

南京大学图书馆电子资源学术信息检索系统研究*

成颖 孙建军(南京大学信息管理系 南京 210093)

张敏(安徽大学管理学院 合肥 230039)

文 摘 对51个中外文数据库,分别就更新周期、同行评议、服务功能、服务方式、帮助、个性化、激励、检索方式、浏览方式/期刊导航、检索字段、检索技术、查全率/查准率、相关反馈、排序/分组、格式/展示、结果输出、项标识/标记方式、分类/主题描述、超链接以及整合检索等进行了调研。结论显示:①中外文数据库的系统特征相似,中文数据库稍显不足;②学术数据库检索系统需作如下努力:进一步改进检索系统输出层面的分析与设计;③扩大整合检索的应用范围;④提高个性化服务的深度与广度;⑤提供增值服务。

关键词 学术信息检索系统 学术数据库 比较研究 调查研究

Investigation and Analysis of Academic Information Retrieval System Based on Electronic Resources of Nanjing University Library

Cheng Ying, Sun Jianjun

(Department of Information Management, Nanjing University, Nanjing 210093)

Zhang Min(School of Management, Anhui University, Hefei 230039)

Abstract: In this paper, the elements of 51 Chinese and English academic information retrieval systems are investigated and analyzed, respectively from aspects of update cycle, peer review, service types, service mode, help, personalized service, stimulation, retrieval methods, browsing/navigation, search fields, search techniques, recall/precision, relevance feedback, sorting/grouping, format/display, output of result, item identification/mark, classification/subject description, hyperlinks, and meta retrieval etc. The results reveal that ① the elements of Chinese and English academic information retrieval systems are similar, but Chinese database need to do more; ② improving the analysis and design of output of retrieval systems; ③ expanding the application of meta retrieval; ④ extending the depth and breadth of personalized service; ⑤ providing value-added services.

Key words: Academic information retrieval system, Academic database, Comparative analysis, Survey

1 引言

学术信息检索系统为人们从事科学研究提供了

便利条件。它将各种学术信息整合为一个高质量的信息平台,科研人员能利用网络迅速有效地获取权

* 本文得到国家社科基金项目“中文学术信息检索系统相关性集成研究”,项目批准号 10CTQ027;教育部人文社会科学研究规划基金项目“面向用户的相关性标准及其应用研究”,项目批准号 07JA870006;中国科学技术信息研究所合作研究项目的资助。

威、及时的学术科研信息,广泛地了解学界和业界的发展动态,科学地评估学科的发展趋势,预测学术研究方向。因而,学术信息检索系统的建设是关系到科学研究的重要基础性问题,然而笔者通过文献调研发现,大多数学者在关注此问题时提出的建议都较为宏观,比如“信息资源收录的学科‘专业化’”、“检索功能和检索界面的优化”、“系统帮助及全文资料保存的便捷化”^[1]等。如果要在微观层面为提高网络信息资源以及学术信息资源的利用效率提供具体的建议,则有必要在现有注重结构模型研究的基础上^[2-5],加强对影响学术信息资源的前置变量的研究,因此就需要对现有学术信息检索系统的系统元素进行系统的调研。本文在此背景下,注重对国内外常用的学术检索系统开展系统的调研,通过对比分析从中找出我国学术信息检索系统存在的主要问题,以及提出具有针对性的建议。

2 调查设计

2.1 调查对象

本文通过南京大学图书馆的网上资源对主要的网络学术信息资源数据库的系统特征进行了调查。调查的对象主要包括南京大学图书馆网站上的51个网络资源数据库(中文数据库10个,外文数据库41个)。

中文数据库:万方、CNKI、维普、超星、CADAL、CSSCI、人大复印资料、南大人文库、南大民国图书、南大民国期刊。

外文数据库:AAAS Science Online、ABI ProQuest、ACM、ACS Publications、AGU AmericanGeoUnion、AIP Scitation、AMS 气象、AMS 数学、APS、CALIS 中心、CALIS 中心 CCC 外文期刊、CUP Cambridge Journal、EBSCOhost、Elsevier SDOS、GSW GeoScienceWorld、IEEE/IET IEEEExplore、Ingenta、Ingram MyiLibrary、IOP institute of physics、ISI MEDLINE、ISI Web of Science、JSTOR、NAS PNSA、Nature、NCBI PubMed、NSTL、NSTL American society of nutrition、OUP Oxford University Press、POA Periodical Archive Online、PQDT 中科、RSC、SAGE Journals Online、SIAM Journal online、Springer、Zentralblatt MATH、SpringerLink、SwetsWise Linker、Taylor & Francis、THIEME、TomsonGale ECCO、Wiley - Blackwell WileyOnlineLibrary、World Scientific。

2.2 调查内容及方式

由于各个数据库的系统特征各有特色,本文的调查主要围绕以下内容展开:更新周期、同行评议、服务功能、服务方式、帮助服务、个性化服务、激励、

检索方式、浏览方式/期刊导航、检索字段、检索技术、查全率/查准率(获得文献的扩展信息)、相关反馈、排序/分组、格式/展示(检索结果)、结果输出、项标识/标记方式、分类/主题描述(检索结果分析)、超链接以及整合检索等19个指标。

本调查采用开放的、探索性的以及迭代的方式进行,即对于每个调查内容当某个学术数据库中出現新的元素之后,则将其纳入调查表中,然后再就该特征进行回溯,筛查先前是否有遗漏。调查的信息源包括从界面上直接可以观察到的内容以及从系统帮助中可以获得的内容,除此之外不能获得的信息则予以排除。

3 调查结果

3.1 更新周期和同行评议

学术数据库的更新周期主要有日更新、周更新两种,不过大部分数据库都没有在站点上说明该特征,调查中获得的信息包括中文的 CNKI 和维普是每日更新,西文的 IEEE/IET IEEEExplore 和 CALIS 中心的 CCC 外文期刊是每周更新,西文的 NAS PNSA 数据库是每日更新。

调查还发现中文学术数据库没有提到同行评议的问题,而外文的 ABI ProQuest、ACS Publications、AIP Scitation、AGU AmericanGeoUnion、GSW GeoScienceWorld、AMS 气象、AMS 数学、CUP Cambridge Journal、Taylor & Francis 等9个数据库提到了其数据库中的文献是经过同行评议的。

3.2 服务功能

针对服务功能,分别调查了离线配套服务、全文提供服务、文献传递服务、人员培训、MARC 数据管理以及 OCR 识别功能等子项。调查可以发现,几乎所有的数据库都没有提供离线服务、人员培训以及 MARC 数据管理等服务;中文数据库除了 CSSCI 是引文数据库之外都提供了全文服务,而外文除了 PubMed 等6个数据库之外都提供了全文服务;OCR 识别主要和阅读器有关,统计表中的中文数据库有两个,CNKI 的 CAJ 阅读器和维普的 VIP 阅读器提供了 OCR 识别功能,由于这两个数据库也都提供了 PDF 格式的全文,而 Acrobat 7.0 版本以上的阅读器都提供了 OCR 功能。因此可以认为现在的西文文献都提供了 OCR 识别功能。

3.3 存取/服务方式

调查发现机构卡、流量计费以及阅读卡等服务模式主要是中文的几个学术检索系统在使用,中文检索系统对于个人用户提供了包括会员卡、充值卡

等丰富的费用支付方式,而外文检索系统仅仅 NSTL 提供了相似的资费支付的方式,所有的数据库主要通过包库模式和镜像模式提供服务。由于数据是根据南京大学图书馆获得,可以推测外文数据库在国内的营销对象主要是图书馆等集团用户,所以主要面向个人用户的会员卡以及阅读卡等方式没有提供。

3.4 帮助服务

大多数检索系统都提供了新手指南/新手帮助等内容以利于新信息用户很快入门,除了主要采用文本以及图表之外,部分检索系统还提供了视频方式的帮助。

相较于商务型网站提供的丰富的在线咨询服

务,学术型数据库提供的在线咨询就弱了很多,中外文数据库分别只有 CNKI 和 NSTL 提供了在线咨询的帮助方式,其他都仅仅提供了电话以及电子邮件等联系方式供信息用户联系之用,帮助的实时性非常弱。

3.5 个性化服务

从表 1 可以看出,学术信息检索系统提供了丰富的个性化服务方式,其中推送服务在中外文数据库中都没有提供。近一半的中外学术数据库检索系统提供多语言平台,而笔者认为检索系统中的术语通常是较常见的单词,对于能读懂外文文献的用户而言,外文界面基本上不会成为障碍,因而其存在必要性有待商榷。

表 1 学术数据库的个性化服务

个性化服务	中文数据库 频次	外文数据库 频次	合计	个性化服务	中文数据库 频次	外文数据库 频次	合计
我的电子书架	6	25	31	我的主页	1	1	2
存储检索记录	3	25	28	用户界面定制	0	2	2
期刊定制	1	26	27	关键词定制	1	0	1
邮件定题服务	0	24	24	推送服务	0	0	0
多语言平台	5	16	21	论文引用提醒	0	11	11
删除检索式	3	10	13	论文内容更正提醒	0	5	5
创建个人账户	5	6	11	将当前文献 email 给朋友服务	0	21	21
存储检索式	2	5	7	期刊新一期出版提醒	0	3	3
存取删除检索史	0	5	5	请求权限	0	14	14
分类定制	1	2	3	重印要求	0	6	6
个人期刊列表	0	3	3				

表 1 的结果表明中文维普和 CALIS 中心提供了个人主页服务,这两个数据库都是国内产品,而国外学术数据库基本上没有提供个人主页服务。该服务属于学术数据库的衍生服务,其对学术数据库的意义是否重大需要深入探讨。其他的服务诸如关键词定制、分类定制、用户界面定制以及个人期刊列表等各数据库提供的比例也偏小,大部分数据库提供的定制服务主要是期刊定制和邮件定题服务,而中文数据库没有提供邮件定题服务。结合国外信息行为研究成果比较丰富的现状,可以认为,这两种面向内容的定制方式更代表了信息用户的行为。而关键词定制以及分类定制粒度偏大,所以提供这两类定制的系统非常少。个人期刊列表功能相对简单,深入讨论意义不大。

表 1 中的存储检索式、存储检索记录、存取删除检索史以及删除检索式等个性化服务对于用户了解

自己的检索历史及之前的文献情况有帮助,通过这些内容可分析用户研究兴趣、信息行为等特征,对改进检索系统性能及提供更适合用户的个性化服务能起到积极的作用。

在同等条件下,检索系统提供的个性化服务的内容以及质量将对用户采纳和接受检索系统有影响。不过,现有的学术信息检索系统还处在一个几乎没有竞争或竞争不充分的阶段,表 1 反映出大部分学术信息检索系统在个性化服务方面还没有投入足够的精力。

表 2 后面 6 项个性化服务方式中文数据库都没有提供,引起学者关注的问题是这些个性化服务方式到底是意义甚微呢,还是国内没有意识到其重要性。

3.6 导航

科学、合理的导航方式能够引导信息用户有效

地利用信息,使其在信息检索过程中不至于迷航,并能够以最佳的方式获得所需信息。目前,学术信息检索系统的导航方式主要有导航菜单、快速链接、面包屑导航、站内检索以及网站地图等。调查显示所有的中外文学术数据库都提供了导航菜单、快速链接导航方式,几乎所有的学术数据库都没有提供站内检索导航,部分学术检索系统提供了网站地图以及面包屑导航。导航方式对于提高系统的易用性有着非常重要的影响,已经成为影响系统采纳与接受的非常重要的因素。

3.7 检索方式

调查发现中外文数据库基本上都提供简单/快速检索以及高级检索方式,也提供自然语言与受控语言检索方式。部分检索系统提供了特色的检索方式,比如化学文摘提供化学结构式检索。而有些系统提供的检索方式较为冗余,比如万方、CNKI 提供的句子检索、作者发文检索、基金检索等,实际上这些功能在简单检索或高级检索中都已经实现,如此做法看似增加多种检索途径,实际上却增加界面的复杂性。外文数据库中文献主要收录的是西文文献,而在中文数据库中存在部分外文文献的现象,因此中文数据库中提供了跨语言检索的检索方式。随着引文分析技术的成熟,来源文献检索逐渐成为重要的检索方式,对于检索字面上不相关而实际相关的文献起到非常重要的作用。专家检索方式对于普通信息用户而言,显然有难度,实际上对于需要使用专家检索方式的信息用户来说,基本上可以通过高级检索实现其检索意图。因此中文数据库只有 3 个提供了专家检索方式,而外文也仅有 6 个数据库提供该方式。跨库检索对于一个数据库提供商有多个数据库的情况还是非常必要,避免了信息用户不断地打开/关闭多个数据库、输入相同检索表达式的操作,节约了用户的时间。

总之,面对当前用户逐渐适应类似 Google 这种简洁界面的潮流,也给学术检索系统的界面设计带来挑战,一些冗余的检索方式是否有必要放在界面上需要关注。此外 EBSCO host 提供的视觉检索,虽然有增加可视化的意图,但是从笔者的使用体验来看,并不比文本方式更具吸引力,反而存在响应速度明显变慢的突出缺点,大规模推广的前景不容乐观。

3.8 浏览方式/期刊导航

调查显示出版物/期刊的浏览方式中最主要的还是按照出版物的名称、学科、专辑或者主题类别的方式。一些有中国特色的浏览方式,比如 CNKI 提

供的根据优先出版期刊、世纪期刊、数据库刊源、期刊荣誉榜以及不同级别的期刊(比如 SCI,SSCI 等)等途径进行浏览。此外,调查还发现中文数据库很少提供按照出版商、受控词表、作者单位、出版时间、出版物类型进行浏览,笔者分析这可能与中国的出版机制有关,以及数据库建设单位考虑有些概念具有很强的专业性,普通用户难以理解等原因。

3.9 检索字段/限制条件

在相关性判断中,几乎所有的文献特征都被作为相关性判断的依据,通过对检索系统检索字段的调研,发现几乎所有的文献特征都成了检索字段或者限制条件,从各系统所用的检索字段的统计可以看出(见后页表 2):

(1) 常用检索字段。作者、出版时间、题名、出版物名称、文摘、全文、匹配方式、关键词、文献编号、ISSN 等是中外文学术数据库主要的检索字段。

(2) 英文特色。主题词、页数/封面报导/附带图像的文章、叙词、系列名称、物种、性别、年龄组、文献编号、团体著者、同行评议期刊、人员、评论、会议地点、过滤、分主题数据库、段落以及标引词等分别只出现在不同外文数据库中作为检索字段。

(3) 中文特色。作者简介、英文题目、同义词、同名作者、期刊来源类别、期刊级别、栏目信息、词频以及被引用次数等在外文数据库中都没有作为检索字段。

(4) 中英文都没有出现。例如在中英文数据库中,书籍类型/适用对象都没有出现。

(5) 中英文在文献处理深度上存在一定差别。相比之下,中文数据库加工层次较浅,如作者简介、英文题目、同义词、同名作者等信息都可直接从文献中获取。反观外文数据库的特有检索字段,如主题词、页数/封面报导/附带图像的文章、叙词、系列名称、物种、性别、年龄组等,需要更多认知投入才能完成。通常认为信息用户的认知投入和检索系统的加工处理二者是此消彼长的关系,数据库服务商的认知投入更多,意味着用户的开销可以减少。

3.10 检索技术

从表 3 可以发现,布尔检索以及组配检索是中外文数据库的主要检索技术,其他还分别提供了相关检索以及引文检索等。相对于中文数据库检索系统而言,外文还根据其语言的特点提供了截词检索、词根检索、嵌套检索(优先算符, nesting)、位置算法、禁用词表、大小写敏感以及智能文本检索等检索技术,实际上在中文检索系统中,这些技术也在使用,例如中文检索系统也提供了通配符检索等。但是中

文的任何一个词在特定的语境之中都有其特定的实际意义,因此禁用词表在中文检索系统中价值不大,其他诸如大小写敏感也是外文独有,中文不存在大

小写区分的问题。总之,在检索技术层面,除了某些外文特色外,中外文检索系统在检索技术层面没有显著差异。

表 2 学术数据库的检索字段/限制条件

检索字段/限制条件	中文数据库频次	外文数据库频次	合计	检索字段/限制条件	中文数据库频次	外文数据库频次	合计
作者	10	40	50	团体著者	0	4	4
题名	6	40	46	支持基金	3	4	7
出版时间	8	38	46	原文出处	1	3	4
出版物名称	9	31	40	索书号	1	3	4
文摘	5	31	36	同行评议期刊	0	3	3
全文	5	30	35	标引词	0	3	3
匹配方式	3	30	33	学位专业	2	2	4
文献编号	0	26	26	特征词	2	2	4
文献类型	2	23	25	提供图书类别	1	2	3
ISSN 等	2	23	25	系列名称	0	2	2
关键词	9	19	28	被引用次数	1	0	1
每页显示结果数	2	19	21	会议地点	0	2	2
卷	1	19	20	段落	0	2	2
主题/分类名	2	17	19	分类表	1	1	2
页码	1	17	18	叙词	0	1	1
机构名称	4	16	20	评论	0	1	1
期、复印期号	3	16	19	过滤	0	1	1
排序	1	16	17	栏目信息	1	0	1
引文	3	13	16	作者简介	2	0	2
语种	1	12	13	英文题目	1	0	1
人员	0	11	11	同义词	1	0	1
任意字段	4	10	14	同名作者	1	0	1
主题词	0	9	9	期刊来源类别	1	0	1
出版地	2	7	9	期刊级别	1	0	1
页数/封面报导/附带图像的文章	0	7	7	分主题子库,数据库	0	1	1
分类号	5	5	10	词频限制	1	0	1
更新时间/最近更新	2	3	5	书籍类型/适用对象	0	0	0
物种、性别、年龄组	0	2	2	从检索结果排除	2	4	6
出版者	4	5	9				

3.11 查询扩展

从表 4 可以发现主要的外文查询扩展途径是通过相关文献以及参考文献两种方式进行,过半数的外文检索系统提供了相关期刊/书目、共引文献/引文、返回检索结果列表、作者其他文献等查询扩展方式。其中作者的其他文献查询扩展方式没有在中文检

索系统中出现。如果一个作者其研究兴趣与领域基本稳定的话,那么该查询扩展方式还是非常有价值的。中文的 CNKI 推出了有特色的相关专家与相关研究机构推荐,不过从笔者的使用效果来看,其算法还有很大的改进余地。与此类似的是有 4 个外文数据库提供了当前期刊中的相似文献查询扩展功能,都是较有

价值的查询扩展功能。中外文数据库都零星地提供了相关文献作者/作者文献、相关检索词、检索结果链接的文献网络图示、相关博文、同行关注文献、相关机构

文献以及分类导航等查询扩展功能,都分别只有1-2个数据库在使用,还不具有普遍性,其效果究竟如何,尚需通过实际应用来检验。

表3 学术数据库的检索技术

检索技术	中文数据库频次	外文数据库频次	合计	检索技术	中文数据库频次	外文数据库频次	合计
布尔检索	8	39	47	禁用词表	0	9	9
组配检索	8	37	45	位置算法	0	10	10
相关检索(扩展或者近义词、同名/合著作者、分类表、相关机构检索)	1	3	4	嵌套检索(优先算符, nesting)	0	11	11
引文检索(或结果相关链接)	1	1	2	大小写敏感	0	2	2
截词检索	0	26	26	智能文本检索	0	1	1
词根检索	0	12	12				

表4 学术数据库的查询扩展功能

查询扩展	中文数据库频次	外文数据库频次	合计	查询扩展	中文数据库频次	外文数据库频次	合计
相似文献	6	21	27	相关博文	1	1	2
参考文献	5	20	25	同行关注文献	1	1	2
相关期刊,书目	1	24	25	相关机构文献	1	1	2
(引文检索)共引文献/引文	1	23	24	中图法文献分类导航	1	1	2
设有返回检索结果列表	0	14	14	检索结果链接的文献网络图示	1	2	3
作者其它文献	0	10	10	相关研究机构	1	0	1
相关文献作者/作者文献	3	3	6	当前期刊中的相似文献	0	4	4
相关专家	1	0	1				

3.12 相关反馈

调查发现,中外文数据库常用的相关反馈技术有二次检索、检索提问修改、相关检索/主题相关、检索词的相关词、词条在工具书中的解释、检索词的相似词、合并检索等。中外文数据库在相关反馈技术层面没有显著性差异,一些使用频次为零的技术对于相关反馈而言都不是核心技术,比如词条的解释仅仅是锦上添花的工具,没有太多实质性的意义,检索词的相似词和相关词的差异也不是十分明显,因此可以合并为一类,合并检索现在大部分检索系统已经不再提供,已经不是非常核心的相关反馈技术。

3.13 结果排序

表5的数据表明,主要的排序策略为日期和相关度。外文有特色的排序依据有设置阈值(高、中、低)、出版者、MATH数据库登录号以及手稿类型等。而中文有特色的排序/分组的依据有新论文、经典论文等。近年来,中外文数据库都注意到了引用频次和下载频次对于文献内容的间接评价作用,从而在中外

文数据库中都增加了针对这两个指标的排序选项。

相对于检索字段,检索系统已经非常充分地利用了几乎所有可能的文献内外部特征,而对于检索结果的排序来说,目前可做的工作还比较多,相关性质量判据中的众多判据可以应用到该领域,也将是笔者今后需要进一步挖掘的研究领域。

3.14 检索结果格式/展示/输出

表6可发现中外文检索系统都提供了丰富的检索结果输出形式以及相关的控制,不过与外文数据库相比,中文没有提供更为便捷的输出服务。例如利用E-mail发送检索结果、直接在网上订购文献全文以及套录输出等,其中尤其是套录输出功能能够显著的加快用户输出检索结果的速度,遗憾的是到目前为止中文数据库还没有提供该功能。其他的诸如请求原文传递以及自定义格式输出的功能中文数据库中只有万方和维普提供了相似功能,而外文数据库分别有10个和7个提供该功能,这些功能对于提高系统的易用性有着非常重要的作用。

总的来说,在检索结果输出方面,中外文数据库及存取性的相关功能,与外文数据库相比,中文数据库基本的常用功能相似,但是对于提高用户便捷性以库还有许多工作有待完成。

表 5 检索结果的分组排序方式

分组排序	中文数据库频次	外文数据库频次	合计	分组排序	中文数据库频次	外文数据库频次	合计
日期	4	36	40	下载频次	1	1	2
相关度	2	31	33	阈值(高,中,低)	0	1	1
文献标题	2	10	12	出版者	0	1	1
来源期刊	1	9	10	新论文	1	0	1
第一作者	1	7	87	经典论文	1	0	1
被引频次	1	4	5	手稿类型	0	1	1
能否消除重复记录	0	3	3	MATH 数据库登录号	0	1	1

表 6 检索结果输出方式

结果输出	中文数据库频次	外文数据库频次	合计	结果输出	中文数据库频次	外文数据库频次	合计
每屏显示量	9	36	45	利用 E-mail 发送检索结果	0	28	28
固定形式导出引用参考文献、文本、XML 等到 Note express、EndNote 等目录管理器	5	35	40	是否提供多种显示格式(如可提供题录、题录+文摘、全记录或选择字段等)	2	18	20
最大输出量:检索结果的最大输出量是多少	7	33	40	能否在结果中对检索词进行突出显示	6	30	36
查看全文/阅读	7	32	39	自定义格式输出	1	7	8
用户是否可以自己定义每屏显示的记录数	4	32	36	直接在网上订购文献全文	0	11	11
请求原文传递	1	10	11	套录输出	0	3	3
下载/打印方式下载数据	4	27	31	该条目的评论	0	1	1
存盘/加入电子书架	2	26	28				

3.15 标记方式

中外文数据库主要设置了逐条选、最大标记量、全选以及逐页选等功能,外文数据库有 2 个设置了任意条选等方式。调查发现,大多数学术数据库(38:51)都设置了最大标记量,该功能是为了规避用户批量下载行为。总体看中外文数据库在标记方式方面没有显著差别。

3.16 检索结果分析

表 7 列举出中外文数据库主要提供的检索结果

分析功能。中文的 CNKI 提供的有特色的分组方式有研究层次,不过就笔者的使用体验来看,该分组方式还有很大的改进余地。

检索结果的排序、聚类分析对于目前信息超载状况的解决具有非常重要的意义,也是能够缩短用户检索过程的非常重要的领域,需要学术界花更多精力使其更加完善的方面。该功能组对于提高学术数据库的易用性以及提高用户的情感感知将起到非常重要的作用。

表 7 检索结果分析功能

检索结果分析	中文数据库频次	外文数据库频次	合计	检索结果分析	中文数据库频次	外文数据库频次	合计
文献作者	1	17	18	出版者	0	3	3
主题聚类	4	15	19	研究资助资金	1	2	3
年代/时间聚类	6	15	21	创建引文报告	0	1	1
出版物类型聚类	1	13	14	国家/地区	0	2	2

来源出版物	1	15	16	分析结果保存(SCI)	0	2	2
学科类别	1	4	5	字顺	0	1	1
作者单位	1	4	5	发表年度	1	1	2
语种	0	4	4	会议标题	0	1	1
检索结果文献计量分析(SCI)	0	3	3	结果类型(研究论文、新闻等)	0	2	2
检索结果论文类型聚类	1	3	4	研究层次	1	0	1
关键词	1	3	4				

3.17 链接关系

目前,超链接已经成为一种重要的导航方式和揭示相关信息的方式,在检索系统中已经得到了广泛的应用。笔者通过调查发现中外文数据库在同一数据库的不同记录、全文数据库以及同一平台的不同数据库之间已经建立了多种形式的链接,除此之外,多数外文数据库与因特网资源之间也建立了链接关系,CALIS 中心的 CCC 外文期刊数据库还与馆藏资源建立了链接关系。

检索结果的排序提供了检索结果的线性排列,检索结果分析提供了检索结果的层次结构序列,而检索结果之间的超链接则提供了检索结果的网状结构呈现形式。三者的交互使用对于用户从不同的视角获得所需信息意义显著。三种方式的具体选择可以从相关性的质量判据以及文献特征判据等充分汲取营养,从而提出更有效的排序、聚类以及超链接的思路。

3.18 整合检索

整合检索指检索提问向多个数据库提交一次,从多个数据库中获得检索结果。从调查来看,中文数据库目前仅仅实现了在同一公司的不同数据库间的整合检索,而外文数据库则实现了同一公司的不同数据库、不同公司的不同数据库以及外部数据库间的整合检索。整合检索对于提高用户检索效率,避免频繁重复劳动具有现实意义,就目前国内数据库的表现,其开放程度还远远不够,整个数据库产业各自为政,同行之间协作与共享的实现还有待努力。

4 研究结论

本文通过中外文 51 个数据库,选取 19 个系统特征进行了调研,针对每个系统特征提出若干观点。通过本文的具体分析,总体可以得出如下结论:

(1)有必要采取多种措施减轻用户的认知负担。从总的方面来说,中外文数据库的系统特征已经没有显著性差异,但是中文数据库还有一定的追赶空间:中文数据库需要更多认知方面的投入,从而减少信息用户的认知负担,间接地提高用户在其他方面的时间与认知等投入,提高教学与科研的整体

效率。例如在导航设计方面、在帮助服务方面,如本文所调查的结果,亟待加强其效果。

(2)有必要改进检索输出的分析和设计。前面已经谈到,在检索入口方面,目前的学术检索系统已经进行了充分的挖掘,但是在检索输出方面,还有较大的发展空间。具体的研究思路为:充分利用现有的相关性判据研究成果^[6],从检索结果的排序、检索结果的分析以及检索结果之间超链接的交互使用等提供线性、树状以及网状结构的呈现形式。

(3)应该扩大整合检索的应用范围。就作者的检索体验来说,为了保证某个主题的查全率,作者需要不断地打开/关闭多个数据库,而在具体的数据库中的操作是相似的,因此,如果能够扩大整合检索的应用范围,那么类似的简单操作则可以节省,从而能够节约信息用户的认知和时间开销。解决该问题的关键是需要制定学术数据库的元数据标准、学术数据库之间的开放接口标准以及数据库商之间的利益分配策略。前两者是技术层面的问题,相对而言易于解决,而后者可能是影响整合检索大范围应用的主要障碍。

(4)需要进一步扩大个性化服务的深度与广度。现有的个性化服务包括非内容层面的用户界面定制以及多语言平台;提供一些便捷提醒,诸如,存储检索记录、存储检索式、删除检索式、存取删除检索史、论文引用提醒、论文内容更正提醒、期刊新一期出版提醒、将当前文献 email 给朋友服务以及我的电子书架等;涉及信息用户内容服务的有期刊定制、邮件定题服务、分类定制、个人期刊列表以及关键词定制等;而难度比较大的推送服务中外文系统都还没有实现。从这些个性化服务可以看出,它们基本上都属于比较浅层次的服务方式,涉及认知投入大的主要以内容为主的个性化服务方式提供得则非常少。突破该瓶颈的必由之路是学术检索系统首先需要了解信息用户个性化的信息需求,只有这样才能提供真正意义上的个性化服务,而这正是学术界目前正在研究与探讨的一个难题,其成功解决还需要时间。

(下转第 36 页)

同的类目联系起来。5 版 G25 类下增添了 24 个类目参见比 4 版多 19 处,《中图法》的类目参见,把那些隐蔽分散的类目关系充分地表现出来,开阔了分类人员的视野,拓宽了分类人员的知识面,使分类人员加深了对类目的理解,并进一步明确了类目的含义和它所涉及的范围,为分类人员选择最恰当的类目提供了方便,更好的体现了《中图法》编制的实用性。

3 类目增改探讨与建议

5 版《中图法》图书馆事业、信息事业类目和其它学科类目一样,经过上述的类目合并增改后,其科学性和实用性较之 4 版有了显著的提高:类目名称更准确;在保证分类法体系稳定性、连续性的情况下,通过恰当的方式增补了新学科、新技术、新主题;类目注释更完善;更符合科学技术的发展趋势,使该类的知识组织更为系统化、规范化。它的启用定将给《中图法》使用人员类分文献带来极大的方便,但还有一些问题有待商榷。

3.1 一些新主题无类可归

如:跟踪服务、国际情报学、密集书库、私人情报馆、馆舍搬迁、智能代理、图书馆文化书店等,这些图书情报方面的新学科、新概念都应在将 G25、G35 两类经整合后,给予增补类名。

3.2 和档案学的关系

从业务处理和研究的对象来看,图书馆工作的对象是书本和专著,档案工作的对象是备查文件。这些工作对象都是一脉相承的仅仅是表征形态不同的信息。在各大大学的系科和专业设置中,情报学系或信息管理系,一般包括有科技情报、图书、档案等专业,而大学系科的设置通常与学科和业务工作的

分类是一致的。这又从另一侧面说明,把这几项工作以及它们的学科拆分入不同的类目是欠妥的。《中图法》将这几项工作及其学科分类入“G25”和“G27”两个大类目之中,是不合适的。通过 5 版整合“G25”和“G35”两大类的有益尝试,我们期待下次《中图法》的修订能把档案学整合进图书馆学,信息学中来。

参考文献

- 1 中国图书馆分类法编辑委员会. 中国图书馆分类法(第 4 版). 北京:北京图书馆出版社,1999
- 2 国家图书馆《中国图书馆分类法》编辑委员会. 中国图书馆分类法(第五版). 北京:国家图书馆出版社,2010
- 3 俞君立,陈树年. 文献分类学. 武汉:武汉大学出版社,2001
- 4 陈久庚. 浅析《中图法》对“情报学、情报工作”的分类. 情报科学,2004(7)
- 5 覃华兵. 《中图法》第 4 版图书馆学与情报学类目设置浅议[J]. 大学图书馆学报,2005(2)
- 6 黄汉年,侯汉青,何琳. 《中图法》图书馆学情报学大类修订构想[J]. 国家图书馆学刊,2008(2)
- 7 李宏. 网络环境下图书情报一体化工作模式的演变[J]. 图书馆理论与实践,2004(6)
- 8 葛驰,罗敏. 图书情报一体化论纲[J]. 现代情报,2003(11)
- 9 周卫妮. 浅析《中图法》4 版中某些注释说明及举例. 图书馆建设,2005(6)

宋力女 1973-, 陕西理工学院南区图书馆采编部,副研究馆员。

(收稿日期:2011-01-13 编发:许桂菊)

(上接第 33 页)

(5) 效仿商务型网站,提供增值服务。比如个人主页服务、学术社区服务等^[6]。

参考文献

- 1 李宝强. 基于 TAM 的网络学术信息资源用户接受与利用研究[D]. 南京:南京大学,2008
- 2 李宝强,孙建军,成颖. 基于 TAM 扩展模型的用户网络学术信息资源利用初探[J]. 情报学报,2008(4):596-606
- 3 李明,雷银枝,李晓鹏. 基于 TAM/TTF 整合的网络信息资源利用效率模型与指标框架研究[J]. 情报学报,2009(1):128-136
- 4 叶晓飞. 基于 TAM 与 TTF 整合模型分析网络信息资源利用[J]. 图书情报工作,2008,52(3):39-42

- 5 董珏. 基于 TTF 模型的网络信息资源利用效率研究[J]. 图书情报工作,2008(4):42-45
- 6 成颖. 信息检索相关性判据及应用研究[D]. 南京:南京大学,2011
- 7 Xie, H. I. (2008). Users' evaluation of digital libraries (DLs): Their uses, their criteria, and their assessment [J]. Information processing and management, 44:1346-1373

成颖 1971-, 博士,副教授。

孙建军 1962-, 博士,教授,博士生导师。

张敏女,1987-, 硕士研究生。

(收稿日期:2011-06-06 编发:王宗义)