

实用检索技术

(第二版)

汪楠 成鹰 主编



科学出版社

实用检索技术

(第二版)

主 编 汪 楠 成 鹰
副主编 李佳洋 肖 倩 李海波
主 审 张 炎

科学出版社

北 京

内 容 简 介

本书是通用性很强的有关信息检索方法和技能的著作。书中阐述了在现代信息技术环境下,信息检索的相关概念、基本原理及常用技术和方法,从多种信息资源检索出发,详细介绍了国内外常用的学术性文献数据库及其他信息资源数据库,系统介绍了国内外著名检索工具的结构和使用技巧,阐述了信息检索的意义和信息资源再生利用的方法,同时,借助大量检索范例,帮助读者有效地掌握现代网络环境下数字化信息资源的检索和利用。

本书结构合理、内容全面、重点突出、注重实用,既可作为高等院校信息检索课程的教学用书,又可作为教学、科研、工程技术和社会各界信息用户检索信息的参考书。

图书在版编目(CIP)数据

实用检索技术/汪楠,成鹰主编. —2版. —北京:科学出版社,2012
ISBN 978-7-03-036003-8

I. ①实… II. ①汪… ②成… III. ①情报检索-高等学校-教材
IV. ①G252.7

中国版本图书馆 CIP 数据核字(2012)第 267805 号

责任编辑:石 悦 胡云志 / 责任校对:钟 洋
责任印制:闫 磊 / 封面设计:华路天然设计工作室

科 学 出 版 社 出 版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

保定市中国画美凯印刷有限公司印刷

科学出版社发行 各地新华书店经销

*

2010 年 8 月第 一 版 开本:787×1092 1/16

2012 年 10 月第 二 版 印张:19 1/2

2012 年 10 月第三次印刷 字数:451 000

定价:34.00 元

(如有印装质量问题,我社负责调换)

第二版前言

当今社会既是一个信息时代，又是一个网络时代。科学技术突飞猛进、社会竞争日益激烈、知识的无限性与个人能力的有限性的社会现实，已经摆在现代人的面前。如何在纷繁复杂的信息海洋中查找到适合的信息，掌握现代信息检索理论和信息检索技术的全新应用是人们的首选。

由于信息社会的网络信息媒体层出不穷，信息检索技术不断更新，新兴的信息检索理念和方法不断涌现，为了适应社会的变化和教学的需要，本书的作者在编著了《信息检索方法与实践》、《数字信息检索》和《实用检索技术》之后，根据读者的反馈，推出了又一新作《实用检索技术（第二版）》。

本书在《实用检索技术》的基础上，本着充实新内容、体现新进展、完善新体系的原则和目标，对前版教材进行了全面的修订和改版。新版教材以信息资源的收集、分析、利用和再生为主线，将信息检索的知识体系进行了重组，使读者在掌握信息检索技术的同时，能更清晰地理解信息资源利用的意义，为终身学习奠定基础。

本书具有系统性、实用性、新颖性的特点，从检索的基本原理出发，针对不同种信息资源的检索需求，注重检索方法和常用检索工具的介绍，并为读者提供了大量的检索实例，详细介绍了检索分析、检索策略的制订、检索结果的甄别及检索系统评价等方法，同时，进一步为读者提供了信息资源的收集、整理及科技文献的写作方法。

本书共8章，包括信息检索基础、常用中文数据库检索系统、网络信息资源检索技术、特种文献检索、国外科技信息资源、综合信息资源、信息资源的利用、文献信息服务。本书融入了新的理念，注重实用，启发思维，理论与案例分析相得益彰，相对前版教材更加突出了检索实践环节，为读者提供了丰富的实训项目。每章设置了“本章小结”、“关键术语”、“知识窗”、“小提示”、“检索实训”等，为学生学习信息检索、掌握信息检索技能提供了较好的学习方法。

本书由汪楠、成鹰担任主编，李佳洋、肖倩、李海波担任副主编，张炎完成全书的审稿工作。王妍、方新儒、于洁、曹辉、潘峰、刘英英参与了本书的校对和审读。

本书在编写过程中借鉴了许多网上资源和相关学者的研究成果，在此谨向相关机构和作者致以诚挚的谢意。

由于作者的水平有限，书中疏漏和不足之处，恳请专家和读者批评指正。

作 者

2012年8月

第一版前言

当今以计算机、网络技术、通信技术为代表的信息技术迅猛发展，加快了信息交流与传递的速度，扩展了人们可利用的信息空间，使我们完全进入一个经济信息化、社会信息化的知识经济时代，信息日益成为社会各领域最活跃、最具决定意义的因素。“网络时代”、“信息高速公路”、“社会信息化”、“知识经济”、“网络经济”、“科技创新”、“数字化时代”等新概念层出不穷。

信息社会对人的挑战表现为对人的综合能力的挑战，而信息素质是信息社会中人的综合能力的重要组成部分。处于信息社会之中的国民是否具备相当的信息素质和掌握足够的信息技术，已成为影响一个国家竞争力的重要方面。信息素质是一种能够发现信息需求、查寻信息、分析判断信息、加工信息、筛选信息、综合利用信息、创造信息的各方面能力的总和，提高了人们的信息素质就培养了人独立自主学习的态度和方法，使之具有追求新知识、运用新信息的能力，从而善于从瞬息万变的事物中捕捉信息、创造新信息。信息素质已成为每个社会成员的一种基本生存能力，更是学习型社会及终身学习的必备素质。信息素质的培养日益成为世界各国教育界乃至社会各界所关注的重大课题。

信息资源是人类进行学习、生活、科学研究和事业发展的基础，而信息资源的有效存储和检索作为人们获取信息资源的主要手段，也已经成为社会所有成员必须掌握的基本知识和技能。目前，“掌握资料查询、文献检索及运用现代信息技术获取相关信息的基本方法”已被规定为大多数专业的人才培养要求。可见，加强信息检索课的教学，增强学生的信息意识，重视培养学生的收集处理信息的能力、获取知识的能力，已经成为当今人才培养的重要任务之一。

在科学技术突飞猛进、文化教育蓬勃发展、社会竞争日益激烈的今天，知识的无限性与个人能力的有限性、课堂教学的局限性与社会需求的多样性、教育的时滞性与生活的现实性等矛盾，已经摆在现代人的面前。

《信息检索方法与实践》、《数字信息检索》自问世以来，受到社会各界的欢迎。由于网络信息媒体的不断出新，新兴的信息检索技术不断涌现，为了适应变化和教学需要，作者对该书内容作了很大的修改和调整，并将书名改为《实用检索技术》。

全书共分9章，内容包括信息检索基础、常用中文数据库检索系统、网络信息资源检索技术、特种文献检索、国外信息检索资源、综合信息检索、科技信息资源、信息资源的应用和检索案例选编。本书具有系统性、实用性、新颖性的特点，从检索基本原理出发，针对不同种类信息资源的检索需求，注重检索方法和常用检索工具的介绍，并为读者提供了大量的检索实例，详细介绍了检索分析、检索策略的制订、检索结果的甄别以及检索系统评价等方法，同时，进一步为读者提供了信息资源的收集、

整理以及科技文献的写作方法。

本书由汪楠、张炎和成鹰编写。汪楠完成全书的统稿工作。王妍、方新儒、章翔、于洁参与了本书的校对和审读。

本书在编写过程中参考了大量的文献资料，借鉴了许多网上资源，如万方数据资源、中国知网、维普《中国科技期刊数据库》系统等，在此恕不一一注明，谨向相关机构和作者致以诚挚的谢意。

由于作者的水平有限，书中疏漏和不足之处在所难免，恳请专家和读者批评指正。

编 者

2010年6月

目 录

第二版前言

第一版前言

第 1 章 信息检索基础	1
1.1 信息与文献	1
1.2 信息检索概述	7
1.3 检索语言	10
1.4 检索工具	14
1.5 信息检索程序	19
1.6 计算机信息检索基础	23
1.7 信息检索的基本方法	27
本章小结	29
检索实训	31
第 2 章 常用中文数据库检索系统	33
2.1 万方数据资源系统	33
2.2 中国知网	46
2.3 维普资讯网	59
2.4 检索实例	64
本章小结	69
检索实训	71
第 3 章 网络信息资源检索技术	73
3.1 网络信息资源概述	73
3.2 全文检索型搜索引擎	79
3.3 地图搜索	87
3.4 移动搜索	91
3.5 图像搜索	98
3.6 搜索引擎选择	103
3.7 检索实例	107
本章小结	111
检索实训	115
第 4 章 特种文献检索	117
4.1 专利文献	117
4.2 中文专利检索	127
4.3 标准文献	133

4.4 标准文献检索	138
4.5 检索实例	146
本章小结	149
检索实训	151
第5章 国外科技信息资源	153
5.1 美国《工程索引》	153
5.2 美国《科学引文索引》	160
5.3 英国《科学文摘》	169
5.4 美国《化学文摘》	173
5.5 检索实例	178
本章小结	189
检索实训	191
第6章 综合信息资源	193
6.1 经济信息检索	193
6.2 法律、法规信息检索	204
6.3 文学、历史、地理信息检索	208
6.4 教育、心理信息检索	210
6.5 美术、体育信息检索	213
6.6 检索实例	222
本章小结	225
检索实训	227
第7章 信息资源的利用	231
7.1 信息资源的收集及分析	231
7.2 信息资源的利用	236
7.3 信息资源研究的评价	244
7.4 检索实例	252
本章小结	256
检索实训	259
第8章 文献信息服务	261
8.1 文献信息服务系统	261
8.2 图书资源检索	268
8.3 中文全文数字图书馆	272
8.4 国内科技信息资源	285
8.5 科技查新	289
8.6 检索实例	294
本章小结	296
检索实训	299
参考文献	301
附录 模拟试卷	302

第 1 章 信息检索基础

学习目标

1. 了解信息的特征、信息检索的概念、信息检索的类型、信息检索的目的、信息检索工作。
2. 理解检索工具的结构及目录、索引、文摘三种主要检索工具的区分及作用。
3. 熟悉检索语言，熟悉分类语言、主题语言的分类方式及存在价值，熟悉规范化语言与非规范化语言的差别。
4. 掌握信息检索的程序。

随着信息时代的到来以及全球市场的形成，信息的瞬息万变给人类社会带来了前所未有的机遇与挑战。在这样多变的环境下，如何利用具有潜在价值的数据资产，及时、准确地获取和传播有价值的信息，是人类社会存在和发展的前提条件之一。

1.1 信息与文献

未来学家托夫勒把人类社会的发展过程划分为三个阶段：以农业经济为基础的农业社会；以工业经济为基础的工业社会；以信息经济为基础的信息社会。在信息社会中，信息资源已被人们公认为是物质资源、能量资源之后的人类可利用的第三类资源，是人类社会领域的基础性资源。

1.1.1 从信息到文献

在当今的信息社会中，只有将自然现象和社会现象的信息上升为对自然和社会发展规律的认识，这种再生信息才能在社会和人类生活中发挥日益重要的作用。

1. 信息 (information)

美国数学家、控制论的主要奠基人维纳在 1950 年出版的《控制论与社会》中是这样来理解信息的：“人通过感觉器官感知周围世界”，“我们支配环境的命令就是给环境的一种信息”，因此，“信息就是我们在适应外部世界，并把这种适应反作用于外部世界的过程中，同外部世界进行交换的内容的名称”。“接收信息和使用信息的过程，就是我们适应外界环境的偶然性的过程，也是我们在这个环境中有效地生活的过程”。

从哲学的观点看，“信息”是物质的一种普遍的属性，它反映不同物质所具有的不同本质、特征以及运动状态和运动规律。它是用来消除人们对客观物质不确定性认识的东

西,“信息”是人们认识事物获取新知识的唯一方式。只有掌握了事物的信息,人们才能对事物的运动规律进行调控,人们在认识世界和改造世界过程中才能处于主动地位。

中国国家标准《情报与文献工作词汇——基本术语》(GB/T4894—1985)中信息的定义为:信息是物质存在的一种方式、形态或运动状态,也是事物的一种普遍属性,一般指数据、消息中所包含的意义,可以使消息中所描述事件的不确定性减少。

信息具备以下基本属性。

(1) 客观性。信息的客观性是指信息是事物本质特征和客观规律的表征,是一种客观存在,不以人们的主观意识为转移。

(2) 传递性。信息在事物之间的相互联系必定在信息的流动中发生。信息的传递性表现在人与人之间的信息交换,人与计算机之间的信息交换,动物界及植物界之间的信号交换,同时,人类进化过程中的细胞选择、遗传也是信息的传递与交换。

(3) 多态性。信息在不同的领域具有多种不同的特性或表现形式,如:客观事物中的各种自然属性;人工设备的技术特征;人类社会的各种社会特征;人脑中反映客观事物认识的思想、知识;人类交流信息过程中的声音、文字、图像以及用各种编码形式记录下来的数据、新闻、情报、消息等。各种形式的信息又常常以综合的方式表现事物的特征,所谓“多媒体”正是信息多样性和综合性的集中表现。

(4) 共享性。同一个信息资源可以被不同用户在不同地点、不同时间中同时利用,不需要任何限制条件,信息量也不会损失。信息不会像物质一样因为共享而减少,反而可以因为共享而衍生出更多信息。

(5) 时效性。信息的时效性是指信息的效能依赖于时间,它既表明了信息的时间价值,又表明了信息的经济价值和社会效益。客观事物不断变化而产生的信息具有很强的时效性。客观事物变化越快,信息的时效性就越强。时间的延误,就是扼杀信息的生命——导致信息价值的衰减,乃至完全消失。时过境迁的信息是毫无价值的。

(6) 价值相对性。由于人们认识的能力与认知的条件不同,信息的接受者获得的信息及信息量也就不同,从这个意义上讲,信息的价值是有相对性的。

(7) 依附性。信息必须通过一定的符号,依附于某一载体中才能被表现。没有载体就没有信息。

(8) 增长性。物资和能源在使用后会有消耗,而信息的使用不但不能减少信息的数量,还能产生更多的信息。

2. 知识 (knowledge)

知识是人们在改造世界的实践中所获得的认识 and 经验的总和,是对客观世界物质形态和运动规律的认识。根据韦伯斯特 (Webster) 词典 1997 年的定义,知识是通过实践、研究、联系或调查获得的关于事物的事实和状态的认识,是对科学、艺术或技术的理解,是人类获得的关于真理和原理的认识的总和。

从信息的观念看,知识来源于信息,是信息的一部分。简而言之,知识是理性化、优化、系统化了的的信息。从反映的内容而言,知识是客观事物的属性与联系的反映,是客观世界在人脑中的相对正确的反映。从反映的活动形式而言,知识有时表现为主体对事物的感性知觉或表象,属于感性知识;有时表现为关于事物的概念或规律,属

于理性知识。

知识是人们在实践活动中获得的关于世界的最本质的认识,是对信息的提炼、比较、挖掘、分析、概括、判断和推论。一般而言,知识具有共享性、传递性、非损耗性(可以反复使用,其价值不会减小)及再生性等特点。

3. 情报 (intelligence)

情报是作为交流对象的有用的知识。“情报是判断、意志、决心、行动需要的能指引方向的知识 and 智慧”、“情报是解决问题所需要的知识”、“情报是激活了的知识”。情报的基本属性是知识性、传递性和效用性。

(1) 知识性。一般而言,情报是一种新的知识。然而称为情报的知识并不排斥时间上的“不新”。这就是说,凡原先不知道的(不论时间多久),而现在又迫切需要的知识,都属于“新”的知识。

(2) 传递性。情报传递的过程包括从情报源进行收集到向用户提供情报的全过程。无论多高深、渊博、新奇的知识,如果不记录、不传递、不交流,都不能称为情报。

(3) 效用性(判断的标准)。情报是一种有价值、有效用的知识。情报均须“有用”,当用户不需要时,百分之百的知识也不能称为情报。

4. 文献 (literature/document)

国家标准《文献著录总则》(GB/T3792.1—1983)给出的定义为:“文献是记录有知识的一切载体。”国际标准化组织 ISO 在《文献情报术语国际标准》(ISO/DIS5217)给出的定义是:“为了把人类知识传播开来和继承下去,人们用文字、图形、符号、音频、视频等手段将其记录下来,或写在纸上,或晒在蓝图上,或摄制在感光片上,或录到唱片上,或存储在磁盘上。这种附着在各种载体上的记录统称为文献。”

从定义中不难看出文献包括四个要素。

(1) 知识和信息,即文献的内容。

(2) 物质载体,如竹简、纸张、胶片、磁盘等,它是文献的外在形式。

(3) 记录知识和信息的符号,有文字、图表、声音、图像等。

(4) 记录的方式或手段,如书写、印刷、复制、录音、录像等,即将知识和信息固化到载体上的手段。

文献是记录有知识和信息的一切载体,或称为固化在某种物质载体上的知识和信息,当文献中记录的知识传递给用户,并为用户所利用时,文献中的知识就转化为情报。因此,文献是记录、积累、传播和继承知识的最有效手段,是人类社会活动中获取情报的最基本、最主要的来源,也是交流传播情报的最基本手段,现在通常理解为图书、期刊等各种出版物的总和。

5. 信息、知识、情报、文献

生活中信息无处不在,无时不有,它们是知识产生的原料,这些原料经过人脑接收、处理后,成为系统化的信息,知识由此产生。从知识管理的层面上看,知识是指可直接用于行动的信息,它使人们可以随时随地作出正确决策。

目前,我国图书情报学界对信息、知识、文献和情报的看法是:知识是对信息加工、吸收、提取、评价的结果,即系统化的信息成为知识,知识记录下来成为文献,文献经传递并加以应用成为情报,情报体现了人运用知识的能力。

1.1.2 信息的分类

信息的分类没有固定的标准,可以根据不同需求从多个角度进行划分。

1. 按信息的内容分类

按信息的内容可分为科学技术信息和社会科学信息。科学技术信息包括自然科学信息、工程技术信息等;社会科学信息包括政治信息、军事信息、经济信息、技术经济信息、社会生活信息等。

2. 按产生信息的客体的性质分类

按产生信息的客体的性质可分为自然信息(瞬时发生的声、光、热、电,形形色色的天气变化,缓慢的地壳运动,天体演化等)、生物信息(生物为繁衍生存而表现出来的各种形态和行为,如遗传信息、生物体内信息交流、动物种群内的信息交流)、机器信息(自动控制系统)和社会信息(人类)。社会信息是指人与人之间交流的信息,既包括通过手势、身体、眼神所传达的非语义信息,也包括用语言、文字、图表等语义信息所传达的一切对人类社会运动变化状态的描述。

3. 按信息所依附的载体分类

按信息所依附的载体可分为文献信息、口头信息等。文献信息,就是以文献为载体的信息,以文字、符号、声像信息为编码信息,是经人们筛选、归纳和整理后记录下来的,它与人工符号本身没有必然的联系,但需要通过符号系统实现其传递。口头信息是原始的没有形成文字、符号、声像的信息,没有文献信息易于流传。

1.1.3 文献的分类

文献信息是信息存在的一种重要形式,其优点是易识别、易保存、易传播,使人类精神文明千古流传。文献是信息的主要载体,大多数规范化的重要信息都以文献形式出现,所以文献的分类非常重要。

1. 按加工层次分类

文献是信息的主要载体,根据对信息的加工层次可将文献分为一次文献、二次文献和三次文献。

(1) 一次文献(primary literature)。人们对自然信息和社会信息进行首次加工而成的文字记载,这是文献信息源的主要部分。例如,专著、报纸、期刊、专利文献、标准文献、会议文献、样本等成品文献属于一次文献。一次文献数量极为庞大,由于在内容上是分散的、无系统的,不便于管理和传播。

(2) 二次文献(secondary literature)。为了控制文献,便于人们查找,对一次文献进行再加工,通过整理、提炼和压缩,并按其外部特征(题名、作者、文献物理特征)和内部特征进行有序化管理,形成另一类新的文献形式就是二次文献。例如,目

录、书目、索引、文摘等。二次文献不是一次文献本身的汇集，而是一次文献特征的汇集，通过它们可以很方便地找到一次文献，或了解一次文献的内容。二次文献内容相对集中、系统性强、便于管理和传播。

(3) 三次文献 (tertiary literature)。利用二次文献，选择有关的一次文献再加以分析、综合而编写出来的第三个层次的文献形式就是三次文献。例如，专题报告、综述、进展以及手册、百科全书、年鉴等工具书。三次文献具有系统性、综合性、知识性和概括性的特点，它从一次文献中汲取重要内容，便于高效率地了解某一领域的状况、动态、发展趋势和有关情况。因此，要在浩瀚的一次文献中查找所需资料，往往离不开二次文献和三次文献。

2. 按文献的载体分类

随着信息记录与存取技术的发展，文献载体形式呈现多样化，如音像磁带、缩微胶卷、光盘等，这些非纸型文献的出现使文献的范围进一步扩大；使文献的生产和传递更加迅速；使知识、信息的存储和利用更加便捷。目前，文献主要有印刷型、缩微型、电子型（机读型）、音像型四种。

(1) 印刷型文献 (paper literature)，是以手写、打印、印刷等为记录手段，将信息记载在纸张上形成的文献。它是传统的文献形式，便于阅读和流传，但存储密度小、体积大，不便于管理和长期保存。

(2) 缩微型文献 (microform literature)，是以感光材料为载体、用缩微照相技术制成的文献复制品，如缩微胶卷、缩微平片。按其外形可分为卷片型和平片型；按对它的穿透力可分为透明体和不透明体，前者需用透射式阅读机阅读，后者指缩微印刷品，用不透明感光纸印制而成，用反射式阅读机阅读。缩微型文献的特点是存储密度大、体积小，便于保存和传递，但必须借助专门的设备才能阅读。世界上许多文献信息服务机构都将欲长期收藏的文献制成缩微品加以保存。

(3) 电子型文献 (electronic literature)，是指以数字代码方式将图、文、声、像等信息存储到磁、光、电介质上，通过计算机或类似设备阅读使用的文献，也称机读型文献，如各种电子图书、电子期刊、联机数据库、网络数据库、网络新闻（如 usenet, mailing list, BBS）、光盘数据库等。其特点是信息存储量大，出版周期短，易更新，传递信息迅速，存取速度快，可以融文本、图像、声音等多媒体信息于一体，信息共享性好，易复制，但必须利用计算机才能阅读。

(4) 音像型文献 (audio-visual literature)，是采用录音、录像、摄影、摄像等手段，将声音、图像等多媒体信息记录在光学材料、磁性材料上形成的文献，又称为声像资料、视听资料、音像制品，如音像磁带、唱片、幻灯片、激光视盘等。音像型文献脱离了传统的文字记录形式，直接记录声音和图像，给人以直观的感觉，又称为直感型资料。其特点是形象、直观，尤其适于记录用文字、符号难以描述的复杂信息和自然现象，但其制作、阅读需要利用专门设备。

3. 按文献的出版形式分类

按照文献的出版形式和内容，可以将文献分为图书、期刊、报纸、特种文献（学位

论文、会议论文、专利文献、标准文献、科技报告、政府出版物、产品样本资料等)。

(1) 图书。联合国教科文组织对图书的定义是:凡由出版社(商)出版的不包括封面和封底在内 49 页以上的印刷品,具有特定的书名和著者名,编有国际标准书号 (ISBN),有定价并取得版权保护的出版物称为图书。

图书是以传播知识为目的,用文字或其他信息符号记录于一定形式的材料之上的著作物;图书是人类社会实践的产物,是一种特定的不断发展着的知识传播工具。它包括专著、教科书、词典、丛书、工具书、百科全书等。

知识窗

ISBN 号

国际标准书号 (International Standard Book Number, ISBN) 是国际通用的图书或独立的出版物代码。一个国际标准书号只有一个或一份相应的出版物与之对应。国际标准书号由 13 位数字组成。前三位数字代表图书,中间的 9 个数字分为三组,分别表示国家代码、出版社代码和书序码,最后一个数字是校验码,从 0 到 9 或 X。

格式举例:《科技信息检索》,ISBN 为 978-7-030-26151-9。

(2) 期刊。期刊也称杂志,是由多位作者撰写的不同题材的作品构成的定期出版物。期刊有固定刊名,是以期、卷、号或年、月为序,定期或不定期连续出版的印刷读物,每期的内容不重复。期刊出版单位出版期刊,必须经新闻出版总署批准,持有国内统一连续出版物号 (ISSN)。

根据期刊的出版周期可将期刊分为旬刊、半月刊、月刊、双月刊、季刊、半年刊、年刊。期刊按用途不同可以分为科普类期刊和学术类期刊两大类。学术类期刊按主管单位的不同,可以分为省级期刊、国家级期刊、科技核心期刊(统计源期刊)、中文核心期刊(北大中文核心)、中文社会科学引文索引 (CSSCI)、中国科学引文数据库 (CSCD)、双核心期刊等。

知识窗

ISSN 号

国际标准连续出版物编号 (International Standard Serial Number, ISSN) 是根据国际标准 ISO3297 制定的连续出版物国际标准编码,是连续出版物的唯一代码标识。该编号是以 ISSN 为前缀,由 8 位数字组成。8 位数字分为前后两组,各 4 位,中间用连接号相连。

格式举例:《中国经济问题》,ISSN 为 1000-4181。

(3) 报纸。报纸也是连续出版物的一种,是以刊载新闻和时事评论为主的定期向公众发行的印刷出版物。它是大众传播的重要载体,具有反映和引导社会舆论的功能。根据出版周期,报纸可分为日报、早报、晚报、双日报、周报、旬报等。

(4) 学位论文。学位论文是指为了获得所修学位,被授予学位的人按要求所撰写

的论文。学位论文是学术论文的一种形式,有严格的格式要求,一般不公开出版。学位论文分为学士论文、硕士论文、博士论文三种。

(5) 会议论文。会议论文是指在会议等正式场合宣读的首次发表的论文。会议论文属于公开发表的论文,一般正式的学术交流会议都会出版会议论文集。会议论文集不是期刊,但是有的期刊为会议论文出增刊。

(6) 专利文献。专利文献是包含已经申请或被确认为发现、发明、实用新型和工业品外观设计的研究、设计、开发和试验成果的有关资料,以及保护发明人、专利所有人及工业品外观设计和实用新型注册证书持有人权利的有关资料的已出版或未出版的文件(或其摘要)的总称。

(7) 标准文献。标准文献是经公认权威机构(主管机关)批准的一整套在特定范围(领域)内必须执行的规格、规则、技术要求等规范性文件,简称标准。

(8) 科技报告。科技报告是记录某一科研项目调查、实验、研究的成果或进展情况的报告,又称研究报告、报告文献。每份报告自成一册,通常载有主持单位、报告撰写者、密级、报告号、研究项目号和合同号等。它按内容可分为报告书、论文、通报、札记、技术译文、备忘录、特种出版物。

(9) 政府出版物。政府出版物是由政府机构制作出版或由政府机构编辑并授权指定出版商出版的文献。常见的政府出版物有报告、公报、通报、通讯、文件汇编、会议录、统计资料、图表、地名词典、官员名录、国家机关指南、工作手册、地图集以及传统的图书、期刊、小册子,也包括缩微、视听等其他载体的非书资料。

(10) 产品样本资料。产品样本资料是指厂商或贸易机构为宣传和推销其产品而印发的资料,如产品目录、产品说明书、产品总览、产品手册等。

1.2 信息检索概述

信息时代一个突出的特点是信息爆炸。人们每天都面临着来自四面八方的庞大信息量,如何快速准确地找出需要的信息,提高信息的利用率,掌握有效的信息检索方法很有必要。

1.2.1 信息检索的概念

对于信息检索,目前尚无公认一致的定义。国内外有关专家从不同角度解释信息检索,较有代表性的观点主要有以下几种。

(1) 信息检索的范围较为广泛。动态信息、静态信息、声频信息、视频信息及各种数值信息均属信息检索范围。如果将信息检索作为一门学科,它应该包括矩阵计数法、概率论、最优化理论、模式识别及系统分析技术等各学科领域的内容。

(2) 信息检索主要是文献检索。信息检索是从大量的文献中查寻与情报提问所指定的课题(对象)有关的文献,或者是包含用户所需事实与消息的文献过程。这里谈到的文献,不仅指文献线索,也包括文献的片断,如章、节、段落以及与事实有关的直接情报等。

本书定义的信息检索是将信息按一定的方式组织起来,并根据用户需求找出相关信息的过程。这是指信息的存储与检索,是针对信息工作者和用户来定义的,如果仅针对用户,信息检索是指在信息集合中找出所需信息的过程。

1.2.2 信息检索的类型

1. 按存储和检索的内容分类

信息检索按存储和检索的内容可分为文献检索、数据检索、事实检索。

(1) 文献检索。文献检索通常指的是在检索系统中检索以二次文献为对象(目录、索引、文摘)的信息,它们是文献信息的外部特征与内部特征的描述集合体。信息用户通过检索获取的是原文的“替代物”。

(2) 数据检索。数据检索指在检索系统中检索数值型数据,如科学技术常数、各种统计数据、人口数据、气象数据、市场行情数据、企业财政数据等,即事物的绝对值和相对值的数字。检索系统提供一定的运算推导能力,例如外推、内插、填补空缺数据,甚至列出曲线图或进行各种分析等功能。信息用户可用通过检索获得的经过核实、整理的数值信息再作定量分析。

(3) 事实检索。事实检索指在检索系统中的原始文献中检索关于某一事物(事件、事实)发生的时间、地点和过程(情况)等方面的信息。它是数值信息和系统数据信息的混合。一般先从系统中检索出所需信息后,再加以逻辑推理才能给出结论。这类信息主要是用于管理决策。

2. 按系统中信息的组织方式分类

信息检索按系统中信息的组织方式可分为全文检索和多媒体检索。

(1) 全文检索。全文检索是指检索系统中存储的是整篇文章乃至整本书。检索时,用户可以根据需要从中获取有关的章、段、句、节等信息,还可以进行各种频率统计和内容分析。随着计算机容量与运算速度的增大和提高,全文检索正迅速扩大到各学科、专业。

(2) 多媒体检索。多媒体检索是对超文本检索的补充。其存储对象超出了文本范畴,融入了静、动态图像(形)以及声音等多种媒体信息。信息的存储结构从单维发展到多维,存储空间范围在不断扩大。

1.2.3 信息检索工作

信息检索主要工作包括以下几个方面。

1. 信息存储

信息的存储方式直接影响信息检索的速度和效率,影响查全率、查准率、误检率等评价检索质量的指标。

信息存储可以利用自然语言或规范化语言。所谓自然语言是指作者的书面用语。在传统的文献检索系统中,为了提高文献的检索效率,文献工作者创建了规范化语言。规范化语言是文献标引和文献检索共同使用的语言,它使文献描述和表达概念具有唯

一性、专指性,便于提高查准率,标引时可以集中相关文献,提高查全率;能显示概念间的各种关系,有利于及时调整检索策略等;检索时能达到较高的查准率,同时提高检索速度。

2. 数据库与文档

信息的存储基本上是靠建立系统的单位自己来组织和构建各类信息的文档。文档的集合称为文献数据库,即通常所指的文献型、数值型、事实型及全文信息的集合。如何合理构建数据库是有效进行检索的关键。有关数据库的研究包括文档编制时对文献的选择、根据检索语言对文献进行标引、数据库文档结构的描述、组织与更新、数据库的质量、使用效果及其评价方法、评价原则、数据库市场及版权问题等。

3. 检索系统

检索系统是一定的目标、工具(印刷型或计算机)、信息资源(数据库)及检索方法与策略的集合体,一切检索都是通过检索系统实现的。检索系统的研究内容是存储信息的广泛性和检索的迅速准确。网络信息检索系统的研究主要包括系统硬件配置、软件的设计与分析、系统的评价、系统的智能化应用等方面。

4. 检索策略

检索策略是在处理信息需求提问的基础上确定检索途径、检索词语,并明确各词语之间的逻辑关系、查找步骤、系统输出顺序等方面的科学安排。检索策略与检索效果密切相关。正确的检索策略可以优化检索过程,以最小的花费、最短的时间获得最佳的检索效果。

检索策略的研究主要包括以查全为目标的检索策略的调节与控制;以查准为目的的检索策略的调节与控制;以最小投入为目标的检索策略的调节与控制;检索策略失误的分析;系统对检索策略的自动修改机制的研究等。

5. 检索工具

信息,资源浩瀚,形式多样,要快速、准确地查询到所需信息,必须借助各种检索工具。要根据不同的检索目的确定不同的检索工具、检索途径和检索方法。

6. 检索服务

检索服务是信息服务的重要环节,包括手工检索服务、机械检索服务、国内外联机检索服务、光盘检索服务、网络检索服务、定时检索服务、定题检索服务、周期检索服务等。

1.2.4 信息检索的目的与意义

1. 信息检索的目的

现在社会信息丰富多样,知识的存储过于庞大和无序,出版物的泛滥使人们难以获得大量准确的信息。知识存储的无序化加剧了“信息污染”的程度,堵塞了通向知识之门的道路,耗费了大量探索自然和社会规律的精力,以至于人们不得不认真学习