

信息检索与利用

马 静 主编

南京航空航天大学

信息检索与利用

马 静 主编

航空工业出版社

内 容 提 要

本书以计算机信息检索为主,帮助读者建立信息意识、提高检索和利用信息的能力。其内容包括信息及信息科学等基础知识、检索基本原理、检索标识系统;中英文核心期刊的利用、中英文检索刊物的利用;计算机检索基础技术、Dialog 国际联机、Firstsearch 联机检索;KROndisc—Ei compendex 光盘检索、ProQuest—Inspec 光盘检索及中刊篇名、中国学术期刊、中国专利光盘检索;Internet 的连接、浏览器、Internet 网上的 FTP、Telnet、E-Mail、Netnews、WWW 功能、Web 搜索引擎及 FTP、Usenet、Business、Meta 搜索引擎,网络资源引路。

本书内容新、涉及知识广、综合性强,可用作研究生、本科生的教科书,同时也可供广大信息管理人员、科研人员参考。

图书在版编目(CIP)数据

信息检索与利用/马静主编。—北京:航空工业出版社,2000.6

ISBN 7-80134-688-2

I. 信… II. 马… III. 情报检索 IV. G252.7

中国版本图书馆 CIP 数据核字(2000)第 33723 号

航空工业出版社出版发行

(北京市安定门外小关东里 14 号 100025)

南京航空航天大学印刷厂印刷

全国各地新华书店经售

2000 年 6 月第 1 版

2000 年 6 月第 1 次印刷

开本:787×1092 1/16

印张:11.75

字数:288 千字

印数:1—3 000

定价:13.00 元

序

近年来,随着计算机网络技术、通信技术的高速发展,网络信息资源的日益丰富,各类网络信息新的查询工具不断问世,获取信息和知识的渠道、方法、手段均发生了根本变化。因此,作为培养高校学生信息意识及获取信息能力的文献检索课,如何适应新形势的需要,及时更新和拓展新的教学内容,已成为当前各高校从事《文献检索与利用》课程教学的广大教师教改的重点。新编教材《信息检索与利用》的出版正适应了这一形势的需要。

《信息检索与利用》一书较全面地介绍了各类信息资源的检索与利用理论及方法。其一,紧扣信息科学理论,使教材结构上能充分反映信息科学和技术发展过程,走出了传统“文献”教材一一罗列检索工具的编著模式,使读者能从信息科学的角度去理解和掌握信息检索与利用技术体系;其二,内容组织上层次清晰,重点突出,按照印刷、电子出版物(光盘)、国际联机、因特网的媒介特征分类介绍,便于读者有针对性、有选择性地掌握信息技术与技能,增强了教与学的灵活性;其三,着重于因特网信息资源查询与利用,比较及时、全面地反映了因特网信息资源及检索技术迅速发展的特征,从而弥补了前些年出版的相关教材之不足。

本教材是以马静为主编的几位长期从事《文献检索与利用》课教学的教师,在总结多年教学经验的基础上,通过反复修改、充实、完善原有教案、教材而逐步积累成的。他们的教学曾得到充分肯定和校级最高奖励。当然,在网络信息技术一日千里的今天,要使一本教材的内容不断更新而能适合时代前进的步伐决非易事。我希望读者在学习使用这本书时,用严格挑剔的眼光来帮助指出其中的缺点和错误,以使本书在再版时能有提高。

陈万寅

2000年5月

前 言

信息化是现代社会最显著特征之一。在计算机技术、网络技术、多媒体技术迅速发展和全面支持下, 社会科技、经济、文化、政治等方面的信息生产和传播达到空前的规模和效率, 迫使人们尤其是高校本科生、研究生及广大科技工作者必须时时刻刻面对信息、感知信息、利用信息、处理信息。据此, 我们以现代信息检索技术为主体, 以实际信息利用为目标, 总结多年文献检索教学实践经验, 吸取各类教材精华, 写成这本教材。

全书分四个部分。首先在总论里介绍了信息、信息科学相关基础, 信息检索原理基础及检索标识系统等基础知识, 随后展开介绍了各类信息资源的检索与利用, 包括印刷型文献信息源、国际联机、光盘资源及 Internet 网络资源的信息查询与利用。其中网络资源利用是本书重点, 从 Internet 概念、网络工作原理、网络的连接到常用网上信息传送功能都作了较为详细的介绍, 并对各类搜索引擎的使用作了重点介绍。

本书第 5 章由邹小筑编写, 第 4 章由王文忠编写, 第 6、8 章由徐清华编写, 第 7 章由徐惠珍、张瑞芳编写, 其余各章均由马静编写。全书由马静负责统稿、定稿。

在本书编写过程中, 得到南航教务处的支持和帮助, 得到黄因慧教授、王连生先生的指导和建议。陈万寅副研究馆员审阅了全书并为本书作序, 在此一并表示由衷感谢。

限于编者水平, 书中不当及误漏之处, 敬请读者指正, 不胜感激。

编著者

2000 年 5 月

1.1.4 目 录

与信息相关概念

第一部分 总 论

第 1 章 信息 信息科学 信息意识	(1)
1.1 信息 知识 情报 文献	(1)
1.2 信息科学	(2)
1.3 信息意识	(3)
1.4 信息化社会的人才素质	(4)
1.5 科技查新工作	(5)
第 2 章 信息检索原理及基础	(7)
2.1 信息来源	(7)
2.2 文献型信息源及其划分	(7)
2.3 文献的产生与演变周期	(11)
2.4 信息检索原理	(14)
2.5 信息检索方法	(15)
2.6 检索工具中信息注册识别	(16)
第 3 章 检索标识系统	(19)
3.1 检索标识系统概述	(19)
3.2 主题法标识系统	(20)
3.3 分类法标识系统	(25)
第二部分 印刷型文献信息源的检索与利用	
第 4 章 中文信息源的检索与利用	(30)
4.1 中文核心期刊利用	(30)
4.2 中文检索工具书利用	(32)
第 5 章 外文信息源的检索与利用	(37)
5.1 外文相关领域核心期刊简介及利用	(37)
5.2 英国《科学文摘》	(39)
5.3 美国《工程索引》	(45)
5.4 外文科技报告利用	(52)
第三部分 国际联机计算机信息检索	
第 6 章 计算机信息检索概论	(56)
6.1 计算机信息检索基础	(56)

6.2	数据库的编排结构与索引	(62)
6.3	联机检索特点及国外主要联机系统介绍	(70)
第7章	DIALOG 联机检索	(73)
7.1	DIALOG 联机检索系统简介	(73)
7.2	DIALOG 数据库组织特征介绍——数据库蓝页	(79)
7.3	DIALOG 联机检索的常用指令	(85)
7.4	DIALOG 系统的检索方法	(87)
7.5	DIALOG 联机检索实例	(88)
第8章	First search 联机检索	(89)
8.1	First search 系统简介	(89)
8.2	First search 检索方法	(90)

第四部分 光盘数据库检索与利用

第9章	中文光盘检索资源	(93)
9.1	《中文科技期刊篇名》光盘数据库	(93)
9.2	《中国专利文摘》光盘数据库	(98)
9.3	《中国学术期刊(光盘版)》全文数据库	(105)
第10章	外文光盘检索资源	(110)
10.1	KR OnDisc 检索系统与 Ei compendex * Plus 光盘数据库	(110)
10.2	ProQuest 检索系统与 INSPEC 光盘数据库	(117)

第五部分 Internet 网络资源信息查询与利用

第11章	Internet 网基础	(126)
11.1	Internet 入门知识	(126)
11.2	网络连接	(128)
11.3	浏览器(Browser)	(129)
第12章	Internet 网上信息传送常见功能	(132)
12.1	电子邮件 E-mail	(132)
12.2	文件传输 FTP	(134)
12.3	远程登录 Telnet	(138)
12.4	网络论坛 Netnews	(143)
12.5	万维网 WWW	(147)
第13章	万维网搜索引擎	(148)
13.1	搜索引擎概述	(148)
13.2	Yahoo(http://www.Yahoo.com)	(148)
13.3	AltaVista(http://www.altavista.com)	(150)

13.4	Infoseek(http://www.Infoseek.com)	(153)
13.5	其他搜索引擎.....	(154)
第14章 其他类型搜索引擎		(159)
14.1	FTP Search Engine(文件服务搜索引擎)	(159)
14.2	Usenet Search Engine(新闻论坛搜索引擎)	(160)
14.3	Business Search Engine(商界信息搜索引擎)	(161)
14.4	MetaSearch Engine(元搜索引擎)	(162)
第15章 Internet 网信息资源引路		(165)
15.1	网上图书.....	(165)
15.2	网上期刊.....	(165)
15.3	网上会议信息.....	(169)
15.4	网上专利.....	(171)
15.5	网上标准.....	(173)
15.6	网上科技报告.....	(173)
15.7	上网的各类数据库.....	(174)
参考文献		(177)

信息特征:

2. 永恒性
1. 客观普遍性
3. 可传递性
4. 共享性
5. 时效性
6. 可转换性
7. 可改性

第一部分 总论

第1章 信息 信息科学 信息意识

在我们这个时代,“信息”已经成为人们生活中最密切相关的概念之一。经济信息、科技信息、市场信息、商品信息乃至就业信息等都在直接或间接地影响着我们的生活、工作和学习,影响甚至决定着每个人生存与发展的机会。如何检索和利用信息,自然地成为信息化社会生活中人们所关心的问题;如何成为善于检索和利用信息的高层次人才,也是高等学校中每一个学生所关心的问题。

为此,我们将从情报、文献检索等学科的基础出发,结合社会信息化过程中不断涌现出的思想、理论、方法、技术和问题,为读者提供一套实用的“信息检索与利用”教学内容。

1.1 信息 知识 情报 文献

信息检索与利用作为一门课程是20世纪90年代从情报、文献等学科中演变发展起来的。在此,有必要首先对若干相关概念加以说明。

信息 这可谓是当代社会使用最广泛的词汇之一,但究竟什么是信息,至今仍众说纷纭。哲学家认为信息是认识世界的依据;数学家认为信息是一种概率;物理学家说信息是“熵”;而通信专家则把信息看作是“不定度”的描述。

比较有影响的定义或解释有以下几种:

信息论的创始人香农(C. E. Shannon)认为:信息是组织程度,它能使系统的有序性增强,减少破坏、混乱和噪声。

美国《韦氏字典》把信息定义为:信息是用来通信的事实,在观察中得到的数据、新闻和知识。

我国《辞海》把信息解释为:信息对消息接收者来说是指预先不知道的报道。

最近,美国图书馆的信息专家作出的解释是:信息是一切思想、事实和富有想象力的作品,它以各种方式进行记录和传播。

其实,我们不妨对信息作出最直观的解释:信即信号,息即消息,信息就是通过信号传递的消息。它不是物质和事件本身,而是物质和事件的反映,是通过声、光、色、符号、语言、资料、数据等传递出物质的存在方式与运动状态。

知识 经过人类社会活动的长期洗炼,相对于信息内容,知识一词的内涵已经是相当稳定、明晰的。按照1999年版《辞海》的解释,知识有三层基本含义:

知识是人类认识世界的成果或结晶。人类在生活中大量接受了来自自然界和人类社会自身的信号,通过循环反复的实验、实践和大脑思维活动,不断将信号进行加工、提炼、分析、综

合,从而形成了认识,随着认识与实践的交互作用,认识不断深化、积淀而构成知识。

知识包括经验知识和理论知识。经验知识是一定时期内人们在社会生活与实践中总结出来的认识;理论知识则是人们对形成的经验认识反复论证和高度抽象后形成的认识。理论知识来自于经验知识、并高于经验知识。简言之,理论来自于经验,又高于经验。

知识借助于一定的语言形式,或物化为某种劳动产品的形式,通过交流和传递,成为人类共同的精神财富,成为一代又一代学习的主要内容。

情报 情报的内涵,人们最早认为是战时关于敌情的报告,后来演变为“竞争对手的情况报告”,并逐渐泛指为“对象的情况报告”。随着人类意识到“交往胜于封闭”,“沟通比敌对更为重要”,情报一词最终演变为“为一定目的而搜集的有价值知识或信息”。今天的情报概念指的就是能为我们所用的知识与信息。

文献 文献的概念更为简明,它是指凡记录有知识或信息的一切载体。以龟甲、兽骨作为材料,记录有“甲骨文”的龟甲、兽骨是文献;用石材刻录文字和图形的碑是文献;竹简和帛文是文献;有铭文的青铜器是文献;纸质印刷品、缩微影像制品及电子出版物等也是文献。

文献是知识和信息的主要载体,是传递知识和信息的工具。

综上所述,信息、知识、情报都是人类认识活动的反映,概念的不同说明了人类认识活动的目的和方法不同,说明了人类认识的不断拓展和深化。“信息”之所以成为现代社会广为使用的词汇,根本原因就是人们意识到,仅仅掌握现有的知识是不够的,是难以适应社会迅速发展的。我们不仅要掌握前人积累的知识,更需要从层出不穷、纷繁复杂的信息中获取崭新的机会、知识。这已经成为当代社会的共识。

1.2 信息科学

一般而言,高等学校的每一门课程都是以一定的科学作为基础,并凝聚着相关的教学思想、教学方法来开展的。但由于科学自身在不断发展,人们对科学体系的认识在不断深化,其科学知识应用在教学中的目的和方式亦会出现变化;高校的课程所依托的科学内容、课程自身的组织结构也将出现相应的变化。信息检索与利用课程也是如此。

在人们对“情报”、“文献”类信息的重视程度高于其他类信息的时代,信息检索与信息利用实际上就定位于文献检索与利用。文献检索与利用课程就是最初的信息检索与信息利用课程。这时,该课程很大程度上是以情报科学为基础的,按照情报科学的原理,该课程甚至可以归属为历史学类。

随着社会信息化进程的加速,人们对信息的认识发生了根本性变化,人们发现,情报、文献类信息对于学生面向历史和学习前人确定的科学理论与知识固然是非常重要的,但其他非情报、文献类信息对于学生面向现实、乃至面向未来,掌握新的理论知识、探索未知,甚至创新新的知识具有更为重要的意义。因此,情报或文献等范畴被纳入内涵更为丰富的“信息”范畴中,情报科学转变为信息科学的一个组成部分,文献检索与利用课程在许多高校中被所指更为广泛的信息检索与利用课程所代替。

综合人们最新的理解和认识,我们可以通过对信息活动的若干层面的分析来理解信息检索与利用课程的学科关系:

信息活动的学习过程,即个体学习他人而从事信息活动的过程。首先是学习自然语言,同

时逐渐开展工具语言的学习,并在此过程中逐步学习具有社会规范的自然语言和具有技术标准
的工具语言,进而学习信息的检索与利用语言,学习制作信息、发布信息的语言,最终才能成
为信息主动的活动者。

信息活动的社会化过程,即个体从事信息活动的行为逐步纳入社会规范的过程。这个过程
是由一定的社会组织实现的,包括制定和发布描述信息的语言标准、传播或对话通信的标准、
信息接受者独享或共享信息的权利以及人们创造、制作、传播信息的道德与经济责任规范。

信息活动的技术过程,即人们创造和利用技术手段(工具)以开展信息活动的过程。这个过
程首先是人们运用自身的器官,然后发明和使用辅助工具,包括接受工具、传播工具及处理工
具。从而不仅能够实现面对面的交流,也能实现跨越时空的交流。

信息活动的科学过程,即人们从科学的高度不断深化对信息活动的认识的过程,包括人们
感知和处理信息的生理机能、风俗习惯、文化传统以及人们从事信息活动的价值观或社会经
济效益的倾向等方面的研究过程,从而最终形成关于信息活动的总体看法和认识。

如此等等的过程,实际上还只是人们对整个信息活动的一部分认识,但我们已经能看到信
息活动是十分复杂的过程,信息科学是一个十分庞大的知识体系,信息检索与利用贯穿于整个
信息活动全过程,其有关的知识与技术又只是整个信息技术与知识的一个部分。

同时必须认识到,由于信息本身纷繁复杂、变化迅速,人们在从事信息活动的过程中,不
仅需要具备优良的自然机能条件,如视、听、记忆、思维、语言等,而且需要掌握多种技术手段
包括机器与机器语言,特别是计算机、网络等方面的技术手段,以便更加准确、更加迅速、更
加大规模地接受和处理信息,同时还需要具备良好的心理素质,以适应我们在信息活动过程中
面临的各种挑战、压力、困惑和矛盾。

需要说明的是,“人工智能”这个全新的科技领域,已成为信息检索技术发展的重要支
撑。表现在:

①人工智能专家有一种共识,认为信息检索是最先从人工智能这种先进技术中获益的机
构,人工智能的搜索技术与信息检索的检索技术实质相同,都是一种认识匹配过程;

②人工智能的专著中,都把数据库的智能检索、自然语言理解和专家咨询等与信息检索关
系密切的领域列为人工智能的首要领域;

③人工智能领域学术信息与信息检索技术信息实际上也存在某种映射关系:人工智能有
什么论题,信息检索也会有相应的论题,反之亦然。

所有这些现象说明了:着重掌握和应用人工智能技术,将有利于我们掌握信息检索技术,
有利于我们更加快速、更加准确、更加大规模地检索和利用信息。

1.3 信息意识

对于学习者而言,信息科学的知识及信息科学的技术只是一种影响其检索和利用信息效
率的工具。能否掌握和利用信息,归根到底取决于我们内在的心态或意识。具备强烈的“信
息意识”,能有意识地去获取信息才是至关重要的前提。所谓“信息意识”就是人们关于信
息的倾向性,是人们获取信息的主观能动性,是人们对信息所具有意义的需求和态度。

任何人都有其信息意识,但又各不相同。有些人对信息及其变化很敏感,有些人则不然;
有些人非常关心现实价值信息,有些人则特别关心长远价值信息;有些人重视具体信息,也有些

人则重视综合信息,有些人对信息非常专注,有些人则非常容易迁移,如此等等。

对于高等院校的学生而言,我们主张的是,应当增加有利于建构和重建学生知识结构的信息意识。按照英国学者布鲁克斯的观点,可作如下描述:

$$K[S]+\Delta I=K[S+\Delta S]$$

$K[S]$ 代表原有知识结构, ΔI 代表吸收的信息增量, $K[S+\Delta S]$ 则代表新的知识结构。

该表达式说明:第一,新的知识结构能否形成,取决于能否在原有知识结构上吸收和增加新的信息;第二,能否吸收新的信息,取决于原有知识结构是否形成,而且是否具有开放性。

例如,儿童没有既定的知识结构,因而不能直接接受抽象的知识,而只能通过接受那些形象具体的信息循序渐进地建立自身的基本知识结构,初等教育阶段的学生必须从数学、语文、物理、化学、生物、社会规范等方面逐步建立起基本知识结构,而难以接受自然和本质性的、深层的知识;高等学校的学生虽然可以在数、理、化等领域获得更加深入的、专业的知识,但对于社会实践中出现的大量动态信息的掌握还存在不少困难;作为专业科技人员和生产者、管理者则应当能从正式和非正式的渠道中获得各种信息,以充实、改善其知识结构。正如美国学者 D. 普赖斯所说,“那些刺激着科学家并使其研究成为可能的信息中,有 80%左右是他在正式报道之前通过非正式渠道从其他研究人员那里得到的,即通过代表会议、讨论会、报告会或通过现在被称为‘无形的集体’的其他一些场合和手段得到的”。

1.4 信息化社会的人才素质

高等学校是社会高级人才的培训基地,高校学生毕业离开高校的同时就立即成为社会一员,成为社会经济生活的重要力量。因此,高等学校的学生必须具备社会所需要的人才素质。

信息化是现代社会最显著特征之一,也是社会对人才的最基本的需求之一。所谓信息化,就是在计算机技术、网络技术、多媒体技术的迅速发展和全面支持下,社会科技、经济、文化、政治等方面的信息生产和传播达到空前的规模和效率,几乎全球每一个角落的活动都将迅速地传播到全球任何区域,从而迫使人们时时刻刻必须面对信息、感知信息、处理信息,及时掌握和发现对他们来说最有意义的信息,排除和修正信息误差并作出迅速的响应和反馈……

信息化社会中,全球每一个角落发生的事件都将给人们带来机会、挑战和压力,使人们的生活方式、思维方式发生变化。

社会信息化对高等教育产生的影响是广泛而深刻的,最直接的影响之一就是必须加强对学生信息意识与技能的培养,具体如下:

(1) 必须加强培养学生捕捉信息的意识。生活在信息社会的人,首先应当敏锐感知各种形式信息的意义,能够对纷繁多彩的信息价值作出反应和判断,并具备捕捉和利用信息的主动意识。在信息社会中,谁取得主动,谁就掌握了获胜的机会。

(2) 必须加强培养学生处理信息的基本技能。信息是现代生产过程的重要资源和结果,善于发现、利用信息资源,并创新信息、发布信息,才能最大限度地实现经济社会效益,是人才必须具备的基本能力。在信息社会中,谁具备信息处理能力,谁就能获得最大的经济效益。

(3) 必须提高学生在信息化社会中的生存与发展素质。信息化意味着不同行业、不同产业、不同职业之间的联系与格局将频繁创新,也意味着人们经济生活形式与形态的不断变化,同时还意味着不同时期、不同区域的经济技术信息同时呈现出来,并在跨越时空的信息交流中

将出现种种混乱和机遇,强烈影响人们的身心状况的变化、调整。在信息社会,谁具备适应变化的素质,谁就能实现最持久的发展。

1.5 科技查新工作

科技查新是信息检索的一个重要组成部分,是为科技立项、成果奖励的严肃性、公正性、准确性、权威性而作的查新咨询,是一种科技信息管理工作。

实际生活中的查新主要表现在如下方面:

(1) 帮助政府主管部门和企、事业单位进行决策。

(2) 开发信息资源,开发知识宝库。我国国家科委系统、国家教委系统等所组建的科技查新中心,其任务就是开展查新服务等。国家规定用户在科研立项时要进行查新,看课题是否重复,看该课题值不值得投资。

(3) 在开发上需查新。例如,某厂家要开发轻便帐篷的生产项目,通过查新,发现国外轻便帐篷与该厂提出开发项目的样式极为相似,则该厂这一项目不能算“首创”或称“独创”。

(4) 在成果鉴定时查新。科研主管部门规定,必须在成果鉴定时进行查新,以确定成果的水平。查新结果则为鉴定成果提供一定的依据。

(5) 技术转让查新。通过查新,帮助评价该项技术。在谈判前,通过查新检索,可以掌握有关技术信息,了解谈判对象情况,了解技术先进程度,了解专利有效年限,以便有效地维护自身利益。例如,我国某单位在引进英国某公司浮法玻璃生产技术时,对方提出该项技术包括 137 项专利,索要入门费 2500 万英镑。经我方检索查新发现,137 项专利中,有 51 项已失效,结果入门费减为 25.5 万英镑。

(6) 申请专利时查新,看其是否具有新颖性。

(7) 在发生专利纠纷时查新。80 年代后期,华北某市医用光电仪器厂指控江苏省光学仪器联合公司生产的透照仪,其原理及用途均同于他们的专利产品——强冷光透视仪,侵犯了产品的专利权,要求江苏省光学仪器联合公司立即停止生产、销售,并赔偿侵权损失费。该省专利局通过江苏省专利局通知该公司,在规定期限内必须应诉。江苏省光学仪器联合公司通过查新检索,找到一篇 1982 年 8 月 28 日批准的美国专利说明书,该说明书详细叙述了同种产品的设计和使用,这比华北某市医用光电仪器厂的发明早 7 年,并已被公用。据此,可以确认原告的发明不具备新颖性,不符合授权条件,因而,原告的申诉不能成立,应请求专利管理机关宣告其专利权无效。在确凿证明面前,原告立即主动要求撤诉。

(8) 实施专利时查新,看其是不是先进的。

(9) 跟外商谈判时查新。除了引进技术、引进设备,还有资金、资信的检索了解,看其信誉度,看其资金是不是落实,看生产的产品能不能外销。如南京一个厂家与某外商谈判合资事宜。该外商注册资金 100 万美元,通过查新检索,了解到该外商在天津已投资 100 万美元,而且他在天津投资的资金还未到位。查新使得南京的这个厂家避免了一次谈判合资的受骗与失误。

(10) 产品销售时查新,查国际行情。

(11) 解决生产技术关键时查新,看人家是怎么解决的。

(12) 在了解国内外动态、确定研究方向时查新。

(13) 查新可以保障在完全了解历史和现状的情况下开题研究。如某单位研究的“石灰加

固处理”，他们说已查过 20 世纪 80 年代、90 年代的文献，未发现此项研究，自信自己的成果是国内外第一。通过查新检索，发现：国外在 60 年代已有这类课题的大量研究报告。

(14) 标准资料查新。上海汽轮机厂运用查新工作，掌握有关国际标准，按国际标准生产产品，获得了我国第一张质量体系认证证书，其产品打入国际市场后，赢得瑞士 ABB 公司、美国 GE 公司等的好评。

除以上情况下需查新外，目前还应加上一种情况即做学位论文时要求作开题查新，随着教育体制的改革，对硕士、博士生的创新能力的培养逐渐提到议事日程，对博士生论文，要求其论文具有创造性成果；对硕士生论文要求有新的见解以表明其具有从事科学研究工作或独立担负专门技术工作的能力。因此开题时查新要求就迫使学生们避免重复研究，真正从事创新研究。

第2章 信息检索原理及基础

信息特征: 客观性、时效性、依存性

2.1 信息来源

可加工性、可传递性、共享性

一切信息来源于自然界,来源于人类社会。人们的生产、科研、生活以及政治活动等是信息产生的源泉。

信息源的形式有口头(语言)信息源、实物信息源和文献信息源。

口头(语言)信息,在情报学领域称之为“零次信息”;实物信息,是以某种实际存在的物品所揭示的信息,如样机、样品等。目前口头信息和实物信息还没有加工存储和提供检索的工具,使用者只能根据自己的需要去专门地、有目的地搜集、加工、分析、利用。文献信息是我们获取信息的主要来源。通过对文献中所记录信息的检索与利用,可以帮助我们打开知识信息宝库的大门。

2.2 文献型信息源及其划分

2.2.1 如果把庞大的文献型信息源按其内容进行划分就可以分成一次信息源、二次信息源、三次信息源。

(1) 一次信息源

一次信息源也称原始文献或第一手资料。凡以作者本人的科学研究或研究成果为依据而创作的原始文献,均属于一次信息源。如期刊论文、科技报告、专利文献、学位论文、会议文献等。

由于一次信息源属于原始创作,直接记录从事科学研究和生产实践活动的劳动成果,用以报道新发明、新创造和新技术。尽管其内容多数并非定论,但都是作者第一次发表关于某一课题的见解、实验条件、原始数据,所以具有很高的直接参考和借鉴价值,受到科技人员的普遍重视和广泛利用。但一次信息源数量庞大,存储分散,不够系统。

(2) 二次信息源

二次信息源是将一次信息源进行了加工、整理、提炼、浓缩、标引、编序后所形成的各种目录、题录、索引、文摘等信息源。二次信息源包括了大量的可供使用的信息检索工具。在信息交流过程中,二次信息源不仅能提供大量信息,而且具有报道和便捷检索的功能。在“信息检索”一类的课程中,二次信息源内容占重要分量。

(3) 三次信息源

三次信息源是对某一专题的一次信息源和二次信息源的有关信息进行检索、筛选、分析、加工,并结合编者的实地调查研究,进行综合分析后重新编制的成果。如书评、专题述评、学科

年度总结、动态综述、字典、词典、百科全书、年鉴、手册、指南等工具书。

三次信息源具有综合性、浓缩性、参考性强的特点。评述性、综述性的文献是重大科学技术决策中的重要参谋,是人们系统了解某一专业的重要途径。参考工具则是人们日常生活和科研活动不可缺少的工具。

2.2.2 如果将文献型信息源按其出版形式进行划分,则可以分为 11 种出版类型。即图书、期刊、会议文献、学位论文、科技报告、专利文献、标准文献、政府出版物、技术档案、产品资料、报纸。

(1) 图书

图书是最早出现的文献类型之一。至今在整个文献信息源中仍占主导地位。

图书的范围广泛,一般可分为两大类型:

其一,阅读类,主要包括专著、论文集、工具书、教科书等。

其二,参考类,包括手册、年鉴、百科全书、字典、辞(词)典、人名录、机构指南、文献指南等。

科技图书大多是对已发表的科技成果、生产技术和经验、或某一知识领域进行系统的论述和概括。它往往以曾经发表过的期刊论文、会议论文、研究报告及其他第一手资料为基本材料,经过作者的分析、归纳、重新组织编写而形成的。有些科技图书也包含一些尚未发表过的科研成果。

图书的特点是内容比较成熟、系统、全面、可靠,有一定的新颖性。它是综合、积累和传递知识的一种重要的信息源。但图书编辑出版的时间往往较长,所报道的知识信息也较期刊论文、特种文献(如研究报告、专利文献、会议文献等)要晚,所反映的内容一般都是 3~5 年以前的研究水平。这是图书作为信息源和信息传递手段的弊端所在。

(2) 期刊

期刊有固定名称、有统一出版形式、有一定出版规律、每期刊载不同著者、译者、编者所编写的论文,是按一定编号顺序连续出版下去的出版物。科技期刊上发表的论文是科技信息的重要来源。同图书相比较,期刊出版发行速度快,报道的内容新。期刊论文一直是科技人员最重视的信息来源。他们一般都有经常阅读期刊的习惯,借以了解学科动态和进展,开阔思路,汲取已有成果。据估计,从期刊等连续出版物方面来的科技信息,约占整个信息来源的 65%。

一般来说,科技期刊具有数量大、品种多,内容丰富多样,出版周期短,报道速度较快,传播、流通面十分广泛,连续性强等特点。伴随着相应的学科发展,科技期刊在科学技术发展中的作用可以归纳为以下三点:①它是公开的科技活动和成就的正式记录;②它是评价科技人员在科技发展过程所作贡献的舆论工具;③它是传递科技信息的主要手段。

科技期刊还起着汇集其它类型文献的作用,即其它类型文献所提供的重要信息常会在期刊中出现。例如,会议论文中约有 35%、科技报告中约 50%、学位论文中约有 20%的成果都会经过改写发表在科技期刊上,重要的专利在期刊中也常有报道。因此,当科研人员难以获得其他类型的文献时,可以利用期刊进行弥补。

(3) 会议文献

这是指在国际和各国国内专业学术会议上所发表的论文和报告。会议文献分为会前、会中、会后出版物,其标志是有会议名称、时间、地点及会议举办单位。目前世界上每年举行的学术会议在 1 万次以上。在会上提供交流的文獻往往代表着某一专业领域最新的研究成果。学术会议是科技人员重要的信息交流场所,传递的信息比较及时、集中,信息的针对性较强,兼有

直接交流和间接交流两种方式的长处,其文献是了解国际及各国的科技水平、动态和发展趋势的重要的信息资料。

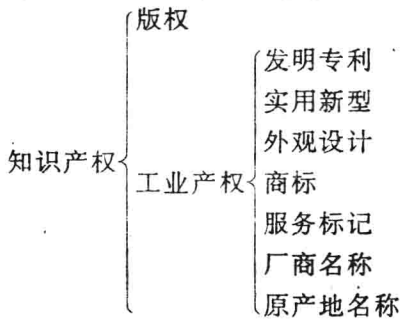
(4) 学位论文

学位论文是本科生、研究生为取得学位资格而撰写的学术性研究论文。其识别标志:在学位论文上往往注明授予学位头衔,授予学位单位的名称,授予学位的地点、时间。学位论文不像公开出版物那样便于广泛流传,它一般在本单位内得到收藏、流通。我国在中国科技信息研究所等单位收藏部分学位论文的复制品。20世纪30年代后期,美国就开始由私人企业建立的复印公司,向社会定期报道每年学位论文的题目和内容提要。日本的学位论文分为两种情况:一种是国立大学的学位论文,由国家图书馆收藏,提供给读者使用;另一种是私立大学学位论文,存放在颁发学位证书院校的图书馆中。英国的学位论文由国家统一规定,存放在国家外借图书馆。其他欧洲的院校,其学位论文一般多存放在本校图书馆内。

学位论文是经过审查的原始研究成果,所探讨的问题比较专门,对问题阐述得比较详细、系统,带有一定的独创性。在指导教师指导下提出的论文,要经过该领域著名学者审查质疑,作者作出答辩通过后,才能授予学位。学位论文的篇幅不限,论文的内容一般从历史、背景、述评开始,详细介绍研究经过、实验记录和具体数据等成果。由于学位论文包括了原始的第一手资料,并具有较高的研究水平,因而是一种重要的信息来源。但学位论文是非卖品,不易收集到,从而影响到流通和利用。博士论文经过改写后在学术会议和期刊上发表,这种现象目前已较为多见。至于硕士论文和学士论文,其水准要求较博士论文要低,但它们也是经作者参考了大量资料,进行系统的概括和研究后的成果,对科研、生产仍有一定的参考价值。

(5) 专利文献

专利是一种知识产权,它包括版权和工业产权。



一切与工业产权有关的文献统称为专利文献,它是一种用法律形式来保护的文献。发明人的创造发明,研究出的某种新技术、新设计,在向政府主管部门提出专利申请,经过政府主管部门审批后,即获得一定年限的专利权,成为可以买卖的商品。专利文献主要指专利说明书。广义的专利文献有专利说明书、专利公报、专利分类表、专利文摘,以及与专利有关的法律文件及诉讼资料等。

(6) 科技报告

科技报告是在二次大战期间及二次大战后迅速发展起来的,它所反映的科研和技术革新成果比期刊论文快。由于它对某一课题研究进展和试验过程如实作了记录,因此,其内容专深具体,完整可靠;又由于其内容具有一定的保密性和专门性,故一般采用出单行本的办法,少的一两页,多的数百页不等。科技报告大致可分为基础理论研究和生产技术两大类。若按使用类