

## 中文期刊论文引文国际化多维分析 ——以图书情报与档案管理学科为例

龚凯乐, 谢娟, 成颖

(南京大学 信息管理学院, 江苏 南京 210023)

**摘要:**【目的/意义】引文国际化能够帮助本土学者以国际视野、国际思维和国际通用的学术范式、理念、方法从事学术研究,提高研究水平,进而更好地在国际上发出中国声音、讲好中国故事。【方法/过程】以图书情报与档案管理(LIS)学科为例,选取中文社会科学引文索引(CSSCI)建库以来每年均收录的16本LIS期刊作为数据源,通过对中文期刊论文引用外文文献的语种、文献类型、来源期刊、学科领域、学术质量、时效性进行多维分析,揭示本土学者在外文文献选择与内化的过程中经历的发展与变化。基于上述指标对施引文献聚类,进一步发掘出引文国际化的主要模式。【结果/结论】被引文献中,外文占比逐年上升且英语占主导地位;被引外文文献中期刊论文占比最高,其次是网络资源、会议论文、图书和学位论文;被引频次较高的外文期刊中,多数来自LIS领域,但也有其他学科的期刊,较少数期刊获得了大多数引用;除本学科外,信息系统、计算机科学、管理科学与商学的外文文献被引较多,跨学科程度稳中有升;被引外文文献的学术质量呈上升趋势;被引外文文献的年龄呈增长趋势。此外,本研究还发现了五种各具特点的引文国际化模式。

**关键词:** 引文国际化;图书情报与档案管理;引文分析;引文模式

中图分类号:G353.1 DOI:10.13833/j.issn.1007-7634.2019.03.021

### Multi-dimensional Analysis of the Internationalization of References in Chinese Journal Papers: A Study on Library & Information Science

GONG Kai-le, XIE Juan, CHENG Ying

(School of Information Management, Nanjing University, Nanjing 210023, China)

**Abstract:** 【Purpose/significance】 The internationalization of references can help domestic scholars engage in academic research with international vision, international thinking as well as international common academic paradigms, ideas and methods. So that, the research quality can be improved and the articles can be published in international journals. 【Method/process】 Taking Library & Information Science as the research object. Selecting 16 journals, which have been indexed by CSSCI from 1998, as the data source. Revealing the development and changes of local scholars in the process of selecting and internalizing foreign-language literatures by analyzing the language, literature type, journal, subject, academic quality and timeliness of references in Chinese journal papers. Discovering the main patterns of internationalization of references by clustering. 【Result/conclusion】 The research finds that the proportion of foreign-language references is increasing year by year and English is dominant; journal articles have the highest percentage in foreign-language references, the following are internet resources, conference papers, books and dissertations; most of the cited foreign journals are in the field of LIS, but there are also journals in other disciplines, very few journals possess the vast majority of citations; in addition to LIS, the foreign-language papers from information system, computer science, management science and business are cited most, the interdisciplinary degree is steadily rising; the academic quality of cited foreign-language literature is on the rise; the age of cited literatures is increasing. Furthermore, the study also finds five patterns of internationalization of references with different characteristics.

**Keywords:** internationalization of references; Library & Information Science; citation analysis; citation patterns

收稿日期:2018-07-27

基金项目:国家社会科学基金项目“施引者引用意向与文献计量视角的学术论文被引影响因素研究”(17BTQ014)

作者简介:龚凯乐(1993-),男,江苏镇江人,博士研究生,主要从事信息计量学研究。

## 1 引言

在中国高等教育“双一流”建设的背景下,学术研究国际化势在必行,其可辩证分解为“引进来”和“走出去”两个维度<sup>[1]</sup>。由于学术研究产生的成果和科学创新发现的知识多以各类文献记载和传承<sup>[2]</sup>,其中期刊论文占据多数,因此在一定程度上,学术研究的国际化最终表现为期刊论文的“引进来”和“走出去”,形式则灵活多样。前者包括中文期刊论文的引文国际化<sup>[1]</sup>、作者和稿源国际化<sup>[3]</sup>、编委与审稿国际化<sup>[4]</sup>等,后者则包括中国学者在国际期刊发文<sup>[5-6]</sup>、中国期刊的国际出版<sup>[7]</sup>、中国期刊被SCI等国际数据库收录<sup>[8]</sup>、中文论文被外文文献引用<sup>[9]</sup>等。基于此,本课题组在前期研究中归纳出了一条从“引进来”到“走出去”再到“引进来”的国际化实现路径<sup>[1]</sup>。其中,引文国际化作为路径起点有着十分重要的作用。

引文国际化是指中文论文引用外文文献(包括期刊论文、会议论文、专著等)的情况,可由“量”与“质”两个维度测度。其中,“量”回答了中文论文引用了多少外文文献的问题,测度指标为外文参考文献的数量与比例;“质”则回答了中文论文引用的外文文献具有何种性质的问题,测度指标包括语种、文献类型、来源期刊、学科领域、学术质量、时效性等特征。

引文国际化在学术研究国际化中的重要作用是由引文的本质所决定的。从知识论关于知识的发展模式来看,文献引用的过程就是在前人知识的基础上知识进化的过程,是知识的选择、遗传和变异的过程,也是知识的生产、传播和应用的过程<sup>[10]</sup>。尽管引用行为的动机多样,其能否完全反映施引文献与被引文献之间的知识关联尚存争议,但由于其具有客观性和可测度性,依旧被作为测度知识流动的有效工具广泛使用<sup>[11]</sup>。当引文流出现于不同的文献集合间时,可通过其定量测度知识在二者间的转移<sup>[12]</sup>,如图1。当集合A、B分别表示两个学科的文集时,可通过引文流测度知识在不同学科间的扩散<sup>[13-14]</sup>;当集合A、B分别为专利与论文集时,引文流则体现出知识由学术界向工业界转移,揭示出学术界对工业界创新的驱动<sup>[15-16]</sup>。因此,当施引文献集合A为中文文献且被引文献集合B为外文文献时,可通过引文国际化(反映为集合A到B的引文流)定量测度科学知识由国际向本土的转移(反映为集合B到A的知识流)。这种知识的转移能够反映出本土学者对国际学术成果的选择、内化与再创造。通过该过程的积淀,本土学者将自觉自发地以国际视野、国际思维、国际通用的学术范式、理念、方法等从事学术研究,这不仅有助于避免闭门造车,不断提升本土研究的质量,还

有助于更好地在国际上发出中国声音,讲好中国故事。

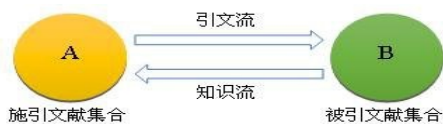


图1 引文流和知识流

前期研究已从“量”的视角分析了图书情报与档案管理(LIS)学科中文期刊论文的引文国际化水平,故本文从“质”的角度展开多维分析。选取中文社会科学引文索引(CSSCI)<sup>[17]</sup>建库以来每年均收录的16本LIS学科来源期刊作为数据源,通过对期刊论文引用外文文献的语种、文献类型、来源期刊、学科领域、学术质量、时效性进行多维分析,定量测度科学知识由国际向本土的转移,这不仅反映出本土学者对外文文献资源需求和利用的变化,还将折射出本土研究自身的发展与变迁。此外,本研究还将基于引文国际化多维指标对施引文献聚类,揭示出中文LIS论文引文国际化的主要模式。

## 2 数据与方法

## 2.1 数据来源与范围

CSSCI是由中国社会科学研究评价中心于1998年创建的引文索引数据库,广泛应用于引文分析、期刊评价等领域<sup>[18]</sup>。由于该数据库的来源期刊目录定期更新,为了保证数据的连续性与完整性,本研究选取1998至2015年持续收录的16本LIS期刊作为数据源(表1)。剔除数据采集时尚未更新全部字段的少量文献后,最终纳入分析的期刊论文共计67724篇,共有参考文献643398条,篇均9.5条。

## 2.2 引文国际化的多维统计

## 2.2.1 语种

CSSCI后台采用数字化方式对被引文献的语种编码<sup>[17]</sup>,编码规则为:中文01,英文02,德文03,法文04,俄文05,日文06,韩文99,译文09,其它07。据此,本研究计算被引文献中各语种所占比例并绘制饼图,统计1998至2015年间各语种文献占当年全部被引文献的比例并绘制折线图。

## 2.2.2 文献类型

CSSCI同样采用数字化方式编码被引文献的类型,编码规则为:期刊论文01,图书02,报纸03,会议录04,学位论文05,档案06,析出文献07,报告08,标准09,法律法规10,电子资源11,其它99。据此,本研究计算被引外文文献中各类

表1 CSSCI持续收录的LIS学科来源期刊

期刊类别			期刊名称		
情报学	情报学报	情报科学	情报杂志	情报理论与实践	情报资料工作
图书馆学	中国图书馆学报	大学图书馆学报	图书馆论坛	图书馆工作与研究	图书馆杂志
图情综合	图书情报工作	图书情报知识	图书与情报	现代图书情报技术	
档案学	档案学通讯	档案学研究			

型占比并绘制饼图,统计1998至2015年间各类型文献占当年全部被引外文文献的比例并绘制折线图。

### 2.2.3 来源期刊

根据国标GB/T-7714《信息与文献——文后参考文献著录规则》,外文参考文献的期刊名称可使用全称、缩写两种形式。因此在统计被引外文文献的来源期刊时,需将缩写还原。由于不同作者采用的缩写方法不一致,因此无法简单地通过《Journal Title Abbreviations》等刊名缩写对照表还原。与此同时,语言障碍、作者疏忽与编辑部把关不严等问题使得中文期刊论文中参考文献著录错误的现象较为普遍,外文期刊名称书写错误便是主要问题之一<sup>[19-20]</sup>。因此,直接以原始数据统计被引外文文献的来源期刊将有较大误差。

为妥善解决上述问题,本研究首先参照国际标准ISO 4《Information and documentation – Rules for the abbreviation of title words and titles of publications》编制外文期刊名称常用词缩写对照表,主要包括刊名中出现的出版类型(如J-Journal, Q-Quarterly, Proc-Proceedings等)、国家名称(如USA-United States of America, Br-British, Int-International等)、机构名称(如Assoc-Association, Soc-Society等)、学科名称(如Med-Medicine, Phys-Physics等)、专业术语(如Inf-Information, Libr-Library, Sci-Science等)等。考虑部分作者未遵循ISO4,因此在利用该表还原单词缩写时将还原标准形式向后0-3个字符的扩展形式。上述过程一定程度上减小了刊名缩写带来的误差,但未被对照表收录的单词与被省略的冠词依然无法还原,拼写错误、漏写多写等问题仍旧存在。

由于Scopus是目前全球期刊覆盖最全的文摘和引文数据库<sup>[21-22]</sup>,基本囊括了国内学者常用的外文期刊。故将缩写还原后的刊名与Scopus来源出版物列表进行近似匹配,从而得到准确且规范的期刊名称。这也将为后续学科领域、学术质量等维度的分析提供可靠基础。近似匹配方法如下:设被引外文文献的刊名集合为 $T=\{t_1, t_2, \dots, t_n\}$ , Scopus来源出版物列表为 $S=\{s_1, s_2, \dots, s_m\}$ , 对任意 $t_i \in T$ , 计算其与 $S$ 中所有元素的相似度, 当存在 $s_j \in S$  满足 $\text{Sim}(t_i, s_j) = \text{Max}(\text{Sim}(t_i, s_1), \text{Sim}(t_i, s_2), \dots, \text{Sim}(t_i, s_m))$  且 $\text{Sim}(t_i, s_j) \geq 0.8$  时, 即可认为 $s_j$  为 $t_i$  的标准名称; 若不存在 $s_j$  使得 $\text{Sim}(t_i, s_j) \geq 0.8$  成立, 且 $t_i$  在刊名纪录中出现的频次大于20, 则回溯元数据后手动检索其准确刊名。 $t_i$  与 $s_j$  的相似度计算方法为(1)。

$$\text{Sim}(t_i, s_j) = 1 - \frac{\text{Levenshtein}(t_i, s_j)}{\text{Max}(\text{StrLen}(t_i), \text{StrLen}(s_j))} \quad (1)$$

其中 $\text{Max}()$ 为最大值函数,  $\text{StrLen}()$ 为字符串长度函数,  $\text{Levenshtein}(t_i, s_j)$ 为字符串 $t_i$ 与 $s_j$ 的Levenshtein距离, 又称编辑距离, 由 $t_i$ 转换成 $s_j$ 所需的最少编辑次数求得<sup>[23]</sup>。

经抽样检验, 上述刊名还原方法的正确率达95.6%, 有效提升了实证分析的准确性。据此, 本研究统计外文期刊中各刊被引频次与占比。

### 2.2.4 学科领域

由于期刊的学科分类体系最为健全, 故分析被引外文文献的学科领域时仅纳入期刊论文。常用的分类标准主要来

自于Scopus和Web of Science, 由于Scopus收录的期刊更全, 故在划分学科领域时主要基于Scopus。另一方面, 国内学者对外文LIS期刊的认知主要基于SSCI, 故LIS学科的期刊取Scopus与WoS的并集。

被引外文文献的学科领域主要由学科专属度与跨学科度两个指标测度。前者体现了被引文献学科来源的专一性, 后者则反映了被引文献学科来源的多样性。计算方法如下: 设被引文献中的外文期刊列表为 $R=[r_1, r_2, \dots, r_n]$ , 有学科集合 $D=\{d_1, d_2, \dots, d_m\}$ 及其所对应的期刊集合 $J_{d_1}, J_{d_2}, \dots, J_{d_m}$ , 则被引外文文献在学科 $k$ 中的专属度为(2), 其中 $\text{Find}(r_i, J_{d_k})$ 的计算方法见(3), 被引外文文献的跨学科度借鉴了Leydesdorff等<sup>[24]</sup>、Wu等<sup>[25]</sup>研究中所使用的香农熵, 计算方法为(4)。

$$\text{Sp}(R, J_{d_k}) = \frac{\sum_{i=1}^n \text{Find}(r_i, J_{d_k})}{n} \quad (2)$$

$$\text{Find}(r_i, J_{d_k}) = \begin{cases} 1, & r_i \in J_{d_k} \\ 0, & r_i \notin J_{d_k} \end{cases} \quad (3)$$

$$\text{Int}(R) = - \sum_{k=1}^m \text{Sp}(R, J_{d_k}) \ln[\text{Sp}(R, J_{d_k})] \quad (4)$$

据此, 本研究首先统计被引外文文献中各学科被引频次与占比(计算时, 属于 $N$ 个学科的一本期刊每获得1次引用, 相应 $N$ 个学科的被引频次均加1)。进一步的, 根据(2)与(4)分别计算1998至2015年间被引外文文献每年的LIS学科专属度与跨学科度并绘制折线图。

### 2.2.5 学术质量

分析被引外文文献的学术质量时同样仅纳入期刊论文。目前, 文献学术质量的测度指标多种多样, 包括被引频次、发表期刊的声誉与影响因子<sup>[26]</sup>、引文网络中的PageRank得分<sup>[27]</sup>、全文数据库中的usage metrics<sup>[28]</sup>以及社交媒体中的Altmetrics<sup>[29]</sup>等。尽管学界对使用期刊影响力代表论文质量的作法尚存争议<sup>[30]</sup>, 但由于本研究数据量巨大且其他复杂指标不易获取, 因此仍使用期刊影响力代表论文的学术质量。由于不同学科在文献发表与引证行为上存在系统性差异, 因此期刊影响因子(IF)等绝对值指标不可跨学科比较<sup>[31]</sup>。为了在分析被引文献的学术质量时消除学科差异, 故本研究采用JCR提供的归一化的、跨学科可比的<sup>[32]</sup>期刊影响力评价指标——Average Journal Impact Factor Percentile (JIFP)度量被引外文文献的质量。JIFP的计算方法为(5), 其中 $N$ 为某一期刊所属学科类别中的期刊总数,  $R$ 是按影响因子降序排列后该刊在其所属学科中的顺序位置。

$$\text{JIFP} = \frac{(N - R + 0.5)}{N} \quad (5)$$

计算时, 被引外文文献的学术质量等于其发表年所在期刊的JIFP, 若该刊当年未被JCR收录, 则 $\text{JIFP}=0$ , 若论文发表年早于1997, 则取1997年的JIFP。据此, 本研究统计1998至2015年间被引外文文献的年均JIFP并绘制折线图。

### 2.2.6 时效性

被引外文文献的时效性仍然仅纳入期刊论文, 由引用半衰期与引文年龄测度。对于发表于同一年的指定施引文献



集合,引用半衰期是指引文达到当前被引文献的50%(较新那部分)所需要的年数<sup>[33]</sup>,计算方法为(6)。其中A为累积引用频次最接近并小于50%的那年的累积引用频次,B为累积引用频次最接近且大于50%的累积引用数,X为累积引用频次为A的年度与观测年的差。

$$HL = X + \frac{(50\% - A)}{(B - A)} \quad (6)$$

引文年龄指施引文献与被引文献的年差。据此,本研究统计1998至2015年间外文期刊论文每年的引用半衰期与年均引文年龄并绘制折线图。

## 2.3 施引文献的聚类分析

为了进一步了解中文LIS期刊论文引用外文文献的具体模式,本研究采用K-means算法<sup>[34]</sup>对施引文献聚类,聚类特征为每篇施引文献的参考文献列表中:①外文文献占比、②期刊论文占外文文献的比例、③外文期刊论文的LIS学科专属度、④外文期刊论文的平均JIFP、⑤外文期刊论文中新文献的比例(其中,新文献是指被引文献与施引文献的发表年差小于等于施引文献发表当年全体文献的平均引文年龄)。

聚类分析的工具为Weka3.8.2,相似度计算方法为欧几里得距离。由于经典K-means算法随机选取初始聚类中心且需要人为指定类簇数量k,会造成聚类质量不佳<sup>[35]</sup>,因此,本研究采用Canopy算法<sup>[36]</sup>选取初始聚类中心,并基于聚类有效性指标——误差平方和(SSE)采用“肘部法”<sup>[37]</sup>确定聚类数量k。计算从k=2到k=10时的SSE值如图2,可知k=5为图中的肘部点,说明将施引文献聚为5类时具有较好的聚类质量。根据聚类结果,本研究统计每一类中各项特征的平均值并绘制聚类结果雷达图,分析LIS学科中文期刊论文引文国际化的5种主要模式。

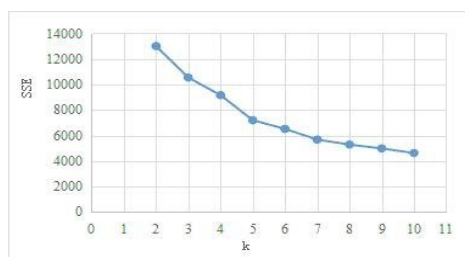


图2 K-means算法的误差平方和

## 3 结果与讨论

### 3.1 引文国际化的多维分析

#### 3.1.1 被引文献语种的时间序列分析

图3反映了1998年至2015年间各语种文献在被引文献中占比的逐年变化情况。

总的来说,外文文献占比呈逐年上升趋势。具体而言,中文与译文文献在被引文献中的占比逐年下降;英文文献的

占比则稳步提升,呈现出赶超中文文献的趋势,而德、日、法、俄等小语种文献的占比则长期处于极低位置。

中文文献占比的逐年降低说明越来越多的中国学者选择将外文文献作为其研究基础与知识来源,研究主题、理论与方法等逐步与国际接轨。译文文献的内容虽源自国外,但其由中文书写并在国内发行,51.9%的降幅说明通过初期CALIS等文献保障体系的建设<sup>[38]</sup>与后期对ScienceDirect等外文全文数据库的大量采购,使得本土学者获取外文原版文献的能力显著提升,同时也反映出了中国学者的外语水平不断提高,越来越多的学者有能力直接阅读外文文献。在被引的外文文献中,绝大多数(99.3%)为英文文献,这是由英文在学术交流中的绝对主导地位<sup>[39]</sup>、中国学者的语言能力和学术资源的可获取范围所共同决定的。

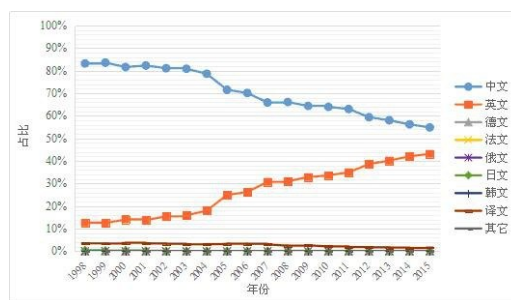


图3 被引文献中各语种占比时间序列图

#### 3.1.2 被引外文文献类型的分析

图4与图5展示了被引外文文献中各类型占比及其随时间的变化趋势。由图4可知,被引外文文献中期刊论文占比最高(47.9%),其余为是网络资源(29.0%)、会议论文(11.0%)、图书(9.8%)和学位论文(0.7%),报纸、档案、标准等其他类型文献占比约为2.6%。

由此可知,中国学者在阅读与利用外文文献时更倾向于选择正式的学术出版物。期刊论文相较于于会议论文经历了更为严格的同行评审,相较于图书又有着更为快速的出版周期,因此在各类正式的学术文献类型中占据最为重要的地位。Glänzel等<sup>[40]</sup>的研究发现LIS学科的期刊论文引用率约为47.6%,与本文结果(47.9%)基本一致,说明LIS领域的中国学者与国际同行在对期刊论文的利用上拥有一致的习惯与偏好。

图5中值得注意的是,随着时间变化,图书与会议论文的占比呈现出大致相反的发展趋势,图书占比明显减少(降幅为77.5%)而会议论文占比则显著增加(增幅达119.2%)。会议论文相较于图书具有更快的出版速度,其内容一般是尚未成熟的新兴研究课题或技术方法,因此也是工程技术领域(特别是计算机学科)重要的学术交流载体;图书具有较长的出版周期,其内容通常为领域内相对成熟的部分,是人文社科学者常用的学术交流工具<sup>[41-42]</sup>。二者占比发生的变化说明LIS领域的中国学者对新内容、新技术、新方法的接纳程度越来越高,更加注重对前沿技术的应用,隐含地体现出LIS学科由重人文向重技术转移的迹象。

外文正式学术文献的使用中存在的主要问题是学位论

文使用不充分。学位论文不仅提供了较为完善的知识体系,还包含了作者的系统性、原创性贡献,因此是全面深入了解某一领域的重要知识来源<sup>[43]</sup>。好在随着2002年ProQuest等学位论文全文数据库的建立<sup>[44]</sup>,一定程度上满足国内学者对外文学位论文的需求,学位论文的占比有所提升,增幅达170.9%。

除正式学术文献类型外,网络资源是中国学者的知识来源中不容忽视的重要组成部分。由图5可知,自1999年起,被引外文文献中网络资源的占比随中国互联网的蓬勃发展而快速增长,并一度在2002、2005和2006年成为占比最高的文献类型。网络资源具有更新及时、易于获取等特点,但去中心、不稳定、非权威等特点使得其质量层次不齐。因此,滥用网络资源将对研究的可靠性产生不利影响<sup>[45]</sup>。不过,随着学科建设不断完善,中国学者也逐渐意识到了网络资源所存在的问题,其占比自2005年达峰值后连续5年逐年递减,并在2011年后逐渐趋于稳定。

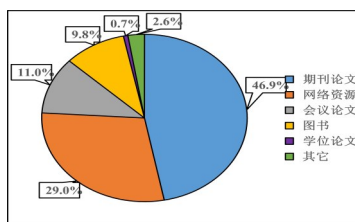


图4 被引外文文献中各类型占比饼图

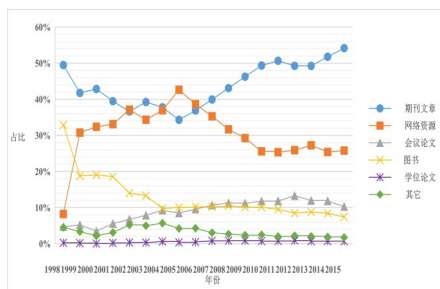


图5 被引外文文献中各类型占比时间序列图

### 3.1.3 被引外文文献的来源期刊

从1998年至2015年,总被引频次大于900的外文期刊共有12种,如表2所示。

其中被引用次数最多的期刊为《Scientometrics》,总被引频次达3402次。但将排名第二的期刊《Journal of the American Society for Information Science and Technology》(《JASIST》)与其更名前的《Journal of the American Society for Information Science》(《JASIS》)的数据合并,其成为被引用次数最多的外文期刊,总被引频次比《Scientometrics》多了近55.6%。由此可见,《JASIST》在中国LIS领域拥有极高的学术影响力,是国内学者最为重要的参考对象与知识来源。但有趣的是,在Scopus的学科分类体系中,《JASIST》与《JASIS》均不属于Library and Information Sciences,这显然有悖于中国LIS领域的引文证据,在一定程度上支持了Wang等<sup>[46]</sup>对Scopus期刊分类的批评,同时也说明了后文中在分析学科领域时将Scopus与WoS中的LIS期刊合并是合理且必要的。

在被引频次大于900的外文期刊中,多数为LIS领域内的期刊,不过也存在着《Strategic Management Journal》、《Research Policy》、《Communications of the ACM》和《Expert Systems with Applications》等来自于管理学与计算机科学的期刊。这说明中国的LIS研究与管理学和计算机科学存在着较多的交叉与融合,这也为前文中会议论文占比较高提供了合理的解释。

通过图6可知,在可识别的4869种外文期刊中,极少数期刊获得了大多数引用,排名前20%的期刊,其被引频次(79113)占据了总被引(88268)的89.6%,表明国内学者对外文期刊的利用很大程度上遵循了“最省力原则”<sup>[47]</sup>,趋向于使用一定范围内熟悉且易获取的外文期刊作为知识来源,这可为我国的外文学术资源保障体系的建设提供一定借鉴与参考。

表2 被引频次大于900的外文期刊

序号	期刊名称	被引频次
1	Scientometrics	3402
2	Journal of the American Society for Information Science and Technology	3041
3	Journal of the American Society for Information Science	2252
4	Information Processing and Management	1845
5	Journal of Documentation	1588
6	Strategic Management Journal	1298
7	MIS Quarterly	1271
8	Communications of the ACM	1055
9	Journal of Academic Librarianship	1022
10	Journal of Information Science	922
11	Expert Systems with Applications	907
12	Research Policy	900

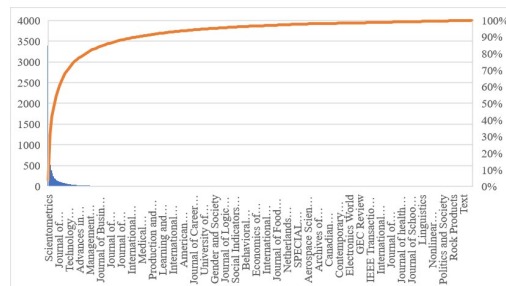


图6 外文期刊被引频次的分布

### 3.1.4 被引外文文献的学科领域

目前,研究问题的复杂性已使得跨学科成为科研与教育进步的代名词<sup>[48]</sup>,与此同时,科学发展的历史也一再表明,重大的科学突破与颠覆性的研究成果往往都离不开学科的交叉与融合,跨学科研究的广度与深度已经成为影响创新的重要因素<sup>[49]</sup>。已有大量研究均表明,LIS学科具有很强的跨学科性,其论文广泛引用其他学科的文献便是重要证据之一<sup>[50]</sup>。

经本研究统计,在Scopus划分的334个学科领域中,中文LIS期刊论文引用了其中的313个,覆盖范围极广。由表3可知,被引频次占比高于2%的学科领域共13个。其中,被引最多的是LIS学科本身,占比为15.36%,此外,信息系统、

计算机科学、管理科学与商学也是外文参考文献所涉及的主要学科,说明我国的 LIS 研究与上述学科交叉频繁、融合较深。进一步提示,通过学习、吸收上述多个学科的国际优秀成果,我国较可能在相关的交叉领域产生较大的发展与突破,其也将是我国 LIS 学科未来更好地走出去的可靠着力点与发力点。

表3 被引频次占比大于2%的学科领域

序号	学科领域	被引频次	占比
1	Library and Information Sciences	34924	15.36%
2	Information Systems	18719	8.23%
3	Computer Science Applications	14131	6.22%
4	Strategy and Management	10826	4.76%
5	Software	7970	3.51%
6	Artificial Intelligence	7878	3.47%
7	Management of Technology and Innovation	7361	3.24%
8	Computer Networks and Communications	6859	3.02%
9	Business and International Management	6594	2.90%
10	General Engineering	5738	2.52%
11	Information Systems and Management	5565	2.45%
12	Management Science and Operations Research	5238	2.30%
13	Human-Computer Interaction	5093	2.24%

图7和图8展示了被引外文文献中 LIS 学科专属度与跨学科度随着时间的变化情况。从图7可知, LIS 学科专属度在大多数年份小于50%, 总体趋势呈现出先下降后上升的趋势。图8中, 被引外文文献的跨学科度自1998年起逐年提升并在2011年前后趋于稳定。图7与图8进一步证实了国内 LIS 领域具有很强的跨学科性, 跨学科程度稳中有升。这一方面说明了本土研究的复杂程度在不断增强, 要求国内学者越来越多地借鉴国际上其他学科的先进理论与方法; 同时也预示着中国的 LIS 学科在与其它学科的不断交叉中很有可能爆发出新的增长点, 展现出良好的发展势头。



图7 被引外文文献学科专属度的时间序列图

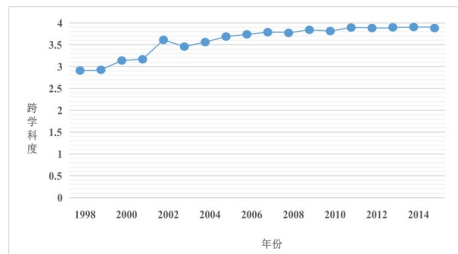


图8 被引外文文献跨学科度的时间序列图

对比图7与图8, 值得注意的是在2011年前, 学科专属度的下降往往伴随着跨学科度的提升; 但在2011年之后, 学科专属度的上升并未导致跨学科度产生较大波动。前者产

生的原因易于理解, 在外文文献被引总量不断增加的背景下, 本学科被引的减少势必会伴随着其他学科被引的增加。后者则值得思考, 其出现的可能原因之一是由于 LIS 学科本身所具有的跨学科性使得大多数 LIS 期刊被归入了多个学科, 隶属学科的多寡又由其发表论文的学科多样性所决定, 综合性 LIS 期刊显然比某一方向的 LIS 期刊拥有更多的学科归属。由于本研究在计算学科专属度时仅考虑被引期刊是否属于 LIS 学科, 而在计算跨学科度时依据隶属学科的数量。因此, 2011年后学科专属度与跨学科度不再同步变化可能是由于中文 LIS 论文越来越多地引用了跨学科程度高的综合性 LIS 期刊。

### 3.1.5 被引外文文献的学术质量

图9反映了被引外文文献的学术质量随着时间的变化情况。总体而言, 被引外文文献的学术质量呈上升趋势, 年均 JIFP 由初期的不足40%逐步增长到50%左右。这说明中国学者越来越注重外文文献的学术价值, 更多地阅读和引用了高水平期刊上发表的论文。Lancho-Barrantes 等<sup>[51]</sup>和 Bornmann 等<sup>[52]</sup>的研究表明: 一篇论文的被引频次与其参考文献的引文影响力有关, 参考文献的引文影响力越高, 施引文献的后续影响力也将更高。姜春林等<sup>[53]</sup>在我国 LIS 领域的实证研究同样证实了上述结论。正如牛顿所说: “我看得之所以比别人远是因为我站在巨人的肩膀上”, 高水平的学术研究同样建立在大量高水平的前期研究之上。因此, 从图9可以预见: 本土 LIS 学者通过“站在巨人的肩膀上”, 不断吸收和内化大量国际高水平研究, 必将带动研究水平的不断提高, 提升本土研究的质量。

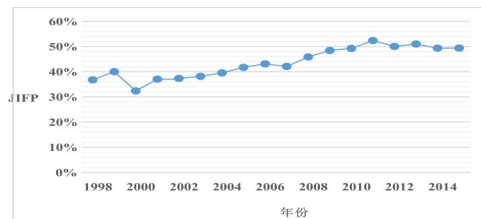


图9 被引外文文献平均学术质量的时间序列图

### 3.1.6 被引外文文献的时效性

从图10中可以发现, 被引外文文献的平均年龄与引用半衰期呈现较为一致的变化规律。二者均从1998年起逐年递减并在2001年前后达到最低点, 之后二者持续上升并于2011年达到最高点后趋于稳定。十八年间, 被引外文文献的平均年龄始终大于引用半衰期, 说明被引外文文献中存在较多发表时间久远的经典文献, 使年差呈偏态分布。

与此同时, 两项指标在2001年之后呈现出的较为明显的增长趋势意味着中文 LIS 论文越来越多地引用了发表时间较久的外文文献。这看起来似乎有违常识, 毕竟多年前就有学者提出, 科学发展的步伐加快会使得科学文献老化的速度变快<sup>[54-56]</sup>。特别是在如今的互联网时代, 各类全文数据库、开放存取期刊、在线出版等数字化方式使得科研人员更容易获取最新的研究成果, 越来越大的引文年差就显得更加违背直觉。不过, Larivière 等<sup>[57]</sup>和 Wu 等<sup>[58]</sup>的研究同样得出



了本研究相似的结果,前者发现自上世纪60年代中期以来,被引用素材的年龄持续增长;后者则证实了学者在互联网普及之后引用了更多的旧文献。至于这一现象产生的原因,可能是由于包含了对期刊进行回顾性收集的在线书目工具经历了爆发式增长,这些工具能够帮助研究人员获取更多的旧文献,相应的,他们也就能够更加频繁地引用这些旧文献<sup>[59]</sup>。图10中的拐点出现在2001年前后,正是中国互联网开始迅速普及且外文电子学术资源建设开始蓬勃发展的关键时期,因此更加可以确定:互联网与外文电子学术资源的发展是2001年后被引外文文献年龄不断增长的重要原因。与此同时,本土研究与国际前沿存在的客观差距也可能造成中国学者频繁引用经典文献。

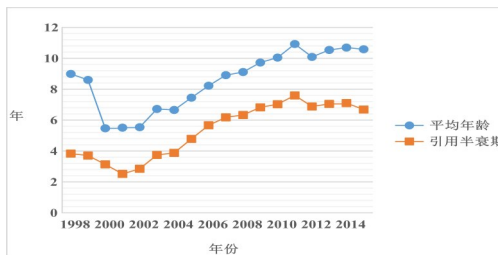


图10 引用半衰期与被引外文文献的平均年龄时间序列图

### 3.2 引文国际化的五种常见模式

图11直观地反映了中文LIS期刊论文引文国际化的五种主要模式,五个类簇在各项指标中的分布特点鲜明。

44395篇施引文献被聚入类簇1,占全部中文LIS期刊论文的65.6%,但该簇中的各项引文国际化指标都非常低,说明极低的引文国际化水平是中文LIS期刊论文中较为普遍的引文模式。

类簇2中有5835篇施引文献,占总量的8.6%。各项指标中,外文参考文献占比很高,达69.9%,但外文期刊论文占比、LIS学科专属度、学术质量和新文献占比都很低。由于后4个指标均基于外文期刊论文测度,说明类簇2中的论文大量引用了非期刊类的外文文献。

类簇3、4、5中分别有5825、7880和3789篇施引文献,分别占总量的8.6%、11.6%和5.6%。它们有着较为相近的外文参考文献占比(约50%),被引外文文献中的期刊论文比例也大体一致(约69%)。三者的差异主要出现在被引外文期刊论文的LIS学科专属度、学术质量和新文献占比中。

类簇3是三者中新文献占比最低的一类,仅为11.2%,其被引外文期刊论文的质量也较为一般(38.1%),LIS学科专属度约为54%,说明类簇3中施引文献的引文国际化水平尚可,其引用了较多质量一般的较为老旧的期刊论文,被引文献没有明显的跨学科特征。

被类簇4中文献引用的外文期刊论文具有最高的新文献占比(81.9%)和较高的学术质量(46.8%),其与类簇5的主要差异在于其LIS学科专属度很低,仅为5.7%,说明该簇引用的外文文献具有较高的跨学科性。因此,综合5项指标认为,类簇4中的施引文献具有较高的引文国际化水平,其跨

学科引用了大量较新的高质量期刊论文。

被类簇5引用的外文期刊论文具有明显高于其它类簇的LIS学科专属度(82.5%)和最高的学术质量(52.4%),新文献占比与类簇4大致相同。因此可以认为,类簇5中的施引文献具有较高的引文国际化水平,其引用了大量LIS学科的较新的高质量期刊论文。

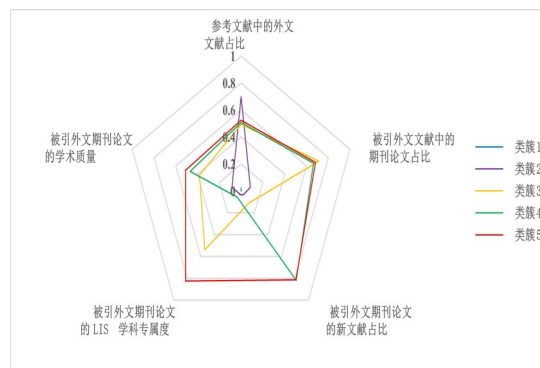


图11 基于引文国际化指标的聚类结果

## 4 结论与展望

本研究选取CSSCI建库以来每年均收录的16本LIS学科的来源期刊数据,通过对中文期刊论文引用外文文献的语种、文献类型、来源期刊、学科领域、学术质量、时效性进行多维分析,揭示本土学者在外文文献选择与内化的过程中所经历的发展与变化。研究结果表明:

- ①被引文献中,外文文献占比逐年上升,其中英语占绝对主导地位;
- ②被引外文文献中期刊论文占比最高,其次是网络资源、会议论文、图书和学位论文;
- ③被引频次较高的外文期刊中,多数来自LIS领域,但也有其他学科的期刊,较少数期刊获得了大多数引用;
- ④除LIS学科外,信息系统、计算机科学、管理科学与商业的外文文献被引较多,跨学科程度稳中有升;
- ⑤被引外文文献的学术质量呈上升趋势;
- ⑥被引外文文献的年龄呈现出先降低后升高的变化趋势。

此外,本研究还基于引文国际化的多维指标对施引文献聚类,揭示出了LIS学科中文期刊论文引用外文文献的五种主要模式。其中,极低的引文国际化水平是较为普遍的引文模式,大量引用非期刊类外文文献是第二种引文模式,大量引用LIS学科较新的高质量期刊论文是第三种模式,跨学科引用大量较新的高质量期刊论文为第四种模式,较多引用质量一般的没有明显跨学科特征的老旧期刊论文为第五种模式。

以本研究为基础,未来还将展开一系列相关研究,包括探索影响引文国际化的相关因素,探究论文的引文国际化水平是否会对学术研究的质量产生影响等等,进而试图更好地了解与阐明引文国际化的价值与意义。

## 参考文献

- 1 龚凯乐, 谢 娟, 成 颖, 等. 期刊论文引文国际化研究——以图书情报与档案管理学科为例[J]. 情报学报, 2018, 37(2): 151-160.
- 2 赵 星, 谭 旻, 余小萍, 等. 我国文科领域知识扩散之引文网络探析[J]. 中国图书馆学报, 2012, 38(5): 59-67.
- 3 Li L. Analysis of factors restricting the development of science and technology journals in China[J]. Publishing Research Quarterly, 2004, 19(4): 45-53.
- 4 Xu L, Fang Q. Internationalization of Chinese STM Journal Publishing[J]. Publishing Research Quarterly, 2013, 29(2): 190-196.
- 5 Zou Y, Laubichler M D. Measuring the contributions of Chinese scholars to the research field of systems biology from 2005 to 2013[J]. Scientometrics, 2017, 110(3): 1-17.
- 6 Moed H F. Measuring China's research performance using the Science Citation Index[J]. Scientometrics, 2002, 53(3): 281-296.
- 7 Wang S, Wang H, Weldon P R. Bibliometric analysis of English-language academic journals of China and their internationalization[J]. Scientometrics, 2007, 73(3): 331-343.
- 8 Ren S, Rousseau R. International visibility of Chinese scientific journals[J]. Scientometrics, 2002, 53(3): 389-405.
- 9 Li J, Qiao L, Li W, et al. Chinese-language articles are not biased in citations: Evidences from Chinese-English bilingual journals in Scopus and Web of Science[J]. Journal of Informetrics, 2014, 8(4): 912-916.
- 10 梁永霞, 刘则渊, 杨中楷. 引文分析学的知识流动理论探析[J]. 科学学研究, 2010, 28(5): 668-674.
- 11 Lockett A, McWilliams A. The Balance of Trade Between Disciplines: Do We Effectively Manage Knowledge?[J]. Journal of Management Inquiry, 2005, 14(2): 139-150.
- 12 Wu C, Yan E, Hill C. Disciplinary knowledge diffusion in business research[J]. Journal of Informetrics, 2017, 11(2): 655-668.
- 13 Yan E, Ding Y, Cronin B, et al. A bird's-eye view of scientific trading: Dependency relations among fields of science[J]. Journal of Informetrics, 2013, 7(2): 249-264.
- 14 Yan E. Disciplinary knowledge production and diffusion in science[J]. Journal of the Association for Information Science & Technology, 2015, 67(9): 2223-2245.
- 15 Narin F, Hamilton K S, Olivastro D. The increasing linkage between U.S. technology and public science[J]. Research Policy, 1997, 26(3): 317-330.
- 16 Callaert J, Grouwels J, Looy B. Delineating the scientific footprint in technology: Identifying scientific publications within non-patent references[J]. Scientometrics, 2012, 91(2): 383-398.
- 17 Su X, Deng S, Shen S. The design and application value of the Chinese Social Science Citation Index[J]. Scientometrics, 2012, 98(3): 1567-1582.
- 18 Wang H, Deng S, Su X. A study on construction and analysis of discipline knowledge structure of Chinese LIS based on CSSCI[J]. Scientometrics, 2016, 109(3): 1-35.
- 19 王 军. 中外十种外科学期刊文后参考文献著录差错对比分析[J]. 情报科学, 2008, 26(1): 101-103.
- 20 朱红梅, 张大志, 任 红. 对我国高影响力医学期刊的中、英文参考文献著录差错分析[J]. 中国科技期刊研究, 2012, 23(5): 784-787.
- 21 Norris M, Oppenheim C. Comparing alternatives to the Web of Science for coverage of the social sciences' literature[J]. Journal of Informetrics, 2007, 1(2): 161-169.
- 22 López-Illescas C, Moya-Anegón F D, Moed H F. Coverage and citation impact of oncological journals in the Web of Science, and Scopus[J]. Journal of Informetrics, 2008, 2(4): 304-316.
- 23 Navarro G. A guided tour to approximate string matching[J]. ACM Computing Surveys, 2001, 33(1): 31-88.
- 24 Leydesdorff L, Rafols I. Indicators of the interdisciplinarity of journals: Diversity, centrality, and citations[J]. Journal of Informetrics, 2011, 5(1): 87-100.
- 25 Wu C, Yan E, Hill C. Disciplinary knowledge diffusion in business research[J]. Journal of Informetrics, 2017, 11(2): 655-668.
- 26 Tahamtan I, Afshar A S, Ahmndzadeh K. Factors affecting number of citations: a comprehensive review of the literature[J]. Scientometrics, 2016, 107(3): 1195-1225.
- 27 Ma N, Guan J, Zhao Y, et al. Bringing PageRank to the citation analysis[J]. Information Processing and Management, 2008, 44(2): 800-810.
- 28 谢 娟, 龚凯乐, 成 颖, 等. 论文下载量与被引相关关系的元分析[J]. 情报学报, 2017, 36(12): 1255-1269.
- 29 Bornmann L. Alternative metrics in scientometrics: a meta-analysis of research into three altmetrics[J]. Scientometrics, 2015, 103(3): 1123-1144.
- 30 Abramo G, D' Angelo C A, Costa F D. Citations versus journal impact factor as proxy of quality: could the latter ever be preferable?[J]. Scientometrics, 2010, 84(3): 821-833.
- 31 Dorta-Gonz, Lez P, Dorta-Gonz, et al. Comparing journals from different fields of Science and Social Science through a JCR Subject Categories Normalized Impact Factor[J]. Scientometrics, 2013, 95(2): 645-672.
- 32 Yu L, Yu H. Does the average JIF percentile make a difference?[J]. Scientometrics, 2016, 109(3): 1979-1987.



- 33 Tsay M Y. An analysis and comparison of scientometric data between journals of physics, chemistry and engineering[J]. *Scientometrics*, 2009, 78(2):279-293.
- 34 MacQueen J. Some methods for classification and analysis of multivariate observations[C]//Proceedings of the fifth Berkeley symposium on mathematical statistics and probability. 1967, 1(14): 281-297.
- 35 龚凯乐, 成颖, 孙建军. 基于参与者共现分析的博文聚类研究[J]. *现代图书情报技术*, 2016, 32(10):50-58.
- 36 McCallum A. Efficient clustering of high-dimensional data sets with application to reference matching[C]// International Conference on Knowledge Discovery and Data Mining. DBLP, 2000:169-178.
- 37 Raschka S. *Python Machine Learning*[M]. UK:Packt Publishing, 2014: 320-321.
- 38 Han S, Shen C, Shen X, et al. A review of the major projects constituting the China Academic Digital Library[J]. *The Electronic Library*, 2008, 26(1):39-54.
- 39 Meneghini R, Packer A L. Is there science beyond English? Initiatives to increase the quality and visibility of non-English publications might help to break down language barriers in scientific communication[J]. *Embo Reports*, 2007, 8(8): 112-116.
- 40 Glänzel W, Schoepflin U. A bibliometric study of reference literature in the sciences and social sciences [J]. *Information Processing & Management*, 1999, 35(1):31-44.
- 41 Earle P, Vickery B. Social science literature use in the UK as indicated by citations[J]. *Journal of Documentation*, 1969, 25(2):123-141.
- 42 Bourke P, Butler L. Publication types, citation rates and evaluation[J]. *Scientometrics*, 1996, 37(3):473-494.
- 43 Zong Q J, Shen H Z, Yuan Q J, et al. Doctoral Dissertations of Library and Information Science in China: A Co-word Analysis[J]. *Scientometrics*, 2013, 94(2):781-799.
- 44 鲁程. PQDD与ProQuest学位论文全文数据库的比较评析[J]. *图书馆论坛*, 2005, 25(1):85-87.
- 45 Chen C, Luo B, Chiu K, et al. The preferences of authors of Chinese library and information science journal articles in citing Internet sources[J]. *Library & Information Science Research*, 2014, 36(3-4):163-170.
- 46 Wang Q, Waltman L. Large-Scale Analysis of the Accuracy of the Journal Classification Systems of Web of Science and Scopus[J]. *Journal of Informetrics*, 2015, 10(2):347-364.
- 47 Zipf G K. *Human behavior and the principle of least effort* [M]. NJ:Addison-Wesley Press, 1949.
- 48 Rhoten D R, Parker A. Risks and Rewards of an Interdisciplinary Research Path[J]. *Science*, 2004, 306(5704): 2046-2046.
- 49 黄颖, 高天舒, 王志楠, 等. 基于Web of Science分类的跨学科测度研究[J]. *科研管理*, 2016, 37(3):124-132.
- 50 Chang Y W, Huang M H. A study of the evolution of interdisciplinarity in library and information science: Using three bibliometric methods[J]. *Journal of the Association for Information Science & Technology*, 2012, 63(1):22-33.
- 51 Llancho-Barrantes B S, Guerrero-Bote V P, Moya-Anegón F. What lies behind the averages and significance of citation indicators in different disciplines? [J]. *Journal of Information Science*, 2010, 36(3):371-382.
- 52 Bornmann L, Schier H, Marx W, et al. What factors determine citation counts of publications in chemistry besides their quality? [J]. *Journal of Informetrics*, 2012, 6(1):11-18.
- 53 姜春林, 刘学, 张立伟. 中文高被引期刊论文的参考文献也倾向于高被引吗?——以图书情报学为例[J]. *情报杂志*, 2015, 34(1):105-109.
- 54 Line, M B. The half-life of periodical literature: Apparent and real obsolescence[J]. *Journal of Documentation*, 1970, 26(1):46-54.
- 55 Line M B. Changes in the Use of Literature with Time—Obsolescence Revisited[J]. *Library Trends*, 1993, 41(4):665-683.
- 56 Price D J. Networks of Scientific Papers[J]. *Science*, 1965, 149(3683): 510-515.
- 57 Larivière V, Archambault É, Gingras Y. Long term variations in the aging of scientific literature: From exponential growth to steady state science (1900 - 2004) [J]. *Journal of the Association for Information Science & Technology*, 2008, 59(2):288-296.
- 58 Wu L, Huang M, Chen C, et al. Citation patterns of the pre-web and web-prevalent environments: The moderating effects of domain knowledge[J]. *Journal of the Association for Information Science and Technology*, 2012, 63(11): 2182-2194.
- 59 Boyce P B, King D W, Montgomery C H, et al. How Electronic Journals Are Changing Patterns of Use[J]. *Serials Librarian*, 2004, 46(1): 121-141.

(责任编辑:张连峰)