

中文人文、社会科学数据库分析探讨

厉 莉 赵 瑾

(中国科学院文献情报中心 北京 100029)

[摘要] 通过对中文人文、社会科学综合性数据库和专题性数据库的评介分析,探讨近几年中文人文、社会科学数据库的发展状况。

[关键词] 中文人文 社会科学数据库 数据库探讨分析

[分类号] G250.74

近几年,我国数据库建设发展很快,人文社会科学数据库在数据库中所占的比例不断加大。这些数据库已成为从事人文、社会科学学习、教育和研究者获取信息的重要来源。

1 中文人文、社会科学综合性数据库

1.1 中国知识资源总库 (<http://211.151.93.134/tpi50/Navigator.aspx?ID=1>)

中国知识资源总库是由中国基础知识设施工程(CNKI)开发,并建立在CNKI平台上的综合大型数据库,其主办方为清华大学等单位。系统收录了CNKI系列数据库和加盟数据库152个,全文和各类信息数据超过5 000万条。涵盖了我国自然科学、工程技术、人文与社会学期刊、博硕士论文、报纸、图书、会议文献等公共知识信息资源。在该知识总库中,自建数据库主要有:

1.1.1 中国期刊全文数据库 收集全文文献1 200万篇,题录1 500余万条。分为九大专集,126个专题文献数据库。有关人文、社会科学的有哲学、人文科学、教育与人才、社会科学等。资料来源为国内公开出版的核心期刊和专业期刊,并以每天上万篇的速度递增。文献收入年限为1979年至今。

1.1.2 全国重要报刊全文数据库 该库以国内重要报纸刊载的学术性、资料性文献为收入对象。目前已累积文献400万篇。分为文史哲、政治军事、法律、文化、艺术、体育及各界人物、经济、社会、科学技术、恋爱婚姻家庭与健康等,43个专题数据库。收入年限为1994年6月至今。

1.1.3 中国优秀博硕士论文全文数据库 该数据库汇集了我国300余所大学自1999年以来的优秀博硕士论文。已收入学位论文19万余篇,分为九大专集,121个专题文献数据库。人文社会科学方面是文史哲、经济政治与法律、教育与社会科学等。

1.1.4 中国重要会议论文全文数据库(CPCD) 收录我国1999年以来国家二级以上学会、协会、高等

院校、科研院所、学术机构等单位的论文集,年更新约10万篇文章。至2005年4月止,累积会议文献27万余篇。与社会科学有关的涉及文史哲、经济政治与法律、教育与社会科学等。

中国知识资源总库中与社会科学相关的还有中国图书全文数据库、中国基础教育数据库(中小学版)、中国企业知识仓、中国城市规划知识仓、中国年鉴全文数据库等。

1.2 新华社综合信息数据库 (<http://info.xinhua.org>)

该数据库的主要特点是以新闻和经济信息为主,并且更新速度快,分为28个数据库100多个子库。主要数据库有:

1.2.1 中文新闻稿库 全文存入新华社每日向国内各报社、电台播发的中文电讯,日均入库200余条。起始日期为1991年。

1.2.2 经济信息库 全文存入新华社信息中心的30种信息产品,设立轻工、家电电子、纺织服装、医药、农业、粮油食品、交通口岸、机械、汽车、金属、能源、化工、建工建材、房地产、经贸、综合金融、新技术新产品、合作与投资、国内商品供求、海外商品供求、中外市场行情、中国市场动态与分析、海外市场动态与分析、中国经济动态、海外经济动态、金融行情、对外中国经济信息等30个子库。

1.2.3 专题资料库 这是一部动态的历史和时事知识词典。数据库选编了国内外政治、经济、外交、文教、军事等领域的各种重要专题性材料。

1.2.4 报刊库 全文存入新华社主办的《半月谈》、《瞭望》、《经济参考报》、《中国证券报》、《上海证券报》、《人民日报》、《经济日报》等报刊内容。还精选国内300余家重点报刊中具有一定参考价值的新闻,年均信息量2万条左右。

1.2.5 组织机构库 收集了我国和世界上100多个国家(地区)的政府机构、政党、社会团体及联合国等

国际组织的基本情况和领导人名单。

此外,还包括收入我国中央和地方近年发布的重要法律和法规文件的法规库;记载国内外体育信息的体育库;证券市场信息的证券信息库,香港库、企业库、人物库等等。

1.3 万方数据库(http://www.wanfangdata.com.cn/src/gywm/wf_gywm_index.asp)

万方数据库由中国科技信息研究所万方数据股份有限公司等单位开发,这是我国最早的中文信息服务系统。该数据库中与社会学有关的子库有:

1.3.1 中国学术会议论文全文数据库 该数据库收录了1998年以来国家级学会、协会、研究会组织召开的全国性学术会议论文。数据范围覆盖自然科学、社会科学诸多领域。收全率达到100%。现有文献量26万篇。该数据库的英文版主要收录在中国召开的国际会议的论文,内容多为西文。

1.3.2 中文学位论文全文数据库 该数据库收集了1980年以来我国各学科领域的博士、博士后和硕士研究生的论文。目前数据库已经收录2.8万篇,每年增长3万篇。

1.3.3 中国法律法规全文库 内容包括自1949年建国以来全国人大、国务院、最高人民法院和最高人民检察院等单位颁布的法律法规、司法解释等,同时还有各部门规章、各地地方性法规和地方政府规章,以及我国签订的国际条约和公约等。

在该数据库中还包括数字化期刊数据库(收录了近5000种核心期刊全文)、商务贸易数据库、参考类数据库、台湾类数据库(内容为台湾地区的科技、经济、法规等相关信息)、机构类数据库(包括我国著名科研机构、高等院校、信息机构的信息)等。

我国社会科学综合性数据库中各高校图书馆和情报机构、数字图书馆建立的书目数据库也是十分重要的信息资源。这些免费信息为社会科学研究者提供了极大的便利。

2 中文人文、社会科学专题数据库

2.1 中国高等教育文献保障系统(CALIS)和国家图书馆开发的专题数据库

中国高等教育文献保障系统(CALIS)是我国高等教育“211工程”总体规划中两个公共服务体系之一,其中的专业数据库系统是CALIS“十五”建设的子项目之一,由武汉大学图书馆牵头组织实施,全国共有61所高校的75个项目获立项。目前,该专业数据库中与社会学有关的子库已有20多个,包括建筑文化、地区性研究、专家学者论著、海

外华人华侨信息等文献数据库。还有古籍音韵学书目库、文学专题数据库、音乐图书馆、孙中山数字图书馆、经济数据库、中国科技史数字图书馆资料库、中国农村问题文献数据库等。

中国国家图书馆从1990年开始开发馆藏特色资源,建立各种类型的专题数据库。“民国中文期刊数据库”即为其中之一。该数据库遵循边建设边服务的原则,在原缩微胶片的基础上进行期刊全文数字扫描,预计将在3年内完成近600万张缩微胶片的数字转换。“地方文献数据库”是将国家图书馆所存清代以前(含清代)的方志资源数字化;“西夏文献数字资源库”包括西夏古籍书口数据125条、原件影像近5000幅、研究论文篇名数据1200余条;研究论文全文扫描9700页。“中文拓片资源库”现已收录数据1.6万余条,影像2.1万余幅。另外,国家图书馆还购买了上海人民出版社和迪志文化出版有限公司合作开发的文溯阁《四库全书》、四部丛刊以及《中国大百科全书》等数据库。

2.2 经济类数据库

中经网(<http://www.cei.gov.cn/>)开发的各类经济数据库包括:

2.2.1 中国经济信息网 这是以国家信息中心为主建立的专业性经济信息网站。该网站开发了数种在我国经济领域比较有影响的数据库。如中国经济统计数据库,涵盖宏观、行业、区域、世界以及经济专题等各个领域,拥有15万指标。中国行业季度报告数据库分为24个大类行业、36个细分行业,共60个行业,每年240篇左右的研究报告。

2.2.2 中国权威经济论文库 提供国内外公开出版的报刊杂志、权威机构专家学者的论文2万余篇。中外经济动态全文库包括中外重要报刊、地区动态信息库、行业动态信息库等。

2.2.3 经济数据特供系统 这是由国家信息中心数据中心与新华在线信息技术有限公司建立的一个数据在线分析系统。该数据库以经济数值型数据为核心内容,于每月、每季、每年第一时间提供最权威的时间序列的经济运行数据,包括宏观观测、地区观测、行业观测、财经观测、产品观测、外贸观测、价格观测、投资观测、世经观测等9个子系统。信息来源于国家发展计划委员会、国家经济贸易委员会、国家统计局、财政部、海关总署、中国人民银行、新华社、人民日报社、国家信息中心等多家权威部门和机构。

2.2.4 中国咨询行 这是香港一家专门从事收集、加工中国商业信息的企业机构。其所建数据库有中

国新闻数据库、中国商业报告库、中国法律法规库、中国统计数据库、媒体库、中国上市公司文献库、香港上市公司资料库、中国企业产品库等等。涉及行业从能源运输、金融保险到外贸科技、食品服务等经济领域的各行各业。形式包括新闻事实、经贸报告、法律法规和统计数据等。目前已拥有近千万篇文献,并提供全文检索。目录检索免费。

2.3 法律类数据库

法律类数据库包括:

2.3.1 “中国大法官”数据库 是由中国法律资源网开发的大型中国法律数据库。包括全国人大法律库,国务院行政法规库,部委行业规章库,地方法规规章库,两高司法解释库,国际条约与惯例,经典司法判例库,规范法律文书库,最新合同范本库等子库。完整收录了自1949年以来我国全国人大、国务院、各部委、各地方人大和政府、最高法院及检察院发布的法律、法规、规章和司法解释全文近14万多条。尤其是近一两年来房地产、金融、证券方面的法律法规规章和最高法院、最高检察院发布的最新司法解释收录得非常全面。

2.3.2 “北大法宝”数据库 由北大法律信息网建立,收录整理14余万篇法律法规,翻译3500多篇英文法规,共分为20个中文子数据库和两个中英文对照数据库。其主要内容有我国全国人大(常委会)、国务院、各部委、最高人民法院、最高人民检察院以及地方政府和港澳台地区颁布的法律法规、规范性文件、司法解释。全国人大常委会、国务院、最高人民检察院、最高人民法院的工作报告,以及由知名学者、法官、检察官、律师为主的法律专业人士撰写的法律论文为主要内容的法学文献库等。

比较著名的法律数据库还有由国家信息中心开发的国家法规数据库和国家法律法规数据库。

2.4 中国历史文化

2.4.1 中国基本古籍库 运用现代信息技术手段对中国传统文化的基本文献进行数字化处理的一个大型数据库。该库分经史子集4个子库,20个大类和100多个细目。收入上自先秦,下迄民国的1万余种历代典籍的全文数据和超过1500余种珍贵版本的图像数据,总计全文约20亿字,图像约2000万页。收录范围涵盖全部中国历史与文化,内容总量约等于三部四库全书。

2.4.2 《文史哲》全文数据库 是山东大学“文史哲”编辑部编辑出版的著名学术理论期刊《文史

哲》的网络电子版。数据库收录该刊自1951年创刊至2002年累计250余期登载的全部文章,共4000多万字。

2.4.3 《国学宝典》数据库(http://www.guoxue.com/gxbd_cd/gxbdzmll.htm)由北京国学时代文化传播有限公司依托首都师范大学组织国内一批文史专家,经过多年资料搜集、精心校勘,利用清华大学的网络技术研制开发的中华古籍全文资料数据库。收录范围为上起先秦、下至清末两千多年的历代经典名著和各学科基本文献。迄今为止,数据库已收入古籍文献3800多部,总字数近8亿字。目前仍以每年1-2亿字的速度增加库容。

3 中文人文、社会科学网络数据库分析探讨

3.1 社会科学数据库发展势头强劲

我国人文、社会科学数据库起步比较晚,上世纪90年代初具规模,2000年以来有了很大的发展。主要表现在:

3.1.1 图书馆和情报系统参与数据库的开发与建设

图书情报系统是文献收藏和信息传播的主要机构,这些机构更多地参与数据库的开发将有利于我国信息资源的共享。近几年我国高校图书馆和大、中型公共、专业图书馆纷纷建立了书目数据库,有的还开发馆藏特色文献数据库。如中国科技情报中心开发的万方数据库、重庆维普咨询有限公司开发的维普经济数据库等。特别是我国高等学校开发的大型综合性、专业性数据库,已经形成一定规模。如中国知识总库已成为从事社会科学研究工作者不可缺少的资源。

3.1.2 建立了法律和经济类数据库 我国比较著名的几个经济类、法律类数据库基本上是由政府部门研制开发的。这些数据库提供权威的、系统的各经济领域和法律法规方面的动态信息、统计数据、研究报告和法律法规条例。有助于用户准确了解经济发展动向、市场变化趋势、政策导向和投资环境;掌握国内外各项法律法规和条例。

3.1.3 数据库检索功能更加完善 一些大型数据库提供多种检索功能。如中国知识总库将媒体性质、数据格式、检索方法、应用环境不同的数据库,使用CNKI平台统一管理起来,集成成为一个整体,实现了跨库、跨平台、跨语言的无缝链接,并向用户提供知识信息分析、数字化学习、研究和情报获取等功能。万方数据库提供我国核心期刊、EI来源刊、SCI来源刊等检索点。还具有免费浏览、下载题录和摘要功能。CALIS系统的特色资源数据库向用户提供公共检索、馆际互借、存档服务等功能。

3.2 数据库开发应多样化

3.2.1 数据库类型多样化 虽然我国近几年社会科学数据库建设有较大的发展,但与国外先进国家相比,还存在一定的距离。从数据库类型看,还比较单一。主要表现在以下几个方面:

- 数据库收入的文献类型单一。目前,我国社会科学数据库的开发多以书目、期刊论文为主,缺乏工具书和数值性数据库。

- 数据库收入资源的载体单一。收入资源多为文本型资源,而实际上,社会科学资源的载体是多种多样的,如录像、声音、图片等。

- 专业数据库的开发不够理想。比如历史、文学、人物等还没有开发出比较大型的专业数据库。

- 回溯型数据库的开发不尽人意。期刊库和论文库收入的文献大多数是当前文献,而社会科学研究其资料的延续性又是十分重要的。

因此,数据库类型的多样化应是我国数据库开发的一个方向。

3.2.2 数据库开发离多样化 目前,我国一些大型数据库的开发主要是由政府、教育机构和图书情报机构完成的,国内一些大型的出版机构和从事社会科学研究机构还没有真正参与进来。而这些机构出版的大量文献资料和科研成果又是数据库开发的重要信息资源。如具有百年历史,以出版社会科学工具书和历史文化丛书闻名的商务印书馆;出版人文科学和资料性图书闻名的中国社会科学出版社等等。今后随着时间的推移,这些出版机构和研究机构若陆续参与到数据库的开发中,势必使我国数据库的类型更加多样化。

3.3 建立多样化的信息检索平台

国外有几个著名的信息检索平台。如美国科技信息研究所基于知识网络 (ISI Web of Knowledge) 建立的 ISI 知识库,这个知识库提供自然科学和人文社会科学的会议录、基本科学指标、期刊引证报告、引文索引、期刊论文目录等社会科学研究方面的信息检索。美国银盘公司 (Silver Platter) 的数据库组使用同一检索平台,检索 21 个不同开发商开发的数据库。其中社会科学数据库有著名的彼得森研究指南、威尔逊传记数据库、国际政治学文摘库等。还有美国剑桥科学文摘、Gale 集团等等,它们开发的大型数据库都是依托于同一检索平台检索各学科的信息。

我国社会科学方面大型检索平台的开发与先进国家有着比较大的差距。目前,我国除清华大学的中国知识总库外,教育和图书情报系统所开发的文献

检索系统中,数据库基本是本系统资源,或购买国外资源。各系统间缺乏相应的合作。在我国一些社会科学研究机构、博物馆、艺术馆、档案馆中收藏了大量有价值的、各种载体形式的社会科学文献资源,并且也开发出了独立的数据库系统。如何将这些资源进行组织、开发,建立集成化的检索平台,还需要各收藏单位、情报系统和开发商进行协调,策划。

3.4 重视中文社会科学免费资源的开发利用

3.4.1 开发利用互联网上中文社会科学免费资源,建立资源导航数据库 互联网上社会科学免费信息资源的来源主要是国家政府机构、国际组织、图书馆、出版发行机构、专业学会和协会、各类教育科研机构、传媒、商业网站、个人及各类非盈利组织等,这些网站提供了大量的新闻事件报道、法律法规、统计数据、财政经济信息、书目查询、网络课程、专业论坛、会议文献等等。近年我国一些大型图书情报系统网站建立了学科导航系统,对互联网上的免费资源进行了挖掘和整理,但是这种导航还只是初步的。社会科学信息资源的开发应重视互联网上的免费资源开发,建立社会科学专业导航库(也可称为学科门户网站)。在学科资源导航库方面,我国自然科学的图书情报系统走在了前面。如中国科学院一些研究所建立的物理、化学、生物学学科门户网站,这为社会学专业门户网站的建设提供了经验。

3.4.2 开发免费学术性社会科学文献数据库,实现信息资源的最大共享 近几年,国外免费社会科学学术性数据库的建设有一定的发展。比如在哲学研究领域,斯坦福大学开发的哲学百科全书数据库就是一个重要的免费网络版哲学百科全书。“Animus”数据库收录了该杂志近 10 年各卷论文的全文,对研究理解西方传统哲学理论和思想提供了便利;“Carleton 大学哲学期刊”提供来自大学本科生和研究生撰写的哲学论文,成为对哲学有兴趣的学者发表自己第一篇哲学论文的园地。在我国,免费社会科学学术性数据库建设刚开始起步。如“公共管理论文资料库”全文收入了行政、政治、社会、法律和经济等领域近 10 年一些大学学报和研究性刊物发表的学术论文。2003 年一些学者又开发出我国第一个预印本数据库——“奇迹文库”。该文库目前已形成自然科学、工程科学与技术、人文与社会科学三大部分,覆盖了主要基础学科,发布各种学术资料 2500 项。国内外这些社会科学免费数据库的开

(下转第 184 页)

模式灵活多样, 人员流动、岗位公开竞争等, 这样更能形成一种新型的公共管理文化。

对员工的培训也是一种激励或者应该说是激励的延续, 是图书馆人力资源管理的重要举措。尽管我们图书馆员工队伍的学历、结构与过去相比已经有了改善, 但还是不能适应当代科学技术发展需要。要使管理者和员工向不同角色转型, 改变人力资源管理理念, 员工的培训和发展尤其重要。

人员发展培训要注意急需掌握高新技术的可用之才, 但总的是要协力为员工创造一个学习环境, 以便他们适应组织不断变化的取向, 同时为他们提供自我完善和岗位流动的真正机会, 并由此提高员工对工作的满意度。

7 绩效与考核

图书馆的人力资源管理同其他一般管理(制度管理)的重要甄别在于从注重过程转向注重结果, 这是一种理念的差别。为了提高人力资源管理职能, 对于人力资源管理进行一定程度的量化是很有必要的。因为绩效是结果, 人力资源管理考核注重的是结果, 而要得到可信的绩效和结果, 量化管理无疑是最好的手段。

由于绩效测量关注组织的任务目标, 因而也有助于提高组织生产力。在图书馆工作流程设计和组织沟通效率, 绩效薪酬制度等改革的可行性, 很大程度上将取决于相关因素是否能被量化。从图书馆的发展或工作的基本点出发, 对成本、时间、服务质量与数量差别进行绩效测量无疑是人力资源管理

量化的必要途径。

基于现代信息网络为用户创造了一种不受时空限制, 虚拟的网络化学习和科研环境, 因此在图书馆现代化、数字化、网络化的进程中, 给图书馆文化的变迁提出了新的挑战。新技术的运用; 信息产品的加工、整合重组为用户提供获取信息, 使用工具的简单易用和可操作性既给图书馆人承担的目标任务增加了新的内容, 也给图书馆人力资源配置或激励都带来了新的课题。

现代化图书馆的人力资源与以往不可同日而语, 合乎图书馆事业发展要求的人力资源结构, 资源配置无不受到市场经济的影响。人的文化差异, 同岗位工作的时间长短、家庭情况、薪酬、待遇等等均给图书馆事业发展所需的人力资源带来了难度。

正如肖知兴所言, 图书馆人力资源管理变革正是“从器具到制度再到文化, 是一个以分析到综合, 从科学到手艺, 从书本到实践, 从照搬搬到变通, 从技术本位和财务本位的管理到以人为本的管理过程”的理念转化和管理科学的实现。

参考文献:

- 1 谭祥金. 图书馆骨干队伍建设思考. 图书馆情报工作, 2005(4):13-15
- 2 甄洁. 中美图书馆人力资源建设的差异与启迪. 图书馆杂志, 2005(4):35-37
- 3 A.Eliasser K 等主编, 王满船等译. 公共组织管理, 2003(7):3-4,164-171

[作者简介] 顾润波, 男, 1948 年生, 副研究员, 发表论文 10 篇; 金真针, 女, 1961 年生, 馆员, 副主任, 发表论文 3 篇。

(上接第 255 页)

发和利用无疑有利于推动科研成果的传播, 促进学术信息的交流。

参考文献:

- 1 [2005-09-27].<http://211.151.93.134/tpi50/Navigator.aspx?ID=1>
- 2 [2005-09-27].http://www.wanfangdata.com.cn/src/gywm/wf_gywm_index.asp
- 3 [2005-09-27].<http://www.calis.edu.cn/calisnew/>
- 4 [2005-09-27].<http://www.cei.gov.cn/>

- 5 [2005-09-27].<http://203.93.24.67/>
- 6 [2005-09-27].<http://cul.sina.com.cn/n/2005-03-04/113544.html>
- 7 [2005-09-27].<http://www.cn-classics.com/chaoshi/>
- 8 [2005-09-27].http://www.guoxue.com/gxbd_cd/gxbdzml1.htm
- 9 葛杭. 网上免费资源综述. 图书馆学研究, 2005(5):
- 10 毛娟. 网上免费资源的来源. 图书情报工作, 2004(12):
- 11 季燕江. 奇迹文库及其理想. [2005-09-27]. <http://www.qiji.cn/eprint/abs/2507.html>

[作者简介] 厉莉, 女, 1959 年生, 馆员, 发表论文数篇。赵瑾, 女, 1966 年生, 馆员, 发表论文数篇。