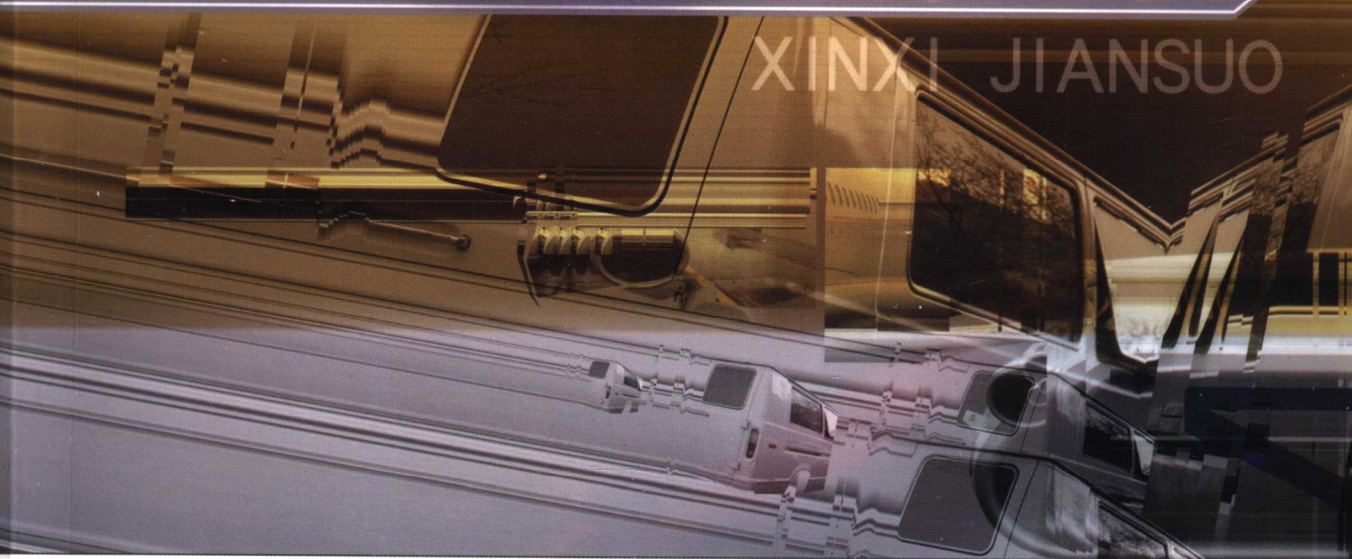


XINXI JIANSUO 信息检索



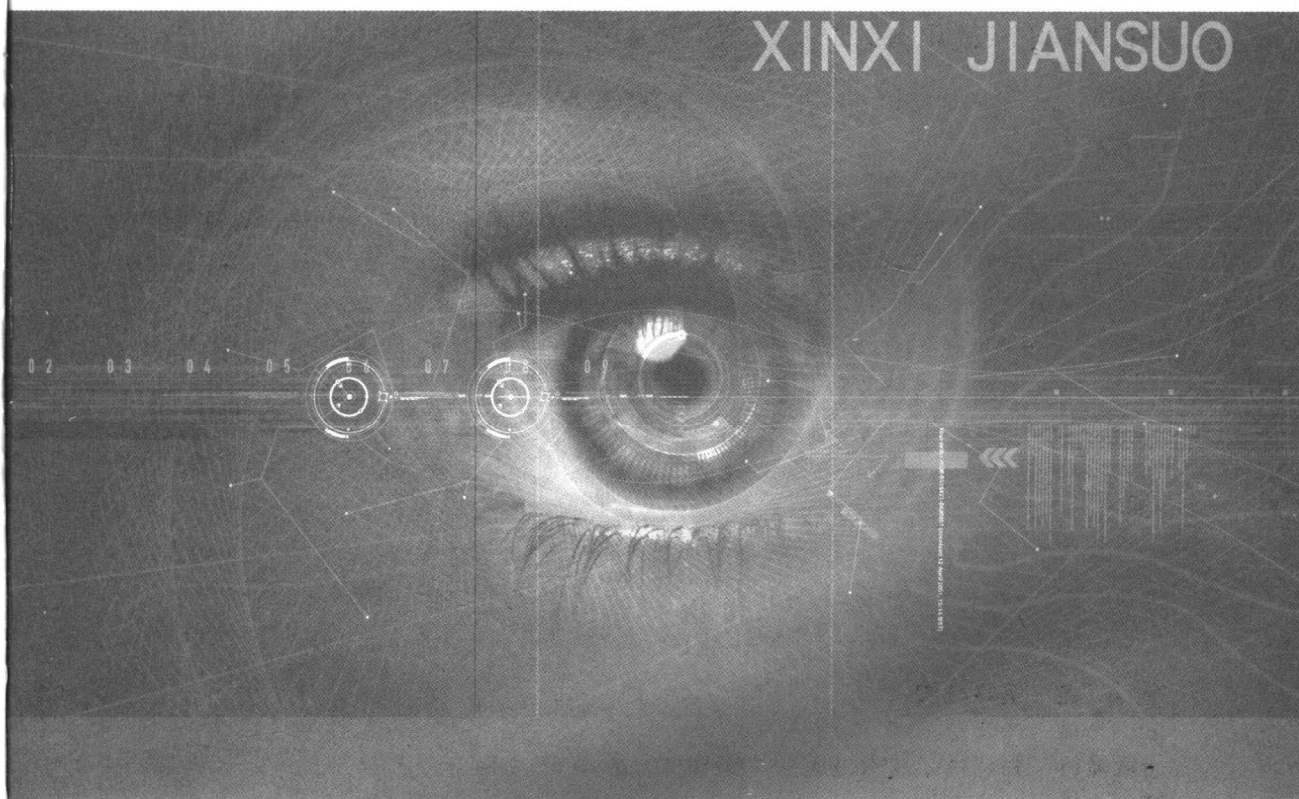
主编 张文德

福建科学技术出版社
FUJIAN SCIENCE & TECHNOLOGY PUBLISHING HOUSE

XINXI JIANSUO

信息检索

XINXI JIANSUO



主编 张文德
著者 王征清 刘 丹 张文德
杨 瑜 肖 春 苏 悦
陈旭华 陈秀莹 陈建芳
林雪英 姚 丹 詹庆东

福建科学技术出版社
FUJIAN SCIENCE & TECHNOLOGY PUBLISHING HOUSE

图书在版编目 (CIP) 数据

信息检索/张文德主编. —福州: 福建科学技术出版社,
2007. 2

ISBN 978-7-5335-2980-2

I. 信… II. 张… III. 情报检索—高等学校—教材
IV. G252. 7

中国版本图书馆 CIP 数据核字 (2007) 第 017045 号

书 名	信息检索
主 编	张文德
出版发行	福建科学技术出版社 (福州市东水路 76 号, 邮编 350001)
网 址	www.fjstp.com
经 销	各地新华书店
排 版	福建科学技术出版社排版室
印 刷	福州晚报印刷厂印刷
开 本	787 毫米×1092 毫米 1/16
印 张	22.75
字 数	578 千字
版 次	2007 年 2 月第 1 版
印 次	2007 年 2 月第 1 次印刷
印 数	1—5 000
书 号	ISBN 978-7-5335-2980-2
定 价	34.00 元

书中如有印装质量问题, 可直接向本社调换

前 言

学会利用信息,掌握获取信息的方法,是信息时代的人所必备的本领。人类有了这一本领,社会的工作效率、服务速度、生活质量和文化享受水平将得到大大提高,工作、生活、学习的方式以至整个社会的面貌将发生根本性的变化。当今,市场竞争的成败胜负,关键在于占有信息的数量和速度,而信息集人类智慧的精华,能及时向人们传播各方面发展的动态。

著者在总结了自己长期从事信息工作的基础上,参考了大量的国内外研究成果,先后出版了《科技文献检索与利用》(云南科学技术出版社1990年出版)、《网络资源与信息检索》(福建科学技术出版社1999年出版)、《网络信息检索》(福建科学技术出版社2002年出版),并在此基础上编写了《信息检索》。本书力求做到理论与实践相结合,内容简明扼要,图文并茂,通俗易懂。希望读者通过本书的阅读,能较好地了解信息检索的基本内容,增强信息意识,掌握打开知识宝库的金钥匙。

本书由张文德教授担任主编,主持编写大纲,并负责对全书进行审订、统稿及组织整理。全书第一章的第1、2、3、4、6节,第三章的第1节、第9节的纸介工具书部分内容,第四章的1、2、4、5节的手工检索部分内容,第四章的第8、9、10节,第六章的第1、2、3、4、5、6节由张文德执笔;第一章的第5节由林雪英执笔;第二章由詹庆东执笔;第三章第9节的网络参考工具数据库部分,第五章的第1、6节由王征清执笔;第三章的第2、5、6、7节由姚丹执笔;第三章的第3、4、8节由陈秀莹执笔;第四章第2节的光盘数据库检索部分内容、第3节,第五章的第7节由陈旭华执笔;第四章的第1节的计算机检索部分内容,第五章的第5节由陈建芳执笔;第四章的第4节的计算机检索部分及第6节由肖春执笔;第四章的第5节的计算机检索部分内容、第7节,第五章第2、3、4节由杨瑜执笔;第六章的第7节由苏悦执笔。附录一、二分别由刘丹、张文德提供。

本书在编写过程中得到了福州大学有关部门领导的大力支持,在此向他们表示衷心感谢!

限于著者的学识水平,加上撰写时间仓促及信息检索的内容变化较快,书中各类错误和缺点在所难免,敬请学界同仁匡谬补正,也欢迎其他读者批评指正。

著者

2007年1月25日

目 录

第一章 网络信息检索基础	(1)
1.1 信息检索的基本概念	(1)
1.2 信息的类型及特点	(4)
1.3 检索工具及索引语言	(8)
1.4 检索程序及获取原文	(20)
1.5 检索效率分析和检索策略探讨	(25)
1.6 联机信息检索系统	(31)
思考题	(46)
第二章 网络资源	(47)
2.1 因特网基础知识	(47)
2.2 WWW 资源	(57)
2.3 电子邮件 (Email) 资源	(72)
2.4 文件传输 (FTP) 资源	(81)
2.5 远程登录 (Telnet) 资源	(88)
2.6 流媒体资源	(96)
思考题	(99)
第三章 中文检索工具	(100)
3.1 检索国内外信息的中文检索工具	(100)
3.2 中国知识资源总库 (CNKI)	(110)
3.3 万方数据资源系统	(127)
3.4 中国人民大学复印报刊资料数据库	(133)
3.5 全国报刊索引数据库	(137)
3.6 中文科技期刊数据库	(140)
3.7 中文科技期刊引文库	(148)
3.8 中国高等教育文献保障系统	(150)
3.9 参考工具书	(154)
思考题	(175)
第四章 外文检索工具	(176)
4.1 美国《工程索引》	(176)
4.2 美国《化学文摘》	(187)
4.3 美国化学学会 (ACS) 数据库	(209)

4.4	美国《科学引文索引》	(211)
4.5	英国《世界专利索引》	(232)
4.6	Elsevier Science Direct Online (SDOL) 全文数据库	(247)
4.7	PQDD 博硕士论文数据库	(258)
4.8	英国《科学文摘》	(263)
4.9	美国《市场与技术预测综览》	(273)
4.10	Emerald 全文数据库	(279)
	思考题	(286)
第五章	网上其他学术信息资源	(287)
5.1	电子图书、电子期刊信息资源	(287)
5.2	专利信息资源	(299)
5.3	标准信息资源	(307)
5.4	学术会议及学位论文信息资源	(310)
5.5	工程技术类信息资源	(314)
5.6	财经类信息资源	(319)
5.7	法律类信息资源	(328)
	思考题	(333)
第六章	信息的利用	(334)
6.1	文献的阅读和鉴别	(334)
6.2	信息的搜集方法	(335)
6.3	如何分析文献获取信息	(336)
6.4	信息素材的积累	(338)
6.5	信息研究成果的主要表现形式	(340)
6.6	获取信息资源的策略	(341)
6.7	如何撰写英文文献	(344)
	思考题	(350)
	附录一 常用网址	(352)
	附录二 DIALOG 系统主要文档简介	(353)
	主要参考文献	(358)

第一章 网络信息检索基础

1.1 信息检索的基本概念

一、基本术语

(一) 信息

维纳在他的《信息控制论》中指出：“信息是人们在适应外部世界并使这种适应反作用于外部世界过程中，同外部世界进行交换内容的总称。”其物理意义是：“信息必须有一定的意义，必须是有意义的载体。”如声、光、热、电流、电磁波以及文字等物质，都可作为有意义的载体，即信息是一种具有实际意义的事物。信息与我们日常生活密切相关，因而也可从各个角度来认识。对信息的定义甚多，从知识的角度来说，“在发生源和吸收源之间，当发生源发出的信号被吸收源所理解，信号就成为信息。”

信息的现代科学含义指事物发出的消息、指令、数据、符号等被包含的内容。

随着科学技术的日益发展，人类的一切活动都离不开信息交换与传输，无论各行各业乃至日常生活的各个方面，信息活动都在发挥着十分重要的作用。因此，人们把当今社会称为“信息社会”。

信息交换不仅存在于人类社会，而且存在于自然界的一切生物之间，尽管有的是本能、原始、低级的。自然界中生物的生存和延续，都有赖于同类体之间的信息交换。如：动物在求偶或遇险时，都会本能地发出感应信息或信号信息，以传递给同类，其传递方式有同类可以判别的“气味”和“声音”等形式。

作为信息，除了必须是“有意义的载体”外，还具有一定的特征，即预先性（如天气预报，必须是提前的，落后则失效）、实用性、时效性、可传递性。

因此，信息是具有一定特征意义的载体，或者说信息是通过某种形式的载体反映出特定的意义。这是同一体的两个方面，相辅相成，缺一不可。

(二) 文献

大凡人类的知识用文字、图形、符号、声频、视频的手段记录下来的东西，统统可称为文献。文献也可称为固化在一定载体上的知识，或者更简单地说，文献是记录下来的知识。作为文献，必须含有知识和载体这两个部分。但载体非仅指纸介而已，也可说文献并不一定就是以书本形式展示在人们面前，它还可以磁带、胶卷等形式展示在人们面前。凡是把知识用文字、图形、符号、声频等手段记录在纸上，晒在蓝图上，摄在感光片上，录在唱片上，存贮在磁带上等，都算为文献。

(三) 资料

资料是固化在一定的实物或载体上的知识。资料是为工作、生产、学习和科学研究等参

考需要收集或编写的一切公开或内部的材料。

二、信息的传递渠道（也称“信息交流”）

信息的传递渠道有两种：非正式传递渠道和正式传递渠道，也可称为非正式过程和正式过程。

（一）非正式过程

（1）科学技术工作者之间就他们从事的研究和研制进行直接对话，如交谈、参加学术会议等。

（2）科技工作者参观同行的实验室、科学展览等等。

（3）科技工作者对某些听众作口头演讲。

（4）交换书信、出版物预印本和单行本。

（5）研究或研制成果在发表前的准备工作，包括发表形式（致期刊编辑的信、工作报告、学术报告等）以及发表的地点和时间的选择。

（二）正式过程

（1）为手稿的发表所作的编辑出版和印刷过程，包括写书评。

（2）科学出版物的发行过程，包括与发行过程有关的书刊商业活动。

（3）图书馆的书目工作和检索工作。

（4）信息工作本身，从收集到检索，包括宣传工作。

我们之所以分为两种，是因为前者带有明显的个体性质，既不能与研究工作分开，也不能由专职信息工作者用正式的方法去完成。非正式过程更多表现为信息制造者和信息利用者个人的行为特征。正式过程则表现为文献流通。

三、信息检索

目前，情报界对信息检索尚无固定的定义，各种定义名目繁多，但均脱离不了文献检索这个范围，即信息检索的最终目的是检索出文献。前苏联学者契尔纳说：“信息检索，就是从大量的文献中查寻与信息提问所指定的课题（对象）有关的文献，或者是包含用户所需事实与消息的文献的过程。”

它包括三个方面：数据检索、事实检索、文献检索，而以文献检索为主。文献检索即提出包含所需信息的文献，而数据和事实检索则是提出包含在文献中的信息本身。

这种将文献检索与信息检索加以区分的看法，固然有一定道理，但不管是检索包含在文献中的信息，还是检索含有信息的文献，均离不开文献这个范畴。因此，二者没有必要严格区分，如有的国家称信息检索，而有的国家称文献检索一样，信息检索可以理解为工作的目的而命名的名词，文献检索是以工作的对象而命名的名词。

四、文献检索

（一）文献检索的定义

文献检索包括两个方面：

(1) 检索系统的建立及检索工具的组织 and 积累。

(2) 文献的查寻, 即根据具体课题的需要查询。主要通过书目、索引、文摘等检索工具, 从众多的文献中, 检出与课题有关的或对课题有用的文献, 并且要求检索工作做到迅速、准确和没有重大遗漏。

(二) 文献检索的类型

人们从文献中获取所需的信息, 通常通过直接检索与间接检索两种检索。

1. 直接检索

所谓直接检索, 是通过阅读原始文献直接获取所需信息。间接检索是通过检索工具的指导再查找原始文献而获取所需信息。

直接检索是人们习惯使用的方法, 他们掌握本专业的核心期刊, 长期养成经常翻阅这些具有信息密度大、质量高的核心期刊的习惯, 这种方法有一定的优点。

(1) 直接检索原文易于掌握文献的实质内容, 可直接判断其信息内容是否符合信息要求, 并可能得到意外的收获。

(2) 直接检索简便易行, 并且时效高。检索工具存贮原始文献信息需要一定的过程, 必然存在时差问题。所以直接检索是及时获取最新信息的有效方法。

但是随着文献的发展, 直接检索已难以达到理想的效果。

目前文献量与日俱增, 在浩如烟海的文献中用直接阅读原始文献的方法来满足信息需求是难以实现的, 此外, 现代文献在文献类型、出版形式、语种上的高度分散都严重地影响直接检索的效果。

2. 间接检索

为了克服直接检索在现代文献状况下的盲目性, 人们开始注重以检索工具为向导的间接检索。间接检索的优越性是由检索工具所具有的检索功能决定的。这种方法有以下优点:

(1) 使盲目的分散检索成为有目的的集中检索。检索工具将分散在不同学科、不同类型、不同语种中, 但主题内容相同的文献集中在一起, 这样就可以避免直接检索的分散性、盲目性、偶然性, 检索效果可大幅度提高。

(2) 文献质量有较好的保证。存贮在检索工具中都是具有一定信息价值的原始文献, 这不仅保证了检索工具存贮文献的质量, 也避免了直接检索时浏览阅读没有信息价值的文献而花去时间和精力。

(3) 加速了检索过程。通过检索工具阅读的不是全文, 也不是全部款目, 而是与检索需求有关主题部分的款目, 因而检索过程可大大加速, 检索效率可大大提高。

(4) 消除语言障碍。科技人员掌握的语种有限, 直接检索必然会漏查许多人们不懂的语种文献所包含的信息, 检索工具由一种文字写成, 只要掌握这一种文字, 就可掌握多语种文献检索。

(5) 提供了广泛的信息来源。直接检索只能从某一图书馆收藏的文献中获取信息, 而某一图书馆所提供的信息来源是极其有限的。检索工具的信息来源非常广泛, 因此, 其查全率、查准率会大大提高。

(6) 提供了有规律的检索途径。检索者只要掌握检索工具提供检索途径的规律性, 就可根据检索需要按图索骥, 从而使既快又准又全的信息检索成为可能。

由此可见, 间接检索是科学的检索方法。

但是，忽视直接检索的作用也是片面的，应该把二者有机地结合起来。实际情况也表明，人们通过查阅信息密度较大的核心期刊及一般性地浏览其他期刊，以了解本专业发展动态的一般情况，在作特定课题研究时利用检索工具，两者结合，互相补充，是一种行之有效的方法。

1.2 信息的类型及特点

信息类型的划分标准很多，由于划分的角度不同，其划分方法也不相同。
下面重点介绍按信息内容和按发表形式划分的信息类型、定义及其特点。

一、按信息加工层次划分

可将信息分成为一次信息、二次信息和三次信息，各信息的定义及特点见表 1.2-1。

表 1.2-1 按内容划分的信息类型

	一次信息	二次信息	三次信息
定义	凡是以作者本人在生产和科研中所取得的成果为依据而创作的原始文献所传递的信息。	是将分散的、无组织的一次信息，按一定原则加工、整理、简化、组织成为系统的、便于查找利用的信息。	在合理利用二次信息的基础上，选用一次信息的内容，根据一定的需要、目的进行分析、综合或浓缩重组而得到的信息。
特点	是信息的基础，是技术前进的标志。具有一定的实用性、新颖性和创造性。 数量大、类型多、学科交叉、获取较困难。	是以目录、题录、文摘、索引等形式提供的检索工具所传递的信息。 它不改变一次信息的内容，仅对一次信息进行压缩和标引编排。	是将一次信息中有价值的信息、事实摘录出来，按性质类别范围重新组织。内容有了很大的变化，比二次信息更为浓缩。
包括范围	期刊论文、专利文献、会议文献、学位论文所传递的信息	各种目录、文摘、题录、索引等所传递的信息。	专题评述、动态综述、年度总结、数据手册、科学大全、百科全书、年鉴等所传递的信息。

二、按信息载体的出版编辑特点划分

按信息载体的出版形式划分为 12 个类型（又称十二大文献源），其中包括：期刊、图书、会议文献、专利文献、学位论文、政府出版物、标准文献、产品资料、技术档案、报纸、新闻稿、工作文稿。其定义、功能及特点见表 1.2-2。

表 1.2-2 按出版编辑特点划分的文献类型

	定义	功能及特点
期刊	采用统一名称, 定期或不定期出版的连续性刊物。 期刊通常也称杂志。“期刊”一词着眼“周期”特征, “杂志”一词着眼“内容”性质。	是信息的重要来源, 是技术成就的正式记录。 能及时反映各学科的发展水平及动向, 出版周期短、速度快、量大、内容新颖、固定、核心强。 质量水平不等, 相差悬殊。
图书	是对科研成果、生产技术和经验的总结性的概括论述。	内容较其他出版物全面、系统、可靠, 有一定知识体系的完整性, 内容比较成熟。 便于人们对某一课题的历史、现状及未来进行研究和探讨。 出版速度较其他文献慢。
会议文献	是研究人员在各种学术会议上, 交流科研成果、新进展及发展趋势的讨论记录或论文等。	不仅是提供信息的重要来源, 而且是迅速获得最新技术信息的一个重要途径。 获得信息直观, 反馈迅速。 新发现、新成果和新见解很多是在学术会议上首先公布。
专利文献	是一种用法律形式来保护的文献。 专利文献包括: 说明书、权利要求书、说明书摘要、说明书附图等。 专利文献的核心是专利说明书。	涉及的技术内容广泛。比较具体可靠, 能较快地反映出世界各国科学技术的发展水平, 是一种重要的信息来源。 有统一的格式, 文字简练。 内容上还具有的先进性、新颖性、创造性及实用性。
科技报告	是关于某项科学研究成果的正式报告, 或是对研究和试验过程中各阶段进展情况的实际记录。	内容比较专深具体。许多研究课题及尖端学科的资料, 往往首先反映在报告中。它不仅叙述了成功的经验, 也记载了失败的过程, 因而对科研人员起到直接借鉴的作用。 能代表一个国家和专业的发展水平与动向。 是不定期出版物, 一个报告为一单行本, 有统一编号。
学位论文	是高等院校研究生、毕业生所写的作为评定学位的论文。	具有独创性, 内容专一, 阐述详细、系统, 尤其是博士论文有一定的参考价值。 是经过一定的审查的原始研究成果。
政府出版物	是指各国政府部门所发表并由政府专设机构统一出版的文献。	对了解某一国家的科技政策、经济政策及其演变情况, 有一定的参考价值。 文献的内容广泛, 数量特别多, 部分文献在列入政府出版物之前, 已被所在单位出版过(如科技报告), 所以有一定的重复。

	定义	功能及特点
标准文献	是对工农业产品和工程建设的质量、规格及其检验方法所做的技术规定，是从事生产、建设的一个共同技术依据，有一定的法律约束力。	它是独立、完整的科技文献，有严格的审批程序，是各方专家集体制定的，内容可靠，其技术信息可直接应用。 标准文献的新陈代谢非常频繁，随经济条件与技术条件的改进，需经常不断地进行修改和补充。
产品资料	是指各国厂商对定型产品的性能、构造、原理、用途、使用方法、产品规格等所做的具体说明，又称产品目录、产品样本和产品说明书。	是生产科研单位研究分析各国技术发展状况和产品水平的重要资料，介绍的技术较成熟，是人们选型、设计和引进国外设备仪器有价值的参考资料。 有一定商业性质，有些产品是试销的，不十分可靠，应注意鉴别。来源不稳定，收集较困难。
技术档案	是指生产建设和科学技术部门在技术活动中所形成的，有一定具体工程对象的技术文件、图样、图表、照片、原始记录的原本以及代替原本的复制本。	是生产过程及研究工作中的用以积累经验、吸取教训、提高质量的重要文献，对科技人员有重要的使用价值。 有的具有保密性。 记载的内容准确、真实。
报纸新闻稿	是指那些阐述问题面广，具有群众性与通俗性，可获得一些重要消息的资料。	是发展远景的展望，是探讨运用某些新发明的可能性、现有的技术与生产工艺的改进方法，以及有关生产组织、合理使用设备、节约原材料等方面的文章或报道。 对科技成果报道不系统，缺乏详细的技术鉴定及理论根据。
工作文稿	一般是准备在期刊上发表或向学术会议提出的论文或研究报告的初稿，也称讨论文稿、研究文稿、工作文稿等。	工作文稿通常被打印出来供征求意见之用。它的价值是在它正式发表之前具有交流信息，它是一种短时效的文献，也是一种难以全面搜集的文献。

三、按信息载体的物理类型划分

(一) 印刷型

传统的印刷型包括铅印、油印、石印、胶印型等。其优点是便于阅读，便于传递，不受时间、地点的限制，不需任何阅读设备。缺点是贮存密度低，收藏管理要占用较大的空间和较多人力。

（二）缩微型

这是一种以感光材料为载体，利用缩微摄影技术为记录手段的类型。它具有节省空间、价格便宜、易于管理等印刷型所没有的优越性。目前最先进的超级缩微胶卷或胶片缩小倍率高达 1/22500，一张全息胶片可贮存 20 万页文献，一般的缩微平片也可存 98 页文献。从而节省书库面积 95%，便于保管和移动，但必须借助于阅读机，不能同时利用几种文献，因而利用不便。缩微型文献及阅读机还待于进一步改进和完善。

（三）电子型

电子型的前身称机读型，是以软盘、磁带、光盘等介质为载体，用键盘输入或光学扫描等手段记录，并通过计算机处理后生成的，包括光盘数据库、电子期刊等。它也称计算机阅读型，其特点是存贮信息量大，并能按一定的程序设计快速输出文献单元及知识单元，缺点是人不能直接阅读，必须借助于计算机。

（四）声像型

声像型，或称直感型、视听型，是一种用唱片、录音带、录像带、电影胶卷、幻灯片来直接记载声音和图像的类型。它脱离了传统的文字形式，给人以直观的感觉，可再现自然界的种种变化发展，帮助人们认识某些复杂罕见的自然现象，如细胞分裂、地震等，有助于联想记忆。而且录音带、录像带可起到快速传递科学信息的作用，缺点是需要相应的设备条件。

四、按信息的性质划分

（一）自然信息

自然信息是自然界发出的信息，它以纯自然物为载体，如生物信息、天体信息等。它是反映客观物质世界过程的原型。

（二）社会信息

社会信息是经过人类利用语言、文字、符号、图像等方式加工过的自然信息，是人间传播的信息，是人类活动的产物。

五、按信息所表征的服务对象划分

经济信息，服务于经济活动；科技信息，服务于科学技术事业；还有教育信息、军事信息、商业信息、金融信息、综合信息等。

1.3 检索工具及索引语言

1.3.1 检索工具

一、文献检索工具的概况和分类

检索工具是二次文献，是人们用来报道、存贮和查找文献的工具。一般来说，检索工具必须具备下列四项条件：

- (1) 对所收录的文献的各种特征（包括外部特征和内容特征）要有详细的描述。
- (2) 每条描述记录（即款目）都标明有可供检索的标识。
- (3) 全部描述记录科学地组织成一个有机的整体。
- (4) 具有多种必要的检索手段。

第一项条件是体现检索工具的报道和存储功能的。第二到第四项条件是体现检索工具的检索功能的。检索工具也有人把它称为“小书库”，把获取的原始文献即一次文献称为“大书库”。

目前，人们依不同的角度来划分检索工具。标准不同，划分的结果也就不同。

按检索方法划分，有手工检索工具和机械检索工具。

按其报道范围的宽窄划分，有包括多个学科或全部学科的综合性和检索工具，或仅包括单个学科的专业性检索工具。综合性检索工具一般由一些历史悠久、具有一定权威性的出版机构发行，提供的检索途径多。美国的《工程索引》、英国的《科学文摘》是利用率很高的综合性工具。专业性检索工具汇集有某专业的多种形式、多种语种的文献线索，适于该专业的科研人员查找有关文献，如《金属文摘》、《生物学文摘》、《化学文摘》等等。

还有按出版形式划分的文献都收录的全面性检索工具和仅收录其中某一类型文献的单一性检索工具。全面性检索工具如《化学文摘》、《科学文摘》、《工程索引》，单一性检索工具有《世界专利索引》、《会议论文索引》、《ISO 标准目录》、《IEC 出版物目录》等等。

二、文献检索工具的类型

各类型检索工具的内容及特点见表 1.3-1。

表 1.3-1 各类检索工具及特点

类型	内 容	特 点
目 录	是对图书或其他单独成册出版的文献特征的记载和描述，用于查找单位出版物文献。	目录所揭示的出版物都是编制目录的单位所具有的。目录有卡片式的、书本式的。只著录每篇文献的名称、著者、出处和文种等，不反映文献内容。报道及时，数量较大。
题 录	是对文献标题和外部特征的记载和描述，用于查找单篇文献。	文献标题（题目）着眼于尽可能快地报道最新发表的文献，不反映收藏单位是否实有。 多按篇报道，著录项目比较简单，对于标题不能或不完全反映内容实质的文献，有漏检的可能，它的检索作用是暂时的，是文摘的辅助部分。

类型	内 容	特 点
文摘	是对文献的摘要及外部特征的记载和描述,分为指示性和报道性文摘。	摘要是对原始文献主要内容语义上相同而又尽可能完备的不加评论和补充解释的陈述。实际上,摘要就是对原始文献的压缩,是在忠实于原始文献的基础上,把原始文献的内容浓缩成一篇语义连贯的短文,文摘通常包括摘要及文献出处等。
索引	是以一定的系统排列揭示文献中的各种知识单元,并指明其出处的检索工具。	通过线索而引得所要查的文献,是检索工具中最常用的一种。 索引一般包括主题、著者、各种号码索引等。 一般附在文摘刊物或书刊、参考工具书之后,起辅助检索作用。

三、文献检索工具的一般结构

文献检索工具的一般结构如图 1.3-1。

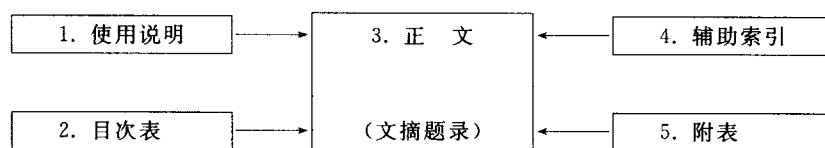


图 1.3-1 文献检索工具的一般结构

一部完整的文献检索工具一般分为五大部分：

1. 使用说明

使用说明是检索工具编制者为检索工具检索者提供的必要指导。它一般包括编辑内容、著录标准、代号说明和使用方法。

2. 目次表

检索工具一般都是按分类的方法组织编排的,所以前面有详简不等的分类目次表,检索者可从分类入手查找。

3. 正文

这是整个检索工具的主体部分。检索工具属于二次文献,汇集的大量内容不是原始文献。编制检索工具时,编制人员将每篇文献的外表特征(如书名、著者、出版单位、出版年月、出版地、页数等)和内容特征(如主题词、分类号、摘要等)提取出来,这个过程叫文献著录。著录后的每篇文献作为一条款目。给每条款目一个固定的序号,注明出处,把大量款目按一定规则(一般是分类)组织起来,即成为检索工具的正文。

4. 辅助索引

为了迅速、准确、全面地获得所需文献,检索者应采用多种检索途径。辅助索引是从主体部分的文献款目编制系统以外的角度,增加检索途径,它不能离开主体而单独存在,是主体的辅助部分。

5. 附表

通常附在检索工具之后, 作为检索工具内容的必要补充, 含有引用文献目录、缩写查全称表、文献使用的代号、缩写术语解释等内容。

四、网络检索工具

网络检索工具是为满足用户查找因特网上急剧增长的信息的需要应运而生的, 其中 Gopher、Archie、WAIS 属于较初期的信息检索工具, WWW 的出现为网络的检索带来了转机和大发展。

根据不同方式可将网络检索工具划分为不同的结果, 通常有以下两种划分结果:

(一) 根据网络检索工具特点划分

有字典型检索工具、索引型检索工具、交互式检索工具。

1. 字典型检索工具

它用于查询网上用户名、Email、URL (统一资源定位器)、服务器地址等, 这类工具有 Internet Yellow Pages、White Pages Directory、Whois 等。

2. 索引型检索工具

它主要是为网上信息源建立索引, 这类工具有 FTP 资源的索引 Archie、Gopher 资源的索引 Verronica 和 Jughead、网上服务器的索引 WAIS 等。

3. 交互式检索工具

它提供类似商用联机检索的网络信息查询服务, 这类工具有 Gopher 和 WWW 两类。

(二) 根据网络检索工具的检索结果划分

有文献型检索工具、数据型检索工具、事实型检索工具。

1. 文献型检索工具

它以文本数据为存储对象, 根据文献不同的著录对象和加工的深度分为书刊机读目录数据库型检索工具、题录和文摘数据库型检索工具、全文数据库型检索工具及引文数据库型检索工具。文献型检索工具依文献加工的深度又可以分为:

(1) 目录数据库型检索工具。它给用户提供一些简单而基本的信息以及原始文献的线索, 指引用户根据文献线索去检索原始文献, 获取充分的信息。目录数据库型检索工具使用最广泛, 包括各种机读版的文摘、题录等形式。

“机读目录”(Machine Readable Catalogue), 简称 MARC, 即机器可读型目录。书刊机读目录数据库主要报道和存储特定图书馆实际收藏的各种文献资料的书目信息和存储地址。读者通过书刊机读目录数据库的查询系统, 能方便地检索到某图书或期刊的藏书位置、出版情况; 图书馆工作人员通过它能实现书刊采访、编目、管理自动化。

(2) 题录、文摘数据库型检索工具。它是单篇文献的名称、著者、所在出版物位置的检索系统。题录型数据库其编排的基本格式是: 题目 (包括论文篇名、专利名称、成果名称、机构名称等)、作者、作者单位、出处 (如刊名、年、卷、期和页码)。题录型数据库检索方法与文摘型数据库基本一致, 但由于检索结果提供的信息量过于简单, 因此, 目前已逐渐被文摘型和全文型数据库所取代。

文摘数据库型检索工具是单篇文献的名称、作者、出处及内容摘要的检索系统。文摘型

数据库与题录型数据库具有相类似的检索功能和编排体例,但由于带有内容摘要,对文献的揭示更深入,可供检索的入口和角度更丰富,因此受到欢迎,是文献型数据库的主要形式。目前,常见的是题录型与文摘型并存的数据库,其中文摘的比例往往大于题录。

(3) 全文数据库型检索工具。此检索工具是存贮文献全文或节选其中主要部分而形成的原始文献数据库。以全篇文献为著录对象,系统揭示原始文献的全部完整信息。全文信息检索系统运行的全文型数据库是目前计算机检索研究和应用的热点,其主要特征是:文献中的任何有检索意义的字符串都可作为检索入口,检索结果能同时提供并输出全文和二次文献,检索结果更直观。

(4) 引文索引数据库型检索工具。引文索引数据库型检索工具是以被引用文献为检索起点的检索系统。引文索引是一种以期刊、专利、科技报告等文献后所附的参考文献的作者、题目、出处等项目,按照引证与被引证的关系进行排列而编制的索引。它在一定程度上表明了科学技术的发展过程,同时能帮助研究人员了解自己著作的被引用率和持续时间,从而估计其影响力。

2. 数值型检索工具

此类型检索工具主要提供数值型数据类信息。包括各种统计数据、科学实验数据、科学测量数据、可观测的数据、预测数据等等。这种数值型检索工具的数值型数据库涉及的范围比较广泛,其数据内容主要是数字和各种特殊的符号。这种数据库不仅可检索诸如物质特性、基本参数、图谱、图像之类的数值数据,而且可与各类应用程序结合起来,提供绘图、统计分析等各种方式的服务。

3. 事实型检索工具

事实型数据库存储的是经过加工的三次文献的信息,一般是用来描述人物、机构、事物等非文献信息源的情况、过程、现象、特性等方面的事实性信息。诸如名人录、机构指南、产品目录、科研成果目录、研究或开发项目目录以及大事记之类。利用事实型检索系统,用户能得到某一具体问题的答案,不必再花费时间去检索原始文献。

五、搜索引擎

在因特网浩瀚无垠的大海里,面对眼花缭乱的信息资源,如何又快又准地找到所需的信息,已成为众多信息检索人员面临的一个必须解决的问题。搜索引擎是因特网上提供公共信息检索服务的 Web 站点,它是随着人们对快速准确地检索网上信息的需求而诞生的一种新一代信息检索工具。在因特网的发展过程中,先后产生了 Archie、Gopher、WAIS 及 Search Engine (搜索引擎) 等多种信息检索服务。在这方面,搜索引擎派上了大用场,它通过提供 Web 信息检索的快捷方式,像图书馆的馆藏目录一样,提供链接路径,让搜索者到相应的网站去寻找相应的信息或资源,以实现信息的快速获取。搜索引擎的发展历史虽然时间不长,然而其影响却十分重大。因特网信息检索的特点为检索范围大、能结合多媒体、界面友好等,但其检索效率还有待改进,信息内容也有待规范。

搜索引擎可以从广义和狭义上去理解。从狭义的角度来说,搜索引擎由搜索软件、索引数据库和检索软件三部分组成。搜索软件从一个已知的文档集中读取信息,并检查这些文档的链接指针,找出新的信息空间,然后取回这些新空间中的文档,将它们加入到索引数据库。检索软件通过索引数据库为用户的查询请求提供服务,即指的是基于技术在整个网上自动执行网页全文搜索的网上指南工具。从广义的角度上讲,搜索引擎是因特网上的一类网