舆情给新刊提供了可开辟的重要空间。当期刊出版单位拟定选题方向后,数据挖掘研究即可为该选题方向提供科研机构的相关信息,体现机构与选题方向的相关度,可从中提取相关科研机构列表和排序情况,分析查找优质稿源。编辑可通过这些信息进一步与科研机构沟通交流,了解机构内具体学术情况,约请稿件。

另一方面,对科研机构的信息挖掘研究还可以反映各机构间的科研亲疏关系及合作程度,追踪组织稿件,策划引导学科发展的出版内容。如世界海洋科研机构NOAA、伍兹霍尔海洋研究所、澳洲CSIRO之间的研究相关性强,可以重点跟踪这些科研机构以及与之合作密切的科研部门,追踪国际学科研究成果,策划特色稿件出版在我们本土出版的科技期刊上。

**2.5 判断地区优势** 一个学科领域的科学研究在不

同地区(区域)有研究层次和研究水平的差异;同一地区的同一学科研究也有子学科科研水平的差异。利用数据挖掘和文献情报学模型<sup>[10]</sup>可计算分析出各地区(区域)学科发展的舆情,辅助办刊人打破思维常规,重新认识和发现学科的地区发展信息情况,掌握国内外行业领域学科优势分布,为学科期刊选题方向和地区稿件的重点优选提供重要信息参考。比如在学科数据挖掘分析中发现,虽然我国海洋学研究总体水平在国际上不处于优势地位,但数据分析显示,我国 engineering marine 科研方向具有很强的区域相对优势,根据数据进行进一步分析,可以具体聚焦到是中国台湾地区顶级论文产出率高,平均科研水平已超过美国等发达国家,成为海洋子学科的地区相对优势。这为海洋科技期刊提供了极为有利的判断依据,重点关注该地区该子学科领域的发展,查找科研成果产出的来源,搜寻地区内的研究机构、作者,以本文 2.3 和 2.4 的方法组织稿件。地区优势判断为期刊主动策划的特色学科出版提供了重要且有价值的学术资源情报。

**2.6 追踪创新前沿** 新刊出版要对学科研究的报道标新立异,做出区别于同行的前沿报道和鼓励科技创新的有关内容。而论文出版往往滞后于科学成果的产生,常规的引文分析可能不能给予及时的数据分析。通过大 2 种方法对数据深度挖掘分析,对前沿创新做出预判。

第 1 种方法是利用基金项目检索库这一重要的数据库资源,因其科学选题有立项评审依据且多数处于科研成果产生的前期,较文献成果发表出来的情况一般会缩短时间因素的影响。利用数据挖掘对基金项目研究方向进行文本分析,将研究内容进行定量研究,掌握时下学科高端科研人员所关注的研究方向和研究类型,据此信息进一步学习,预估项目即将产生的成果而重点关注组织稿件。

第 2 种方法是利用词频变化分析文献数据库,重点关注词频的增长趋势,初期是细微的由小到大的量变,直至出现反曲点(拐点)可能预示着学科即将产生新的研究课题,比如 2015 年初词频分析开始出现“智慧海洋”,随后数量持续增长,至 2016 年下半年在可视化分析中出现拐点,这时《海洋信息》发现这一现象,在当时正在全面改刊的方案中纳入这项前沿技术并鼓励创新。至今 1 年时间里,沿海各地已纷纷开展智慧海洋重大工程项目,新版《海洋信息》率先拥有了这项前沿创新的学术资源。跟踪学科隐形信息为新刊出版提供了掌握创新前沿的可能。

**2.7 策划专题报道** 学科信息深度挖掘可通过对科研专题的文献分析及相关性分析拓宽选题思路,充实

学术资源,弥补办刊人知识贮备的不足,有效组织开发稿源,为新刊出版提供专题策划的特色学科信息。如在对 2018 年 1 月巴拿马籍桑吉号油轮在长江口以东国际航道起火燃烧的海上搜救事件的特别专题策划中,出版单位利用信息挖掘方法将“海上搜救”的文献数据进行文本挖掘和共现分析,结果不仅有专业搜救、水上交通、辅助决策系统、船舶污染等常规可联想到的关键词,还通过可视化关系获得了海面风、扇形搜索、海浪、海流、溢油应急、实况分析等文献节点。这为期刊特色专题设计提供了可拓展的极大空间,不仅可涵盖传统搜救研究,更可将学科领域内微观层面的专业研究点纳入其中。期刊专题策划紧紧抓住了这些学术层面的选题点,以风、浪、流的预测为切入点加强专业策划,选取海洋动力学模型计算研究搜救船只坐标和行动轨迹的稿件进行报道,分析角度新颖、针对性强,特别能满足学科读者的求知心理,出版策划取得了较好效果。通过期刊微信公众号的优先发布,其中 1 篇稿件的作者研究团队还受到了权威电视媒体的采访。这充分说明了以学科文献数据进行专题资源挖掘的应用性效果。

### 3 结束语

学科信息深度挖掘研究可为期刊提供创办、更名,全新改刊出版初期的专业技术支持,辅助搜寻行业信息和学术资源,提供开展特色出版的科学指导。同时,该研究也可为新刊成长及其可持续发展提供较大助益:1)有助于充分占有学科期刊市场地位。学科信息挖掘研究指导的出版实用性强,立足专业领域发展,满足学科市场阅读需求,不仅填补学科领域期刊市场空白,更助推学科建设,引领读者市场关注方向,加快对学科期刊市场的占有速度。2)有助于科学构筑专业学术交流平台。学科信息挖掘研究促进对学术资源的认知,使专家、作者、科研机构这些科研主体不再各自孤立,而连结成相互间的网络关系和动态关系,加快资源共享与共建,提供产生学术共鸣、讨论的机会,为学科交流提供更大空间,从中可源源不断地挖掘学科领域新课题方向。3)有助于积极发挥期刊载体中介作用。期刊是实践基础与上层建筑之间的纽带<sup>[11-12]</sup>,对成果的推介、普及与转化发挥重要的科技生产力价值。学科信息挖掘分析国家政策-科研-产业之间的数据与文本关系,传导科技政策对学科研究的方向指引,塑造作品传达学术观点,推送利好科技项目,指引行业领域科研成果的落地转化。

科学在发展,期刊在发展,以科学技术支撑的新刊出版需要在新形势下与时俱进,充分重视文献情报学

的研究价值,探索学科信息深度挖掘研究对期刊出版的科学指导作用,以大数据研究分析抢先掌握学科科研热点,发现科研创新前沿,查找专家作者和研究机构,优选出版主题,判断区域优势,策划特色专题作品,以高起点、高水平的稿件作品面向读者受众,努力构筑学术界的交流平台,产业界的桥梁纽带,不断占有学科期刊市场的优势地位,将新刊打造成为能够面向未来发展、更具行业竞争力的学科优秀出版物。

感谢中国知网中国科学文献计量评价研究中心肖宏、伍军红老师给予的学术指导及赵一方博士对本研究提供的技术支持。

#### 4 参考文献

- [1] 邓的荣. 传统期刊产业结构性调整与发展趋势[J]. 中国报业, 2013(17): 58
- [2] 秦洁雯, 段艳文. 从2015年度期刊创办和期刊更名看我国期刊发展趋势[J]. 科技与出版, 2016(3): 24
- [3] 刘俊, 张昕. 大数据视域下的科技期刊数据库建设[J]. 编辑学报, 2014, 26(1): 59
- [4] 党亚茹, 彭丽娜. 基于CSSCI的区域社会科学学科优势分析[J]. 情报杂志, 2012, 31(4): 22
- [5] 李雪. 期刊大数据与学科发展测度研究: 以海洋科学期刊研究为例[J]. 科技与出版, 2017(1): 98
- [6] 胡德华, 种乐熹, 邱均平, 等. 国内外知识检索研究的进展与趋势[J]. 图书情报知识, 2015(3): 93
- [7] 方熠. 新办期刊的四个定位[J]. 出版科学, 2007(6): 42
- [8] 陈楚湘, 郭晓峰, 陈冰, 等. 基于CitsSpaceIII的系统工程学科知识图谱初探[J]. 时代教育, 2015(3): 198
- [9] 宿瑞芳. 基于引用认同的学科作者研究: 以语言学为例[D]. 南京:南京农业大学, 2011: 7
- [10] 岳洪江. 各地社会科学学科优势及结构研究[J]. 图书情报工作, 2008(1): 66
- [11] 闫蓓, 严谨, 肖宏. 搭建科学与大众的桥梁: 谈科技期刊与大众媒体的新闻报道合作实践[J]. 编辑学报, 2009, 21(4): 325
- [12] 尤小龙. 期刊要成为实践基础和上层建筑的桥梁纽带[J]. 中国农民合作社, 2014(6): 29

(2018-03-13 收稿;2018-04-24 修回)

### 编辑颂

霍振响/西北农林科技大学学报(自然科学版)编辑部,712100,陕西杨凌

十年前,像一滴水汇入海洋,  
我进入了这个向往已久的行业,  
带着飞蛾扑火般的倔强。

初时的仓皇、迷茫、无助、彷徨,  
在深夜柔和的灯光里,  
在白昼匆匆的脚步中,  
渐渐变成了淡定、从容、笃信、自强。

这里有坚持奋战在一线的耄耋老人,  
这里有肩负重任的中坚力量;  
为了钟爱的职业,为了期刊的强国梦,

他们邀约过无数次夜半的钟响。

十年后,也许中国的期刊已经屹立世界,  
也许还蛰伏在春天,  
但我知道:  
有这样一支怀揣远大梦想,脚踏实地奋斗的队伍,  
梦想终究会启航!

所以,我要为亲爱的同行们唱首颂歌!  
为理想,为中国,  
我们一直在路上!

后记:拜读了游苏宁、陈浩元、冷怀明的《砥砺前行 实现科技期刊强国梦》(刊登于《编辑学报》2018年第4期)一文后,深受触动,写就了上述文字。

时下,在科研人员追逐SCI、影响因子、核心期刊的热潮中,越来越多的人开始冷静思考其弊端。有的从办刊角度,讨论中国期刊的发展历史、国际影响力及在新形势下应有的发展自信;有的则从科研角度,指出了唯国外期刊的评价机制对我国科研人员及科研经费指向的深远且不利的影响。这些直达问题本质、击中要害的论述,让我等小辈读来,深有“直抒胸臆”的酣畅淋漓感!中国科技期刊将走向何方?如何建成科技期刊强国?这是值得很多领域共同思考的问题,更是广大编辑孜孜不倦的探索方向!