

# 近三十年域外汉籍研究的现状与展望——基于文献计量分析和知识图谱绘制

唐榕骏<sup>1</sup>, 彭志峰<sup>\*1,2,3</sup>

<sup>1</sup> 暨南大学文学院 <sup>2</sup> 暨南大学汉语方言研究中心

<sup>3</sup> 广东省岭南数字人文实验教学示范中心

{moniyi}@yeah.net, {pengzhifeng}@jnu.edu.cn

## 摘要

域外汉籍研究对中华文化传播意义重大、成果丰富，但缺少文献计量分析。本研究的目的是通过CNKI优质期刊论文，分析中国近30年域外汉籍研究的演进趋势。通过知识图谱、普赖斯理论、Leydesdorff引文理论、斯皮尔曼等级相关系数与关键词共现、聚类、突现、时线分析研究学科演进趋势。研究发现，“汉籍整理与数字化研究”“文化交流与文化传播”为前沿热点，揭示出领域内研究工作已从回归普查转向应用转化。本研究为学科发展提供了参考数据。

**关键词:** 域外汉籍；知识图谱；文献计量；共现分析

## Extraterritorial Chinese Texts in the Last Thirty Years: Research Advances and Future Perspectives Based on Knowledge Mapping and Bibliometric Analysis

Tang rongjun<sup>1</sup>, Peng zhifeng<sup>\*1,2,3</sup>

<sup>1</sup> College of Arts, Jinan University <sup>2</sup> Institute of Chinese Dialects, Jinan University

<sup>3</sup> Provincial Demonstration Center for Lingnan Experimental Digital Humanities Education

{moniyi}@yeah.net, {pengzhifeng}@jnu.edu.cn

## Abstract

The study of Extraterritorial Chinese Books (ECBs) holds great significance for the preservation and dissemination of Chinese culture, however, there is a dearth of bibliometric analyses in this field. This study aimed to investigate the evolutionary trends of ECBs research in China over the past 30 years. Base on CNKI database, the research advances of ECBs were analyzed using knowledge graph visualization, Price's law, Leydesdorff's citation theory, and Spearman's rank correlation coefficient, along with keyword co-occurrence analysis, clustering, and timeline analysis. The hotspots identified include "ECBs arrangement and digitization" and "cultural exchange and dissemination". It suggests that ECBs research has evolved from a focus on cataloging and inventory-taking towards the application. This study informs the future development of the discipline.

**Keywords:** Extraterritorial Chinese books, knowledge map, bibliometrics, co-occurrence analysis

\*通讯作者: 彭志峰 (pengzhifeng@jnu.edu.cn)

## 1 引言

域外汉籍指域外所藏中国古籍、域外刻本及域外人士撰写的汉文典籍(金程宇, 2010)。历史上,我国大量珍稀古籍流入海外。新中国以来,我国政府部门与学术机构高度重视域外汉籍的回归和研究工作。值得关注的是,相关政策重点发生显著的历时迁移:“十一五”时期《国务院办公厅关于进一步加强古籍保护工作的意见》提出“大量珍贵古籍流失海外,加强古籍保护刻不容缓”,“十三五”时期《“十三五”时期全国古籍保护工作规划》以“建立海外中华古籍回归工作机制,有计划、分步骤开展海外中华古籍回归工作”为目标,2020年《国家新闻出版署关于编制“十四五”国家重点图书、音像、电子出版物出版规划和2021—2030年国家古籍规划的通知》倡导“对收藏海外中国古籍进行系统性整理,加强古籍数字化开发利用”,2022年《关于推进新时代古籍工作的意见》鼓励“充分利用海外文化平台开展古籍对外宣传推广活动”。

中国域外汉籍研究始于1986年“中国域外汉籍国际学术会议”(张伯伟, 2006),已有三十余年历史,成果丰富。但目前该领域的研究综述主要依托短时段样本开展定性研究(奎欣怡, 2023),缺乏基于历时数据的量化客观分析,如傅荣贤、华建铭和李琳分别综述了1949年至2019年中国古典文献学(傅荣贤 and 华建铭, 2020)和古代文学(李琳, 2014)的发展的演变历程,其中片段式地简要概述了域外汉籍研究。金程宇、刘泰廷分别综述了2000-2009年(金程宇, 2010)、2010-2014年(刘泰廷, 2017)阶段的域外汉籍学术研究成果,却未对域外汉籍研究内部的发展沿革进行量化梳理和剖析。因此,该领域目前面临着政策转向显示出的社会需求升级,与学术研究全景的未经量化、不可捉摸现状之间的信息不对等。这不利于学者把握研究方向和规划发展路径。因此,本研究目的是系统分析中国知网(CNKI)近三十年域外汉籍研究优质期刊(含辑刊)论文数据,总结归纳出发展沿革与演进趋势。为此,拟通过定量与定性结合的方式,基于知识图谱绘制和文献计量综合分析筛选的809篇文献。这将为本领域的未来研究方向提供依据。

## 2 研究方法

### 2.1 发文信息分析理论与算法

#### 2.1.1 普赖斯曲线

普赖斯曲线认为科学文献的增长与时间成指数函数关系。其公式为:  $F(t) = a \times e^{bt}$ ,  $F(t)$ 表示预测发文章量,  $a$ 为初始年发表量,  $e$ 为常数,  $b$ 为时间常数,表示期刊持续增长率。 $b$ 值的计算公式为  $b = \sqrt[n]{\frac{A}{a}} - 1$ ,其中 $n$ 为统计年限,  $A$ 为发文总量。本研究用于拟合实际发文情况,分析发文时序发展健康程度。

#### 2.1.2 平均值、方差、标准差

发文章量机构年发文章量的平均值、方差、标准差,可以用于评估机构发文章量的变异程度。其公式依次为:

$$\text{Mean} = \frac{\sum \text{number of publications}}{\text{number of institution}}$$

$$\text{Variance} = \frac{\sum (\text{number of publications} - \text{Mean})^2}{\text{number of institution}}$$

$$\text{Standard Deviation} = \sqrt{\text{Variance}}$$

#### 2.1.3 斯皮尔曼等级相关系数

斯皮尔曼等级相关系数是用于衡量两个变量相关性的非参数统计方法(白雪涵, 2022),本研究使用其验证刊文量前20期刊与CNKI影响因子高低的相关度。其中,  $\rho$ 表示斯皮尔曼相关系数,  $d_i$ 表示两个变量在排序中的差异,  $n$ 表示样本数量。具体公式为:  $\rho = 1 - 6 \sum d_i^2 / n(n^2 - 1)$ 。

#### 2.1.4 普赖斯核心作者定律与综合指数公式

普赖斯定律公式用于计算核心作者范围。公式为  $m = 0.749 * (n_{\max})^{1/2}$ ,其中 $n_{\max}$ 为发文最多作者的总发文章量(丁学东, 1993)。确认范围后,进一步通过综合指数法测定核心作者,其中 $Z_i$ 是第 $i$ 位候选人综合指数值,  $X_i$ ,  $Y_i$ 分别是第 $i$ 位候选人的发文章量、被引量,  $X$ ,  $Y$ 分别是候选人平均发文章量、平均被引量。公式为  $Z_i = X_i/X * 100 * 0.5 + Y_i/Y * 100 * 0.5$ 。

## 2.2 研究热点前沿分析理论与算法

### 2.2.1 Leydesdorff长期、短期引用理论

引文分析依据Leydesdorff提出的长期引用、短期引用理论(Loet et al., 2016), 其认为文献根据发表到成为高被引文献的时间长短, 可以分为长期、短期引用文献。短期引用说明被引文献处于研究前沿; 长期引用说明被引文献对于学科基础理论的形成作出贡献。

### 2.2.2 高频词汇边界算法

高频关键词根据高频词边界公式为 $T = (-1 + \sqrt{1 + 8I}) \div 2$ 确定(Donohue, 2014)。T为高频关键词个数, I为频次为1的关键词个数。

### 2.2.3 共现分析

共现分析的基本原理是: 两两统计一组词在文献中的出现频次, 通过共现次数测度其亲疏关系。共现次数多说明这组词关系更密切, 语义上更易产生联系(李杰and 陈超美, 2016)。分析过程如图1所示。

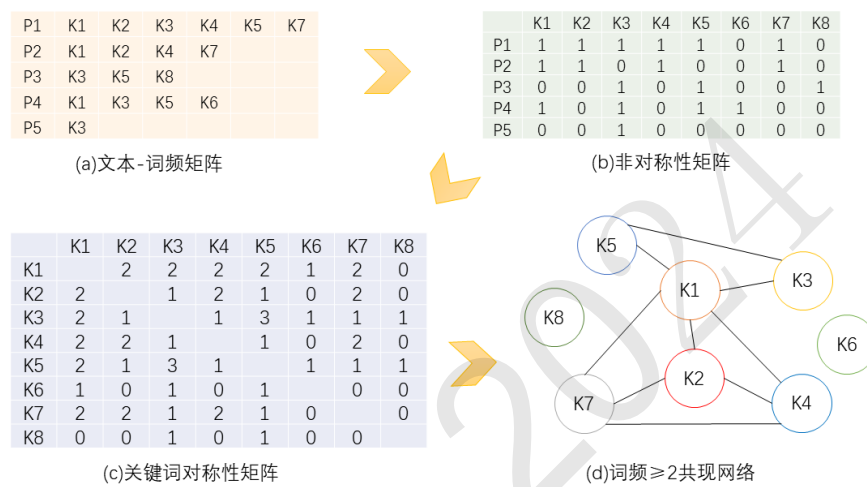


图 1: 共现分析原理示意图

## 2.3 数据来源

“域外汉籍”简称“汉籍”, 为避免以“域外汉籍”“海外汉籍”为主题检索时有遗漏, 或以“汉籍”为主题检索时结果宽泛, 采取“系统+人工”方式筛选样本。首先, 在CNKI以“主题=汉籍”进行检索, “期刊来源类别=北大核心+CSSCI+CSCD+AMI+SCI+EI”以保障论文优质, 并纳入两本域外汉籍专门辑刊中的相关研究文献。时间范围为1992年至2023年。检索时间为2024年3月1日。其次, 人工去除“西学汉籍”“汉族典籍”“汉代典籍”“英译汉典”等无关文献及干扰信息。共得到809篇样本文献。

## 2.4 研究流程

本研究基于CNKI采集文献数据、CiteSpace与VOSviewer绘制知识图谱、Microsoft Excel绘制图表、Note Express辅助词频分析。具体流程如图2所示。

## 3 发文信息分析

### 3.1 发文时间分布与拟合分析

论文数时序变化是衡量领域发展的重要指标(邱均平等, 2012)。如图3所示, 年发文量时间分布整体呈上升趋势。按照年发文量与发文增速阶段差异, 可将其分为“萌芽期”(1992-2005)、“成长期”(2006-2012)和“高速期”(2013-2023)。其年均发文量依次为10篇以下, 20篇以下, 20篇以上。为研究发文时间分布的健康程度, 分别绘制发文量曲线和普赖斯曲线。结果显示, “萌芽期”两条曲线拟合度高, 但此后年发文量低于普赖斯曲线。

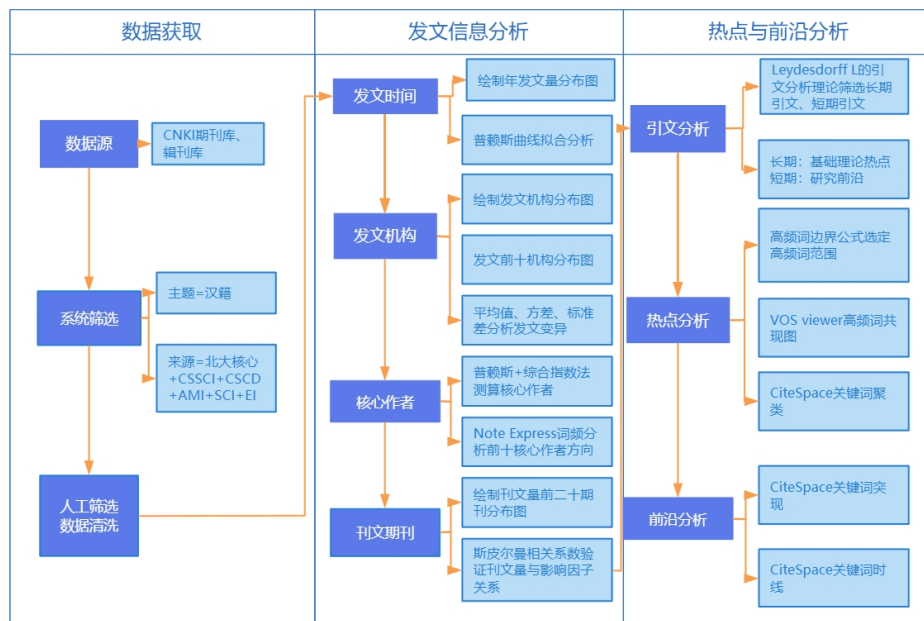


图 2: 研究路线流程图

这说明该领域学科发展的干扰因素可能较多。分析发文峰谷年份的发文环境可能有助于了解干扰因素。对关键峰值2006年分析发现，2005年创办了《域外汉籍研究集刊》，2006年《学习与探索》开设域外汉籍专栏，说明专业刊物对学科的积极影响。此外，2020年出现明显发文量低谷，这可能是受新冠疫情影响引发的线下交流与调查困难导致。但情况在一年后快速好转，说明该领域已形成一定的抗风险能力。

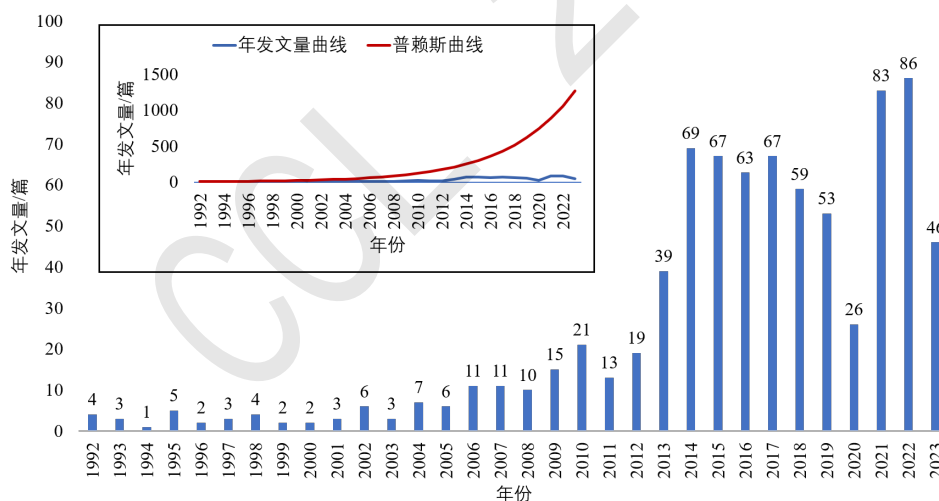


图 3: 域外汉籍研究发文时间分布与拟合曲线

### 3.2 研究机构分布

分析发现，域外汉籍研究机构以文学和历史院系为主（图4a）。提取机构信息，按照机构一级单位归并编码，并生成发文量前十机构逐年发文量图（图5）。统计发现，发文最多的机构是南京大学（89篇），其次是山东大学（28篇）、复旦大学（28篇）。高产机构大部分为“双一流”高校。南京大学于2000年2月成立了域外汉籍研究所，山东大学2013年起承担文化部“全球汉籍合璧工程”，复旦大学2022年获批国家社科基金重大项目“东亚汉籍版本学史”。上述因素可能对其成果产出有增益。在“双非”高校中，云南省蒙自市红河学院发文量为全国第11（9篇），

其在2008年成立越南研究中心。全球汉籍分布GIS系统<sup>0</sup>统计，与我国临近的越南等国汉籍资源丰富（图4b）。红河学院的高产出，说明边境省份高校开展域外汉籍研究地缘优势显著。



图 4: 发文机构分布与全球汉籍分布GIS系统越南汉籍地图示意

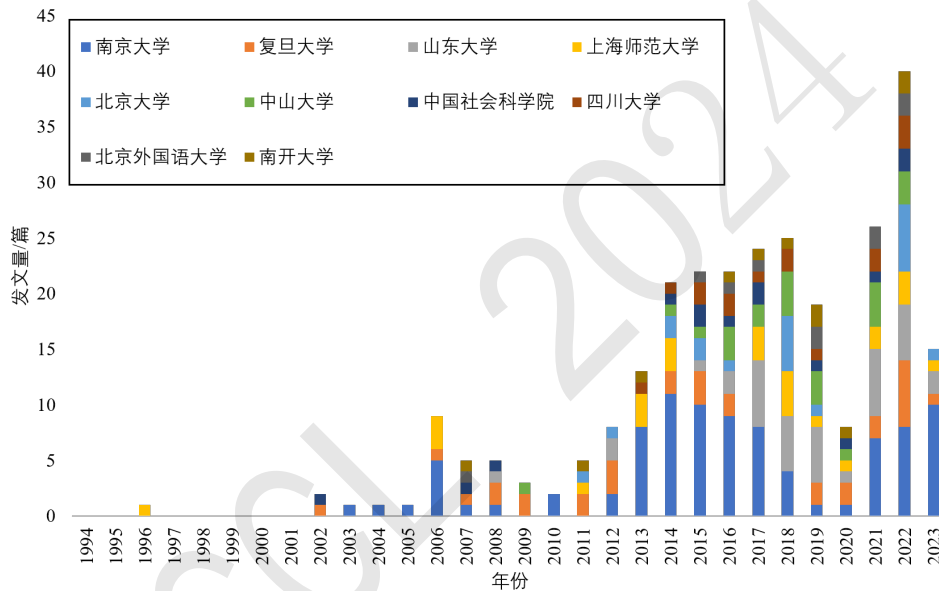


图 5: 发文前十机构逐年发文量

为评估机构发文量的变异程度，进一步计算发文前十机构的年发文量的平均值、方差、标准差（表1），结果发现，2014年标准差最高，为3.11，可知这一年各机构发文量差异最大。且此后标准差整体处于较高水平，说明“高速期”各机构成果产出差异变大。

### 3.3 核心作者分析

分析核心作者及其研究方向，有助于快速把握该领域的研究热点和最新动态。为确定核心作者，通过综合指数法进行测定。首先，使用普赖斯核心作者公式计算 $m=2.8\sim 3$ ，说明发文量在3篇及以上的为核心作者候选人。结果显示，366位作者（图6， $N=366$ ）中有34位候选人，合计发文159篇，占总成果量的19.6%。其次，将发文量和被引量权重值分别定为0.5和0.5（钟文娟，2012），计算综合指数值（附表B）。最后，使用Note Express对综合指数前10的核心作者成果进行词频分析（表2）。结果发现：核心作者关注日本、越南、朝鲜半岛地区，目录书、使臣著作、东亚汉诗类作品，数字化、传播策略相关策略经验。

<sup>0</sup>全球汉籍分布GIS系统.<https://www.wenxianxue.cn/gis/>.

年份	1994	1995	1996	1997-2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
平均值	0	0	0.1	0	0.2	0.1	0.1	0.1	0.9	0.5	0.5	0.3	0.2
方差	0	0	0.09	0	0.16	0.09	0.09	0.09	2.69	0.25	0.45	0.41	0.36
标准差	0	0	0.3	0	0.4	0.3	0.3	0.3	1.64	0.50	0.67	0.64	0.60
年份	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
平均值	0.5	0.8	1.3	2.1	2.2	2.2	2.4	2.5	1.9	0.8	2.6	4	1.5
方差	0.45	1.16	5.81	9.69	7.56	5.76	6.24	4.05	1.49	0.36	5.04	4	8.45
标准差	0.67	1.08	2.41	3.11	2.75	2.40	2.50	2.01	1.22	0.60	2.24	2.00	2.91

表 1: 各机构年发文量平均值、方差、标准差

另外，图6显示，海外高校作者表现积极，如陈翀（日本广岛大学）、水上雅晴（日本中央大学）。核心作者间合作较少，但存在多组跨国合作，如金程宇（南京大学）与衣川贤次（日本花园大学）、河野贵美子（日本早稻田大学）与葛继勇（郑州大学）。



图 6: 作者分布与合作共现图

### 3.4 刊文期刊分布

样本文献分布于161本期刊中，学科分布以文献学、文学、图书情报学为主，说明该领域目前文献辑佚、整理相关成果较为丰富，为后续研究打下坚实基础。期刊主要集中在社会科学领域，说明自然科学学科汉籍研究成果较少。《关于推进新时代古籍工作的意见》专门提出“梳理挖掘古典医籍精华”的要求，但当前研究缺口较大。统计刊文量前二十的期刊（图7），发现刊文最多的期刊为《域外汉籍研究集刊》（394篇），其次为《汉籍与汉学》（43篇），《文献》（20篇），其中前两者为学科专门性辑刊。

斯皮尔曼相关系数可以验证刊文量前二十的期刊与CNKI影响因子高低的相关度大小。将刊文量与影响因子整理为表3（辑刊无影响因子数据，不计入统计），计算发现两组数据的斯皮尔曼相关系数  $\rho = -0.4419$ ，表示该领域刊文量前二十期刊与CNKI综合影响因子之间存在一个较弱的负相关关系，说明刊发该领域论文数量的多少与期刊的综合影响因子并没有强烈的直接联系。

### 3.5 域外汉籍研究的热点与前沿

#### 3.5.1 基于引文分析的基础理论与前沿成果

Leydesdorff的引文分析理论将文献根据发文到成为高被引文献时间的长短，分为长期和短期引用。长期引用体现领域基础理论的形成，短期引用体现研究前沿性。将短期引用年限设为5年，分析被引频次前10的高被引文献(胡一尘, 2020)。结果发现，9篇为长期引用，分析其选题方向发现，域外汉籍在东亚汉字文化圈研究范畴已形成基础理论；1篇为短期引用（《中国

序号	姓名	高频词
1	张伯伟	方法; 文学; 书目; 东亚汉文化圈; 朝鲜
2	巴兆祥	日本; 地方志
3	卞东波	诗; 东亚; 日本; 诗人
4	金程宇	日藏善本书志; 中日韩戏剧; 东亚汉文小说; 东亚汉诗
5	陈舫	日本; 空海; 《日本国见在书目录》; 唐诗
6	刘玉	越南; 使臣; 燕行诗
7	王志强	日本; 朝鲜; 越南; 《往津日记》
8	毛建军	古籍数字化; 数据库
9	任晓霏	蒙学; 教育; 传播
10	叶少飞	大越史; 汉文小说; 安南使臣

表 2: 前十位核心作者研究方向高频词表

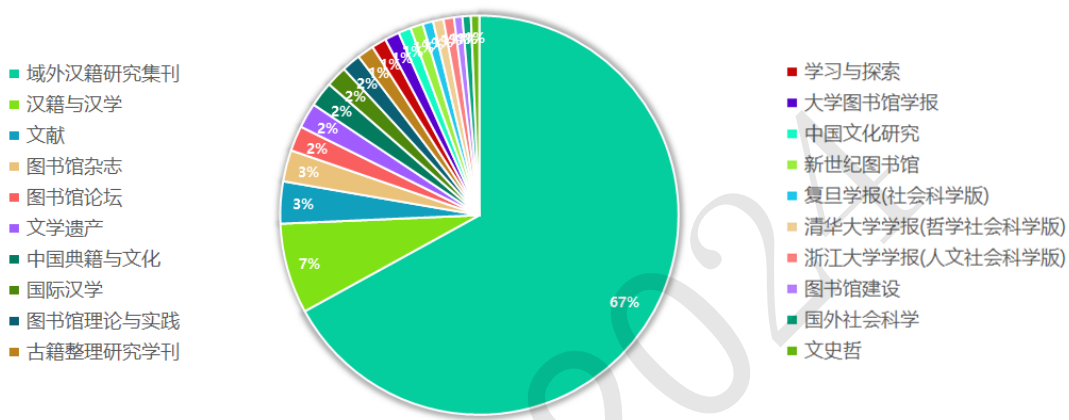


图 7: 刊文量前二十期刊分布图

古代蒙学典籍海外传播和影响研究》)，说明“蒙学教育”与“传播策略”是学者近年关注的热点（高被引文章列表可见于附表A）。

### 3.5.2 基于词频、聚类、共现的研究热点分析

关键词的频率、聚类、共现可以分析研究热点。样本文献包含487个关键词（图9，N=487），根据高频词边界公式 $T=27$ ，因第27个关键词出现频次为3，故将出现频次大于等于3的关键词认定为高频关键词，合计41个。使用VOSviewer生成高频词共线图，与使用CiteSpace生成的聚类图比对分析可以归纳出三个重要研究热点（图8、9）。在描述热点及内部成员发生的时间节点时，结合时线图、突现图进行分析（图10、11）。

#### 热点一：“图书馆+目录”模式下的汉籍普查研究

综合观察图8图书馆、古籍目录等共现词，图9聚类#2图书馆与聚类#6目录学，结合突现与时线图分析。结果发现，从1992年起，以图书馆为依托、以馆藏目录为抓手的汉籍普查已是域外汉籍研究的重要热点。研究对象主要分布在欧美和东亚。一般采用实地调查与数据检索结合的研究方式。欧美汉籍普查研究聚焦法国、英国、俄罗斯、梵蒂冈和美国。欧洲汉籍收藏历史悠久且存量丰富。法国国家图书馆、巴黎汉学研究所图书馆与《皇家图书馆写本目录》为关键普查线索(刘蕊and 岑咏芳, 2019)。英国目前所见的14种英国图书馆所编汉籍目录，涉及大英博物馆、牛津大学博德利图书馆等8间图书馆(谢辉, 2015)。俄罗斯汉籍主要藏于俄罗斯国立图书馆等4间图书馆(郑诚, 2023)，但其馆藏汉籍目录尚未出版。梵蒂冈图书馆的重要目录是伯希和所编《梵蒂冈图书馆所藏汉文写本和印本书籍简明目录》及高田时雄《梵蒂冈图书馆所藏汉籍目录补编》(谢辉, 2022)，后者编目采用汉语拼音并标注汉字，更易读。德国、奥地利、瑞典等国图书馆亦存少量汉籍(谢辉, 2016)。北美汉籍收藏始于1867-1869年中美之间的图书交换(杨

期刊名	文献	图书馆杂志	图书馆论坛	文学遗产	中国典籍与文化	国际汉学
刊文量	20	15	12	12	12	10
影响因子	0.345	1.747	2.448	0.424	0.145	0.305
期刊名	图书馆理论与实践	古籍整理研究学刊	学习与探索	大学图书馆学报	中国文化研究	新世纪图书馆
刊文量	9	8	7	7	6	6
影响因子	1.417	0.09	1.589	2.457	0.465	1.071
期刊名	复旦学报(社会科学版)	清华大学学报(哲学社会科学版)	浙江大学学报(人文社会科学版)	图书馆建设	国外社会科学	文史哲
刊文量	5	5	5	4	4	4
影响因子	1.202	1.173	1.173	2.091	1.833	0.944

表 3: 刊文前二十期刊刊文量与影响因子对照表

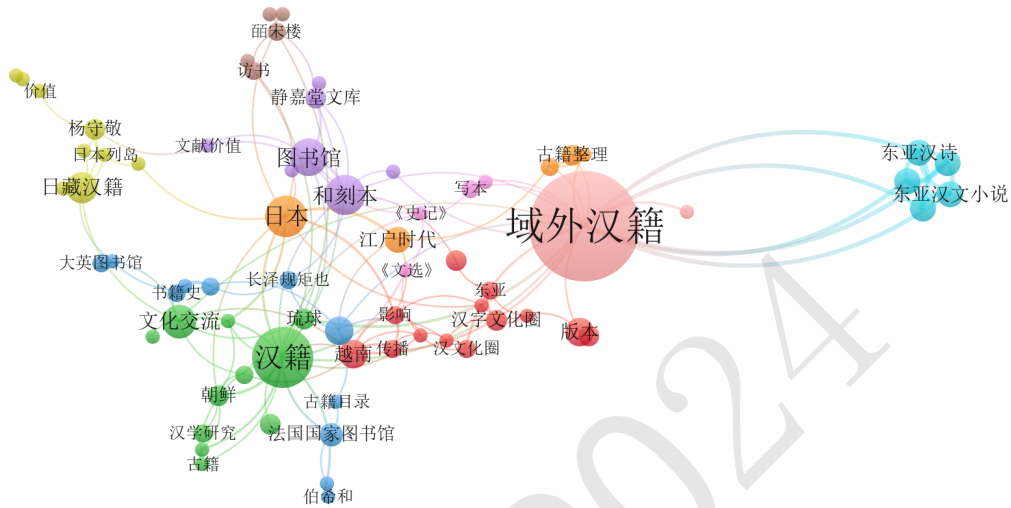


图 8: 高频关键词共现图

海峰, 2022)。主要藏于美国国会图书馆、哈佛大学燕京图书馆等数十间图书馆(严润开, 2001), 其中26间的汉籍普查已完成<sup>1</sup>。东亚汉籍普查研究聚焦于日本、韩国和越南。日本静嘉堂文库和韩国奎章阁热度最高。《“日本所藏中文古籍数据库”目录》《奎章阁图书中国本综合目录》研究最多。越南汉喃研究院和越南国家图书馆从2013年起成为热点(图11)。国内图书馆亦有回流汉籍, 如北京北堂图书馆(今西什库教堂附属图书馆)、北京大学图书馆、国家图书馆、天津图书馆和大连图书馆“大谷文库”。虽然海外汉籍普查至今尚无系统明确的统计数据公布, 但全球汉籍合璧工程项目组在2023年表示已基本已掌握境外汉籍存藏情况, 并通过课题方式逐步落实普查工作(陈灏et al., 2023)。因此, “图书馆+目录”模式下的汉籍普查研究, 可能在未来几年成果加速产出后结束热点。

**热点二：版本考证与版本流布研究**

综合观察图8和刻本、写本等共现词, 图9聚类#1汉籍(聚类成员版本研究、日本本、写本等)与聚类#16汉学(聚类成员安南本、出版等), 结合突现与时线图分析。结果发现, 版本是域外汉籍研究的热点, 关注版本考证和版本流布研究, 以1992年和刻本研究为起点, 在2016年至2020年热度最高。版本考证研究重视域外中国写本、抄本及孤本。研究对象存藏较为集中的有法国、英国劫购的敦煌遗书。其余孤本存藏分散, 仅日藏本研究成果相对较多。如日本所藏稀见明词研究(周明初and 彭志, 2014)。版本流布研究关注域外刻本。其中日本和韩国的“和刻本”“朝鲜本”研究热度最高, 尤其是江户时代和朝鲜时代。主要是因其汉籍流入时间早, 版本丰富, 有助于从版本学视角分析“和而不同”的东亚汉文化圈。成果以个案研究为主, 尤其关注佛典儒典(#1汉籍)。近十年, 越南“安南本”研究热度提升。研究材料获取难、部分文献社会价值不足是该方向的研究难点, 可能是其2020年之后前沿性不足的原因(图10)。完整系统的域外汉籍目录与数据库有助于解决该问题。因此, 在汉籍普查完成后, 该研究可能再次成为前沿热点。

<sup>1</sup>全球汉籍合璧工程.境外藏书机构联系进展情况一览表. <https://hanjihebi.sdu.edu.cn/info/1010/1032.htm>.



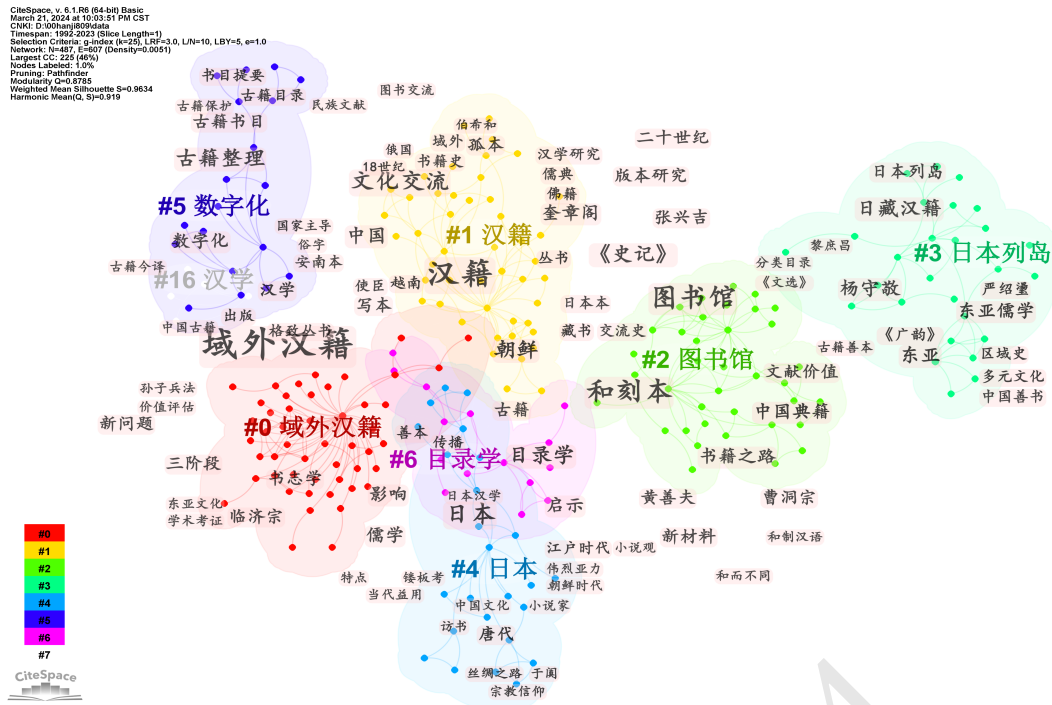


图 9: 关键词聚类图

### 热点三：名家名作专题研究

综合观察图8《史记》《文选》、长泽规矩也等共现词，图9聚类#1汉籍（聚类成员《史记》等）与聚类#4日本列岛（聚类成员严绍璁、《广韵》《文选》等），结合突现与时线图分析。结果发现，域外汉籍名家专题研究中，杨守敬、严绍璁（国内），以及伯希和、岛田翰、长泽规矩也（国外）是研究热点。其中杨守敬研究起步最早，持续最久，为1993年至2019年。严绍璁研究最晚，始于2009年。陆心源、王重民、董康等人的研究热度稍次之，但前沿性更显著，其中对董康的研究热点在2018年至2022年。国内名家研究关注其藏书、访书情况与代表作品。如杨守敬日本访书回传情况、陆心源藏书流失日本的情况、董康日本访书及其《书舶庸谈》、王重民欧洲访书及其所获敦煌遗书和太平天国文献、严绍璁及其《日藏汉籍善本书录》。国外名家研究关注其获取、整理汉籍。如日本学者岛田翰与宋楼藏书去日情况、长泽规矩也中国访书经历及成果、法国学者伯希和在华所购汉籍与海外收藏。名作研究重点关注作品在海外的经典化路径，热点有《广韵》《史记》《文选》《世说新语》，其中《广韵》研究热度开始最早。“世说学”仍在前沿（图11#0），相关研究有日韩对《世说新语》经典化研究、《世说新语》衍生作品研究(王子成, 2023)。总体来看，名家专题研究热点的热度持续时间大都较短，前沿性不显著。名作专题研究热点的热度持续时间较长，且“世说学”研究仍具有一定的前沿性。

### 3.5.3 基于突现、时线的研究前沿与展望

#### 前沿一：汉籍整理与数字化研究

基于数字化的汉籍整理出版研究是前沿热点之一，尤其聚焦于数据库平台和数字化技术两方面（图11#5数字化）。数据库平台研究关注数据库建设和应用研究。数字化技术研究关注汉籍数字化整理与转化。“数字化”突现首年是2018年（图10），当前处于研究前沿。

数据库建设研究重点关注日本汉籍数据库。其原因是日本相关建设比较成熟。“日本所藏中文古籍数据库”<sup>2</sup>检索系统是日本目前参与机构最多的汉籍数据库系统。在建设规范和实践路径上有重要参考价值，包括色彩平衡与数字标尺、采集粒度、存储参数与格式、特殊情况处理方面（古籍碎片、无字页、书内夹抄、附属物件）(朱本军, 2020)。欧美汉籍数据库研究热度较低，因其汉籍数据库建设投入不足、发展较慢(张敏and 董强, 2011)。数据库应用研究起步不

<sup>2</sup>日本所藏中文古籍数据库. <http://kanji.zinbun.kyoto-u.ac.jp/kanseki>.

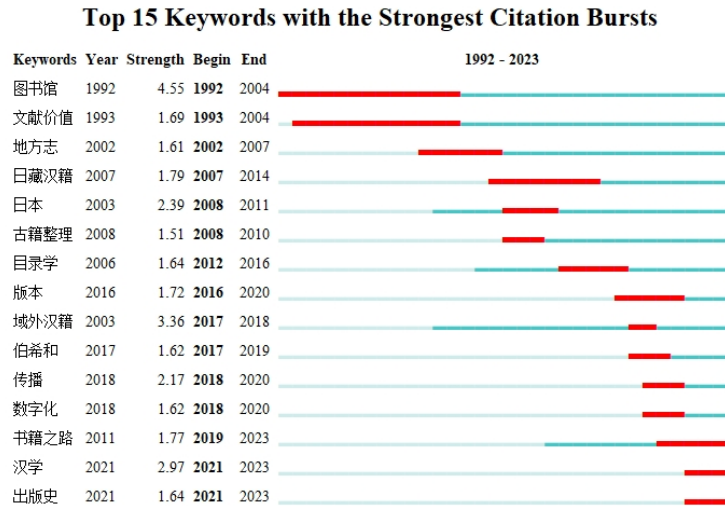


图 10: 域外汉籍研究关键词突现图

久。如2023年《中日汉籍出版中的“鳌头本”》使用国内外数据库检索查考出40余部“鳌头本”(潘超, 2023)。“中国古典文献资源导航系统”<sup>3</sup>数据库索引为汉籍研究提供便利。但整体来看, 汉籍数据库建设仍有不足, 大多仍需配合纸本使用。

汉籍数字化研究主要服务于回归汉籍的整理工作。目前汉籍以影像回归方式为主, 因此影像采集阶段的清晰度、速度、还原度, 以及采集后的影像自适应智能处理技术是影响数字化质量的关键。例如, 古籍印章识别技术就是古籍副文本智能处理研究的经典案例(顿泽栋, 2022)。但相关技术研究成果多为影像采集阶段, 智能处理阶段成果依然不足(罗渝and 黄璜, 2018)。“全球汉籍合璧”工程等项目显示出, 国家主导下汉籍数字回归加速(图11#5), 其伴生的汉籍数字化整理工作方兴未艾。更重要的是, 数字化整理、加工是汉籍规模化“用起来”的重要前提, 因此, 本方向的研究在未来较长时间内都将是前沿趋势。

### 前沿二: 文化交流与传播研究

汉籍承载着中华文化之精神, “书籍之路”即是中华文化流播之路。“传播”“书籍之路”首现时间为2018年和2019年(图10), 图11中多个聚类时线最新的成员词都与传播相关, 如#0文化传播、#2唐诗传播、#4传播途径等, 说明文化交流与传播研究的前沿性趋势。与汉籍普查、回归、整理等“引回来”研究不同, 该方向关注“走出去”, 尤其是“再次走出去”的相关研究。包括三类内容: 汉籍与文化传播史、汉籍与文化交流融合以及当代传播策略。

汉籍与文化传播史研究, 关注汉文化圈国家的汉籍受容。国别上关注日本、韩国、越南; 时代上关注中国的唐代、日本的江户时代、朝韩的朝鲜时代。传播途径上关注赏赐(求书)、贸易、劫购、访书等。传播者关注来华使臣、留学僧(生)、贸易商、汉学家。具体研究内容上有蒙学、道教、汉字、儒学、佛学等某一门类汉籍与文化的传播史, 以及《三字经》《文心雕龙》《欧苏手简》等某一汉籍及其文化义涵的受容与变容。少量欧美国家汉籍与汉文化传播史研究关注欧洲的18世纪“中国热”时期。汉籍与文化交流融合研究, 关注汉籍中外互动研究。研究内容包括“汉籍拟效研究”“融合互鉴研究”。“汉籍拟效研究”有三个方向。其一是海外对汉籍编辑方法拟效, 如英国汉学家伟烈亚力以中国传统四分法编著《中国文献解题》。其二是仿写续写研究, 如日本《两假名杂字尽》对中国杂字类蒙书的仿作、17至19世纪日本朝鲜笈释摹仿《世说新语》热潮研究。其三是科技典籍仿效应用研究, 如中国《农政全书》及其日仿本《农业全书》对日本农业文明与技术发展影响的研究。“融合互鉴研究”关注书籍环流中的多元文化共生, 如《纪效新书》等中国武术古籍在朝鲜和日本传播后, 其原著及衍生作品回流中国, 其作品中展示出的东亚三国文化和而不同的多元融合(蔡艺et al., 2022)。当代传播策略研究, 关注如何利用汉籍开展对外交流和文明互鉴。在汉籍译介和学术交流等传统路径外, 依托新媒体的汉籍转化传播开始受到关注(姚委委et al., 2021)。但仍属于新兴领域。

<sup>3</sup>中国古典文献资源导航系统.<https://www.wenxianxue.cn/view/1.html>.

CiteSpace, v. 5.10.R6 (64-bit) Basic  
 March 19, 2024 at 4:00:58 PM CST  
 CML: 0.300(0.05)Data  
 Timespan: 1992-2023 (Slice Length=1)  
 Selection Criteria: g-index (w=25), LRF=3.0, LN=10, LBY=5, m=1.0  
 Network: N=487, E=607 (Density=0.0051)  
 Largest CC: 226 (46%)  
 Nodes Labeled: 1.0%  
 Pruning: Pathfinder  
 Modularity Q=0.8785  
 Weighted Mean Silhouette S=0.9634  
 Harmonic Mean(Q, S)=0.919

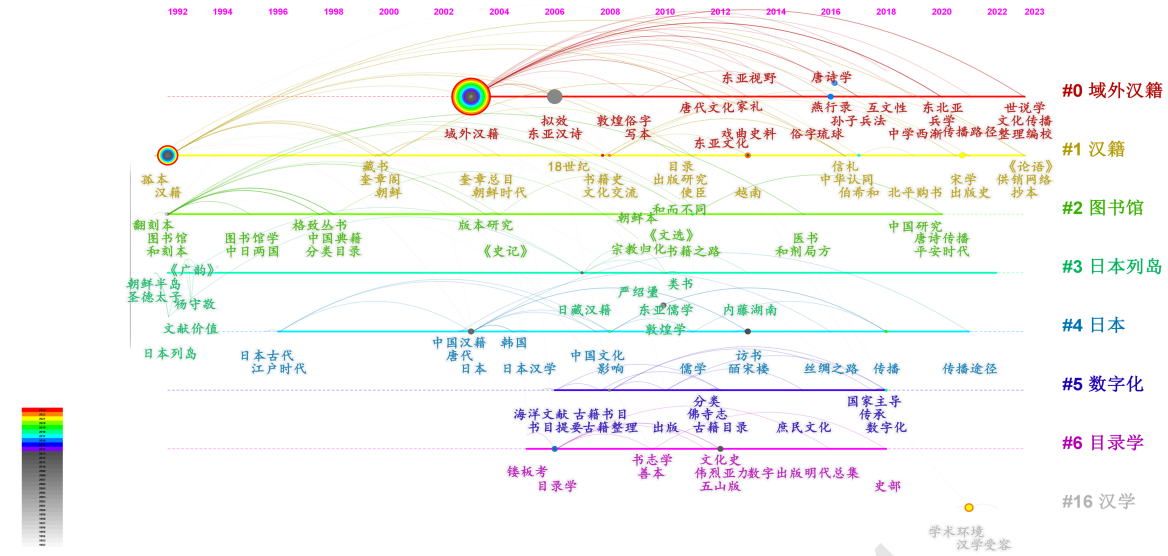


图 11: 域外汉籍研究关键词与时线图

#### 4 总结与展望

综上所述，通过定量与定性结合的分析方法，分析以往传统定性研究无法实现的长时段、大数据样本，对域外汉籍研究情况有了更宏观和全景式的把握，并精确发现：域外汉籍研究在近几年已形成一门有核心作者群和鲜明研究方向的学科，且热点与前沿方向的发展变化揭示了领域内研究工作已经出现显著的学科转向。其从“汉籍回归普查”的研究，转向为“汉籍应用转化”研究。这表明该领域已经逐渐从“引回来”研究，过渡到“用起来”和“再次走出去”的相关研究。因此，如何基于数字化手段充分展开域外汉籍开发整理、如何将域外汉籍转化为中华文化海外传播的重要材料都是应当关注的重要方向。本研究仍然存在不足，如本研究样本不包括硕博论文、专著，以及主题不符但实际研究内容为汉籍的文献，但整体来看仍能反映出学科研究的现状，为学科未来发展方向提供参考数据。

#### 参考文献

Joseph C. Donohue. 2014. A bibliometric analysis of certain information science literature. *Journal of the American Society for Information Science and Technology*, 23(5):313-317.

Loet, Leydesdorff and Lutz, Bornmann ,et al. 2016. Citations: Indicators of Quality? The Impact Fallacy. *Frontiers in Research Metrics and Analytics*.

白雪涵. 2022. 基于深度学习的语义理解算法及应用研究. 北京邮电大学.

蔡艺and 谭飞and 唐群. 2022. “书籍之路”与东亚武籍交流. *体育学刊*, 29(06), 31-36.

陈灏and 萧海川and 袁敏. 2023. 一个工程，一群学者，一场对海外珍稀汉籍的寻觅. 新华社, [http://www.news.cn/local/2023-07/11/c\\_1129744215.htm](http://www.news.cn/local/2023-07/11/c_1129744215.htm).

丁学东. 1993. 文献计量学基础. 北京大学出版社.

顿泽栋. 2022. “汉籍合璧”工程中古籍印章的识别与研究. 山东大学.

傅荣贤and 华建铭. 2020. 新中国70年(1949-2019)中国古典文献学研究综述. *图书馆杂志*, (02), 4-33.

胡一尘. 2020. 基于Web of Science大规模文献数据的高引论文的影响因素研究. 西南大学.

奎欣怡. 2023. 2022年中国域外汉籍研究述评. *中国文学年鉴*.

- 金程宇. 2010. 近十年中国域外汉籍研究述评. 南京大学学报: 哲学. 人文科学. 社会科学, 47(3):14.
- 李杰and 陈超美. 2016. CiteSpace:科技文本挖掘及可视化:text mining and visualization in scientific literature. 首都经济贸易大学出版社.
- 李琳. 2014. 新时期以来中国古代文学的文献研究. 复旦学报(社会科学版), (02), 57-64.
- 刘蕊and 岑咏芳. 2019. 伯希和1933年北平购书与法兰西学院汉学研究所的藏书. 国际汉学, (03), 152-159, 207-208.
- 刘泰廷. 2017. 中国近五年域外汉籍研究述评. 图书馆理论与实践, (01), 67-72.
- 罗渝and 黄璜. 2018. 中文古籍数字化建设的实践与思考——以“域外汉籍数字服务平台”为例. 出版发行研究, (07), 39-42.
- 潘超. 2023. 中日汉籍出版中的“鳌头本”. 古代文明, 17(02), 141-155, 160.
- 邱均平and 杨思洛and 宋艳辉. 2012. 知识交流研究现状可视化分析. 中国图书馆学报,38(02), 78-89.
- 王子成. 2023. 文化传播视域下《世说新语》的经典化. 明清小说研究, (02), 180-195.
- 谢辉. 2015. 英国图书馆所编汉籍目录初探. 新世纪图书馆, (02), 75-79.
- 谢辉. 2022. 梵蒂冈图书馆收藏中国传统古籍的历程与价值. 图书馆论坛, 42(10), 150-160.
- 谢辉. 2016. 欧洲图书馆所编早期汉籍目录初探. 图书馆理论与实践, (02), 96-100.
- 严润开. 2001. 美国著名的华文汉籍文献收藏的特殊文库. 情报杂志, (04), 99-100.
- 杨海峥. 2022. 顾盛与北美汉籍收藏. 图书馆论坛, 42(01), 131-136.
- 姚委委and 金长善and 程钰情. 2021. 新媒体时代中韩文化交流与传播路径探析——推荐《域外汉籍传播与中韩词学交流》. 新闻记者, (10), 97.
- 张伯伟. 2006. 域外汉籍研究:域外汉籍研究——一个崭新的学术领域. 学习与探索,(02), 159-160.
- 张敏and 董强. 2011. 数字时代海外中华古籍的回归. 图书馆建设, (09), 38-40, 44.
- 郑诚. 2023. 俄罗斯国立图书馆藏中文古地图研究. 自然科学史研究, 42(02), 170-190.
- 钟文娟. 2012. 基于普赖斯定律与综合指数法的核心作者测评——以《图书馆建设》为例. 科技管理研究, 32(02), 57-60.
- 周明初and 彭志. 2014. 日本所藏稀见明词续补. 南京师范大学文学院学报, (02), 173-179.
- 朱本军. 2020. 海外汉籍数字化加工现状与实践研究. 古籍保护研究 (01) 18.

## 附录A.域外汉籍研究高被引文献

## 附录B.核心作者综合指数表

序号	文章标题	关键词	发表时间	被引量
1	域外汉籍与中国文学研究	原文无	2003	39
2	中日韩越古医籍数据的比较研究	汉字文化；古医籍；定量分析；比较研究；医学本国化；共同现象	2010	29
3	从“汉籍”到“域外汉籍”	书籍之路；和制汉语；和而不同；和刻本；“域外汉籍”	2011	28
4	近十年中国域外汉籍研究述评	原文无	2010	27
5	中国古代蒙学典籍海外传播和影响研究	蒙学典籍；传播；海外	2019	26
6	域外汉籍研究——一个崭新的学术领域	域外汉籍；东亚汉文小说；东亚汉诗；中、日、韩民族戏剧；拟效；汉籍东传、域外访书	2006	24
7	早期的汉外字典——梵蒂冈馆藏西土语文手稿十四种略述	汉外字典；西方汉语研究史；梵蒂冈馆藏汉籍	2007	21
8	日藏稀见汉籍《中兴禅林风月集》及其文献价值	原文无	2004	20
9	越南汉籍《往津日记》及其史料价值评介	《往津日记》；阮述；中越关系；轮船招商局	2010	20
10	“燕行录”戏曲史料的学术价值初探	燕行录；戏曲史料；域外汉籍	2013	20

序号	作者	发文量	平均发文量	被引量	平均被引量	综合指数
1	张伯伟	14	4.68	99	17.5	432.4298
2	巴兆祥	4	4.68	51	17.5	188.4493
3	卞东波	12	4.68	20	17.5	185.348
4	金程宇	5	4.68	44	17.5	179.1331
5	陈祚	12	4.68	15	17.5	171.0623
6	刘玉	6	4.68	28	17.5	144.1026
7	王志强	3	4.68	37	17.5	137.7656
8	毛建军	3	4.68	33	17.5	126.337
9	任晓霏	3	4.68	32	17.5	123.4799
10	叶少飞	9	4.68	9	17.5	121.8681
11	王晓平	3	4.68	29	17.5	114.9084
12	钱婉约	4	4.68	25	17.5	114.1636
13	陈恒新	4	4.68	24	17.5	111.3065
14	左江	6	4.68	12	17.5	98.38828
15	宋莉华	3	4.68	22	17.5	94.90842
16	孙卫国	4	4.68	18	17.5	94.16361
17	宋丽娟	3	4.68	21	17.5	92.05128
18	严明	3	4.68	20	17.5	89.19414
19	河野贵美子	5	4.68	10	17.5	81.99023
20	陈东辉	4	4.68	12	17.5	77.02076
21	水上雅晴	3	4.68	12	17.5	66.337
22	国威	6	4.68	0	17.5	64.10256
23	叶杨曦	5	4.68	1	17.5	56.27595
24	定源(王招国)	4	4.68	4	17.5	54.16361
25	颀飞	3	4.68	7	17.5	52.05128
26	付星星	4	4.68	0	17.5	42.73504
27	羽离子	3	4.68	3	17.5	40.62271
28	俞士玲	3	4.68	3	17.5	40.62271
29	杨洋	3	4.68	2	17.5	37.76557
30	何仟年	3	4.68	2	17.5	37.76557
31	童岭	3	4.68	1	17.5	34.90842
32	吴伟明	3	4.68	0	17.5	32.05128
33	刘芳亮	3	4.68	0	17.5	32.05128
34	刘一	3	4.68	0	17.5	32.05128