

昌吉学院学术著作出版基金 资助

《回族文学》

Huizu Wenxue
Suoyin Xinxi Jiansuo
yu Yingyong

索引信息 检索与应用

任景华 著



中国地质大学出版社
ZHONGGUO DIZHI DAXUE CHUBANSHE



作者简介

任景华，新疆昌吉学院副研究馆员，1995—1998年就读于新疆师范大学外语学院，攻读英语教育专业，2007—2009年就读于武汉大学新闻与传播学院，攻读网络传播专业，获硕士学位，目前在武汉大学新闻与传播学院，攻读传播学博士学位。研究领域包括情报学、文献编目和传播学。已在国内图书情报类和新闻传播类核心期刊公开发表论文30余篇，主持教育部人文社会科学课题1项、省级课题2项、省教育厅课题1项、新疆昌吉回族自治州科技局计划课题2项、校级课题2项，此外，作为子课题负责人和参与者的国家和教育部课题有3项。

《回族文学》索引信息 检索与应用

任景华 著



中国地质大学出版社
ZHONGGUO DIZHI DAXUE CHUBANSHE

图书在版编目(CIP)数据

《回族文学》索引信息检索与应用/任景华著. —武汉:中国地质大学出版社, 2013. 12

ISBN 978-7-5625-3312-2

I. ①回…

II. ①任…

III. ①《回族文学》—索引

IV. ①Z89:I207.913

中国版本图书馆 CIP 数据核字(2013)第 307375 号

《回族文学》索引信息检索与应用

任景华 著

责任编辑: 阎 娟

责任校对: 张咏梅

出版发行: 中国地质大学出版社(武汉市洪山区鲁磨路 388 号) 邮政编码: 430074

电 话: (027) 67883511 传 真: 67883580 E-mail: cbb@cug.edu.cn

经 销: 全国新华书店 <http://www.cugp.cug.edu.cn>

开本: 880 毫米×1 230 毫米 1/32

字数: 288 千字 印张: 10

版次: 2013 年 12 月第 1 版

印次: 2013 年 12 月第 1 次印刷

印刷: 武汉三新大洋数字出版技术有限公司

印数: 1—500 册

ISBN 978-7-5625-3312-2

定价: 28.00 元

如有印装质量问题请与印刷厂联系调换

序

“这是一个最好的年代，也是一个最坏的年代；这是一个智慧的年代，也是一个愚昧的年代。”狄更斯在其小说《双城记》开篇语中这句宿命式的名言，似乎总适用于每一个发展变化中的大时代。网络的不断发展，信息的获取和利用成为人类赖以生存的本能。信息技术的高速发展，文本数据已经不再是人们日常生活中主要的信息来源，人们可以通过 Web 页面、数字图书馆、多媒体百科全书等载体共享获取，信息世界变得丰富多彩。我们每一个人面对海量信息，困惑于两种境况，一是自己想要的信息找不到，二是检索到的信息都是自己不想要的。再者就是每一个人都经常遇到这样的问题：我们如何在网络上找到自己想要的信息？有哪些工具可以利用？这些网络工具或电子资源如何使用呢？等等。故此，本书立足于介绍现代信息检索的方法和技术，在对信息检索基础知识进行全面阐释的基础上，重点介绍《回族文学》30 年的索引，为读者提供信息检索知识，同时展示地方特色刊物 30 年来的成果。

《回族文学》索引包括“分栏分类索引”和辅助索引“著者索引”两部分。“分栏分类索引”突出栏目特色，按栏目出现时间先后顺序编排，每个栏目下面依内容分类排列。“著者索引”按著者姓名的汉语拼音音序排列，以第一著者为主，国外译作将译者集中收录。《回族文学》索引的著录格式根据国家《检索期刊条目著录规则》进行著录。其著录项有：顺序号、文献题名、责任者、年、期、页码。

《回族文学》索引首次采用分栏分类的方法编目，从而反映出栏目设置、发生变化的过程，并保留了原刊栏目的特征。在索引信息检索时采用了检索语言表达各种概念及其相互关系，普遍地应用了概

念逻辑的原理,有效利用了“概念的划分与概括”和“概念的分析与综合”这两种方法来建立自己的结构体系。用检索语言将文献的内容特征和外表特征表述出来,形成信息标号,比如分类号、题名等,然后将标引记录存放在系统中,以供用户检索使用。

该书观点新颖,资料翔实,系统性和科学性是该书的质量保证。《〈回族文学〉索引信息检索与应用》注重理论与实践相结合,全面倡导昌吉州的本土文化,促进社会和谐发展。应用性是该书的主要特点。该书的出版将对弘扬昌吉本土文化,打造昌吉地域文化名片产生深远影响,尤为重要,对发展地方经济具有极其重要的现实意义和战略意义。该书的出版改善了学者们研究《回族文学》及其回族文化研究资料的匮乏困境。由于时间紧、任务重,书中的缺点与不足之处在所难免,尚需从事信息检索的同仁匡正。

目 录

第一章 绪 论	(1)
第一节 研究缘起	(1)
第二节 研究的目标和方法	(2)
第三节 全书的基本框架	(3)
第二章 信息检索	(5)
第一节 信息的含义与特征	(5)
第二节 信息检索的概念和类型	(8)
第三节 信息检索系统	(10)
第四节 信息检索技术	(12)
第三章 检索语言	(15)
第一节 检索语言的概念	(15)
第二节 检索语言的功能、要求和分类	(16)
第三节 分类检索语言	(21)
第四节 主题检索语言	(33)
第四章 索引信息检索	(37)
第一节 索引的概述	(37)
第二节 索引的创建	(40)
第三节 索引的使用	(47)
第五章 《回族文学》索引设计	(51)
第一节 《回族文学》索引——1979(创刊)至 1989 年	(51)
第二节 《回族文学》索引——1990 至 1999 年	(120)

第三节	《回族文学》索引——2000 至 2009 年	(179)
第四节	《回族文学》著者索引	(226)
第六章	《回族文学》索引检索系统	(302)
第一节	系统的总体设计	(302)
第二节	系统的实现及主要功能的使用	(302)
参考文献	(308)
后 记	(312)

第一章 绪 论

第一节 研究缘起

《回族文学》是昌吉地区乃至全国唯一一份以“回族”命名的文学期刊。《回族文学》创刊于1979年,原名《博格达》,1985年第一期更名为《新疆回族文学》,2000年第五期定名为《回族文学》,由《回族文学》杂志社编辑出版并在全中国发行。因其办刊的质量和影响力,2007年该刊获新疆优秀期刊奖称号。《回族文学》作为在昌吉本土创刊历史悠久、影响面广,特别是以培育回族作家为主体的期刊园地,理应成为昌吉乃至新疆学人课题研究的首选。透过《回族文学》这个窗口,可以触摸回族文化的沧桑更迭,触摸回族及其他少数民族、汉族几代作家的文学创作轨迹,正因如此,《回族文学》的资料性价值越来越引发研究者的兴趣,而借助于计算机科学对《回族文学》索引进行全面的整理和著录就显得尤为必要和迫切。

随着计算机技术、通信技术和网络技术的迅速发展,信息高速公路的建设与利用为大规模信息系统、图书馆自动化系统的发展提供了环境和条件。检索系统的开发和研究是典型的多学科综合的研究领域,其发展变化迅速,新思想、新技术、新应用等层出不穷。

检索系统不仅仅等同于一个拥有信息管理工具的数字化馆藏,它更是一个将收藏、服务和人融合到一起以支持数据、信息乃至知识的创造、传播、使用和保存整个生命周期的环境。文献检索系统是一种信息服务,在这种服务中,所有的信息资源都能利用计算机格式,通过使用数字化技术可实现信息的获取、存储、保存、检索、访问与显

示功能。

许多学科背景不同的人都在探究文献检索系统的内涵、意义和应用价值,也为检索系统的发展提供新的源泉、新的领域、新的方向。身处热点中心的图书馆学界更有责任深入、系统地研究检索系统的理论和实践问题。近三年来,在中国科技文献学会的组织和推动下,检索系统的建构在我国得到了越来越多人的关注。特别是年轻学者,敏锐地抓住了这一发展动态,进行了具有不同特色的探索和研究,并且产生了一批重要的研究成果。这对活跃我国文献学的学术气氛,促进文献学理论研究的深入,都很有益处。

第二节 研究的目标和方法

一、主要研究目标

首先,从信息理解的角度来探讨检索系统的问题,更有利于深入揭示检索系统的内涵和本质;其次,把检索系统的研究扩展到任意信息集合(不限于网站),系统地探讨和阐述基于检索系统的过程和方法,对建设《回族文学》检索系统和篇目的索引的著录格式有重要的参考价值;再次,采用“分栏分类索引”和辅助索引“著者索引”的形式进行编辑,全面完整地将《回族文学》30 年来的期刊第一次编制成索引,为本领域的理论建设做了很有意义的工作。

通过网络提供高速横向跨库连接的多媒体信息存取服务,促进社会各类信息高效、经济地传递,从而极大地方便读者的学习、交流和生活。它将为 90% 的在校大学生和 100% 的教师、科技研究者提供信息服务。

信息标引是一个群体行为,只有共同依据检索语言,才能保证标引信息的一致性。

本书概括出了信息构建的一般过程,即由信息检索的基础理论、

概念设计、数据、检索语言及方法、检索技术设备五个阶段组成的完整流程。检索系统的进步,尤其是英特网的发展,带来了信息传播与存储方式的极大变革,对信息检索产生了深刻的影响。《回族文学》检索系统的研究与开发注重理论与实践相结合,全面倡导昌吉州的本土文化,促进社会和谐发展。

二、主要研究方法

本书主要采用了跨学科研究方法、文献调研法、分析综合法、模型法、分析归纳法开展研究,对不同的研究对象分别用其中一种或几种方法来进行研究,力求获得科学、客观的结论。

采用跨学科研究方法来研究理解信息,检索语言,索引这些问题时,分别利用了诠释学、信息学、符号学等学科对该问题的研究成果;文献调研法主要用于《回族文学》30 年期刊中的每一期的“分栏分类索引”和辅助索引“著作索引”两部分进行研究;采用信息抽样法和模型法建立了信息检索系统。

第三节 全书的基本框架

第一章为绪论,主要介绍本书研究的背景与意义,以及本书的研究目标、研究方法和主要框架。

第二章论述信息、信息检索的基本概念,阐释了信息检索系统的类型,信息检索包括的主要要素;介绍了信息检索最常用的方法:布尔逻辑检索。

第三章论述检索语言的建构问题。作者从四个方面展开检索语言的研究,即检索语言的概念、检索语言的功能和分类、分类检索语言和主题检索语言。

第四章论述索引信息检索的基本概念、索引信息检索、索引的设计和使用、数据库索引。

第五章论述《回族文学》索引设计方案,分三个阶段,每一阶段为10年,分别为1979至1989年、1989至1999年、1999至2008年分栏分类的编目。

第六章论述回族文学检索的功能,教会读者使用检索《回族文学》的检索使用方法。

第二章 信息检索

第一节 信息的含义与特征

进入 21 世纪以来,通信技术、计算机技术和网络技术飞速发展,信息日益成为社会各个领域关注的焦点研究问题。信息已经渗透到社会生活各个领域,成为人类社会不可缺少的重要资源。每一个劳动者在从事创造性研究过程中,运用观察、比较和推理的能力首先来自于广泛获取文献信息,才能在前人已经取得的成就基础上进行新的探索。《回族文学》30 年来的“分栏分类索引”和辅助索引“著者索引”检索系统的建构方便了广大读者的检索。信息在这里起到了关键性的作用。本章介绍了信息的内涵、信息检索的概念和类型、信息检索系统以及最常用的一种检索方法:布尔逻辑检索。通过这些信息检索的基本知识学习和掌握一种常用的检索方法以了解信息检索,从而更好地帮助自己获取所需信息。

一、信息的含义

信息作为现今社会使用频率最高的词汇之一,显现着时尚,蕴涵着古老。人类很早就开始了信息活动,古波斯人塔顶设置的“喊话站”和古罗马人的“悬灯”讲述着古老信息活动的故事。我国远古的“结绳记事”和殷商的“烽火告警”,同样展示着信息存储和传递的方式。

对信息的概念进行科学探讨开始于 20 世纪 20 年代。1928 年,哈特莱(R. V. L. Hartley)在《信息传输》中将信息理解为“选择通讯

符号的方式”。1948年,信息论创始人申农在《通讯的数理理论》中将信息定义为“信息是用以消除随机不确定性的东西”。1950年,控制论的创始人维纳将信息的概念引入了控制论,他在《人有人的用处——控制论与社会》中指出:“人是通过感觉器官感知外部世界”,“我们支配环境的命令就是给环境一种信息”。维纳在《控制论》一书中指出:“信息就是信息,既非物质,也非能量。”王春梅认为信息是客观事物存在方式及其运动规律、特点的外在表现。该定义言简意赅,基本概括了信息的一些特征,其强调的是信息的外在表现。

目前对信息这个概念描述有很多,但是却不能涵盖信息的本质特征。信息是事物的一种普遍属性,可用一个简单的词说明信息的本质:记录。具体来讲:信息是事物在相互作用中所“刻画”出的记录。它可以编码、存储、传递、加工、分享,而其意义是不变的,这一表述可以说明信息产生、存在、运动、产生进一步影响等多种问题。是按照事物的发展而变化的。当人们一提到信息时往往潜意识里就认为是一个概念群。因为,信息本身不是一个孤立的概念,它是一个动态性的,你可以理解成一个名词、动词或是形容词,因为信息本身就包含很多语义。

关于信息的定义,目前在学界还没有一个统一的定论,不同领域的研究者从不同角度出发,对信息有不同的理解和认识。结合信息检索的特点,我们认为信息的含义有广义和狭义之分,广义的信息指自然界和一切人类活动所表达出来的信号和消息,是事物表现的一种普遍形式;狭义的信息指经过搜索、记录、处理和储存的可供检索的文献、数据和事实。

综合以上的文献综述及先前研究者的探索,笔者认为信息经典的定义有两种——申农和维纳的表述。本书对信息的定义为:信息是以物质介质为载体,传递和反映世界各种事物的存在方式和运动规律的表征。它是一种状态,因为有些信息是人们创造出来的,如政策法规、条款、数字的统计等。

二、信息的特征

信息的特征是指信息区别于其他事物的属性。

尽管从不同的角度出发对信息存在不同的定义,但是信息的一些基本属性还是得到了共识。

(1)普遍性:只要有事物的地方,只要有事物运动,就必然地存在信息。信息在自然界和人类社会活动中广泛存在。

(2)客观性:信息是客观存在于事物和运动状态的方式表现,信息是客观现实的反映,不随人的主观意志而改变。如果人为地篡改信息,那么信息就会失去它的价值,甚至不能称之为“信息”了。

(3)动态性:信息是一个动态的变化,它是随着事物不断变化发展的,使得信息的内容、形式、容量都会随着时间的变化而改变。

(4)时效性:由于信息的动态性,一个固定的信息的使用价值必然会随着时间的流逝而衰减。因此只有及时地将信息收集、加工,才能充分利用信息。

(5)可识别性:人类利用信息的前提是通过感官和现代技术等方式获取、整理和认知信息。

(6)可传递性:信息可以跨越时空和地域进行无限传输,通过各种媒介在人与人、人与物、物与物等之间传递。

(7)可共享性:信息能够同时为多个使用者所利用,同一信息可以在同一时间被多个主体共有,而且还能够无限地复制、传递。这是信息与物质、能量显著不同的地方。

第二节 信息检索的概念和类型

一、信息检索的概念

信息检索起源于 19 世纪下半叶,到了 20 世纪 40 年代,索引和检索已成为图书馆参考咨询的一个独立部门为用户进行信息检索服务。计算机技术的发展使传统的信息检索逐步进入计算机领域并与信息检索理论紧密结合起来;脱机批量情报检索系统、联机实时情报检索系统、文献信息检索相继研制成功并商业化。20 世纪 60 年代到 80 年代,在信息处理技术、通讯技术、计算机和数据库技术的推动下,信息检索在教育、军事和商业等各领域高速发展,得到了广泛的应用。Dialog 国际联机情报检索系统是这一时期的信息检索领域的代表,至今仍是世界上最著名的系统之一。在网络环境下,传统的情报检索研究和文献检索研究紧密结合。信息检索研究就更具兼容性的概念。

信息检索又称情报检索,有广义和狭义之分。从广义的角度讲是指将信息按一定的方式进行组织、加工、整理和存储起来,再根据信息用户的需要找出有关信息的过程。或者是指信息存储通过对大量无序信息的选择和收集、著录、分类和标引等方法加以处理,使得信息有序化、系统化,并按照一定的技术进行编制建立信息检索系统,使之成为有序的信息集合过程。例如,查找有关“突发事件信息”方面的文献,就属于信息检索。信息检索是最主要、最基本的形式,它是一种相关性的检索,是给用户所需相关信息的文献,取舍依据用户的需求。传统的手工信息检索有一系列的信息线索,如文献的题名、作者、出版者、出版日期、文献收藏地址等信息,用户可以通过这些信息索取所需的文献资料。

狭义的信息检索是把已经收集到的信息储存好,再从信息集合

中找到信息用户所需的相关信息查找过程。狭义的信息检索包括 3 个方面的含义:了解用户的信息需求、信息检索的技术或方法、满足信息用户的需求。具体来说,狭义的信息检索指通过一定的方法,从已经存储的信息中检索出与用户需求相关的文献、数据和事实过程。

本书所讲的信息检索主要是指查找信息的方法和手段,使用户能从海量的信息中迅速、准确、全面地查找到所需信息的过程。可以说信息检索对人们的学习、生活和工作等方面都有很大的作用。

二、信息检索类型

(一)按信息检索对象分类

(1)数据信息检索(Date Information Retrieval)。数据信息检索是以特定的数值型或非数值型数据为检索对象的信息检索。例如,查找电话号码、银行账号、年度报表、各种统计数据。数据信息检索一般利用的工具是年鉴、百科辞典、名录、指南等。

(2)事实信息检索(Fact Information Retrieval)。事实信息检索是指用户为了确定某一个事件或事实等方面的信息而进行的检索,一般指某一具体的事实,知识具体且狭小。例如,查找什么是新闻学?什么是传播学?这都是事实信息检索,常用的工具有词典、辞海、手册、名录、百科全书以及对应的数据库和网络资源。

(二)按信息检索方式分类

(1)手工检索(Manual Retrieval)。手工检索是一种传统的信息检索方法,即在图书馆使用卡片目录检索信息,较直观,不需要借助设备即可检索信息,但是检索速度很慢,查全率较低。

(2)计算机检索(Computer-based Retrieval)。计算机信息检索又称机检。又分脱机检索、联机检索、网络信息检索。计算机检索是指利用计算机系统,从存储于计算机的大量数据中检索出用户所