

《信息管理科学研究丛书》

文献标引概论

王金祥 黎盛荣 谢群宗 主编

陕西科学技术出版社

内容提要

本书是专为广大文献标引工作者进行分类标引和主题标引的参考书。全书共分八章，约18万字，内容新颖，理论联系实际，举例说明多，实用性强。对新版《中图法》和《汉语主题词表》以及新版《中国分类主题词表》的体系结构、使用方法及其分类标引与主题标引的一般原则和方法都作了较为详细地论述。书末附有国家标准 GB3860—83 文献主题标引规则和文献分类标引规范。对图书情报工作者、大专院校师生以及计算机检索工作者广大用户，均有较大参考价值。

前 言

文献标引概论是《信息管理科学研究丛书》系列中的一门重要的二级学科。信息管理科学是随着社会进入信息时代而产生的一组科学群，研究的范围极其广泛。大而言之，整个宇宙就是一个信息整体，是一个信息交流的巨网；小而言之，传播、新闻、广播电视、文献、情报、信息、档案、博物、标准、专利、咨询、开发、数据网络以及一切有关信息活动的领域都有信息管理问题，博大精深，都在研究之列。从当前发展的水平看，其研究的范围大致包括若干门二级学科，如：信息管理学基础理论、信息管理学史、信息载体形态学、信息文献采访学、信息文献标引学、信息服务学、信息咨询学、信息检索学、信息管理技术等等。其各二级学科之下还有三级学科、四级学科若干门。文献标引学是其二级学科中的一门。

随着现代科学技术的高度发展和电子计算机在图书情报工作中的广泛应用，文献检索从手工检索开始向机械检索过渡。因此，作为计算机检索系统前处理工程的文献标引工作就显得更为重要了。文献标引是揭示文献内容和提供检索途径的重要手段。标引的质量直接影响到信息的传递与利用，影响到为生产、为科研和为现代化服务的效果。因此，文献标引在图书情报工作中是一项极为重要的基础工作。

目前，我国专门研究分类标引和主题标引的专著为数不多，为了向广大图书情报工作者介绍有关文献标引方面的基础知识，编者根据在实际工作中的体会，编写了这本《文献

标引概论》，供广大图书、情报、档案工作者和本专业的大专院校师生参考。全书共分八章，约18万字。

本书内容新颖，理论联系实际，举例说明多，实用性强。对分类法语言和主题法语言的一般原理、文献标引的意义、类型与方法都作了较为系统的论述。同时，本书还探讨了文献检索语言分面组配化、文献标引自动化以及分类法与主题法一体化的发展趋势。

本书由湖南省科技信息研究所，陕西省科技信息研究所，西安公路交通大学图书馆，西安理工大学图书馆，海南大学图书馆，湖南衡阳医学院图书馆，邮电部第四研究所，陕西中医研究院图书馆，西安地质学院图书馆，西北政法学院图书馆，西安电子科技大学图书馆，西北轻工业学院图书馆，西安外国语学院图书馆等协作编写。本书从1994年暑假开始计划，落实编写任务，至12月收稿，由主编、副主编审阅，定稿。陕西科学技术出版社出版，西安市雁塔区唐都航天书店经销。

由于我们的水平所限，缺点错误在所难免，敬请同仁和读者批评指正。

编者 1994年12月

目 次

第1章 文献标引原理	(1)
§ 1-1 文献的含义、类型及其发展趋势	(1)
§ 1-2 文献标引的意义与作用	(8)
§ 1-3 检索语言的定义与特点	(9)
§ 1-4 文献标引与文献检索	(11)
第2章 文献分类标引的历史沿革与特点	(13)
§ 2-1 我国古代文献分类法发展概况	(13)
§ 2-2 我国近代文献分类法的演变	(15)
§ 2-3 我国现代文献分类法介绍	(18)
§ 2-4 国外文献分类法简介	(32)
第3章 分类标引工作	(54)
§ 3-1 分类标引的意义与特点	(54)
§ 3-2 分类标引的原则	(57)
§ 3-3 各类文献分类标引的方法	(60)
§ 3-4 期刊分类标引	(87)
§ 3-5 新学科分类标引	(92)
第4章 主题法的形成与发展	(98)
§ 4-1 《汉语主题词表》简介	(98)
§ 4-2 国外主题法语言发展概况	(106)
§ 4-3 国外主题词表简介	(109)

§ 4-4 《中国分类主题词表》简介	(116)
第 5 章 主题标引工作	(134)
§ 5-1 主题法的意义与特点	(134)
§ 5-2 主题法的类型与作用	(136)
§ 5-3 主题法的基本要素与性质	(138)
§ 5-4 主题标引的一般原则与方法	(143)
第 6 章 各类文献主题标引	(148)
§ 6-1 马列主义与毛泽东著作和哲学著作主题标引	(148)
§ 6-2 社会科学文献主题标引	(150)
§ 6-3 自然科学文献主题标引	(158)
§ 6-4 工业技术文献主题标引	(162)
§ 6-5 特种文献主题标引	(170)
第 7 章 主题目录的组织	(175)
§ 7-1 主题目录的意义与特点	(175)
§ 7-2 主题目录的类型与组织	(180)
§ 7-3 主题目录的编制原则	(182)
§ 7-4 主题目录导卡设计	(185)
第 8 章 计算机检索语言自动标引	(186)
§ 8-1 自动标引的意义与特点	(186)
§ 8-2 自动标引的一般原理与方法	(188)
§ 8-3 我国自动标引研究的现状及其发展趋势	(192)
附录 1. 文献主题标引规则(国家标准)	(196)
附录 2. 文献分类标引规范(送审稿)	(206)

第1章 文献标引原理

文献是人们从事生产活动和科学实验的总结,是人类精神财富的重要组成部分,它记录和反映着科学技术的成就与水平;预示着未来科学技术发展的趋势与方向,它随着科学技术的发展而不断发展,是社会发展到一定阶段的产物。

§ 1-1 文献的含义、类型及其发展趋势

1. 什么是文献

人们通过生产实践和科学实验获得了大量反映客观规律性的知识后,将这些知识记录在一定形式的载体上,便构成了文献。因此,文献是记录着知识内容的物质载体。文献的本质属性是它的知识性、记录性和物质性。

知识性——人类在认识和改造客观世界的实践中,不断地有所发现和发明,获得了经验和认识,从而使自己的知识体系不断地向深度和广度发展。所以,文献是表现科学技术知识赖以存在和传播的媒介;是人们得以继承前人知识、不断发掘和创造未来知识的重要工具。由此可见,文献的本质属性之一是它的知识性。

记载性——人类在认识世界和改造世界的过程中,虽然获取了大量的经验和知识,如果不以一定的形式记录在一定的物质载体上,则不能构成文献。因此,文献的又一本质属性就是它的记载性。

物质性——人类通过生产斗争和科学实验所获得的许多关于自然知识和社会知识其所以未能全部或广泛记载下来,最重要的

原因之一,就是记录知识的载体很落后,如古代文献的载体都是粗笨的龟甲、牛骨、石头以及竹木等。只有发明了纸、胶片、磁带、光盘等先进的人造物体时,人们才能大量地记载知识。因此,物质性也是文献的本质属性之一。

知识性、记载性和物质性是文献的基本属性,三者缺一则不能构成文献。

“文献”一词在我国最早见于《论语》,孔子说:“夏礼,吾能言之,杞不足征也;殷礼,吾能言之,宋不足征也;文献不足故也。”宋代朱熹解释:“文,典籍也;献,贤也。”我国第一部以文献命名的著作是宋末元初马端临的《文献通考》。他在该书序言中说:“凡叙事,则本之经史,而参之以历代会要,以及百家传记之书,信而有征者从之,乖异传疑不录,所谓文也。凡论事则先取当时臣僚之奏疏,次及近代诸儒之评论,以至名流之燕谈,稗官之记录,凡一语一言,可以订典故得失,证史传之是非者,则采而录之,所谓献也。”但是,“文献”这个词发展到今天,其含义就有了很大变化,它不只是限于一般典籍,凡是用文字、图形、符号、声频、视频等为手段记载一切知识的载体都统称为文献。1983年国家标准《文献著录总则》对“文献”一词作了简明扼要地解释:“文献”是“记录有知识的一切载体。”由此可见,知识与载体是文献的两个最基本的因素,记载是文献的基本特征。科学知识载体是构成文献的内涵,不同的知识内容,不同的载体和不同的记载形式是构成文献的外延。

总之,文献是用文字、符号、图形、声频和视频等方式记载人类科学知识的一种信息载体,是人类脑力劳动成果的一种表现形式。知识是文献的实体、载体是外壳,记载方式是两者的联络体;文献信息是人类的“第二资源”,是“无形的财富”,它和能源、材料是支撑现代科学技术的三大支柱。

2. 文献的类型

文献的类型大体上有以下几种:

(1)按文献的载体分

①印刷型,其中包括铅印、油印、胶印等,这是一种传统的印刷形式,其优点是便于阅读,可以广泛传递,缺点是保存不方便,占用的面积很大。

②缩微型,包括缩微胶卷、缩微胶片等。它是以感光材料为载体,利用光学记录的技术,使其文献缩小到若干倍的一种现代化文献形式。其优点是体积大大缩小了,便于保管,占用的面积与空间很少,缺点是阅读起来不方便,要借助于阅读机才行。

③计算机阅读型,这种文献主要是通过编码和程序设计,把文献变为数学语言或机器语言后,再输入到计算机数据库里,存贮在磁带或磁盘上,阅读时,再由计算机将它输出来为用户阅读。它的优点是存贮的文献量大,检索速度快,缺点是必须借助于电子计算机才能使用,价格比较昂贵。

④直观型资料,包括唱片、录音带、录像带、光盘、电影、幻灯片等,这种文献的特点是耳闻目睹,直观性很强。

(2)按文献内容、性质和加工程度分

①一次文献,即原始文献。它是作者以生产或科研成果为依据而创作的文献。其中包括专著、期刊论文、研究报告、专利说明书、会议论文等等,译文也属一次文献范围。

②二次文献,即对一次文献进行加工整理后的文献。如目录、文摘、索引等工具书。

③三次文献,是指在一次和二次文献的基础上,进行综合分析、加工整理后而编辑的文献。如科技动态、情报综述、决策参考等等,情报调研报告属三次文献。

(3)按出版形式分

①图书,图书的范围较广,一般可分为科技图书、科普读物、教科书和工具书等等。它的内容比较系统、全面、成熟,是生产技术和科研成果和经验总结与概括。

②期刊,是指定期或不定期的连续性出版物,其特点是有固定的刊名和出版形式,内容新颖,能反映当前科技水平,报导文献信息快、品种多、数量大。一般有学术性期刊、检索性期刊、消息报导性期刊等等,它是情报的主要来源。

③科技报告,是科研单位和政府机关或生产部门关于科研成果的总结报告。其特点是一种连续性出版物,有机构名称、统一编号、独成一册、发行不规则、内容新颖、时间性较强。

④学术会议文献,是指国内外各种专业学术会议所发表的论文。它反映一门学科或某一专业的最新研究成果和国内外科学技术发展水平动向,是很重要的科技情报来源。其特点有以下几点:反映新成果快;文献质量高;专业性很强;文献量大。

⑤学位论文,是指高等院校毕业生和研究生所提出的学位论文。它是在导师的指导下通过专题研究和经过审查后的学术性研究论文,所以一般说是带有一定创造性的一次文献。其特点和学位论文基本上相同,对科研工作有一定的参考价值。

⑥政府出版物,是指世界各国政府部门及其专设机构所发表的文件和学术报告。其特点是:文献范围广、品种多、数量大;对了解一个国家的科学技术和各项政策有一定参考价值。

⑦专利文献,是指专利说明书、专利公报和其他专刊文献。它是十八世纪随着资本主义的发展而形成的专利制度的产物,是申报专利人向政府专利机关递交的说明书,由专利主管部门审查批准后给予公布。其特点是专利文献能反映出现代科学技术发展的最新水平,专利说明书一般都附有图表、数据和较为详细地文字说明。我国专利文献包括发明专利说明书、实用新型说明书、发明专利公报、实用新型专利公报、外观设计专利公报等。

⑧标准文献,是指有关产品和工程的质量、规格、生产过程、检验方法等方面所作的技术文件。它具有一定的法律约束力,是从事生产设计的技术依据。标准一般分国家标准、部颁标准、专业标准、

企业标准。其最大特点是内容随着科学技术的不断发展而经常修改,所以标准的新陈代谢比较频繁。

⑨产品样本,是各厂商向用户推销其产品的连续性出版物。是指对定型产品的性能、构造、用途、使用方法和产品规格等所作的说明。其特点是:产品一般有样本照片、直观性强;设计技术成熟、数据可靠;对新产品的设计与选型都有一定参考价值。

⑩技术档案,它是生产和设计部门在技术活动中形成的,有一定具体工程对象的科技文件、图样、图表、照片等原始记录的原本或复制本等文献。其中包括任务说明书、协议书、技术指标、研究计划、方案和技术措施等。它具有保密和不公开使用等特点。

⑪特种文献,如内部刊物、报纸、新闻通报等等。这类文献报导快、涉及面广、内容新等特点。

3. 现代文献的特点及其发展趋势

1946年,世界上第一台电子计算机的问世,标志着现代信息革命的开始。这场革命将导致人类社会结构、生产结构和科学结构发生巨大的变化。因此,文献也正在产生重大的变革,从而对文献标引带来了一定的难度。所以,掌握现代文献的主要特点及其发展趋势,是搞好文献标引工作的关键。

现代文献的主要特点及其发展趋势表现以下几个方面:

(1)文献信息量大、增长速度快、更新期短

随着现代科学技术的飞跃发展,科学成果与文献层出不穷。据联合国科教文组织的统计,当前自然科学领域内,仅基础学科就有538个主要专业;技术科学有412个专门分支;全世界每年发表的论文达500万篇左右;出版的图书大约有50万种以上,其中科技图书有12万种,国内外出版的期刊达数万种以上;全世界申请的专利文献达2000万件。我国建国以来,文献出版量达60余万种,61亿册,人均6册,全国出版的图书品种从1979年的17212种,发展到1985年的45000种,平均每天出书100多种,出版量进入

了世界前列。科技文献年平均增长率,60年代为9.5%;70年代为10.6%;80年代为12.5%;90年代为15%;预计下一个世纪初将突破20%。一般文献每10年翻一番,尖端科学的文献3—5年翻一番。因此,科学技术的更新期正在逐步缩短。

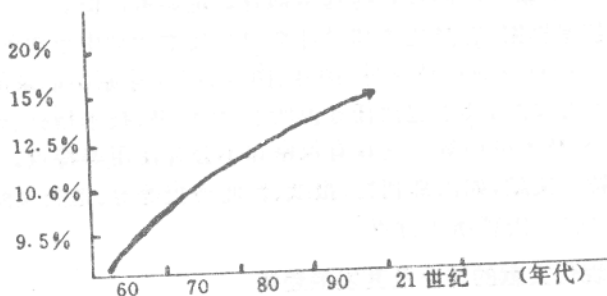


图1 科技文献增长示意图

(2)文献信息种类多、类型杂

当前,由于新技术革命的挑战,世界各国出版物,不仅种类多,而且文献的载体也发生了巨大的变化。据统计,全世界出版的文献光文种就有:“英、俄、德、法、日等66种,其中英文占50%;从出版的形式分,有以下几种:印刷型;缩微型;计算机阅读型;直观型,如唱片、录音带、录像带等等。

(3)文献内容交叉、重复、科学之间相互渗透

随着现代科学技术的发展和各门学科之间的相互渗透,导致了文献内容交叉重复,各门学科向高度分化与高度综合化延伸。因此,交叉科学、边缘科学和综合性科学不断涌现,如数学早在三百多年前就开始向自然科学渗透,如公元1666年牛顿力学的诞生就与欧几里德的《几何学》有着直接的关系。

本世纪五十年代,由于控制论、信息论、系统论的影响。数学开

始向社会科学渗透。如图书馆计量学、图书馆管理学、教育统计学等等都是数学与这些学科相互渗透的结果。

至于文献的重复现象就更为惊人,据统计,加拿大出版的专利说明有 87.2% 与其他国家重复,其中 2/3 与美国重复。我国在 1985 年进口的期刊有 25000 多种,其中重复订购的就有 5000 种,复本平均达 10 册之多。

(4) 知识老化快、利用率低

由于现代科学技术发展的速度很快,文献信息量正处于“信息爆炸”时代。全世界出版的文献信息量日益剧增。据统计,本世纪 60 年代为 72 万亿字符,70 年代为 232 万亿字符,80 年代为 500 万亿字符,全世界出版物信息量总共 4000 亿字符,每人平均为 100 字符。因此,随之而来的是科学技术更新期大大缩短,知识老化的速度也逐步加快,这是文献信息新陈代谢的必然规律,而文献信息的有效利用期一般为 5—10 年,真正有价值的文献只不过占 50%,最高利用率为约 30% 左右,其利用率的比数也正在下降,据日本统计:1960 年 23.5%,1968 年为 15.9%,1970 年为 15%。由此可见,文献信息正处于不可抗拒的自然淘汰和新旧交替的状态。

年 代	更 新 期
18 世纪	80—90 年
19~20 世纪	30—40 年
本世纪初至 50 年代	15—20 年
20 年代以后	5—10 年

图 2 文献更新期示意图

综上所述,我们不难看出,今后文献信息的发展趋势是:文献量将会继续不断地成倍增长;文献载体是多样化,印刷型、缩微型与听频型、视频型并存;文献信息传递手段现代化、网络化和资源共享已成为今后发展的必然趋势。

同时,随着缩微技术和电子计算机技术的发展和“录音资料”的出现,有人预言:今后“听书”代替看书,“有声图书馆”、“手提图书馆”、“电子图书馆”将会逐步发展起来。甚至还有人认为将来“纸书”消亡,人类将进入“无纸社会”。

§ 1—2 文献标引的意义与作用

1. 文献标引的意义

在文献工作中,文献标引是实现情报存贮和情报传递的重要手段。

众所周知,任何情报工作者,要搞好情报的分析研究工作,首先必须掌握国内外的各种文献。因此,文献工作是科技情报工作的基础,而文献标引又是文献工作的基础。

如果我们说文献馆和资料室是存贮文献的宝库,那么文献标引则是打开这个宝库的钥匙。特别是在当前,整个人类社会面临着信息革命的挑战,作为传递信息手段的标引工作,就显得更为重要了。

所谓文献标引,顾名思义,就是根据文献内容的分析,把某些具有检索意义的语言或符号记录下来,并作为存贮与检索依据的文献处理过程。文献标引包括分类标引和主题标引两个方面。

分类标引和主题标引都是以揭示文献内容为其特征,前者是标引分类号的过程,后者是标识主题词的过程。但两者是相辅相成、互为补充的两种标引方法。

文献标引既然是揭示文献内容和提供检索途径的重要方法,那末,标引的质量将会直接影响文献资料的传递与利用,影响到文献为生产、为科研、为教学和为社会主义现代化服务的效果。由此可见,文献标引工作在科技情报工作中的重要意义就不言而喻了。

2. 文献标引的作用

文献标引是搞好情报服务工作的基础。它对整个文献的搜集、整理、传递和利用都有着密切的联系。文献标引的作用概括起来有以下几点：

(1) 文献标引是揭示馆藏的手段

文献门类多,数量大,如果没有科学的管理方法和系统地揭示馆藏,读者和用户要直接找一篇文献就等于大海捞针,不知所向。譬如一个生物学家要了解有关植物方面的文献,就可通过“Q94 植物学”这个类目,找到 Q942 植物细胞学、Q943 植物遗传学、Q944 植物形态学、Q944.5 植物解剖学、Q945 植物生理学和 Q946 植物生物化学等各方面的文献。

(2) 文献标引是建立目录体系的重要基础

文献标引可以从各种不同的角度揭示一书的内容,建立各种不同类型的目录。如分类目录,主题目录等等。读者和馆员通过这些目录可以了解馆藏,从而为查找各种文献提供了依据。

(3) 文献标引是开展手工检索与机械检索的必要条件

目前,文献检索正处于手工检索向机械检索过渡的新阶段。文献标引工作,不论对手工检索或机械检索都是不可缺少的基础工作。否则,不仅手工检索不可能,就是万能的电子计算机,没有输入,也就无从输出了。所以,文献标引是搞好情报检索的必要条件。

总之,文献标引是搞好情报工作的基础,文献工作的各个环节都离不开文献标引工作,文献标引的好坏直接影响到文献工作和情报工作的开展。

§ 1—3 检索语言的定义与特点

检索语言是根据文献检索需要而制定的一种人工语言,这种语言又称情报语言、标引语言或索引语言等等。它是沟通文献标引

和文献检索的桥梁。所谓检索语言,顾名思义,就是根据文献的内容分析,把某些具有检索意义的语言或符号记录下来(如主题词、分类号、篇名、作者、专利号等等),并作为存贮与检索依据的文献处理过程。而文献检索,简单地说,就是指文献的存贮与查找的全过程。

目前,国内外常用的检索语言分两种类型,即分类法语言与主题法语言。前者是标引分类号的过程;后者是标引主题词的过程。分类标引和主题标引都是以揭示文献内容为其特征,这是两者相同的地方,不同之处,说明如下。

所谓分类法语言,就是指以数字或拉丁字母与数字相混合作为标识符号,以类目概念作为基本词汇,以类目之间的等级关系来表示复杂概念的一种检索语言。这种语言的特点是以科学分类为基础,能充分体现学科的系统性和学科之间的逻辑关系,可以满足族性检索的要求。

例如:

S 农业科学
S5 作物学
S51 禾谷类作物
S511 稻
S511.1 水稻

所谓主题法语言,是以反映文献内容的主题词为基本单元,以规范化的自然语言为基本词汇,以词汇概念的组配方式为基础,以表达主题概念之间的逻辑关系的词组作为文献检索的一种检索语言。这种语言的特点是用揭示文献内容的主题词按其字顺系统进行检索标识,其灵活性、直观性与专指性很强,可以满足特性检索的要求。例如:医疗器械、医疗事故、医疗体育、医疗用灯、医疗用纸

按“医”字汉语拼音或笔划次序检索。如医疗器械是族首词,其下位概念有病人监护器、超声波诊断机、脑电图机、心电图机和注

射器等等。

分类法语言与主题法语言各有千秋,概括地说,分类法语言的优点,正是主题法语言的缺点,反之,分类法语言的缺点,正是主题法语言的优点。所以两者是相辅相成的两种不同的检索语言。因为分类法语言的基本体系结构是按学科门类的逻辑次序,从总体到个别、从一般到特殊、从低级到高级、从简单到复杂,逐级展开的层累号码制,其文献检索的标识是分类号。而主题法语言的体系结构是以规范化的主题词来表达一篇文献的内容特征,是按主题词的字顺进行排列的,其文献检索的标识是主题词。而主题词之间的逻辑关系与排列顺序无关,它不受学科逻辑体系的约束,专指性很强。所以,使用不同的检索语言标引文献,所得的检索标识形式不同,其检索效用也各有差异。

§ 1—4 文献标引与文献检索

文献标引和文献检索是两个不同的概念,标引是手段,检索是目的,标引是为检索服务的。如果从广义上说,文献检索是指文献的标引、加工、存贮以及针对读者与用户的需求查出所需各种文献的全过程。

文献检索过程主要包括两个基本环节:一是将搜集的文献按一定的形式进行标引、加工、存贮和编制各类检索工具,从而构成文献检索系统;二是根据读者和用户的需求,利用手工检索或机械检索的方式,为他们提供各种原始文献,从而形成文献检索输出系统。

文献检索必须要依赖于文献检索系统。而建立文献检索系统,其中包括手检系统和机检系统,都需要对大量的文献资料或数据进行前处理。所以,文献的前处理工作,实质上是提出文献检索标识,即文献标引工作。

所谓检索标识,就是以简练的语言形式来揭示文献内容的特