

中医药类英文科技期刊国际同行评议专家数据库建设*

鲁 珊 张 韞 玉 邓 杰 邹 龙 彭 清 华

湖南中医药大学《数字中医药(英文)》编辑部,410208,长沙

摘 要 学术期刊同行评议过程是保证学术论文质量的重要环节。目前的同行评审公共平台所包含的专家学者均以全球性学科为主。而中医药是我国独具特色的卫生资源,也是我国少数几个拥有自主知识产权且在国际上具有竞争优势的行业之一,因此,建立一个全球范围的、专门的、包含大量针对传统医药相关专家的同行评议数据库就显得尤为重要。本文建立了一个信息量大、更新及时、交互式管理的中英双语国际中医药同行评议专家数据库系统(TCMPEER),并以 Web Service 方式设计公共开放接口,提供实时在线专家遴选服务。该数据库能实时掌握更新国内外医学界专家学者动态、研究近况,为评审专家遴选工作提供信息服务和决策支持。该数据库重点解决以专家数据信息为中心,从展示、关联、交互、服务等多维度辐射扩展,为专家带来完全精准匹配的增值服务。本研究构建的专家数据库着重服务于中医药类英文期刊,形成优势公共资源,未来将与行业学会对接,首批实现以我国创办的中医药类英文期刊集群共同进入实质性业务运营,进而辐射世界,补充替代医学领域相关专家数据库。

关键词 中医药英文科技期刊;同行评议专家库;大数据存储系统;专家本体系统;专家遴选;中医药英文期刊群运营服务

Establishment of international peer-review expert database for English-language journals on traditional Chinese medicine//

LU Shan, ZHANG Yunyu, DENG Jie, ZOU Long, PENG Qinghua

Abstract Peer review is a key link to ensure the quality of academic articles. The current international public academic platforms include experts and scholars focusing on global disciplines. Traditional Chinese Medicine (TCM) is a unique health resources in China, and also is one of few industries in our country with independent intellectual property rights and competitive advantage worldwide. It is particularly important to establish a global, specialized peer review database containing a large number of experts related to TCM. In this project, a Chinese - English bilingual international peer-review expert database system which contains a large volume of information, receives updates regularly, and is managed interactively for the TCM field, named TCMPEER, was developed, and a public open interface was designed as a web service to provide real-time online expert selection services. This database can obtain and update information on experts and scholars in the medical field worldwide and their recent research developments in real time, providing information services and decision-making support to facilitate the selection of

reviewers. This project was focused on collecting data and information about experts, the developed system radiates and expands from multiple perspectives (e. g. display, correlation, interaction, and service), providing completely and accurately matched value-added services for experts. The expert database established in this study is focused on providing services to English journals in the TCM field and creating valuable public resources. Based on this expert database, knowledge service platforms for relevant fields can be further developed. In the future, it will connect with industry associations first to realize Chinese medicine English journals cluster to enter substantive business operation, so as to serve relevant journals and experts in the field of complementary and alternative medicine worldwide.

Keywords English science and technology journals of TCM; peer-review expert database; big data storage system; ontological expert system; expert selection; English journals group of TCM operation service

Authors' address Periodical Press, Hunan University of Chinese Medicine, 410208, Changsha, China

DOI:10.16811/j.cnki.1001-4314.2021.05.013

同行评议,是科学界对项目进行评审和对科研成果进行评估的一种基本方法,亦是科技期刊学术出版中的一个重要环节,是对学术论文科学价值和学术质量等进行较为准确评价与控制的机制^[1]。同行评议的本质和前提是需要有足够数量与质量的专家。尤其作为国际公开发行的英文学术期刊,单纯一本期编委会的外籍专家远远难以适应同行评议“随机盲法”的需求,且随着激增的学术研究成果的发表,目前公用的同行评议模式已面临专家数量饱和,评审效率和学术评审质量下降的局面^[2]。对比国外专业领域期刊、出版集群、运营商,数据资源是具有竞争优势的重要组成部分^[3],因此在新兴交叉学科,尤其基于本土具有竞争优势的学科当中先行进行数据积累是具有前瞻性的开创性工作^[4]。中医药是我国独具特色的卫生资源,也是我国少数几个拥有自主知识产权且在国际上具有竞争优势的行业。中医是基于人体复杂系统的整体科学,而“中医药科技化”则是中医药走向信息化、标准化、客观化、微观化的道路。目前的同行评议公共平台所包含的专家学者均以全球性学科为主;对于中医药学这一非全球性学科,随着中医药现代化、国际化

* 中国科技期刊卓越行动计划选育高水平办刊人才子项目——青年人才支持项目(2020ZZ111017)

发展趋势与需求,建立一个全球范围的、专门的、包含大量针对传统医药相关专家的同行评议专家数据库就显得尤为重要。

目前,具有中医药数字化研究背景和能力的学者比较缺乏,精通英语的中医专家则更少,而中医药当中的部分学科分支在国际上已得到较多国家不同程度的认可;因此,中医药相关行业的研究者和传播者都应抓住这个契机,注重中医药的科技化发展和国际化传播。本研究设想在《数字中医药(英文)》(《Digital Chinese Medicine》)杂志建设的基础上,通过构建基于本体系统的专家数据库,为中医药期刊(尤其英文期刊)精准发现、筛选、联系同行评议专家提供服务,从而全面提升同行评议速度和质量。

1 构建中医药同行评议专家数据库的意义

中医药行业在实现现代化、国际化进程中,急需借助国际通用的大数据分析与机器学习技术,在国际上集成展示祖国文化,以及推送中医药行业最新研究成果。在此过程中,同行评议环节对学术发现和学术质量评价均具有重要意义:1)通过本研究构建中医药英文科技期刊同行评议专家数据库(TCMPEER)的过程,将汇聚全球范围内中医药现代化研究相关的专家学者,且能够对所发表的研究成果进行客观的、高质量的学术评价。2)对期刊编辑及运营人才的培养。在项目的组织、策划、实施过程中,期刊编辑和办刊人通过利用先进技术手段搭建平台,提高出版服务国际化眼界,提高组织调动与整合相关资源的能力,进而培养国际一流科技期刊编辑和办刊人才。3)基于本项目构建的专家数据库,可实现与相关领域学会等学术交流平台的相互引流,进而拓展基于整合国际优势资源相关领域知识的服务平台。

2 TCMPEER 的设计思路

本研究构建的 TCMPEER 系统,通过多途径获取专家信息数据源,并进行结构化清洗,形成规范数据格式入库,总体功能框架分为专家本体系统、专家注册系统和审稿专家遴选系统。

专家资源库将作为科技期刊运营“基础设施”进行持续的更新、维护。一方面,建立数据更新、入库的机制;另一方面通过在资源库的架构上,同多方系统集成,拓展应用场景,通过系统能力记录学者行为信息,对学者进行多维度的、动态的肖像刻画。

项目重点解决多数据源自动获取,并基于自动分类聚类技术,构建基于本体系统的专家图谱,实现在线交互、精准推送,以及第三方系统集成,实现数据库系

统的高扩展及高延伸。系统更着重将数据信息公开开放地提供给其他第三方系统与数据库,以 Web Service 方式设计公共开放接口,提供实时在线专家遴选服务。

3 TCMPEER 建设的实施

3.1 专家本体系统建设

专家本体系统本质上是一个基于信息工程系统的任务信息存储与管理库。专家信息获取、信息表示、信息存储和信息运用是此系统的核心内容。专家信息的结构分为基本信息、履历信息、论文信息及关联信息等。其中,基本信息包括姓名、性别、工作单位、年龄、联系方式、通信地址、研究领域等;履历信息主要包括专家的就职及兼职信息;论文信息则指专家以往发表的文献信息;关联信息则是系统基于知识图谱自动构建的与专家相关的信息(图1)。

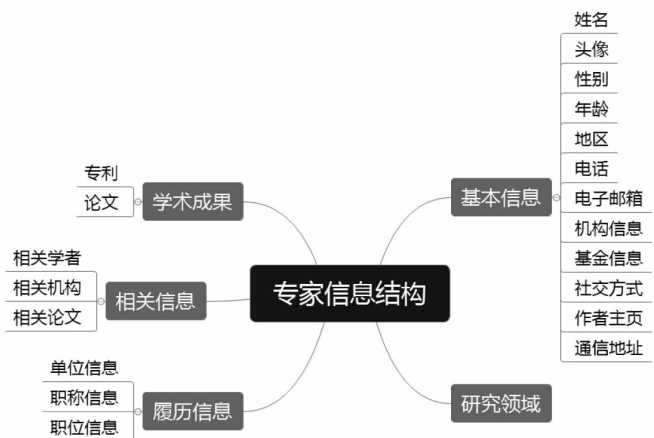


图1 专家信息结构

所构建的国际同行评议专家数据库按学科分支、研究方向等信息标签精细分类并交叉整合,以便稿件精准送达相应专家。专家系统以信息库中存储的信息进行推理来获得问题的解决办法,其主要通过专家信息进行获取和加工处理实现。

3.1.1 专家数据获取 专家数据获取主要是基于文献进行专家信息提取及清洗。专家资源沉淀的具体来源和信息获取的途径如下:

1)基于 DCM 杂志对文献的 XML 全文数据解析。将论文中的学者信息进行解析入库,形成学者的基础信息,包括学者姓名、研究领域、工作单位、电子邮箱、简介等^[5]。同时,支持对论文中解析的作者进行信息人工消歧等处理,保障作者信息的唯一性。

2)基于网络精准抓取获得领域拓展专家。分别对中医药研究领域高影响力文章和刊物进行学者相关信息分析和提取,进而基于大数据图谱采集机构信息进行关联分析。

3)对接相关行业学会,获取会员信息,并进行如上信息解析、入库及关联信息挖掘和信息网络构建。在获取专家数据时,我们着重基于行业细分关键词的组合和拆分对从事中医药现代化研究的学者进行精准匹配和发现。

3.1.2 数据加工处理 数据收集后,对各途径获取的数据进行加工处理,包括数据的消重(根据唯一的Email或手机号信息)与融合,数据格式规范整理、去重等。基于对专家基本信息进行规则制定,对论文入库解析的专家信息进行系统自动消歧。同时,针对自动消歧的数据结果进行人工识别去重,确保专家信息的唯一性和准确性,确保网站信息学者的权威性。对已获取的专家数据还要经过行业细分以达到精准入库。因为部分学者是文献来源的参与者,可能有大部分是从事行业研究的阶段性学者,所以要不断对数据进行过滤、清洗,以达到对专家学者进行动态管理和调整的目的。

3.1.3 专家图谱呈现 根据专家个人信息,以及文章的被引信息,分别形成2个维度的专家关系图谱。每篇文章会自动匹配关联系统内其他专家的文章,并关联显示文章及文章作者(其他专家)。专家文章可设置私人可见或公开可见,公开的文章可以平台内实名讨论。另外,基于对专家的基本信息、履历信息、论文信息等内容的关联处理,利用知识图谱能力,自动构建专家与机构、论文,专家之间的关联关系,例如可以自动构建合著关系、同行专家等。

3.2 专家数据库应用

本文构建的TCMPEER主要包含专家注册、专家精准检索,可支持未来可拓展功能的应用,如专家个人空间建设、学术交流、内容服务等全过程。

3.2.1 专家注册 专家库实名邀请注册。对专家本体系统形成的专家库成员进行自动注册,并以Email形式自动分发账号信息,邀请专家实名认证登记,并补充维护个人信息。专家图谱作为中医药专家总和,可对接用户管理系统,基于应用场景对专家进行认证、激活。专家本体库作为中医药专家总库,除可进行编辑部内部创建及第三方数据库对接以外,同时支持领域内专家自主注册方式,由此通过运营策略吸引更多的专家进入到专家库中,实现专家信息的持续更新维护。

3.2.2 专家检索 专家检索作为最基础的应用场景,可为相关领域的刊社、科研单位提供专家检索服务,可支持按照领域、单位、基金、科研项目等属性进行导航检索,同时也支持按照专家基本属性进行精确查询。

3.2.3 专家个人空间建设 基于专家的认证和运营,可以为专家开辟个人空间,打造专家“学术博客”,展

示专家学术信息。在运营过程中,系统通过专家的行为、活跃度等,延展用户画像的维度,支撑丰富的应用场景和运营策略。

3.2.4 学术交流

1)管理员同专家的交流。基于专家本体库可以构建留言板功能,为专家与管理员之间搭建一个在线沟通交流的平台,方便专家与管理员之间进行沟通。专家可以基于微信、网站等渠道进行网站留言,管理员则可以通过管理后台进行留言的答复。

2)专家之间的学术交流。基于专家库可以搭建专家之间的学术交流平台,专家之间可以互相关注并添加好友,方便专家之间的学术交流,互相关注对方的最新学术成果。

3.2.5 内容服务 对专家进行内容推送。针对推送的公共信息,如相关会议资讯、新闻资讯、最新同行文章等,系统根据专家信息自动分类匹配关联,为专家精准推送。针对专家个人订阅信息,如文章被引信息推送,系统自动监测专家的文章的被引情况,当有新被引时,自动通过Email和微信公众号推送到个人。通过在线学术交流,可以积累和整理出科技工作者共享的学术资源,提高科学交流的范围、质量和效果。

3.3 审稿专家遴选系统

目前,现有的专家遴选系统多是采编系统的一个模块,对编辑而言实用性不强,数据对比度不高且可视化效果不好^[6]。本项目利用数据库平台技术优势,打造了集智能专家遴选、专家审稿反馈、信息统计于一体的审稿专家遴选系统。

3.3.1 智能专家遴选 专家遴选系统能够支持智能专家遴选、专家选择、信息统计等功能。智能专家遴选模块支持综合检索和高级检索模式。检索的范围是基于前期建立的专家本体数据库。综合检索则根据输入的关键词从数据库中所有字段进行模糊查询。高级检索设置“and”“or”“not”等检索条件,便于对库内专家进行精准检索。高级检索的结果更加准确,且效率更高^[7],可以在库内按照姓名、学科领域、研究方向、工作单位等各种基本信息进行检索。

3.3.2 专家审稿反馈 我们可以对选出的专家发送审稿邀请,专家一方面可以选择是否接受审稿邀请,专家评审可以就作者研究问题的理论或方法的创新性、语言表达的流畅性、研究课题的价值等进行综合性评判^[8];另一方面可以对文章进行自由评价,重点评价文章的学术创新点和存在的不足,以及评审人对该文章的具体处理意见,如accept(录用)、minor revision(小修)、revise(修改)、major revision(大修)和reject(拒稿)。

3.3.3 信息统计 该模块由专家拒审率、审稿时长、审稿质量、审稿结论4部分组成。拒审率统计主要分析不同职称、年龄审稿专家的拒审率及拒审原因。审稿时长主要统计专家数据库内成员稿件平均审稿时长,以及统计专家年龄、职称等因素与审稿时长的关系,以便对影响专家审稿时长的因素进行探究。审稿质量统计主要研究专家职称、研究方向、学科背景等与审稿质量的关系,为组建一支优质的审稿队伍提供参考建议。审稿结论统计主要分析不同职称审稿专家的退稿率,以及不同审稿结论分布比例,不同年龄段审稿专家的各审稿结论所占比例等^[3]。

4 专家数据库建设成果及运营策略

本研究建立信息采集和动态更新机制,形成信息采集的持续机制,基于信息源定义信息采集频率和策略,实现信息的持续更新和动态维护,具备持续运营能力。系统通过严格抓取中医药相关行业高校、研究院所及医院机构网站、开放域等多种数据库资源,完成中医药领域专家数据的获取。利用Email、手机号、规范数据格式等,对专家进行清洗关联,目前共采集入库了1.1万名中医药领域专家的信息,并将持续更新维护。通过系统的“检索”及“高级检索”功能,用户可深入了解专家所在单位、个人简介、发表文章、承担项目等信息,可实现中医药领域审稿专家的遴选工作。系统通过专家个人信息、发表文章数据分析和工作单位线索发现等多维度展示专家图谱,为每一位专家开通配有独立URL的个人空间,可用于专家对外宣传,也便于查询者全方位了解专家信息。专家信息的展示包括专家姓名、头像、所在机构、职务、发文期刊、论文关键词、共事学者、共同作者等。其中,学术成果支持年份、类别过滤、时间与引证数排序等多种组织方式,提供更为友好的交互界面。文章的被引数据根据系统采集数据计算,用户点击文章题目,可跳转至该文DOI解析地址,做进一步阅读。

然而,本数据库事实是一个较大的项目和目标,此事业的完成和运营将会对中医药科技期刊在学术内容质量提升上有重大作用。通过全球范围内相关行业专家的招募和汇聚,也将对中医药在国际上的知名度和影响力有一定的提升作用。本研究已完成专家数据库建设,并已完成专家注册、遴选工作测试,以及端口外接、相互引流测试,随时可以开通预设功能进入运营阶段。然而在初始运营上,本数据库还需谨慎开展。由于目前仅由本刊单刊实施运营,力量单薄,达不到预期效果,很少有协作或合作的关系。并且,相关平台较少

且功能单一,缺乏对同行评议的多样性、同行评议协作、增强型同行评议的支撑^[9]。这是资源的浪费,违背该项目计划实施的初衷。招募的专家一旦进入使用,平台需要即刻活跃起来,用户体验好才是留住专家的最好途径,因此,未来我们将该平台对接中华中医药学会中医药英文科技期刊集群,实现实质性业务运营。

5 结束语

专家资源是科技期刊运营的核心资源。专家数据库建设一方面重视对现有资源的维护,另一方面基于期刊业务过程中,文献线索、引文线索等逻辑,借助大数据信息采集技术,逐步增加、完善专家信息。本研究从服务学科的角度,先行具备数据、平台积累,形成数据优势公共资源,对中医药行业发展及加速中医科技化、国际化均具有重要意义,且具有一定前瞻性。未来基于实体专家数据库可拓展搭建相关领域知识服务平台,形成基于优势公共资源中心,且通过该项目的完成和使用,培养期刊编辑利用大数据、互联网等新技术和新工具,以及对信息采集、信息加工、信息传播和信息评价的整体思路和能力。

6 参考文献

- [1] 胡克兴,刘徽,卢珊,等. 开放科学环境中的科技期刊同行评议研究[J]. 编辑学报, 2019, 31(6): 610
- [2] NASSI C L. In time: publons seeks to attract reviewers and improve peer review [J]. Rev Paul Pediatr, 2017, 35(4): 367
- [3] CITROME L. Peer review and Publons: enhancements for the reviewer[J]. Int J Clin Pract, 2016, 70(5): 364
- [4] MONDAL H, MONDAL S. Pattern of peer review in dermatology domain: an analysis of reviewers' profile from publons.com[J]. Indian J Dermatol, 2021, 66(2): 225
- [5] 谢文亮,郑添尹. 大数据时代学术期刊信息数据库的建设[J]. 现代信息科技, 2020, 4(14): 126
- [6] 于红艳. 基于 axure RP 的智能审稿专家遴选系统原型设计[J]. 中国科技期刊研究, 2019, 30(9): 976
- [7] 闫蕾,周珑,赵义强. CNKI 检索模式结合关键词选取在检索中的应用探讨[J]. 技术与创新管理, 2020, 41(5): 463
- [8] 于化东. 英文数学期刊国际审稿人专家库创建探索: 以《数学研究及应用》为例[J]. 编辑学报, 2021, 33(1): 89
- [9] 常唯,袁境泽. 国际学术出版中的同行评议进展与展望[J]. 中国科技期刊研究, 2020, 31(10): 1181

(2021-06-21收稿;2021-08-26修回)