

国际联机情报

检索原理与方法

龚国伟 编著

湖北科学技术出版社

国际联机情报检索原理与方法

龚国伟 编著

湖北科学技术出版社

内 容 提 要

本书全面系统地阐述了国际联机情报检索系统的基本工作原理、系统结构与设备配置、检索策略的编制与技巧的运用,分别介绍了Dialog、ESA、STN、BRS和Orbit等五大国际联机检索系统的发展历史与特点、数据库及其文档结构、检索功能与技法及其87、88两年的最新发展等,并对比分析了这些系统的优缺点,提出了改进的方法与建议。书中既有各类具体的检索实例分析,又有多种指导学员上微机实习联机检索模拟系统的操作指导书。书末还附有“国际联机情报检索”课考试试题及参考答案。

本书可作为高等院校图书馆学、情报学、经济管理学和计算机科学等专业以及其他专业的研究生和高年级本科生“情报检索”课的教材,也可供一切想掌握现代化信息检索方法与技巧的科技人员、经营管理人员、大专院校师生和图书情报工作者学习参考。

国际联机情报检索原理与方法

龚国伟 编著

湖北科学技术出版社出版发行 新华书店湖北发行所经销

湖北省出版印刷物资公司印刷厂印刷

787×1092毫米 16开本 16印张 430千字

1989年2月第1版 1989年5月第1次印刷

ISBN 7-5352-0386-1/TP·7

印数,2,500 定价,9.80元

序 言

龚国伟副教授编著的《国际联机情报检索原理与方法》一书问世了,我感到由衷的高兴,并表示热烈的祝贺。

美国Dialog、Orbit、BRS,以及欧洲的Esa和Stn等联机情报检索系统,是当代规模最大、数据库最丰富、用户终端覆盖面最广的商业化国际情报服务系统。它们是广大科技、商业经济、社会政治等各界从事研究、生产、经营管理和决策的人们获取情报的重要渠道。在这些系统中,有数以亿计的文献题录、文摘、全文、事实、数值与图象信息资料。借助于现代通信设施,检索者可同这些“远在天边”的系统进行如同“近在咫尺”的会话式检索,实现最新情报的获取和时间跨度可长达几年~几十年的追溯式文献资料普查。国际联机情报检索系统及其向广大公众提供灵活多样的服务,是当代信息化社会的产物和标志。八十年代以来,我国已先后开通了上述系统的终端服务,在北京、广州、上海、武汉等30多个大、中城市都可以对上述系统进行检索。近年来,情报界、科技界和工商企业界的广大用户对国际联机检索的兴趣与日俱增,对学习与掌握这些系统的功能和检索技巧的愿望日益迫切。正是在这种背景下,龚国伟同志的这本书应运而生。它的出版是非常适时的。

虽然在此书出版之前,我国科技情报界已发表与出版了若干有关国际联机检索方面的文章与书籍,但这本书却有着它的独到之处。首先是其论述的全面系统性、材料的翔实丰富性和分析的精辟透彻性。它不仅叙述了以上五个联机系统的历史、数据库及其文档结构、检索功能与特点、最新发展动态等,而且对它们作了相互间的比较分析,特别是对这些系统存在的缺陷和不足之处进行了剖析,并提出了改进的意见。这些都是较国内同类著作为优的。此外,它对BRS和STN系统所作的详细分析介绍也是以前国内同类著作中所少有的。第二,此书理论与实践相结合,它不仅叙述有关事实与检索实例,而且对国际联机检索的原理、检索策略及其调整的理论技巧有专章论述。这样,便于读者能从实践基础上向情报检索的理论高度思考,以检索理论提高检索实践的效果。因此,此书不仅有教程和手册的功能,而且超过了这种功能,可视为情报检索理论专著之一。第三,此书是一本供高校图书馆学、情报学、经济管理学和计算机应用等专业“国际联机情报检索”课程使用的良好教材。书中的Dialog、ESA和STN等模拟检索系统的上机实习操作指导书,颇便学生在微型机上练习。书末所附的“国际联机情报检索”课考试试题与参考答案也编写得十分精采。可以预言,此书的出版必将促进我国图书馆学、情报学与经济学等专业《国际联机情报检索》课程的普遍设立,丰富这些专业的教学课程体系。

此书作者是一位较为年轻但富有成果的图书馆学、情报学与经济信息管理学的研究工作者和教师,他先后出版了《微型计算机在书刊情报资料工作中的应用》、《联机检索策略》(上、下册)等多部著作,发表过大量的系列研究论文,并有多年的教学与实践经验。此书是在他的1984年交通部全国情报专业干部进修班讲义《计算机联机情报检索方法》和同年的全国信息组织高级信息员培训班讲义《国际联机情报检索原理》的基础上,经过大幅度修改和更新而重新编写的。作为初稿,他在1985年全国情报精英高级讲习班(苏州)、1987年东北三省文献检索课专家讲习班(哈尔滨)和1988年全国高校“文检”、“机检”课师资培训提高班(武汉)等高层次的研讨班上先后讲用过,受到同行专家教授们好评。此书并在武汉水运工程学院和华中师范大学图书情报学专业的历年教学实践中得到检验。因此,此书是作者多年辛勤耕耘的力作,是有较高价值的。为此,我向情报界、图书情报学专业教育界以及广大对情报检索有兴趣的科技人员和企业管理工作者推荐此书。

陈光祚 1989年3月

编 著 者 说 明

国际联机情报检索是当代信息化社会的重要标志,也是现代图书情报工作者、科技人员以及管理决策者必须掌握的一种工具与手段。然而,遗憾的是在我国这种先进的技术手段未能得到充分的利用与重视,例如,我国的使用率(年检索课题量)仅为美国的0.3%、英国的1.5%,以至于因信息不灵造成决策失误、重复科研、乃至在外贸与引进谈判中吃亏上当之事经常发生,从而给国家造成巨大损失与浪费。我国的国际联机情报检索工作难以深入开展与普及的原因是多方面的,但其中教育与培训工作跟不上需要也是一个重要原因。据一项调查资料表明,目前我国绝大多数科技人员和大专院校师生都不了解这项新型的情报检索技术,甚至80%的情报专业人员都还没有掌握这项先进的检索技法。

因此,在高校图书馆学专业、情报学专业和经济管理专业以及其他专业的研究生和高年级本科生中开设“国际联机情报检索”专业课程是十分必要的。由于以往“科技文献检索”课仅仅将联机情报检索或计算机情报检索作为一个较短的章节进行概念性的介绍,而没有涉及到具体的检索系统及其检索方法与技巧;而“计算机情报检索”课或“情报检索自动化基础”课,又多侧重于机检理论、数据结构、检索算法和系统评价等设计方面的内容,没有从用户的角度对系统的应用技术进行深入探讨。为了弥补这方面的不足,使学生掌握完整的情报检索理论和实用的联机情报检索技法,以适应社会发展的需要,作为尝试,笔者于1984年3月在武汉水运工程学院情报学专业(干部进修班)中正式开出了“国际联机情报检索”课(作为“科技文献检索”课的后续课程)。1986年,笔者所在的华中师范大学图书情报学系正式将该课列为84学时的专业必修课,本书就是在这些课程讲义的基础上丰富发展起来的。为了完善“国际联机情报检索”课的教材体系,笔者正计划编写出版一本与本书配套的习题集(包括答案、部分考试题以及上机实习题),并计划分三册出版商业经济、工程技术、生物医学等三个专业的联机情报检索技法的教材(其中商业经济类分册即将出版)。

本书在编著过程中,得到著名情报学专家、国家科委全国计算机情报检索领导小组技术顾问、武汉大学图书馆学情报学研究所所长陈光祚教授的指导和帮助,还得到我国著名的情报学专家教授:曾民族、于宏、高崇谦、赵宗仁,以及华中师范大学校领导与图书情报学系、校科研处和教务处等部门有关负责人的热情支持,在此一并表示衷心的感谢!

此外,本书在编写中参考引用了一些很有价值的文献(见参考引用文献目录),也特向这些文献的作者们表示谢忱!

由于编写时间仓促,加上笔者水平有限,书中一定存在不足乃至错误之处,恳请读者批评指正。

高 强 伟 1989年2月

于武汉华中师范大学图书情报学系

目 录

序言

编著者说明

第一章 国际联机情报检索绪论	1
第一节 国际联机情报检索的产生与发展概况	1
第二节 国际联机检索的服务方式	5
第三节 国际联机情报检索的优点	6
第四节 我国的国际联机情报检索工作概况	10
第二章 国际联机情报检索系统的构成	15
第一节 国际联机检索的通信线路与设备	15
第二节 国际联机检索系统用户终端设备	21
第三节 我国国际联机检索终端的分布及其通信线路	29
第四节 国际联机检索的主计算机系统	32
第五节 联机数据库	34
第三章 国际联机情报检索的基本原理	56
第一节 计算机情报检索的基本原理	56
第二节 联机情报检索系统最基本的检索功能	57
第三节 计算机检索的逻辑处理基本过程	61
第四章 DIALOG系统及其检索方法	65
第一节 DIALOG系统简介	65
第二节 DIALOG系统的联机通信过程	68
第三节 DIALOG系统的常用指令	73
第四节 DIALOG系统的常用算符	88
第五节 检索字段的限定	101
第六节 其他限定检索范围的方法	106
第七节 主要辅助检索功能	110
第八节 DIALOG系统的其他功能与指令	120
第九节 DIALOG系统的最新进展	124
第十节 DIALOG系统的辅助检索参考工具书	133
附录4.1 美国DIALOG模拟检索系统上机实习操作指导书	136
附录4.2 DIALOG系统数据库一览表	142
附录4.3 DIALOG系统文档专业主题分类索引	157
第五章 ESA/IRS、STN、BRS和ORBIT系统及其检索方法	160
第一节 ESA/IRS系统及其检索方法	160
第二节 STN系统及其检索方法	169

第三节 BRS系统及其检索方法	179
第四节 ORBIT系统及其检索方法	199
第五节 DIALOG、ORBIT和BRS系统检索的不一致性及其标准化问题.....	207
附录5.1 ESA/IRS系统联机检索模拟软件——QUESTSIM使用说明书	217
附录5.2 STN系统数据库目录及其特色文档简介.....	218
附录5.3 STN联机检索模拟软件——MENTOR使用说明书	224
附录5.4 ORBIT系统特有数据库简介	226
附录5.5 BRS系统数据库分类目录	229
第六章 联机检索策略与技巧	231
第一节 了解用户需求和分析检索课题.....	232
第二节 选择合适的数据库和联机检索系统.....	235
第三节 选择检索用词和编制逻辑检索式.....	238
第四节 分析误检或漏检原因随机调整检索策略.....	245
附录6.1 《国际联机情报检索》课考试试题与答案	247
主要参考与引用文献	249

第一章 国际联机情报检索绪论

第一节 国际联机情报检索的产生与发展概况

什么叫联机情报检索？它是指情报用户利用检索系统网络的终端设备，通过通信线路，运用一些特定的指令和检索策略与存贮了大量情报资料的计算机数据库检索系统的主机进行人机对话，从后者存贮的数据库中查找出所需情报资料并将检索过程与结果显示和打印（或存贮在磁盘上）出来的全过程。而国际联机情报检索就是指情报用户利用终端设备，通过国际卫星通信网络与地球上任何地方的大型计算机情报检索系统的主机进行直接的人机对话式的联机检索的全过程。简单地说，国际联机情报检索即国际化的全球范围的联机情报检索。

一、国际联机情报检索产生的时代背景

1. 计算机国际联机情报检索是对当今世界文献量剧增的直接反应，它是人类走向信息化社会的必然产物。

当今世界上的发达国家正处在从工业化社会向“后工业化社会”和“信息化社会”过渡的时代。所谓信息化社会，简单地讲，就是大量产生知识，依靠知识的生产、积聚和传播来推动社会的进步，推动社会生产力和经济发展的社会。也就是说，信息已成为社会的一种重要资源，成为社会生产活动、科技活动、乃至日常生活中不可忽缺的要素。由于现代科学技术的飞速发展，社会经济活动越来越活跃，使得科技文献和经济信息资料的数量爆炸性地增长。据1986年8月30日的《南京日报》报道，当今世界每一分钟产生一本图书、每三十秒钟产生一篇专利文献，每5~6秒钟产生一篇论文。根据美国报刊文章的统计报道，全世界科学技术杂志的数量从1965年到1981年已翻了一番（由28,000种增加到65,000种）。加上分支学科和边缘学科不断繁衍，各学科间交叉渗透现象日益突出，使得文献的分布十分分散。这种状况给搜集情报资料、查阅文献的工作带来了极大的困难。在这种情况下，如还用传统的手工方式在浩如烟海的文献资料中查找所需的情报，将会像大海捞针一样困难。同时，由于现代科技情报日新月异，经济市场信息瞬息万变，使得文献资料的寿命（或称有效期）越来越短。西方国家认为，当今世界80~90%的科技文献其有效期为3~5年。经济情报的寿命则更短，80~90%的经济情报其有效期为1~4周，有的甚至仅90秒钟。这就对情报检索的速度提出了十分严格苛刻的要求。在这种情报资料的作用和需求与其获得和利用之间的矛盾日趋尖锐的现实状况下，即使象美国这样高度发达的国家（其信息技术及设施居世界之首），因情报信息不灵而造成的损失平均每年高达25亿美元（例如，美国在电子领域中研究课题的重复率就达35%）。这种严峻的现实迫使情报服务部门必须采用以计算机为主的现代化装备和手段。

2. 计算机技术、高密度海量存贮技术以及现代通信技术的飞速发展，为国际联机情报检索的产生与发展提供了物质基础。

在这里,让我们简单地回顾一下计算机情报检索的发展历程(有关计算机情报检索的详细信息历史,请参阅笔者编译的《图书情报人员机检须知》一书):

五十年代初、中期,由于计算机尚处在第一代(电子管)时期,用于情报检索有很大的局限性,因此几乎没有实用系统,有的也只是内部的实验性或半实验性系统,而且只能脱机检索。据记载,最早的计算机情报检索系统雏型是美国海军军械中心(NOTS)于1954年研制的。该中心在IBM-701型计算机上建立了世界上第一个科技文献检索系统,但该系统仅能把文献号码和少量索引词存贮在计算机中,检索的结果(命中输出的)也只是文献号码。因此,它只是一种试验性的内部系统。到了五十年代末和六十年代初,由于第二代计算机(半导体分立元件)的大量涌现,计算机系统的软、硬件都有了很大发展,使得计算机文献处理与情报检索有了很大的活力,从而进入实用化的脱机批处理*检索方式阶段。1961年,美国化学文摘社(CAS)第一个公开将计算机应用于文献型(书目型)情报的处理中,由此产生了世界上第一个机读字顺主题索引——《化学题录》(Chemical titles,简称为CT)。它索引了包括化学文摘在内的600多种最有影响的化学化工杂志。由此产生的上下文关键词(KWIC)索引,是对出版物中的题名自动进行抽取关键词,从而节省了在概念标引中人的智力负担。这种索引原理代表了自然语言和全文检索的原始形式,此后在机检系统中得到了广泛的应用。1963年,美国国家医学图书馆也开始使用计算机来编制世界医学文献的检索刊物《医学索引》(Index Medicus)的月度主题索引。其后(1964年),为了适应用户的需求,开始以批处理方式对数据库进行追溯检索。到了六十年代后期和七十年代初期,由于第三代计算机(集成电路)的产生与发展和高密度海量贮存器——硬磁盘及磁盘机的出现,以及数字通信技术的发展和分组交换公用数据通信网的普及,使得情报检索从脱机批处理阶段进入联机检索阶段。单台计算机通过通信线路可联接多个检索终端,利用分时技术,若干个用户可同时直接与主机以“对话”方式进行检索。最早出现的联机情报检索系统是美国洛克希德导弹与宇航公司下属的一个实验室(后发展为DIALOG情报服务公司)于1965年至1966年间建立的DIALOG系统(其详细情况将在第四章中介绍)。紧随其后的是美国国家医学图书馆的MEDLINE(Medical Analysis and Retrieval System On-line即联机医学分析与检索系统),1967年秋,美国国家医学图书馆与系统发展公司(SDC)签订合同,将该公司的ORBIT(On-line Retrieval of Bibliographical Information Time-shared)软件装入该馆的系统,并使用神经病学方面的小型数据库,进行联机情报检索的试验。1970年6月,该馆利用IBM 360/67计算机,通过TWX通信网络(Teletypewriter Exchange Network)连接美国各地90个医学机构的远程终端,开展对该馆《医学索引》数据库的联机检索服务。到了七十年代中、后期至八十年代初期,由于第四代和第四代半计算机(大规模集成电路和超大规模集成电路)的出现,以及卫星通信技术和光纤通信技术的实用化,使得计算机情报检索冲破时间和空

*脱机批处理——从事这种服务的是一些专门学科领域的情报中心,它们从数据库生产者那里购买机读磁带装入自己的计算机中,并编制必要的软件,以进行现期通报和追溯检索的成批式处理。从机器时间来说,批处理是经济的,但有两个严重的缺点。第一是迫使用户在收到检索结果方面有一段时间(有时甚至是几个星期)的延迟,因为批处理方式须积累到一定数量的检索提问后才能进行,以达到经济的目的。第二,检索者对检索的过程与结果是“盲目”的,因为,检索策略是由情报中心的工作人员(检索分析员)代他们拟订的,检索者无法从检出文献中及时进行取样,以检查其相关性并予以及时的反馈调整。

间的限制,而进入国际联机检索阶段;情报检索超出一个地区、一个国家的范围而进入国际信息空间领域。过去千里迢迢不能找到看到的情报资料,现在利用国际联机情报检索终端,只用几分钟就可以查遍几年~几十年的世界各国几万种期刊中的有关资料;利用终端,用户可检索出前一天甚至前几分钟的世界经济和金融市场信息。正如现在世界流行的、人们形容国际联机情报检索的一句话: A world of information at your fingertips (世界信息就在您的手指尖上)。国际联机情报检索的成功运用,为快速全面地获取全球性科技情报资料和经济信息提供了非常方便的条件,从而极大地提高了情报资料的可获得性和利用价值,充分实现了人类情报资源的共享。所以,在1975年美国情报学会的年会上,人们一致把联机情报检索称为过去十年间人类最重大的革命。

二、国际联机检索的发展现状与趋势

目前世界上国际联机情报检索已发展到相当高的水平,联机网络和检索终端几乎遍及世界所有的地区和国家。开展国际联机数据库检索服务的机构(又称作卖主)已达二百家左右(其中美国116家,欧洲51家,加拿大19家,日本3家)。其中属于商业性的规模较大的著名国际联机检索系统有:①DIALOG系统(美国DIALOG情报公司经营,现有370多个数据库,文献量已超过1亿6千万篇,文献专业范围涉及到所有的学科领域,用户终端超过7万家,几乎遍及世界各大、中城市。年收入近3千万美元);②BRS系统(由美国文献题录检索服务公司经营,现有128个公用数据库和40多个私人数据库,报导范围涉及各个学科,尤以产品信息和工业标准与技术规范方面的数据库独具特色。其全文数据库多达23个以上,全文检索服务优于其他联机系统。其用户终端已有2万5千多家,遍及世界各地);③ORBIT系统(原属美国系统发展公司,现改由英国Pergamon ORBIT Infoline公司经营,有近80个公用数据库,以专利、石油与能源方面的数据库独具特色,其用户终端约3万家);④MEDLINE系统(美国国家医学图书馆经营,约有30个生物学方面的数据库。该系统有时也称为NLM系统,它是美国资格最老的联机检索系统);⑤NYTIB系统(美国纽约时报社经营,有10多个数据库,主要提供纽约时报、世界邮报以及世界各地60多种主要报纸和杂志的详细文摘。其中GLOBE AND MAIL——世界邮报数据库和NEW YORK TIMES ONLINE——纽约时报联机资料库可提供全文检索,其AMI——销售广告信息库和MEDAB——中东和英国信息资料库亦颇具特色。该系统所有的数据库都是每天更新);⑥DOW JONES系统(由美国道·琼斯有限公司经营,其每天更新的商业/经济新闻库可提供全文检索,其美国与世界股票市场行情报告数据库每90秒钟更新一次);⑦OCLC系统(由美国国际电脑图书馆中心经营,该系统在世界各国拥有6千多个图书馆用户,7310个终端。除了提供数据库检索外,该系统还提供联机编目、馆际互借和各成员馆的馆藏信息。该系统年收入高达2700万美元以上);⑧DMS/DR1系统(美国国防部和数据资源公司经营,有15个国防、兵器工业、航空航天工业以及防务市场与军事标准等专业的数据库);⑨I.P.SHARP系统(由加拿大雅博深联合会经营,有80多个金融、经济、能源、航空、保险与法律等专业的数据库,文献量达5千5百万篇以上,用户终端遍及22个国家);⑩BLAISE系统(英国国家图书馆经营的自动化情报服务系统,其服务分为两个部分,一部分为BLAISE—LINK,它连接到美国国家医学图书馆的NLM系统,检索其生物学方面的数据库。另一部分为BLAISE—LINE,提供20多个非生物学数据库的检索服务。此外,该系统还可提供联机编目服务);⑪Pergamon InfoLine系统(英国

培格曼公司经营,它除了提供检索本公司生产的地质和力学二个文摘数据库外,还提供检索约30多种其他数据库生产者的产品);⑫DIANE系统(由欧洲共同体组织经营的欧洲联机情报检索网,参加该网的有40多个联机情报检索系统,通过欧洲EURONET数字信息传输分组交换网向广大用户提供约400多个数据库的检索服务);⑬ESA/IRS(Quest)系统(由位于意大利弗拉斯卡蒂的欧洲航天局情报检索中心经营,有120多个数据库,涉及各个科技领域,文献量已达5千万篇以上,用户终端在5千家以上,遍及世界50多个国家);⑭STN系统(由美国化学文摘社和西德卡尔斯鲁厄能源、物理、数学专业情报中心以及日本国际化学协会三家联合经营,约有50多个科技方面的数据库,该系统以化学、建筑、纺织与能源专业数据库最具特色);⑮DIMDI系统(西德医学文献情报所经营,约有40个生物医学及其有关的数据库);⑯Data-Star系统(瑞士无线电公司经营,约有50多个各种专业的数据库);⑰Telesystemes-Questel系统(法国远程通信公司经营,约有40多个科技、经济、商业方面的数据库);⑱PASCAL系统(法国科技文献中心经营);⑲JOIS-II系统(日本科技情报中心经营,该系统有30多个数据库,其中以“日本科技文献文摘”、“日本国内医学文献文摘”、“日本工业产业新产品、新技术情报”、“日本公共资料情报——政府研究报告”以及“日本科技研究情报”等数据库最具特色。此外,该系统还可提供一些非日文数据库的联机检索服务。它可允许用户以假名、日语汉字和英语字符进行联机检索。该系统的文献量已达3千多万篇);⑳PATOLIS系统(日本专利情报中心经营,提供各种日本专利信息的联机检索服务);㉑WSR系统(奥地利经济和社会科学计算中心经营,提供经济和社会科学方面的数据库检索服务);㉒Finsbury系统(英国Finsbury数据服务公司经营,提供英国、西德、法国和日本的主要报纸报导的新闻);㉓AUSINET系统(澳大利亚联邦科学与工业研究组织和国家科技情报中心经营);㉔BELINDIS系统(比利时经济数据处理中心经营,提供十几个经济、法律和核情报等专业数据库的检索服务)等等。就连埃及、印度这样一些发展中国家也建立了自己的大型联机情报检索网络系统。

不但国际联机检索系统日益增多、其功能越来越完善(可以检索图象与文献全文),而且联机数据库的规模和数量也增长迅速。现在世界上绝大多数文摘或索引刊物都有机读版本,其中80%以上可在世界范围内供联机检索。而且由于纸张与印刷费用的不断涨价,没有对应的硬拷贝版本(主要指印刷本)的电子数据库也逐渐增多。据欧洲科学情报与提供中心协会(EUSLIDIC)的“数据库指南”一书统计,1984年世界商业性的比较大的数据库数量已达2509个,估计现在已达3500个以上。小型数据库的增长则更为迅速,有人作了一个统计,说仅美国平均每一个工作日(五天工作制,全年200多个工作日)就要新增加2~3个数据库。此外,国际联机检索终端用户的数量也越来越多。例如,DIALOG系统现已有7万多家用户终端,分布在世界80多个国家和地区的400多个大、中城市。BRS系统也拥有2万5千多家终端用户。仅美国1979年就有300多万台联机检索终端,预计到1995年将增加到1千万台。法国1982年有4万5千台联机检索数据终端,1986年发展到24万7千台。国外甚至将国际联机检索终端带到了海上钻井平台上。美国不少大学教授、公司经理和高级技术专家家里都装有联机检索终端。

国际联机情报检索服务除了面向科技领域外,还面向经济、管理和社会科学,甚至深入到社会公用事业服务与家庭生活的各个领域之中。联机情报检索服务已发展成为一种新兴产业——联机检索工业,它作为当今信息产业的一个重要组成部分,不但产生了巨大的社会效益,而且产生了巨大的间接经济效益,其自身的经济效益也十分可观。据一份日本资料统

计,八十年代初,美国数据库服务的总营业额已达2,520亿日元,欧洲为1,510亿日元,日本为440亿日元。另据美国Nielsen公司调查,1984年全世界情报产业的收入达到136亿美元(不包括报纸、杂志、无线电广播部门以及计算机硬件和软件部门的收入),其中60%为联机检索工业的收入。由此可见,国际联机检索已成为世界科技和经济发展中必不可少的重要工具和手段。其重要性可以从检索量的惊人增长得到证明。据资料统计,1974年全世界国际联机检索量为70万次,1975年发展到100万次,1976年为120万次,1977年为200万次,1980年已达到700万次。仅英国1978年一年的检索量就超过10万次。在西方发达国家,使用国际联机检索最多的用户是工商企业,其次是学术团体和大专院校,公共图书馆相对少些,但近几年有所改变,公共图书馆和情报部门的使用量正在不断增加。

总之,随着每秒处理几亿条指令的超大规模集成电路巨型计算机的出现和高质低价的新型微型计算机的大量涌现与普及,高密量的激光盘存贮器的实用化及其大量投入市场,以及众多大容量通信卫星的发射与应用,必将加速国际联机情报检索服务沿着社会化、商品化、产业化以及公用事业化和家庭化方向深入发展。预计,到2000年,各种联机检索系统将通过通信卫星和数据传输网络联接起来,实现全球性的大联网,人类的全部知识几乎都将存贮在这种网络的数据库群中。那时,人们可以随时在家中利用个人计算机或电视机(利用Videotex技术)检索包括声音、图象、文字等各种形式的全世界的新闻和各种情报资料。

第二节 国际联机检索的服务方式

国际联机情报检索系统的主要功能,即主要服务方式有四种:

一、追溯检索 (Retrospective Search),简称RS。这种检索不仅能查找最新情报资料,而且可以追溯(回溯)查找数据库存贮年限范围内的某一段时间之内的或特定时间以前的有关文献。例如,查找1970年~1985年的有关文献,或查找1983年~今的有关文献,或查找1980年以前的有关文献等等。用户在检索式中既可以限定时间范围,也可以不限定时间范围。不限定时间范围的检索,即按用户提问式(检索式)的主题要求去查找数据库建库以来的全部新老文献,只要符合用户检索课题要求的文献都将被挑选出来。

追溯检索并不是用户事先将课题的提问式存入计算机系统中,而是临时根据科研生产和社会经济活动的需要提出的课题,对以往一段时间积累收藏的文献资料进行全面的检索。这种服务方式对用户要着手研究某一个新课题非常有用。对用户已进行了一段时间的研究课题,遇到了问题,要进一步查找有关文献也很有用。用户通过这种检索,可以得到对口文献的索引、文摘或者全文,便于科技人员了解自己所从事专业的国外动向和水平。

总之,由于追溯检索一次就能掌握相当长时期以来所积累起来的某个专业的全部资料,因此它特别适合于申请专利、撰写评论文章以及从事新课题研究等需要全面系统地掌握有关文献资料的用户要求。因为追溯检索可供用户查找文献的年限长,数据库的文献量必须相当大,所以这种检索只有具备有分时系统与高速度的大型计算机,且拥有大容量数据库存贮设备的大型联机检索系统才能办到,而一般小型计算机系统是很难满足这种要求的。

二、定题情报检索服务 (Selective Dissemination of Information)。直译为情报或信息有选择性地传播,简称SDI服务。它是指针对相对固定的用户课题提问要求,对新到磁带

文档进行检索并定期地向用户提供所需情报资料的服务。这种检索也叫现刊检索,它相当于手工检索时查找现刊上的文献。这种检索方式是在1958年由美国一位叫芦恩(H.P.Luhn)的工程师首先提出来的,之后在计算机检索中使用得越来越普遍。其具体做法是根据用户面交或寄来的检索提问单所提出的课题检索要求,选择出合适的检索主题词,按布尔算符或位置逻辑算符组配成检索提问式(逻辑式),按照用户要求定题服务的期限,第一次检索后,把提问式长久地保留在计算机内,以后每当数据库文献更新期(不同的数据库,其文献更新周期也不尽相同,一般有一天、半周、一周、半月、一月、一季或半年等几种更新周期)一到,计算机系统就会自动地按用户的需求进行检索并输出检索结果,定期向用户提供。

SDI服务对用户掌握某一学科最新水平和发展动向特别有用,而且情报提供及时,费用也较低。

国际联机情报检索提问单的例子见第四章第七节(第117页)。

三、联机订购原文。联机检索到的情报信息,大多数是原始文献的索引、题录或文摘,即所谓二次情报或二次文献(也有少数数据库提供一次文献,特别是经济类的文档)。用户如需要原始文献而在国内馆藏中又查不到时,可以通过检索终端向联机服务系统订购文献原文的复制件。同样,用手工检索到的文献线索,也可通过国际联机检索终端订购原文的复制件(分复印件和缩微平片两种形式)。国际联机订购原始文献提问单的例子见第四章第八节(第123页)。

四、传递电子邮件(Electronic Mail)。国际联机终端通常提供的是数据库检索服务,在“人一机对话”过程中,一边是能够灵活选择指令和编写检索策略式的用户,另一边则是按编制好的软件进行应答的机器。在联机检索过程中,用户往往会遇到各种各样的问题需要向系统询问,对这种随机提出的问题机器显然不能回答,只能由管理系统的人来回答;另一方面,用户之间也常常需要相互交流经验和互通信息。为了适应这种需要,世界许多大型联机检索服务系统在不断增加数据库和完善检索软件的同时,又设计了一种用高速数据终端传递的、作用相当于用户电报和加急快递信件电子邮件服务。例如,ESA—IRS系统设立主机值班员直接与终端用户对话;BRS系统设计的MSG系统可以在用户~系统、用户~用户之间快速传递电子邮件;DIALOG系统的DIALMAIL以菜单形式向用户提供丰富多彩的传送信息等方式。使用国际联机的高速数据终端传递电子邮件比用国际电传终端传递用户电报的费用低得多,前者单位字符的通报费一般仅为后者的几十分之一。

第三节 国际联机情报检索的优点

与手工检索和其他形式的计算机检索相比,国际联机情报检索具有以下优点:

一、检索速度快。国际联机检索系统的主机一般都是每秒钟能执行几千万条指令以上的大型计算机(例如DIALOG系统的两台主机NAS—AS9080,其运算速度为20MIPS,即每秒钟可以执行2千万条指令),因此,其检索速度之快是相当惊人的。过去用手工检索需耗时几周、甚至几个月的检索课题,使用国际联机检索只要几分钟便能完成。例如,检索DIALOG系统或STN系统中的美国“化学文摘”数据库(前者记录中不带文摘,后者带文摘),一般一个课题仅需5~15分钟(DIALOG系统的“化学文摘”数据库收录的文献从1967年至今已有800多万篇,是从近150个国家和地区的56种文字的14,000多种期刊、会议录、科技报告、图书、专利中摘录下来的),如果用手工检索《化学文摘》从1967年到现在的累积索引

或卷期索引,以及查出文摘记录并复印或抄录下来,至少得花费几天的时间。1983年,有人通过下面的实例对国际联机检索与手工检索的速度进行了比较:

例如,有一个名为“溶氧分析”的课题,用户选用ESA系统的CHEMABS(化学文摘)库。经过分析课题的题意,确定了检索词为DISSOLVED、OXYGEN、ANALYSIS,填写提问单,用户要求文献量不超过一百篇,最好检索词都在题目中出现。

联机检索操作人员按用户要求,开通终端机,接通国外的计算机系统和数据库,然后键入上述三个检索词,终端机显示屏上一步一步地显示,当键入dissolved时,显示屏显出:

1 7480 DISSOLVED

这表示第一组含有检索词dissolved的文献有7480篇。键入的词显示为小写字母,机器给出的回答以大写字母显示。

同样,键入oxygen和analysis后屏幕显示:

2 117304 OXYGEN

3 362391 ANALYSIS

考虑到用户需要的文献应同时包括上面三个检索词,于是再键入指令C 1 * 2 * 3(这些指令及逻辑运算符将在下面几章中介绍),得到:

4 241 C 1 * 2 * 3

这就是计算机告诉用户,三个词都包含的文献数目是241篇,显然,用户嫌它太多,要精选一下。于是,键入L 4/TI,CT,UT,也就是从241篇文献中,选出那些在题目中或受控词字段或自由词字段(字段以后再讲)中包含这三个检索词的文献,计算机回答:

5 142 4/TI,CT,UT

也就是说有142篇文献,用户感到还太多,只要其中那些题目中含有这三个检索词的文献就满足了,于是再键入L 5/TI指令,计算机接受指令后立即回答:

6 46 5/TI

这就仅剩46篇文献了,符合用户需求,立即键入t 6/1/1—46指令,即告诉计算机把上面4⁶篇文献对应的印刷本文摘号立即联机打印出来(只打印文摘号,是为了节省打印费),计算机回答:

T 6/1/1—46

并立即执行,在终端打印机上打出46个文摘号码。用户用它们就可对照美国《化学文摘》上的文摘来阅读。整个机检过程前后历时才2—3分钟。

如果用户以手工检索上述课题作为对照,也采用美国《化学文摘》(与CHEMABS是一样的内容,但它是印刷本),就需要从第66卷开始查,一直要查到第101卷(与联机检索的年限范围一样)。对比情况如下:

已知《化学文摘》从第66卷到95卷已有五年累积索引8th、9th、10th三套每套,查一本计3本,第96卷到100卷这五卷有卷索引,每卷查一本计5本。101卷卷索引尚没到货,仅能查期索引,每卷26期索引,要查26本,总计手工检索要查34本索引。从34本索引中查出的文摘号需要进一步核对,又需要翻阅上百本的文摘。也就是说,这个课题查下来,检索者就要从书架上找出上百本书搬到桌上,看完又要搬上书架,一个人往往要花费好几天时间,才能翻完这上百本书,找出与联机检索相同的文献。

由此可见,国际联机检索的速度之快、效率之高是手工检索根本无法比拟的。

二、检索范围广而全面。国际联机检索系统不仅主机的运算速度快，而且内、外存贮器的容量相当大。例如，DIALOG系统外存贮器有380多台磁鼓机，总容量高达25万兆字节以上。因此，它们一般都能容纳众多的数据库，可收录几千万甚至上亿篇文献记录。其专业范围几乎包括所有的学科与主题。世界上公开出版发行的文献中，90%都可以通过几种主要的国际联机检索系统查到。不任何地、何部门，只要与这些联机系统签订使用合同并拥有终端设备，就可以通过通信卫星和数传网检索这些相距几万公里之遥的联机数据库中的文献。特别是对于一些交叉学科的检索课题（其文献分散在各种专业的杂志中），用手工或其他形式的计算机情报检索系统检索，一次只能查一种专业的工具书或一个数据库，而由于国际联机检索系统具备有跨文档检索（Crossfile Search）功能或一次检索（One Search）功能，它可允许用户用一种检索策略（检索提问式）一次查多个不同的数据库，从而一次性地将分散在各个学科数据库中的有关文献（切题文献）检索出来。

三、检索途径多、检索质量高，随机调整检索策略方便。由于国际联机情报检索系统的检索途径（数据库记录中的可检索字段*）相当多，所以它既可以满足要求高检全率的课题用户，也可以满足要求高检准率的课题用户。例如，DIALOG系统一般数据库都可向用户提供10种以上的检索途径（可检字段），有的数据库甚至多达96种。最常见的可检字段有：篇名、文摘、叙词（规范标引词）、孤立的叙词（标引的规范化单元单）、自由标引词、孤立的自由标引词、机构名称、更新日期、著者、出版年、期刊名称、期刊通告、期刊缩称代码、文献类型、语种、出版者、专利号、分类号、报告号、存取号、国际标准书号（ISBN）以及国际标准连续出版物编号（ISSN）等。用户可根据需要选定文献记录的任何可检索字段作为检索途径，利用系统提供的逻辑组配和文中查找功能，进行“人机对话”式的检索，随机分析判断检索结果，根据需要扩大或缩小检索范围，随时调节检索深度，不断地调整修改检索策略，直到获得满意的检索结果。

四、检索内容新、实时性强。这是国际联机检索最突出的优点。众所周知，印刷版手工检索工具书的发行比相应的机读版文献数据库的发行至少滞后1个月以上的时间（国内滞后时间则至少在2个月以上），所以国际联机检索能查到手工检索查不到的最新文献。特别是检索那些时效期极短的商业经济、金融市场信息，国际联机检索更能显示出其独特的快速获取最新信息的优点。例如，DIALOG系统的Businesswire（610号文档）、McGraw-Hill News（600号文档）PR Newswire（613号文档）可以向用户提供15分钟内更新的信息，以使用户可以随时掌握最新的商业经济、科研技术进展、金融和社会动态情报。此外，该系统还有不少每天或每周更新一次的数据库，例如COMMERCE BUSINESS DAILY（195号文档，美国商务日报索引）、CENDATA（580号文档，世界经济和人口统计报告）、NEWSEARCH（211号文档，每日新闻索引）、STANDARD & POOR'S NEWS（132号文档，美国公司经营新闻）、UPI NEWS（261号文档，时事新闻评论）、WASHINGTON POST ELECTRONIC EDITION（146号文档，华盛顿邮报电子版）以及WASHINGTON PRESSTEXT（145号文档，华盛顿通讯社全文库）等可提供当天或前一天的商业贸易信息和各种社会新闻。而DOW JONES系统的“Stock Quote Reporter”（美国与世界股票市场行情报告）数据库甚至可向用户提供每90

• 字段(field)——计算机运算中，为存取某一特定信息而对计算机记录的某一特定部分给出的标识。例如在文献数据库中，字段可以是每篇文献记录的特定数据项（如篇名、著者、文摘、出版日期……）。如果把机读数据库中的一篇文献记录比作文献著录中的一条款目，那么一个字段就相当于这条款目中的一个著录项。文献的内容特征和外表特征是通过不同的字段来表达的。

秒钟内更新的世界金融市场情报。国际联机检索系统这种快速提供各种最新信息检索的功能是手工检索或脱机检索系统根本无法办到的。此外,国际联机检索系统一般都是昼夜服务的。例如,DIALOG系统除了每周的星期六美国东部时间12时至星期日3时停机外,其它时间都对外提供联机服务。当用户遇到某些紧急的检索课题(如商业谈判、医疗急救或者某些政治与外交方面的紧急课题等等)时,随时都可利用终端向国际联机检索系统提出紧急咨询检索,及时获得答案。例如,美国一家医院的值班医师获悉:救护车正载着3个遭雷击的男孩向医院疾驰,其中一个已处于极度危险状态。这位医师从未治疗过这类损伤,于是用医院终端向NLM系统紧急咨询检索,当救护车到达医院时,这位医师已知道了该如何处理,3个孩子全部得救。又如,当印度博帕尔毒气泄漏事故发生后,成千上万的人受到毒害,救护小组的专家们利用NLM系统紧急检索咨询,在NLM系统的帮助下,专家们在数小时内详尽地收集了全世界范围对付这种毒气的一切已知资料,及时飞赴事故现场指导抢救。

五、检索辅助功能完善,使用方便。国际联机检索系统一般都具有成熟完善的辅助或引导用户检索的功能,以人机对话的方式辅助用户,使那些不太熟悉系统的用户能够正确地进行检索操作,从而获得较为满意的检索结果。例如,国际联机检索系统具有帮助用户选择数据库、选择检索用词(通过显示联机词表或机内可检词词典)、修改检索策略、保留检索策略(保留检索提问式,以便查另外数据库时调用)、回顾检索策略(重新显示前面进行过的检索步骤和查找结果)等功能。对于新用户,多数系统有一个“Explain”或类似的指令,在该指令后输入有关指令、数据库或系统的名称,系统则显示有关的信息,以便解答用户对于系统的种种疑问。当用户在检索过程中遇到困难,需要系统帮助时,可输入“Help”或类似指令,系统即列出系统信息的选择单,用户可跟随选择单一步步地获得其所需的信息。

六、检索结果输出方式灵活、实用。国际联机检索的结果可以当时联机显示阅读与联机打印出来,也可采用脱机打印方式(比较省钱,脱机后让系统集中打印然后寄给用户。一般,欧洲系统打印件一星期左右、美国系统打印件10天左右用户便可收到),还可利用微机终端和套录(或称转录)技术*联机快速地将检索结果存贮在微机终端的磁盘上,然后脱机后再按用户的需要进行编辑整理后打印出来,或者加入用户自建的专题数据库中。

国际联机检索系统的记录输出打印格式有多种供用户选择。除了系统提供的各种固定的记录输出格式外,用户还可根据需要自定输出打印格式,既可将完整的文献记录打印出来,也可只打印记录中的某个或某几个字段(如存取号、著者、篇名、文摘、文献出处、标引词……)。此外,大多数系统还具有排序输出打印的功能。例如,ORBIT系统不仅允许用户根据需要对输出文献记录的内容及篇数进行限定,而且允许用户选择命中文献记录中的某几个字段,并按其内容的字顺排列打印输出。利用这种功能,用户可以很方便地建立各种索引(如主题词索引、著者索引、篇名索引、公司名索引、产品索引、化学物质登记号索引……)和小型专题数据库,或编制各种信息资料通报、目录与手工检索刊物。

*套录(downloading或offloading)是指在联机情报检索过程中,利用微机终端从数据库中套取某一方面或某一主题的情报(文献记录),将其存贮在磁盘上,然后在脱机后进行倒置、合并、删除、格式转换等重新处理,以建立本机上的专业或专题子库,供今后重复多次检索使用的一种新型的情报检索与贮存技术。有关套录技术的问题请参阅笔者撰写的《套录技术及其对联机情报检索工业的影响》一文(载《情报学刊》1987年第6期)。

第四节 我国的国际联机情报检索工作概况

一、我国国际联机检索终端的设置

1980年3月,在中国建筑技术开发中心文献部的倡议下,九个专业部委情报所联合在香港租用了一台数据终端机,利用香港大东电报局的国际通信线路与美国DIALOG和ORBIT系统联机,与日本相比,只略晚几个月。该终端1980—1981年共检索了526个课题,对国内科技工作的发展起了很大作用。以后,有的专业部情报所不满足于这种由香港转手的检索服务,如石油部情报所和北方科技情报所(现为国家机械委210研究所)分别于1981年3月和12月先后在北京用电传终端和当地电信局的国际电报线路与美国DIALOG系统、ORBIT系统联机,这样既解决了由香港转手的麻烦,又可进行人机对话式检索,比以前方便多了。但由于他们采用的是电报线路,电传费用昂贵,从而影响了用户的积极性。

1983年第一季度,在联合国教科文组织协助下,我国邮电部与意大利邮电部签订协议租用其ESANET网络,并合作改造了我国与意大利之间的24路载波电报所占用的两条9600波特卫星专线国际话路,将其由原来的频分调制(FDM)改为时分调制(TDM)的方式,因而挤出了一条话路专用于两国间的数据通信。于是中国科技情报研究所开始与欧洲宇航局情报检索服务部(ESA-IRS)联机,使用ESA-IRS系统。

1983年9月15日中国科技情报所联机检索服务部成立,10月2日以后正式对社会服务。由于采用数据通信传输,通信费迅速下降,比电传终端的电报通信费减少近9/10,深受用户欢迎。同年年底,中国科技情报所国际联机检索服务部通过ESANET网与美国Telenet网挂钩,开始检索DIALOG系统和ORBIT系统。

1985年,辽宁省科技情报所、中国科技情报所重庆分所已可通过北京中国科技情报所国际联机检索部终端分机时,在当地使用国际联机终端检索课题了。

与此同时,北方科技情报所、水利电力部科技情报所都已直接利用邮电部网络及线路,与意大利ESA和美国DIALOG和ORBIT系统联机。北方科技情报所并通过内部电话线路,给南京、石家庄、包头、长沙、长春、重庆等地本系统的单位分机时;水电部情报所也通过内部电话线路,给郑州、武汉等地本系统的单位分机时,从而使这八个城市也可直接对国外ESA、DIALOG和ORBIT系统联机检索。

到1986年10月底,全国已有北京、上海、广州、南京、济南、嘉兴、杭州、宁波、福州、厦门、深圳、郑州、武汉、南昌、合肥、长沙、重庆、西安、太原、成都、绵阳、沈阳、长春、哈尔滨、大连、保定、烟台、无锡、天津、丹东、石家庄、呼和浩特、包头等33个城市设置了国际联机终端60多台(见表1·1)。

二、国际联机情报检索服务在我国四化建设中发挥的重要作用

我国自1980年3月开展国际联机情报检索服务以来,至1988年10月底,据不完全统计,各终端服务部检索课题总量累计已超过3万个,为我国经济建设、科学研究、技术进步、管理与经营决策以及文化教育等工作提供了大量的信息,取得了较大的经济效益和社会效益。其主要表现在以下几个方面:

1. 向研究所和大专院校科研人员提供了与其研究课题有关的国外研究动向、成果和发展水平的文献资料,加速了科研项目的开展。