

科技文献检索

顾茂生 徐克敏

国家图书馆出版社

科技文献检索

赖茂生 徐克敏 等编

赖

本书据北京大学出版社 1985 年 3 月第 1 版排印

前 言

科技文献检索是大学图书馆学专业和情报学专业的专业基础课之一,并且即将作为理工科学生的一门工具课纳入教学计划之中。自北京大学图书馆学系 1974 年正式开设这门课以来,本书的编者曾先后参加过该课程的教学工作,并于 1980 年初集体编写成《科技文献检索》讲授提纲。之后,经过修改和补充,于 1981 年底编成《科技文献检索》讲义。在近两年的试用期间该讲义得到了本系和外界同行们的好评。这次正式出版之前,编者根据实际教学效果和各方面的意见,对讲义又进行了全面的修改和校订。

全书共分为九章,内容包括三部分:第一部分(一至五章)基本原理和基本方法,主要介绍科技文献检索的基本知识,检索工具和检索系统的一般构成机理,索引语言和索引方法。第二部分(六至八章)基本工具,重要介绍分析国内外一些有代表性且适用面较宽的检索工具的结构、特点和使用方法。第三部分(第九章)应用与评价,主要介绍如何综合运用基本原理方法和工具去开展检索工作,以及如何评价和改进检索工具或系统的性能提高检索效果。鉴于手检与机检关系极为密切,它们的方法和研究成果可以互相利用或借鉴,而且机检在我国已成为现实。因此,本书内容虽然主要立足于手检,但也在某些章节中适当地增加了机检方面的内容。

本书的主要参考文献一般都在各章末尾给出。

参加编写工作的人员有:徐克敏(第六章第四节),申嘉廉(第七章第二节),白光武(第八章第二节),刘明起(第八章第三节),孙琦(第八章第四节),赖茂生(第一章至第六章第三节,第七章第一节,第八章第一节,第九章及附录,并负责全书的统稿工作)。

本书付印前,承蒙北京大学图书馆学系周文骏副教授、侯汉清、利求同同志,化学系孙亦梁教授,生物系冯午教授,物理系虞丽生副教授等审阅有关部分。另外,在编写过程中还得到彭淮源、林尧泽、杜宝荣、赵凤仪、秦铁辉、祁延莉、赵文等同志的指导与协助。在此一并致以深切的谢意。

由于水平所限,难免有错误或不当之处,望读者指正。

赖茂生

1983 年 12 月于北大

目 录

第一章 绪论	(1)
第一节 文献检索概述	(1)
第二节 文献检索发展概况	(10)
第二章 科学技术文献	(18)
第一节 科技文献概述	(18)
第二节 科技图书	(22)
第三节 科技期刊	(23)
第四节 科技报告	(28)
第五节 科技会议和会议文献	(31)
第六节 专利文献	(33)
第七节 标准文献	(39)
第八节 其他一次文献	(41)
第九节 科技文献的发展状况	(43)
第三章 检索工具及其编制方法	(49)
第一节 检索工具与检索系统	(49)
第二节 文摘性检索工具	(53)
第三节 文摘编写方法	(66)
第四节 文摘刊物的编制	(76)
第五节 目录性与题录性检索工具	(82)
第六节 检索工具的质量评价	(88)
第四章 索引和索引法	(92)

第一节	索引导论	(92)
第二节	索引的结构	(95)
第三节	索引语言	(104)
第四节	索引编制方法	(109)
第五章	索引的类型	(118)
第一节	先组式字顺主题索引	(118)
第二节	元词索引	(129)
第三节	叙词法及其在索引中的应用	(141)
第四节	关键词索引	(149)
第五节	计算机自动标引和辅助标引	(158)
第六节	引文索引	(171)
第七节	著者索引和专用索引	(188)
第六章	专业性检索工具选介	(192)
第一节	美国《工程索引》	(192)
第二节	英国《科学文摘》	(206)
第三节	美国《生物学文摘》	(222)
第四节	美国《化学文摘》	(243)
第七章	综合性检索工具选介(一)	(292)
第一节	科技报告检索工具	(292)
第二节	专利文献检索工具	(318)
第八章	综合性检索工具选介(二)	(362)
第一节	我国科技文献检索刊物体系	(362)
第二节	日本《科学技术文献速报》	(384)
第三节	法国《文摘通报》	(403)
第四节	苏联《文摘杂志》	(422)
第九章	检索服务与检索效果评价	(445)
第一节	检索服务的方式和方法	(445)
第二节	检索技术、检索策略与检索方法	(454)
第三节	检索效果评价	(459)

附录一:常用检索工具选目	(471)
附录二:我国收藏的国外文献检索磁带目录	(480)
附录三:字母音译对照表	(482)

第一章 绪 论

第一节 文献检索概述

人类进入二十世纪以来,科学技术以前所未有的高速度向前发展,物理学和生物学相继发生了深刻的革命,化学、天文学、地学等领域也取得了惊人的进展。尤其是四十年代以后,由于战争和国防的需要,一些国家大力动员科学技术力量为国防服务,开展了大规模的综合性研究,出现了许多新的学科和研究领域。原子能工业、电子计算机和空间技术的出现,开始了全面的、影响深远的第三次技术革命。科学技术的高速发展,研究规模的迅速扩大,各种研究成果大量涌现,导致科技文献的数量急剧增长。这主要表现在:科技期刊和其他非公开出版的文献的数量迅速增加,出版物的篇幅也在增加,各学科的文獻愈来愈分散,而且学科之间互相渗透。另一方面,由于条件的限制,人们吸收和利用情报的能力并未得到相应的提高,人们对汹涌而至和川流不息的文献洪流感到难以应付,许多有用的情报还未来得及发现和利用就自生自灭了。文献量的增长和人们利用情报的能力两者之间出现了越来越大的剪刀差。传统的情报处理和检索方法已不能适应了,科学交流的渠道出现了严重的障碍。对这种社会现象,人们惊呼为“情报危机”或“情报爆炸”。

面对这种危机和障碍,再加上某些重大国际事件(如 1957 年

苏联发射人造卫星成功)的发生,使许多国家的政府和科技界开始重视情报问题,迫切需要更有效的情报技术和方法,需要培养情报专家。因此,情报检索作为一个专门的研究领域,逐步纳入美苏等许多国家高等学校的教学计划中。各种与情报检索有关的研究开发活动、培训活动和学术交流活动越来越频繁,规模也越来越大,有关的书刊和文章也纷纷涌现。

一、文献检索的定义

为了更确切地解释文献检索这个概念的含义,我们先解释一下什么叫情报检索。

情报检索(Information Retrieval)——是指将情报按一定的方式组织和存储起来,并根据情报用户的需要找出有关的情报的过程。所以,它的全称又叫“情报存储与检索”(Information Storage and Retrieval)。这是广义的情报检索。狭义的情报检索则仅指该过程的后半部分,即从情报集合中找出所需要的情报的过程,相当于人们通常所说的情报查寻(Information Search)。

作为检索对象的情报,它有不同的形式,有的以文献的形式出现,有的以数据或事实的形式出现。根据检索对象的形式不同,情报检索又分为文献检索和数据检索。凡以文献(包括文摘、题录或全文)为检索对象的,就叫文献检索(Document Retrieval)。同理,凡以数据或事实为检索对象的,则是数据检索(Data Retrieval; Fact Retrieval)。可见,文献检索只是情报检索的一部分,但它是其中最重要的一部分。

文献检索还可以分为手工式文献检索(简称“手检”)和计算机化的文献检索(简称“机检”)。两者的关系是:手检是基础,机检是发展方向。这种划分方法也适用于数据检索。

从性质上说,文献检索是一种相关性检索,系统不直接解答用户所提出的技术问题本身,只提供与之相关的文献供用户参考。

例如,某用户需要有关建造压水堆式核电站的技术资料。这是属于文献检索范畴的问题。而数据检索则是一种确定性检索,系统要直接回答用户提出的技术问题,即直接提供用户所需要确切的数据或事实。而且检索的结果一般也是确定性的,要么是有,要么是无;要么是对,要么是错。例如,某用户询问目前全世界有多少座核电站,或者询问世界上有哪些国家建有核电站,年发电量分别是多少,有效利用率如何,等等。这些都是属于数据检索范畴的问题。文献检索与数据检索除有上述的区别之外,两者在原理、方法和实践方面并无其他本质区别。相反,它们之间倒是存在密切的关系。例如,美国化学文摘社(CAS)生产的数据库当中就包括文献型数据库和非文献型数据库,检索时这两种数据库可以配合起来使用,最终都是检出化学文献。

另外,有些教科书或文章中常常单独提到“事实检索”(或事项检索)这样一个概念,这里需要说明一下。如果它所指的是从情报集合中找出用户所需要的事实的过程,那么,这种事实检索就是上面所说的数据检索。因为事实也是一种数据,即非数值性数据,所以本讲义把事实检索放在数据检索这个概念里头,不单独列出来下定义。但是,如果它所指的是通过对情报集合中已有的基本事实或数据进行逻辑推理,然后输出新的(即未直接存入情报集合中的)事实的过程,那又另当别论了。有人认为,后一种所谓的“事实检索”,实际是一种逻辑运算过程,不属于情报检索的范畴。我们认为这种看法是有一定道理的。

二、文献检索在科学交流中的地位

要恰当地认识文献检索在科学交流中的地位,还得把它放到科学交流或“情报流程”中去观察,看看它在其中处于一种什么样的地位。

科学交流的过程就是情报流动的过程,即情报从情报生产者

到情报用户之间的传递过程。社会实践中每时每刻产生着和利用着的各种情报的总和,称为“情报流”。这是情报从某一点向另一点传递时,由于不断地大量地传送而形成的一种流态。情报流所流经的路线就称为“情报流程”,它相当于通讯技术中的信息通道。情报流程是否合理简便,反映了科学交流的效率。

在研究工作中,每个科技人员都需要跟他的前辈或同时代的人交流情报(包括单向交流和双向交流),而且人们总是希望通过最简便的途径来交流情报。当社会上科技人员数量较少,情报生产量也比较小的时候,用简单的交流渠道,如直接交谈、通信或出版几部著作,大体上就可以满足科学交流的需要了。但是,随着人类社会的不断进步,科学技术越来越发展,参加科学技术活动的人数越来越多,情报的生产量愈来愈大,形式愈来愈多样化,人们对情报的需求量也愈来愈大,愈来愈迫切。原来那种简单的交流渠道渐渐地不能满足科学交流的需要了,原有的情报传递路线也难以完成传递情报的任务了,常常发生通道过分拥挤甚至堵塞的现象。这时,人们就会要求变革科学交流的方式,增设新的传递渠道,或扩大原有通道的容量,以便使情报交流的渠道重新畅通起来。于是一些新的交流渠道和交流手段就相继出现,一部分情报就通过新的渠道传递到情报用户手中。今天的情报检索和情报检索系统就是在科学交流方式的不断变革过程中出现的。这里,我们试用一幅情报流程图(见图1-1)表示各种常见的科学交流渠道。

从图1-1可以看出,包含着文献检索的情报检索,它是传递情报的一条重要渠道,是联系情报生产者和情报用户的一个主要中间人,是人类为了合理地分发情报、充分地利用情报和提高劳动效率而采取的一种重要的交流方式。文献检索的出现和不断发展,使科学交流从非正式的、无组织的状态逐渐演变为正式的、有组织的状态。它使原来的情报流程大大简化了,通道的容量也大

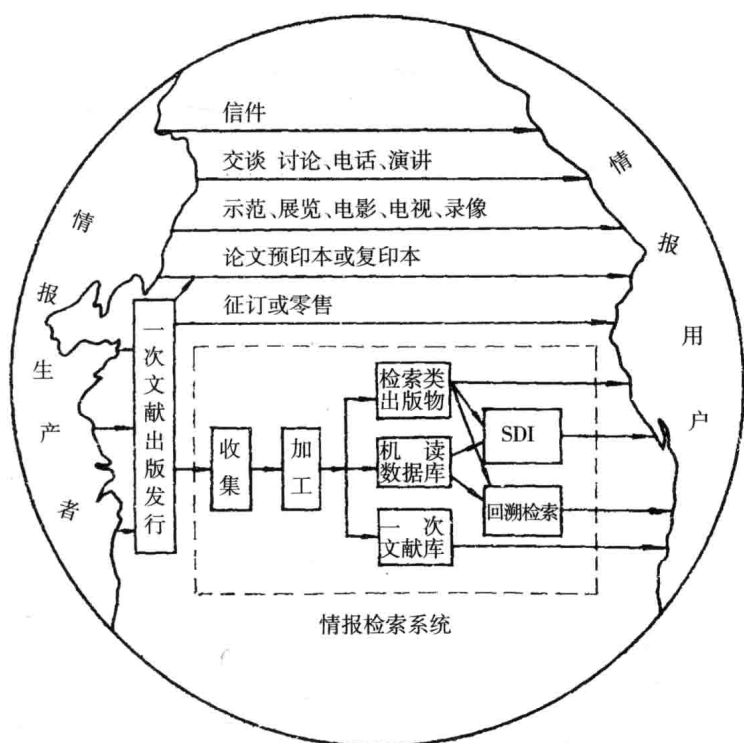


图 1-1 情报流程图

大扩充了,科学交流也显得更加便利和完善了。科学交流方式的这种演变,是社会进步的一种表现。

三、文献检索的重要作用

文献检索在科学交流中的突出地位,已经显示了它在社会生活中的重要性。对于它的直接作用,人们有各种不同的说法。在这里,我们结合有关事例将文献检索的作用归纳为下列四个方面。因为情报工作的核心和基础是文献检索,所以,尽管下面有些例子未直接提到文献检索,但实质上都是与它有密切关系的。

(1) 促进智力资源的开发和两个“文明”的建设

历代流传下来的和目前正源源不绝地涌现着的文献,是一个巨大的知识宝库,是一种如同能源、材料和劳力一样的重要资源。能否充分地利用这种资源,直接影响到一个国家的科学、教育、文化和经济的发展。有人说,现在国际之间的较量主要是知识的较量,是靠知识去创造产品价值,争夺国际市场,甚至控制别国。情报事业是否发达,是衡量一个国家是否发达的重要标志之一。一般来说,不发达的国家,情报事业也不发达。这种不发达主要表现为:情报意识薄弱,情报技术落后,人们不能及时地便利地获得全面而准确的情报,社会和个人的知识不能及时地得到补充和更新,所以科学技术和文化就落后,科学人才就匮乏。我国要实现四个现代化,建设社会主义的精神文明和物质文明,这都离不开知识,离不开智力资源的开发和利用。从全社会来说,文献检索是人们打开知识宝库的一把金钥匙,是开发智力资源的有力工具。它能帮助人们传播知识和利用知识,使知识转化为社会物质财富或创造出更多的精神财富,为建设两个“文明”贡献力量。

(2) 协助管理者作出正确的决策

情报虽不能确保决策正确无误,但它是决策的基础。《孙子兵法》云:“知己知彼,百战不殆。”打仗是如此,搞管理也是如此。在今天这个复杂的社会,管理问题显得尤其重要。从一个国家(地区或单位)来说,要发展什么,限制什么,引进什么,都需要有准确、可靠和及时的情报,以便作出正确的决策。日本引进大量国外先进技术,促使其经济“起飞”,重要原因之一是技术情报搞得好,及时准确地掌握了国外科技发展动向。而我国前几年由于情报不灵,盲目引进,使一些引进技术和设备不适用,造成了惊人的损失和浪费。再举一个具体的例子,比如,有个发达国家想和我有关部门做一笔生意,要求我方在很短时间内提出一个包括几十万种产品的价格单。我们的人答复说,报价格单可以,但需要半年时

间。对方说等不及,结果这笔生意就吹了。可见,作为一个管理者,如果平时重视情报工作,善于利用情报检索等有效工具,就有可能避免重大的损失和浪费。

(3)帮助研究人员继承和借鉴前人的成果,避免重复研究或走弯路

整个科学技术史表明:积累、继承和借鉴前人的研究成果,是科技发展的重要前提。没有继承就不可能有创新。所以,牛顿说:“假如我比别人看得略为远些,那是因为我站在巨人的肩膀上。”

在研究工作中,任何一个项目从选题、试验研究或设计、到成果鉴定,每一步都离不开情报。只有充分掌握了有关情报,知道哪些工作前人已经做了,哪些目前正在做,什么人在做,进展情况如何,这样才能避免重复,少走弯路,保证工作在高水平上进行,并缩短研究周期,获得预期的经济效果。例如,日本高能物理研究所由于借鉴了外国的经验及其它因素,它的第一台高能加速器投资比国外同类装置节省 $1/2$,人力节省 $3/4$ 。苏联在研究美国航天科技情报的基础上制成的登月 1 号和登月 2 号,其中的电气自动控制系统投资,分别只有美国的 $1/20$ 和 $1/50$ 。

相反,如果继承和借鉴工作做得不好,“闭门造车”,就容易造成重复研究;或者采用了较差的实验方法或程序,使研究工作进展缓慢甚至失败。这种例子是很多的。例如:美国的几家研究机构为研究继电线路的综合方法,花了大约 20 万美元。事后发现苏联已有这种装置在运行,有关其结构的情报也已公开发表。美国曾用专门的发生器向大气喷射气溶胶进行人工造云试验,花了 25.6 万美元,并公布了试验结果。但是,另外一组美国研究人员不知道这一试验,也进行了这种试验,又花了大量的研究费用。又如,日本在五十年代由于未掌握确切的情报,向外国申请了大量落后的专利,造成的经济损失占全部专利申请费的 65%。类似的例子许多国家都有。据有的材料透露,在五、六十年代,美、苏、日等国的

重复研究率都比较高,造成的经济损失达数十亿美元以上。在欧洲共同体的研究开发活动中,由于忽视已有的成果,情报不灵,使15%的研究工作要么重复,要么不能受益。我国的情报工作一直比较落后,所以重复研究的现象一直较为严重。一方面是重复研究国外已有的技术,另方面,国内各机构之间相互重复研究的现象也不少。某市有几位科研人员长期废寝忘食地搞出一项成果——熔点仪,满怀希望准备申报国家发明奖。但是,在查阅专利文献时,发现几年前别人就已完成同样的工作,并获得了专利权。又如,某研究所花了十年时间研制成功“以镁代银”新工艺。可是,美国某公司早在1929年就已经申请了这方面的专利,而且该项专利的说明书就收藏在附近的情报所。

(4) 节省研究人员查找文献的时间

文献量过分庞大和迅速增长,加重了研究人员搜集情报的负担。据苏联有关部门分析,研究人员在进行一项研究时,查找资料的时间“由过去的几小时,增加到几天,甚至几十天以上”。我国科技人员也有同感。有一个搞生化研究的研究员有一次向情报人员诉说:四十年前,我搞一个不太大的课题,到图书馆查两个星期的资料,就可以动手做实验了。隔一段时间再到图书馆去翻一翻,看一看有什么新东西,以便实验能顺利地进行。六十年代以后,信息爆炸,东西太多,过去的老办法不行了。在图书馆查了一个月,好像差不多了,可以开始工作了,其实不然。因为后面东西越来越多,就是一年到头,从早到晚坐在图书馆里,恐怕也掌握不了课题的全面情况。前两年有些转机了,化工部买了CA的检索磁带,开展定题服务,初步解决了我们的困难。

据国内外有些材料估计,研究人员在查文献上花费的时间是相当多的,一般约占本人工作时间的1/3左右。如果有完善的检索设施和周到的检索服务,无疑会节省研究人员的大量时间,腾出更多的精力搞研究,提高科研效率。

由于文献检索具有上述多方面的作用,所以,许多国家和单位积极支持文摘索引工作和检索服务工作,尽量采用新技术新设备去武装它,不断开拓新的服务领域,提高服务质量,以支持科学研究工作和其他开发活动。

四、文献检索的研究对象、目的和方法

文献检索作为一个学科或研究领域,是属于情报学的一个重要分支。它是在文献工作、图书馆学、目录学等领域的交叉点上发展起来的比较年青的学科。它植根于文摘索引工作和文献检索服务工作,并且越来越多地从计算机科学、数学、语言学等学科引进一些科学方法和技术手段,在此基础上逐步形成了自己的专门研究领域。目前,这个领域的轮廓已基本明朗化。它的研究对象大致可以归纳为下列四个方面:

(1)检索对象,即科学文献,包括它们的性质、特点、用途以及专门的查找工具;

(2)检索工具和检索系统,包括对其整体的全面考察和对某种工具(或系统)的研究;

(3)索引语言和索引方法,这是文献检索的核心问题;

(4)检索策略、检索技术、服务方式以及检索效果评价。

研究上述四个方面,总的目的是为了总结继承前人管理文献和利用文献的宝贵经验,建立和不断完善文献检索的理论,发展新的技术、新的方法和新的服务领域,指导文摘索引工作和检索服务的实践,使庞大的科学文献得到有效的控制和充分的利用。

文献检索的研究方法目前处于发展阶段,还有待完善。在研究过程中,人们较常用的方法是分类描述和结构分析等方法。分类描述法是在收集许多具体事例的基础上,加以分类整理,以认识事物的特征和相互联系。结构分析法是对具体的系统或方法的结构成分进行分析,以掌握其构成机理和功能。五十年代末,人们开