

中国索引学会 编著

ZHONGGUOSUOYINXUELUNWENJI

中国索引学论文集

1991—2011

- ◎ 索引理论与方法
- ◎ 索引事业与索引史
- ◎ 索引语言与索引软件
- ◎ 索引评介与研究综述

上海辞书出版社

中国索引学会 编著

CHONGGUOSUOYINXUELUNWENJI

中国索引学论文集

1991—2011

- ◎ 索引理论与方法
- ◎ 索引事业与索引史
- ◎ 索引语言与索引软件
- ◎ 索引评介与研究综述

上海辞书出版社

图书在版编目(CIP)数据

中国索引学论文集: 1991~2011/中国索引学会编著. —上海: 上海辞书出版社, 2012. 9

ISBN 978 - 7 - 5326 - 3789 - 8

I. ①中… II. ①中… III. ①索引—文集 IV. ①G353.21 - 53
中国版本图书馆 CIP 数据核字(2012)第 201024 号

责任编辑 张亚芳
特约编辑 沈毅骅
编 审 王有朋
装帧设计 汪 溪

中国索引学论文集 1991—2011

上海世纪出版股份有限公司 出版、发行
上 海 辞 书 出 版 社
(上海市陕西北路 457 号 邮政编码 200040)
电话: 021—62472088

www.ewen.cc www.cishu.com.cn

上海图字印刷有限公司印刷

开本 787 毫米×1092 毫米 1/16 印张 24.5 字数 460 000

2012 年 9 月第 1 版 2012 年 9 月第 1 次印刷

ISBN 978 - 7 - 5326 - 3789 - 8/G · 785

定价: 128.00 元

如发生印刷、装订质量问题,读者可向工厂调换
联系电话: 021 - 55032807

目 录

一、索引理论与方法

现代的索引就是数据库	张琪玉 (3)
关于索引学研究和索引工作开展的设想与建议	
——为中国索引学会成立大会而写	张琪玉 (7)
论索引项	张琪玉 (16)
数字索引学要旨	刘 炜 (20)
信息化条件下引文索引的信息精选与学术评价功能	叶继元 (29)
引文索引的功能与科学评价	
——以美国《基本科学指标》引文数据库为例	邱均平 马瑞敏 (36)
索引作为文献计量学分析工具的科学性与局限性	王知津 王 璇 马 婧 (53)
再版索引	
——信息揭示与信息组织、文献计量与文献评价的新方法	赵呈刚 (67)
关于索引的选题问题	傅德华 (74)
索引与知识发现	史 睿 (78)

二、索引事业与索引史

新世纪我国索引事业的发展	葛永庆 (91)
--------------------	------------

繁荣我国索引事业的目标与对策	罗春荣	曹树金 (95)
知识时代我国索引事业可持续发展战略探讨	何绍华	朱红涛 (102)
ISO 999 的编制思路及其对我国制订索引标准的启示	仇 琛	(108)
制订《新编地方志索引标准》的可行性分析及技术性建议	衡中青	(114)
论续修志书推广索引体裁的必然性	夏侯炳	(125)
关于书后索引现状的几点思考	周柏康	(134)
古籍书后索引述论	时永乐	门凤超 (139)
上海图书馆在书目索引工作中的贡献	张贤俭	(150)
中国索引发展史略	印永清	(156)
索引的历史发展与未来趋向	戴维民	(163)
20 世纪的索引编纂与研究	潘树广	(175)
创新发展, 开拓前进		
——新中国成立 50 年来的索引编纂与出版	卢正言	(188)
西方索引早期在中国的传播研究	平保兴	(198)
钱亚新先生对我国索引事业的贡献	罗友松	(204)
我国第一部语词索引《老解老》研究	王雅戈	侯汉清 (215)

三、索引语言与索引软件

从人工语言到自然语言	侯汉清	吴建中 (227)
网络环境信息标引的测评与比较研究	刘 竟	朱书梅 侯汉清 (231)
受控语言在国内网络数据库中的应用与展望	李 琳	华薇娜 (240)
《中国图书馆分类法》索引编制技术的演变		
——评《中国图书馆分类法》第二、三、四版索引	李 华	徐青青 (249)
分类: 不可小觑的检索途径	华薇娜	(258)

索引语义扩展的探讨	王 梅 (266)
近年来国内知识组织研究热点综述	卜书庆 刘华梅 王广平 (275)
中国索引软件的开发与应用	王彦祥 (294)
国内外索引软件比较研究	郭丽芳 温国强 (302)

四、索引评介与研究综述

从油印本、印刷版到电子版

——论《古今图书集成索引》的编制	林仲湘 李 龙 (311)
关于《四库系列丛书目录·索引》的编纂	吴 格 (323)
德温特公司“世界专利索引”体系的发展及其对我国索引编制工作 的启示	马海群 (329)
确保质量是我们编纂《申报索引》的准则	黄秀文 (335)
查找民国文献的有效工具 ——《民国时期期刊篇名数据库》	吴佩娟 (339)

特色数据库的内容揭示

——以基于 TPI 平台的张謇研究特色数据库为例	端木艺 (347)
用途广泛而大有可为的图像索引	朱赛虹 (352)
索引研究与文献形式变化的相关性 ——评 10 年来我国索引研究状况	曲静涛 (358)
国内索引研究现状与发展的文献计量分析	汪徽志 (368)
中国索引学会历届理事会名单	(378)
后记	温国强 (385)

一、索引理论与方法

SUOYINLILUNYUFANGFA

现代的索引就是数据库

张琪玉

本文试图从索引工作现代化、数据库与传统索引在结构与功能上的比较、数据库推动了索引工作的现代化、数据库扩大了索引原理的应用、数据库对索引学发展的贡献这几个方面,说明“现代的索引就是数据库,现代的索引工作者就是数据库建造者”这一观点,进而指出“当前索引事业发展的重点是数据库建设”。

一、索引工作现代化和现代索引的概念

索引工作现代化的实质就是索引编制和使用的计算机化。用计算机编制索引是索引技术发展的高级阶段。在一些信息技术发达的国家,各种索引都利用计算机编制。在我国,利用计算机编制索引也已相当普及,用手工编制索引已越来越少。

用计算机编制索引有两种方式:一种方式是手工编制索引稿,再输入计算机编排和生产各种索引产品;另一种方式是直接在计算机上制作索引数据和生产各种索引产品。

用计算机生产的索引产品有多种载体形式,其中以数据库(数字化索引)和印刷型索引为多见。特别是数据库,由于有许多无与伦比的优点,因而发展迅速,数量已远远超过了印刷型索引。80年代以来,我国索引悄悄地、越来越多地以数据库的形式出现,这是为什么在当今信息时代,我国出版的印刷型索引反而越来越少见的原因。

现代的索引就是数据库,现代的索引工作者就是数据库建造者。

二、数据库与传统索引在结构与功能上的比较

数据库在功能上相当于传统索引的一个索引体系。数据库包含许多字段,一部分

字段相当于文献款目的各种著录事项,另一部分字段相当于文献的各种检索标识项(如分类号、主题词、题名、著者等)。后者一般是每一字段生成一个索引,通过索引对数据库进行检索(但也可不通过索引直接对数据库的相应字段进行检索)。所以,一个含有分类号、主题词、题名、著者字段的数据库相当于分类索引、主题索引、题名索引、著者索引四套卡片式索引,或相当于一种按详细分类排列正文并附有主题、题名、著者三种索引功能的检索工具。数据库的一个重要特点是数据的最少冗余,例如,在传统索引中,题名和著者既作为文献著录的项目又作为文献检索的标识时必须重复著录,而在数据库中则无必要重复。在传统索引中,诸如出版地、出版年等是不可能作为检索标识使用的,而在数据库中,必要时也可作为可检字段提供检索(一般是作限定检索)。数据库在检索上最主要的优点是可以用多个同一种类的检索标识或不同种类的检索标识进行组配检索(多种条件的联合检索),这是传统索引所难以做到的(比号索引和比孔索引除外)。此外,数据库还可用于文献计量和情报研究。

可以说,传统索引的全部检索功能,在数据库中都能实现;而数据库有许多检索功能,却是传统索引所不能实现的,数据库比之传统索引有更多的检索功能。数据库是比传统索引更为高级、更为先进的索引。

三、数据库推动了索引工作的现代化

数据库这种现代的索引形式,其编制和使用技术的广泛应用,推动了索引工作的现代化,具体体现在下列三个方面:

1. 在机编索引(先建立数据库,再由数据库生产印刷型索引)方面:

(1) 利用计算机编制索引可以一次输入、多次多种输出。索引数据一次性输入并校对正确后,可以根据具体需要生成不同检索途径、不同范围、不同格式、不同载体和份数多少不限的各种索引产品。

(2) 提高索引质量,诸如提高索引的标引深度,规范款目格式,减少手工编制时抄写、排序、打字或排版过程中的差错等。

(3) 加快编制速度。由于减少了抄写、校对等工序,以及加快了排序、打字或排版、累积以及编制轮排款目等的速度,因而可大大缩短由一次文献(原始文献)到二次文献(索引产品)的时差。

(4) 索引更新(增补)和累积(编制累积本)十分容易,这是使索引产品长期保持使用价值的一个重要条件。

(5) 作为机编索引副产品的数据(或直接建立的数据库)可开展各种索引服务。

(6) 可以编制各种手工难以编制或无法编制的新型索引。

2. 在自动标引方面: 利用计算机可进行自动抽词和自动赋词、自动赋分类号(一种自动分类方法), 使索引编制过程达到很高的自动化程度。此外, 目前正在普及的全文数据库, 还可免除标引工序。

3. 实现检索自动化: 利用数据库进行计算机检索和网络检索, 不但可千百倍地提高检索速度, 而且还可使用各种各样的检索技术, 大大提高检索效果。

四、数据库扩大了索引原理的应用

数据库的检索原理其实就是索引原理的新发展。目前数据库的类型极多, 可按其性质、专业内容、语种、载体、使用技术等予以分类。按其性质, 大体可分为文献数据库和非文献数据库。文献数据库又可分为文献目录数据库(包括机器可读目录、题录数据库、期刊目次数据库、文摘数据库、引文数据库等)和全文数据库; 非文献数据库可分为数值数据库、事实数据库、图像数据库、多媒体数据库等。不但在图书馆、情报、档案专业领域普遍利用数据库, 而且各行各业也广泛利用数据库进行管理和服 务, 可以说, 在当今信息社会, 数据库是一种最基本的管理和传播信息的工具。从某种意义上说, 数据库大大扩大了索引原理的应用。

五、数据库对索引学发展的贡献

数据库的出现极大地丰富了索引学的内容, 推动了索引学的发展, 具体体现在下列三个方面:

一是推动了索引编制技术的发展, 创造了许多新的索引编制方法, 特别是自动抽词、自动赋词和赋分类号等索引编制新技术。

二是推动了检索方法的进步。自数据库出现后, 创造出了适用于数据库的许多新的检索方法, 如布尔逻辑检索、加权检索、扩检、缩检、改检、二次检索、各种标识联合检索、截词检索、模糊检索、成批检索、SDI 服务、回忆检索过程、保留检索课题表达式、检索对话, 等等。

三是推动了索引用语言(情报检索用语言)的创新和改造, 最重要的是索引语言的组配化和自然语言的应用。

六、当前索引事业发展的重点是数据库建设

既然现代的索引就是数据库,索引事业发展的重点就应放在数据库建设方面。虽然印刷型索引的某些种类(如书后索引)仍应大力发展,但就整体而言,印刷型索引已失去发展的有利条件,例如,出版补贴缺乏就是印刷型索引发展的一个很大的障碍。

数据库是因特网发展的重要支柱之一。随着因特网的发展,数据库建设已成为非常迫切的问题。因特网将成为整个社会信息化的一个重要因素。但是,如果没有大量数据库的支持,就好比造好了信息高速公路,但没有载着货物(信息)的车辆在上面行驶,因特网也就形同虚设,不能充分发挥其信息资源共享的通信设施的作用。我国通信设施建设发展很快,但因特网上信息资源(中文信息资源)太少,也就是说联网的中文数据库太少,是一个亟待解决的问题。对我们索引工作者来说,这是一种历史任务,也是一种严峻的挑战。

七、推动传统索引与数据库的结合

我们强调数据库的重要性,认为索引工作的重点应放到数据库建设方面,并不意味着抛弃传统索引及其原理和方法。

传统的索引著作,有些至今仍有使用价值,可以将其转换成数据库,使其继续发挥作用甚至发挥更大的作用。

传统索引的某些原理和方法,如某些古籍索引的原理和方法,也可引进数据库,以开拓数据库的应用领域。

参考文献:

[1] 张琪玉. 关于索引学研究和索引工作开展的设想与建议. 江苏图书馆学报, 1993(1): 3-7.

[2] 张琪玉. 推广文献索引计算机编制法是促进我国索引事业发展的一项重要措施. 图书与情报, 1996(2): 33-36.

作者单位: 南京政治学院上海分院信息管理系, 上海, 200433

原载: 《图书馆杂志》2001年第12期、《中国索引》2003年第1期

关于索引学研究和索引工作 开展的设想与建议

——为中国索引学会成立大会而写

张琪玉

一、索引是开发、利用文献资源的一种重要工具

索引是对某种文献或某一文献集中所包含的各篇文章，或所讨论的各个局部主题，或所涉及的各种事项(如地区、人物、机构、事件、生物、矿物、产品、设备、公式、数据、著作等)以简明的方式分别著录标引，即确定其检索标识和指出其所在位置，并将款目按一定的可检顺序排列和组织，以方便检索的一种工具。

“索引”一词可以是：(1)指某种书刊的一个组成部分(不管它是否作为一种独立的著作出现)，摘记书刊中的知识单元或事项为条目，标明出处，并按一定次序编排，以方便查检该书刊内容的附属性资料，如各种专书、专刊索引；(2)指某种检索工具或某个检索系统的一个组成部分，以简明的方式提供与该检索工具的正文部分或检索系统的主文档部分不同的检索途径，如美国《化学文摘》的各种索引提供了与该文摘正文部分不同的多种检索途径；(3)指独立于某批书刊之外的一种简明检索工具，如各种群书索引(《十三经索引》等)、群刊索引(《全国报刊索引》等)和专题论文索引。第三种情况很难与文献目录截然区分。

索引作为一种检索工具的特点在于：(1)它是一种高深度标引的检索工具，以文献中的一个局部内容或事项，或期刊中的一篇作为一个标引单位；(2)它以极简洁的形式(一般仅有检索标识和出处)起指引的作用，而不是对文献进行登记和报道(报刊论文索引除外)；(3)它总是提供与书刊的目次或检索工具的正文(或检索系统的主文档)部分不同的检索途径；(4)它以便于查检的顺序编排其款目。

索引的功用在于：(1) 方便查检，可大大节约查阅文献的时间；(2) 可增加查全或查到所需资料的可能性；(3) 提供与文献正文不同的另一种甚至多种查检途径；(4) 浏览索引时，可使读者发现某些他所未想到的有用资料；(5) 某些书虽非工具书，有了索引，在一定程度上也可起到工具书的作用，其使用价值就可大大提高；(6) 一种检索工具或工具书配备了索引，可使它的检索功能或参考功能成倍增强，甚至可以说，某些检索工具的核心是它的索引体系；(7) 某些索引还有特殊功用，如引文索引、杂原子索引、化学结构索引、等同专利索引、某些古籍索引等。

所以，索引是开发、利用文献资源的一种重要工具。索引在现代人的研究、工作、学习、生活中的作用是十分重要的。索引法就是节省时间、提高效率的一种方法。

二、索引的研究和编纂早已被纳入许多学科和专业的范围

目前有许多学科在研究索引。

由于报刊索引和专题论文索引绝大多数是篇目索引，也可以说是论文目录，同时，图书目录也需要配备索引，以提高其检索功能，所以，目录学总是把索引法作为自己的一个组成部分，纳入自己的学科范围。

文摘刊物在过去一直是情报传递的主要工具，而文摘刊物要充分发挥其检索作用，必须配备索引，所以，索引法成为文摘学的一个重要组成部分，现代索引中许多新类型的出现，与文摘的发展有密切联系。

索引作为一种检索手段，作为检索系统的重要组成部分，其原理、结构、性能、检索效率等都是情报检索理论的研究对象。

研究情报检索系统语言保证问题的情报语言学，其主要的应用领域就是索引编制，所以，情报语言学也必须研究索引法，从而使索引法与情报语言学难分难解。作为情报语言学主要研究对象的情报检索语言在欧美使用的名称之一，就是“索引语言”。

索引作为图书的一个组成部分，也是图书编辑学的一项研究内容。

索引作为一种重要的检索工具，又是正在形成中的“文献检索与利用”学科不可缺少的研究内容。

图书馆、情报机构、出版社以及研究和教学单位的资料室等都把编制索引作为自己的专业工作内容之一。

三、数据库就是信息时代的索引

用计算机检索文献,是情报、图书馆、档案工作现代化的核心。国际联机检索使人们可以在世界各地任何装有检索网络终端的地方从上亿条的文献目录数据中以一二十分钟的时间找出任何问题的资料线索,并且如果需要,还可以在两三个星期内得到订购的原文复印件。

数以千计的各种类型的数据库是国际联机检索的支柱。

无论是文献目录数据库还是其他类型的数据库,一般都是依靠各种索引来查检的。有的数据库虽然没有索引,其实也是利用索引原理来查检使用的。可以说,没有索引法,也就没有数据库。而且,在数据库中,索引方法更是别出心裁,日新月异,其功能使人惊叹!

目前,全世界投入检索服务的数据库约五千个,至于一些单位或个人自建自用的数据库更是不计其数。从某种意义上讲,数据库就是信息时代的索引。

四、索引学作为一门新学科怎样认定自己的领域

尽管索引这一事物的出现已有不短的历史,索引方法和技术已有长足的进步,但索引学作为一门学科还是处于幼年时代。

关于索引的知识,过去称为索引法。“索引学”这个名词出现的时间还很短,还很少使用。到目前为止,还没有看到一部以索引学作为题名的专著。“索引学”一词甚至在词典中还找不到。从这个意义上说,索引学是一门新学科。

由上面的叙述可以看出,索引学在其出世之前,它的研究对象已被许多学科纳入自己的研究范围,索引学几乎没有其专辖的领土了。在这种情况下,索引学作为一门新学科该怎样认定自己的领域呢?

我认为,索引学既不可能把一切与索引有关的事物都归于自己的领域,也不要把自己的领域划定得太狭窄。应当承认,索引学与情报检索理论、情报语言学、目录学、文摘学、图书编辑学等学科是“你中有我、我中有你”的,无法划清界限,但索引学应当有自己的研究重点。

我认为,索引学的研究应包括索引原理、索引结构和设计、索引编制技术、文献微观标引、索引计算机化、索引法应用和索引使用法、索引评价、某些类型索引的专

门研究以及索引发展史等几个方面。其中应以索引这一事物的基本原理、基本方法和基本技术作为自己的研究重点,对各种类型的索引进行统一研究。有些与其他学科有交叉而且看来与其他学科关系更加紧密的部分,则不必再作为研究重点。

我们研究索引学,要放开视野,注意吸取相关学科的观点、原理和方法,为我所用。同时,索引事业在今天已不可能成为界限分明的独立事业(或许在过去也是这样),推进索引学研究和索引事业发展,需要各有关学科和专业共同努力,携手协作。我们的索引学会应当促进这种交流和协作。

五、索引学研究要密切结合索引实践

索引学确实只能算作一门新学科。目前,中文索引学专著仅《索引和索引法》、《古籍索引概论》、《索引的概念与方法》、《索引编制工作手册》等屈指可数的几种,就是真正研究索引法(而不是评介现有索引)的论文,数量也十分有限。

目前的情况是,索引实践走在索引学的前面,我们遇到许多很好的、新型的索引,却未见关于它们的编制方法和性能分析的详细记载。所以,索引学的第一个任务,就是要总结索引实践来丰富、充实自己。从学科理论建设的角度去分析研究现有的索引,从中提取基本的原理、方法和技术,应是索引学研究的重要课题。

既然各种索引方法都是适应索引编制任务的需要而被创造出来的,那么,使索引学研究与索引实践密切结合,无疑是推动索引学前进的有效措施。

国外索引比较发达,索引知识已有相当多的积累,特别是在索引工作现代化方面,已远远走在我国的前面,我们应当十分重视这一点,积极吸取国外的研究成果。

六、索引编制工作要讲求质量和效益

效益一般是指一项工作所投入的人力、物力、财力与所得的效果和利益之比。索引工作效益的绝对评价是索引产品所起的社会作用,即是否满足了社会对索引的需要。如果一种索引编制出来因种种原因而很少被人使用,那就是效益很小或者没有效益。

影响索引工作效益的主要因素有选题、质量、成本等方面,具体是:

1. 选题是否符合需要。这是首要的。选题不符合社会的迫切需要、大量需要或长期需要,则索引编制质量虽高,也不可能较大的社会效益。

2. 索引的收录范围。这是指第二类和第三类索引而言。如果收录率低,其使用价值就不会高。

3. 索引标引的质量。指标引深度、正确性、遗漏率、规范性、参照度等,这些都会影响索引的使用价值。

4. 索引结构或功能的完备程度。检索途径单一的索引,因其功能有限,一般来说总是效益较小的。

5. 索引的时效性。如果索引编制不及时,就会大大影响其社会效益,甚至毫无用处。

6. 索引的印制质量。包括版面设计、印刷装订质量等,对索引的使用价值也会有一定的影响。

7. 重复编制,特别是重复编制低水平的索引,是很大的浪费。

8. 索引编制技术是否先进,会影响索引编制的成本,同样也会影响索引工作的效益。

我国建国四十多年来编制的索引不少,但能发挥现实作用和具有长期使用价值的索引并不很多,这说明索引工作的效益不高,这是值得我们重视的问题。

七、索引选题要适应时代需要

索引选题具有决定意义,因为选题是索引工作能否产生效益以及效益大小的第一个环节。

选题要从三个方面考虑:

一是社会需要方面。虽然社会对索引的需要是多方面的,各行各业的工作、科学研究、学习、日常生活乃至文学欣赏等都需要索引,索引工作应当为多方面的社会需要服务,但这些需要总有个重要性大小的问题,有个需要量多少的问题,这是进行索引选题时必须考虑,必须进行调查研究的。总之,要根据现实需要、长期需要和社会大量需要,有计划、有比例地来确定选题,并尽量避免重复。在当前,索引选题尤应注意多为经济建设服务,多为科学技术进步服务,这是时代的需要。

二是索引对象(即文献)的价值方面。要多为使用价值大的文献编索引,不要为没有多大使用价值的文献编索引。

三是文献资源及其他条件方面。例如,一个某方面收藏不丰富的单位,要编好该方面的索引就比较困难。此外,人员素质和数量、设备、经费等,也是选题时应予考