

国际联机检索概论

李 平 源

重庆大学图书馆

前言

教育部(84)教高字004号文件“关于在高等学校开设《文献检索与利用》课的意见”中指出:“让学生具有掌握知识情报的意识,具有获取与利用文献的技能,是培养学生能力的一个重要环节”。根据教育部的这一指示精神,近年来我国不少高等院校已在本科生和研究生中相继开设了“科技文献检索与利用”课程,并编写了相应的教材。遗憾的是这些教材内容均以手工检索工具和手工检索方法为主,缺乏通过现代通信手段传播的国际联机情报检索系统的知识及其检索方法的内容。本教材就是为弥补目前科技文献检索教材的这一不足而编写的。

电子计算机情报检索系统具有存贮信息全、检索速度快、传输距离远、服务功能多等特点。利用电子计算机特别是国际联机情报检索系统进行检索,在检索速度、查全率等方面都是传统的手工检索无法比拟的。因此,由手工情报检索转向利用电子计算机进行情报检索的意义十分重大,也是科技情报检索发展的必然趋势。

当代情报工作发展的重要特点,是情报资源共享。情报资源采用电子计算机存贮并通过通信卫星等现代化通信手段加以传播是情报工作现代化的主要方面。经济发达国家从六十年代开始先后建立了计算机化的情报检索系统。到七十年代已发展为联机检索的普及使用时期,现在全世界已有一百多个较大的联机检索系统。

当前我国情报机检工作尚处于初始阶段。一九八三年十月二十五日中国科技情报研究所国际联机情报检索服务部正式开业。该服务部通过国际通信卫星网络,与欧洲最大的联机检索系统 ESA—IRS (有六十多个文档,存贮六千万篇文献)及美国最大的联机检索系统 DIALOG (有二百三十九个文档,八千万篇文献)相连,随时向全国情报用户提供上述两大联机检索系统的国际联机检索服务。不久还将与美国 ORAIT 系统(有八十个文档,六千万篇文献)连接。

一九八五年四月十八日,中国科技情报所重庆分所国际联机情报检索终端正式开机,从而向重庆市及西南各省提供了国际联机检索的最方便条件。

由上可知我国联机检索起步较晚,利用国外联机检索系统的国际联机检索服务也刚刚开始。广大科技人员,包括高等院校的教师和学生由于缺乏这方面的必要知识,因而使我国已建立的现代化的国际联机检索服务未能得到充分应用。

本教材系统的介绍了国际联机情报检索系统及其检索功能,并重点地介绍了联机情报检索过程和检索策略。本教材除了作为我校科技文献课程补充教材之外,亦可供理工科高等院校专业教师及广大科技工作者进行国际联机检索时参考。

目 录

前 言

第一章 国际联机情报检索的一般问题	(1)
第一节 计算机情报检索系统及服务功能	(1)
第二节 情报检索与计算机网络	(2)
第三节 国际联机情报检索系统	(7)
第四节 我国使用的国际联机情报检索系统	(10)
第二章 数据库及其检索途径	(14)
第一节 数据库	(14)
第二节 文档的构成及检索途径	(20)
第三章 联机检索系统的基本检索功能	(23)
第一节 布尔逻辑算符	(23)
第二节 位置逻辑算符	(25)
第三节 字段限制	(27)
第四节 截词算符及其它功能	(28)
第四章 联机检索的方法步骤及基本指令	(30)
第一节 联机检索的方法和步骤	(30)
第二节 检索结果处理	(31)
第三节 基本指令	(34)
第五章 联机检索的辅助检索功能	(38)
第一节 查询功能	(38)
第二节 保留检索和定题服务(SDI)	(45)
第三节 联机订购原始文献	(47)
第六章 检索策略	(49)
第一节 构造检索策略的几个主要环节	(49)
第二节 调节检索策略的若干方法	(52)
第七章 国际联机检索实例	(58)
第一节 一般文献检索实例	(58)

第二节 专利文献检索.....	(67)
参考文献.....	(72)
附录一、 <i>DIALOG</i> 系统操作指令.....	(73)
附录二、“ <i>DIALOG</i> 系统总索引”的文档主题分类.....	(76)
附录三、 <i>ESA</i> 系统数据库简介.....	(84)
附录四、 <i>DIALOG</i> 系统数据库简介.....	(100)

第一章 国际联机情报检索的一般问题

第一节 计算机情报检索系统及服务功能

计算机情报检索系统，就是利用电子计算机，根据某种目的，在一定时间内，从经过整理并已贮存在计算机内的情报中，检索所需的情报的系统。简单地讲，情报检索系统就是情报存贮和检索的技术。科技文献的检索就是包括把文献加工为资料档的存贮技术和根据人们需要从资料档中找出所需文献的检索技术两个部分的内容。犹如查电话号码，首先要有本电话号码簿，然后按照号码簿的索引目录来寻找所需要的电话号码。电话号码簿就是情报的存贮（它按一定方式方法印刷在书本上），检索就是翻阅分类索引，然后翻到指定的一页，逐一查找的方法。这是众所周知的人工查电话号码簿的方法。计算机情报检索关键在于把情报合理地编成文献资料档（File），再有效地从中检索出所需要的情报。上面的例子中，电话号码簿，就可认为是一种“文献资料档”，存贮电话号码及户名信息的介质是纸张，记录在上面的是文字；而计算机的存贮介质是磁带、磁盘等，记录在上面是一些由二进制数“0”或“1”所表示的信息。

图1—1是计算机情报检索系统的简略示意图。图中情报收集和情报加工部分是由人工来完成的。通过人工加工完成的文摘、索引等文献资料的输入是以数据的形式出现的。数据

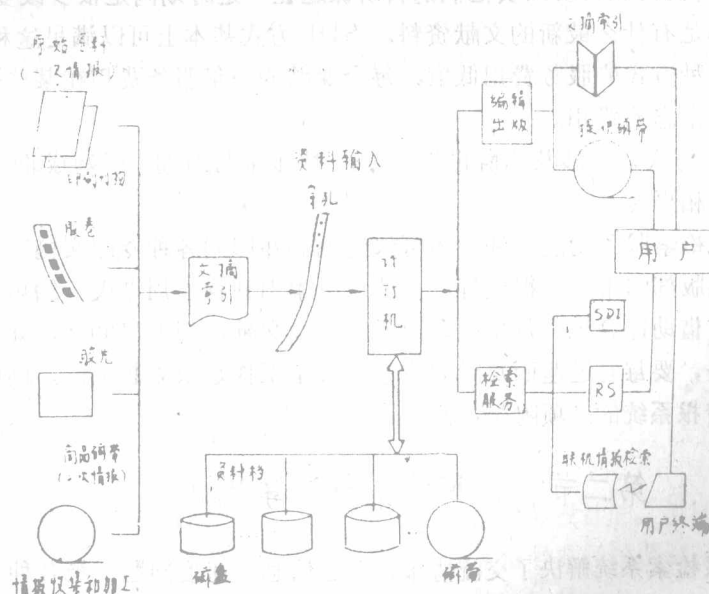


图1—1 计算机情报检索系统示意图

要能在计算机与人之间通信,就必须把印刷的信息形式转变为机器语言。并通过卡片机或其它输入设备输入到计算机中贮存。由于计算机的内存贮器的容量有限、大批资料档(文摘、索引)都要放在大容量的辅助存贮器,即外存贮器中。目前计算机常用的外存贮器为磁盘、和磁带。

计算机情报检索系统对用户提供服务,一般有以下几种方式:

1. 定题情报检索,就是常说的 SDI (Selective Dissemination of Information), 也称现刊检索。相当于手检时查找现刊上的文献,具体做法是根据用户提出的要求及说明分析研究课题的内容,编写成逻辑提问式,在情报检索系统中建立户头(用户一般是固定的,提问式一般也是固定的),系统定期(每周或每月)对新到的资料按用户要求进行检索,并将检索结果及时分发给用户。

2. 追溯检索 RS (Retrospective Search), 也称过刊检索,根据用户的随时要求(作成提问式),对过去累积的资料档进行专题追溯检索,并针对这一用户给出检索结果。资料档一般累积五年至十年的资料。

以上两种服务方式可采用比较简单的脱机方式,计算机成批处理,即用户不直接使用计算机,也不能立即获得检索结果。用户将提问单交给专职情报人员,由系统的工作人员把一定数量的用户提问单集中起来,按系统的规定格式输入到计算机进行处理,通过计算机查找,然后把检索结果一起打印出来,再由系统的工作人员整理后将结果分发给各用户。

3. 联机情报检索,即用户使用计算机终端设备直接和计算机对话,通过会话问答方式提问和获得结果。在“联机”的情况,要求计算机尽快给出答复,所以只能将文献资料档存贮在磁盘上。

定题情报提供(SDI)方式,虽然比较简单,但很实用,直到现在还是主要的服务方式。因为大多数科研机构和科研人员他们的科研课题在一定时期内是很少改变的,因此大多数科研人员关心的是有什么最新的文献资料,SDI方式基本上可以满足这种要求。还有很重要的一点就是这种方式的服务费用低廉,每个课题的一年服务费用在某些国家只有几个美元,相当于买一本杂志的费用。

追溯检索(RS)方式,所涉及的情报量很大,情报量与追溯的时间成正比。追溯时间要考虑与系统存贮量相平衡。

计算机除了上述检索服务功能之外,还可快速编制并印刷各种专门文摘、索引,它给读者和情报索引刊物出版部门带来了很大好处。使出版索引刊物的周期大大缩短,成本下降。象美国《化学文摘》借助计算机,每年编印出版50万篇文献,用人1100名,如果不是利用计算机这样先进的设备,要每月迅速而及时完成这样大量工作是很难想象的。因此,可以说编制索引也是计算机情报系统的一项附带产品。

第二节 情报检索与计算机网络

电子计算机情报检索系统解决了交流情报,传递信息的速度问题。把各种各样的情报信息 and 数据存储在电子计算机系统中,确实大大加快了检索速度,查找的时间大大缩短了。但也只仅仅解决了检索速度问题。用通信线路将各自具备独立功能的计算机联接起来组成计算机

网络,情报用户在检索系统网络的终端上就能对网络上各计算机存贮的不同文献资料档进行检索。这就使计算机情报检索系统的功能大大提高。

一、计算机网络的功能

计算机网络就是以能够相互共享资源(硬件、软件和数据)的方式连接起来,并且各自具备独立功能的计算机系统之集合。

计算机网络系统包括数据通信网,终端设备及计算机在内的总的系统。计算机网络的建立是为了提高计算机的处理能力,提高可靠性、扩展性、分配负载,但共享资源是主要的目标。

计算机网络的功能是:

1. 将地理位置分散的资源实现实时的集中管理。
2. 共享数据库资源。可以避免同样的文献资料在多处设置的浪费现象。这点对于科技文献情报检索特别重要。
3. 可访问多个系统。
4. 可以共享硬件资源。可以使用本系统没有的设备,由多个用户共享硬件以后,可提高利用率。当某个计算机系统负载已满时,新的作业可通过网络转送其他系统处理。
5. 共享软件资源。目的是避免软件研制上的重复劳动。在使用别的系统软件时,只要送去必要的参数,即可获得需要的结果。在计算机网络系统中积累软件资源,即意味着积累社会的知识。
6. 提高可靠性。当某个计算机系统因故障停止工作时,可以考虑由别的系统代为接收,继续运行下去。
7. 由于网络的特点,可以产生大量各种各样的服务系统。如情报检索,医疗、教育等领域中的服务。
8. 易于扩展。在扩展现有的小规模系统时,只要接入网络,本身作为远程成批终端,便可以达到目的。

二、计算机网络的结构

1. 计算机网络的组成

计算机网络的基本组成部分如图 1—2 所示。

由图 1—2 可见,计算机网络由四个基本部分组成:

- 1) 计算机;
- 2) 终端设备;
- 3) 数据通信网(包括调制解调器和通信线路);
- 4) 通信控制器和终端控制器。

目前大量的传输信道(通信线路)都是按传输模拟信号设计的,但计算机所需传送的大量数据都是数字信号,因而在传输过程中发送端的数字信号要经过调制器变成适合于信道传输的调频波,然后在接收端经过解调器还原为原来的数字信号。

通信控制器是计算机与远距离终端或网络中其他计算机进行通信的控制设备。通信控制

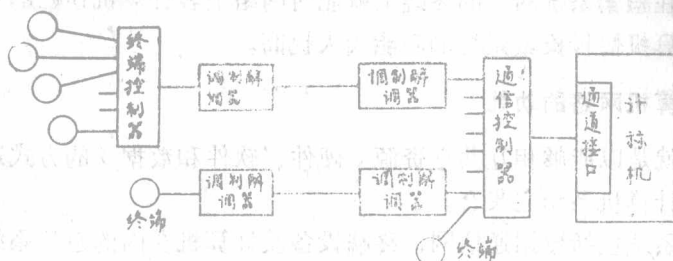


图 1—2 计算机网络的基本组成框图

器是联机系统的一个组成部分。

终端控制器是控制多个远距离终端用的，通过终端控制器把多个远距离终端集中到用一条通信线与计算机通信，以降低通信线路的费用。

2. 计算机网络的类型

构成计算机网络的基本成份为计算机，终端设备，数据通信网和网络处理器。此外还把通信控制器，终端控制器以及相应的软件功能等称为“网络处理器”（NPU—Network Processing unit）。上述四种结构成份组合起来通常可构成以下三种网络结构：

1) 星式（或称集中式）结构

如图 1—3 所示，这是把运算处理能力集中到中央主计算机的方式。图中 T：终端设备；NPU：网络处理器；HOST：主计算机。

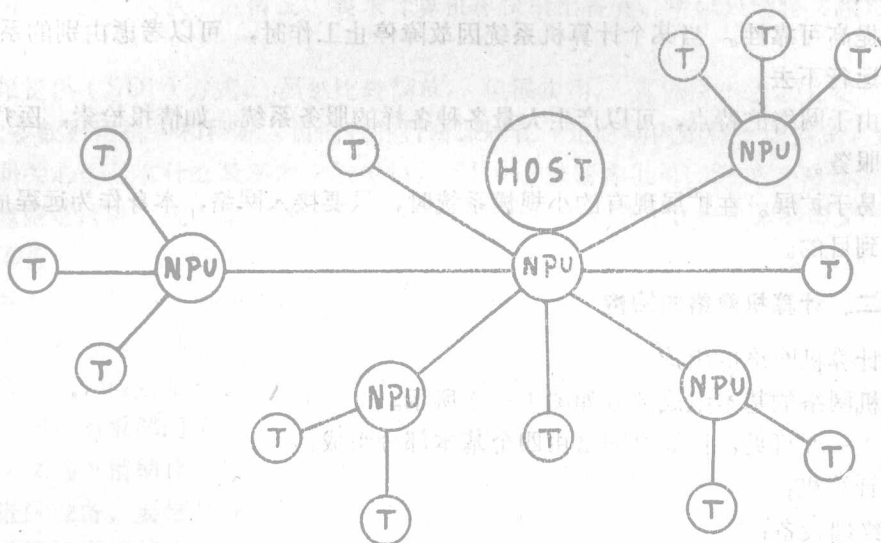


图 1—3 星式结构的计算机网络

星式结构的优点是不仅通信线路的费用比较便宜，而且主要设施，资源和人员等也集中起来共同使用。所以比较经济。

缺点是故障的影响容易波及整个网络，当中央的主计算机停机时，整个网络都无法使用，即使是通信线路发生故障，也无替代的通信途径可用。

2) 环式结构

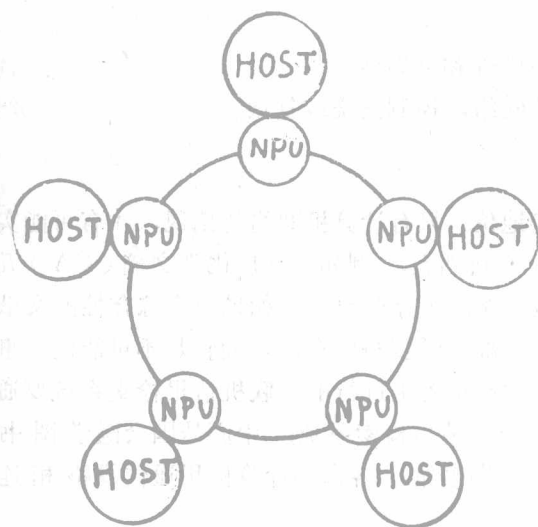


图 1—4 环式结构的计算机网络

这是星式结构与环式结构的组合，如图 1—5 所示。这种结构一方面可相互补充两者的

如图 1—4 所示，采用附加地址的方法把数据从环内的任一计算机送往其他任意的计算机。环上的任一 NPU 都具有将全部数据送到最终目的地的能力。

环式结构的优点是即使有一部分通信线路发生故障，还能利用另外的替代通信途径。假如有一台主计算机不能使用，与其他主计算机之间的通信仍可照常进行。这种结构就网络整体而言，具有很高的可靠性。

缺点是 NPU 等的软件较复杂，同时因数据中继功能而引起的辅助操作也要比星式结构多。另外由于增加了替代通信途径等冗余性，通信费用比较高。

3) 分布式结构

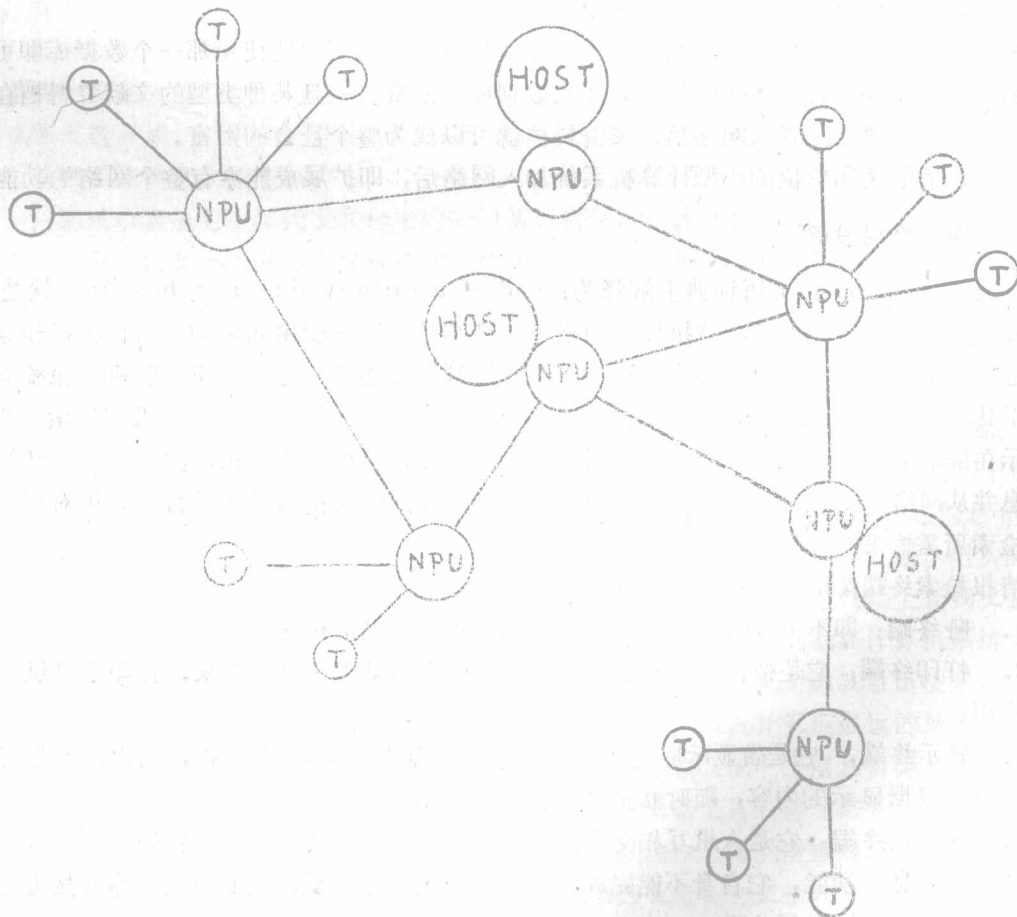


图 1—5 分布式结构的计算机网络

优缺点,另一方面可以兼顾经济性和可靠性。

分布式结构的优点是在各个方面有很高的灵活性和可靠性。

缺点是要维持星式和环式这两种不同性质的网络,控制系统较复杂。

三、情报检索与计算机网络

计算机网络在情报检索系统中占极其重要的地位。只有计算机网络的出现,才有可能实现情报检索的联机系统。情报检索的一个特点是资料档大,例如要存放化学文摘(CA)五年的文献资料,就要有几十台 100MB 的磁盘组才能满足存贮要求。如果一个综合性的文献情报检索系统,要在一个地方一台计算机上解决全部文献资料的存贮,简直是不可能的。事实上,文献情报的收集和检索都是在各个不同的地理位置上进行的,联机情报检索系统要通过计算机网络才能实现。例如美国国会医学图书馆的情报检索系统,中心是国会医学图书馆,网络的下一级是各大区以及州的图书馆或大学图书馆,各点的计算机用通信线路相连接,构成联机检索系统。

情报检索通过计算机网络可以实现:

1. 读者在终端上直接进行检索,也可以直接从终端获得结果。系统又能集中进行控制管理。系统必须提供给用户使用的检索语言。

- 2) 网络内的数据库都可以共享。用户在使用时,只需要说明使用那一个数据库即可,因此系统不必重复投资,造成软件、硬件及数据库的浪费。一旦某种类型的文献资料档在某个计算机系统上建立并加入网络后,该资料档就可以成为整个社会的财富。

3. 各个地区和单位的小型计算机系统加入网络后,即扩展成能享有整个网络的功能。

四、终端设备

终端即 terminal,英语词典上解释为: end of a railway line or a bus line 就是铁路或公共汽车线路的终点站,情报检索网络上的终端是从这里引喻而来的。所谓情报检索终端,是一台小型设备。终端设备通过通信线路与计算机相连接,是用户使用联机情报检索系统的工具。用户通过终端上的键盘,把自己的检索提问输入计算机,而计算机把检索结果立即显示在屏幕上或由打印机打印出来供用户使用。因此,也可以说情报检索终端是向网络发送信息并从网络接受信息的一种设备。用户通过它可以同情报检索系统进行“人机对话”以达到检索所需情报的目的。

情报检索终端按用途不同,分为以下几种:

1. 哑终端:即不带打印设备的终端,供一般了解商品价格之用。

2. 打印终端:它是带打印设备的,能把用户检索结果迅速打印下来,打印件可供用户长期使用。

3. 显示终端,它是带显示屏的终端,它能及时显示检索结果,指示用户输入是否有误,用户可根据显示的内容,随时修正自己的检索策略。

4. 交互式终端:它是人机互相交换信息用的一种终端,其工作方式较简单,只有输入信息及输出信息的功能,它自身不能完成任何信息处理动作。受信或发信只能靠人按键或中央主计算机发来的信息所启动。

5. 成批处理终端：它含有小型处理机，输入输出设备。它可以在远距主计算机的地方，把不同用户分散的检索要求收集整理成一批，在适当时候送往主计算机处理。再把从主计算机得到的处理结果进行整理分解，以用户所需的形式输出。

6. 智能终端：由于智能终端自身带有处理机，因而具有信息加工能力。它有如下的功能：删除有明显错误的数据库，汇集数据并保存在通信线路上，向中央计算机发出中断请求，对同一通讯上的几个输入输出数据流实行多路复用与鉴别等。

7. 7×8 吋的袖珍式终端：它是一种小型终端，供旅游和野外作业使用。

此外，按作用距离则有近程终端和远程终端之分，按速度快慢则有高速终端和低速终端之分。

第三节 国际联机情报检索系统

一、国际联机检索的由来

国际联机情报检索系统是指情报用户在检索系统网络的终端上，使用一些特定指令和检索词，按布尔算符进行逻辑组配，从计算机检索系统中查找用户自己研究课题所需要的文献资料。我国科技人员和情报人员坐在终端旁，甚至坐在自己有终端的办公室内，即可联机浏览从系统中检索出来的文献和数据，自行决定取舍。总之，所谓国际联机检索就是利用电子计算机和人造卫星，远距离（超越一国）进行人机对话，随机判断，查找美国或其他国家存储在计算机系统内的文献资料。

国际联机检索是七十年代发展起来的一门新兴科学，具体地说是1963—1964年从美国搞起来的。先是在美国国内建立联机检索系统网络，然后逐渐形成国际联机检索系统，一般说只有20年左右的历史。

国际联机检索的出现，首先是具备了物质前提，这就是：电子计算机诞生（四十年代）和人造卫星上天（五十年代）。

这两项科学技术发展成果，使得进行科技情报工作的方法、手段发生巨大变化，使人们交流情报，传递信息的方法，手段有可能现代化了。人造卫星的出现，并把它应用于情报检索工作中，更是一件有划时代意义的事情，这表明人类已进入了利用空间传递信息的时代。原来电子计算机的诞生，只帮助人们解决了交流情报，传递信息的速度问题。但实际检索起来，还是受地理位置远近的限制，检索的人只能在离计算机不很远的终端上查找文献，例如终端只能设在离计算机系统几十米，几百米、几公里远的地方。终端离计算机系统太远了，检索起来就会发生困难。人造卫星的出现就解决了这个问题。利用微波通讯技术，经过调制解调器，通讯线路由直流电改为交流电，这样用户的终端设在计算机很远的地方也可以检索了。有了卫星全世界数以万计的终端就可以分享国际检索系统中的信息资源了。

国际联机检索发展的另一方面原因，是对这种现代化检索工具的客观需要。当今科技文献数量激增，类型复杂，重复交叉。有人把我们所处的时代称之为“情报爆炸”或“知识爆炸”的时代。科技人员面对浩如烟海的文献，要找到自己所需要的文献，犹如大海里捞针。国际联机检索就是为了适应这种客观需要而产生和发展的。

早在电子计算机产生的初期,就开始了情报检索自动化的实验研究工作。在国外,从五十年代起就有人进行了利用计算机编制文摘、索引和自动分类的尝试。1954年到1964年计算机用于存贮和检索取得了初步成效,这叫做脱机检索时期。1965年到1972年是所谓的联机检索时期,出现了若干个联机检索系统。从1973年到现在是联机检索普及时期,世界上建成了一些著名的国际联机情报检索系统。

目前世界上较大的联机检索系统有100多个,数据库有1100多种,存贮的机读文摘和索引超过1亿条,计算机网络和联机终端已遍及世界主要国家和地区。最大的电子计算机国际联机情报检索系统是美国DIALOG国际联机检索系统,其次是美国的ORBI系统。ESA系统是欧洲最大的联机检索系统,也是世界第三大的联机检索系统。下面将分别简单介绍,这三个系统。

二、世界著名国际联机检索系统简介

1. Dialog 系统

它是目前美国和世界上最大的联机检索服务系统。Dialog是一个英文词,就是“对话”的意思,因研究人机对话的检索系统软件而得名。该系统建立于1963~1964年间,原为美国洛克希德导弹与空间公司(Lockheed Missiles & Space Company)下属的情报科学实验室。该系统开始建立时,其规模比较小。

Dialog检索系统总部设在美国加利福尼亚州帕洛阿尔托(Palo Alto)市,1981年6月Dialog系统正式宣布为Lockheed公司的子公司,独立开展一切业务活动。

1966年Dialog系统与美国国家宇航局(NASA)的研究实验室图书馆合作,通过一条电话线路,使用一台终端,查找宇航局图书馆的数据库中的文献资料。当时数据库存储的文献比较少,只20多万篇。

进入七十年代以后,特别是七十年代末期,随着电子工业的迅速发展,检索系统硬件和软件迅速发展,该检索系统也随着发展起来。到1972年开始成为国际联机检索系统,为全世界用户服务。

Dialog系统的数据库(或称文档)逐年增多,现列示如下:

1972(年)	2(文档)
1973~1974	14
1975~1976	39
1977~1978	72
1979~1980	101
1981	130
1982.1	178
1982.7	202
1983.1	216

随着时间的推移,Dialog系统的数据库的数量还会逐年增加。

目前Dialog检索系统拥有两台最新、最大型的电子计算机分别与卫星通讯网联系,其中:

IBM 与 Tymnet 卫星通讯网络联系;

NAS—9000 与 Telenet 卫星通讯网络联。

总的运算能力为14MIPS (million of instructions per second), 就是说每秒钟可处理1400万条指令。外存设备有150台 IBM—3033 1型磁盘机, 存储总容量7万兆字节, 还有14台 IBM—2321磁卡机, 总容量为5600兆字节。有3台快速打印机为用户打印输出的文献, 每天分三次成批打印。

Dialog 系统存储的文献在8000万篇以上, 占世界机存文献总量的50%以上。包括40多种语种, 6万种期刊, 占目前世界发行期刊总量10万种的60%以上。

该系统联机检索的终端数量也逐年增加:

1967年	23个终端
1981年	15000个终端

现在 Dialog 系统联机检索的终端, 遍布于全世界70多个国家和地区的200多个城市, 亚洲、欧洲、拉丁美洲各国大城市如伦敦、巴黎、波恩、东京、香港等都有其终端用户, 现在在美国和其他国家的许多工商企业、科研机构、情报中心、大型图书馆、高等学校, 甚至学者和教授的办公室或家里都设有终端, 一旦需要什么情报资料, 与打电话一样, 随时可开机查找。Dialog系统每天22小时为全世界用户服务。

2. ORBIT 系统

它是目前美国和世界上第二个最大的联机检索系统, 属于系统发展公司 (System Development Company), 简称 SDC。ORBIT 是 Online Retrieval of Bibliographic Information Time-shared 的缩写。意思是文献信息分时的联机检索。即是说以该系统的功能和作用来命名该系统的。该系统全称的5个词的第一个字母拼成 ORBIT 而成为一个有意思的英文词, 这是一种巧合, 我们有的人称它为“轨道”系统, 并不是该系统名称的本意。

ORBIT 系统的总部设在美国加利福尼亚州圣莫尼卡 (Santa Monica), 其发展速度也是相当快的。

1960年开始搞联机检索软件, 并获得了成功。

1965年在美国国内开展联机检索。

1974年也开始成为国际联机检索系统, 通过卫星网络, 为美国和加拿大以及欧洲、亚洲的用户服务, 现在伦敦、巴黎、多伦多、东京和香港等大城市, 也有许多终端与 ORBIT 联机检索。联机时间过去每天为12小时, 现在也改为每天22小时。

ORBIT 系统拥有数据库 (即文档) 也逐年增多, 最近三年的情况如下:

1980年	52个文档 (files)
1981年	65个文档
1982年3月	80个文档

ORBIT 系统每月更新文献可达20万篇, 数据库内存储文献量为5500万篇, 占世界机存文献总量的25%以上。

3. ESA—IRS

ESA—IRS是欧洲空间组织情报检索中心的简称,位于意大利首都罗马附近的弗拉斯卡蒂(Frascati)于1966年为适应欧洲尖端工业的发展而建立起来的。自1969年开始用NASA文档开展服务以来,学科范围不断扩大,文档数量增加很快,设备几经更新,发展比较迅速。ESA-IRS是欧洲最大的联机检索系统,也是世界上最大的联机检索系统之一。

ESA-IRS目前有60多个文档可供使用,其中有大型美国化学文档CHEMABS储存文献600万篇,综合性法国PASCAL文档储存文献450万篇,还有专业性很强的ALUMINIUM文档,原材料价格等的数值数据库PRICEDATA。该系统储存的文献总量达3000多万篇。在世界各国拥有3500个终端用户。

第四节 我国使用的国际联机情报检索系统

一、检索系统的构成

1980年4月国务院所属的9个部委(国家建工总局、铁道、交通、一机、石油、化工、地质、煤炭、冶金部)联合写了开展联机检索的申请报告,经国家计委、科委和财政部批准,通过香港大东电报局,与美国DIATOG系统和ORBIT系统联机,为国内用户查找各种研究课题所需的文献。

到1983年已发展到15个单位,北京大学、建材部、电力部、南海石油勘探指挥部、上海市和北京市情报所等都参加。参加的单位每年交付2万元港币,年终结算,多退少补。80—81共检索了526个课题。82年共检索1245个课题。

由于广大用户在国内,而检索终端却设在香港,存在两个严重的缺点:

1) 联机检索周期太长。在终端上查找的速度是很快的,一般查找一个课题只要几分钟或10多分钟就够了,但实际通过香港检索一个课题,北京和香港之间邮件往返时间大多在一个月以上,有的课题由于这样那样的原因,检索周期竟长达2~3月之久。如果用户有紧急需要,要求快速查找,虽然检索系统本身效率很高,但终端不在身边,远水救不了近火,无法满足用户的需要。

2) 不能充分发挥人机对话,随机判断的优点。检索系统是很先进的,其功能是很强的。在香港终端上的操作人员,只是根据国内寄去的提问单进行人机对话,帮助用户检索,由于不可能熟悉各部专业,有时又不能融合贯通理解用户每个课题的检索意图,所以有的课题检索质量不好。

为了缩短联机检索的周期,同时为了提高检索质量,从1982年9月起冶金部,石油部、化工部和兵器工业部已分别与美国DIALOG系统和ORBI系统鉴订了合同。安装电传机,不通过香港终端,直接利用电传终端检索美国上述两个系统的文献。

这样首先检索周期大大缩短了,查到的文献一般只要一周时间就可从美国航寄到北京,用户急需的文献,很快就能拿到手。其次检索质量提高了。用电传检索时,可根据系统的回答随机判断,特别当用户在场,可以随时征求用户的意见,灵活地增删检索词或改变检索策略,并根据用户的要求规定脱机打印多少篇文献。

但是由于国内电传费比较贵,每分钟7.60元。检索一个课题一般得上百元,如脱机打印

文献多，得花几百元。是香港终端的两倍。

中国科学技术情报研究所在调查了建立国际终端的必要性和可能性的基础上，对各种可能的方案及其经济性和适应性进行了分析比较，认为在目前情况下，象国外那样建立公用数据网进行检索国外数据库的条件尚不成熟，象国内某些单位那样租用专线，与国外专门的数据库相连接，又费用太高，在经济上不够合理。因此，只有采用分组交换技术，建立一种小型的公用数据检索系统才是既经济又适用的方案。

在邮电部门的支持和帮助下，采用了通过意大利公用数据网连接到欧洲和美国的数据交换方案。1983年2月，我国与意大利国际电报局（ITALCABLE）达成了开展中意两国公用数据业务（分组交换业务）的协议。按照该协议，在中国的分组交换业务的控制中心（北京电报大楼）设立2套（一主一备）9600BPS（比特/秒）的CODX96调制解调器。该调制解调器的一端通过卫星线路与罗马的分组交换中心相连，另一端有4个端口，每个2400BPS，其中一个接2400BPS的ESA—IRS高速终端，另一端口接一个统计型多路复用器（TDM），为中国科技情报研究所和其他部门提供用户接口，可连接4个1200BPS和4个300BPS的用户终端。调制解调器尚有两个端口留作备用。

该系统如示意图1—6所示。中国科技情报所目前已安装2套300BPS的用户终端，其中一套为DEC公司的LA—34打印终端，另一套为同一公司VT100显示终端加LA—120打印机。待有关设备齐备后，将再扩充两个1200BPS终端，并计划安装一套2400BPS的ESA—IRS高速终端。其他4个终端将安装在其他部门和地区。该系统具有扩充能力，通过增加多路复用器或安装符合X25的集中器可增加异步终端64台。

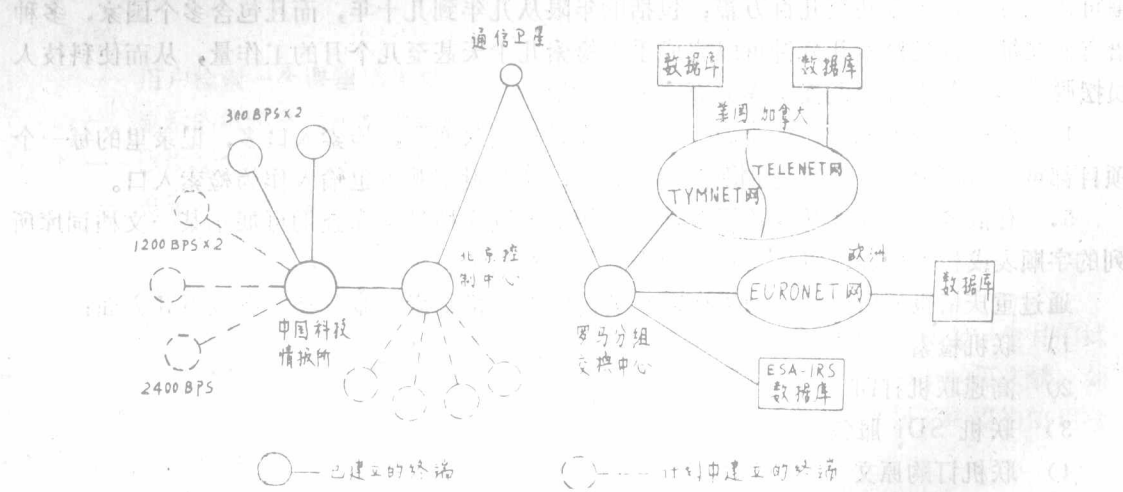


图 1—6 我国使用的国际联机情报检索系统构成示意图

图 1—6 中，在意大利分组交换中心里有专线与 ESA—IRS 数据库相连，并通过网络接口系统与欧洲的 EURONET 网和美国的 TYMNET 网、TELENET 网相连。所以中国的终端用户除检索 ESA—IRS 数据库外，还可检索 EURONET、TYMNET、TELENET 网上的各种数据库。1983 年 9 月 1 日，中国科技情报所首先与 ESA—IRS 联机。接着又与美国 DIALOG 数据库联机。目前提供检索这两个数据库的服务，不久还将提供检索美国

ORBIT 及其他数据库系统的服务。

我国首次采用分组交换技术建立起来的公用数据检索终端开始使用后,受到了各部门各单位的热情支持和帮助。在1983年5~6月份由联合国教科文组织、欧洲空间组织情报检索中心和中国科技情报所、邮电科学研究院联合举办的“系统设计和联机检索学习班”上进行了检索试验,共检索了800多个课题。对其中200个课题的检索结果进行调查分析表明,其中83%的课题用户满意和比较满意,不太理想的只有17%。不太理想的原因大多数是由于查找人员刚刚学习终端的使用技术,对选词、选库和使用检索策略都不熟练。说明检索效果与使用人员的业务水平关系也很大。因此今后应进一步提高广大科技工作者的国际联机检索知识和技能,以便更好地发挥我国国际联机检索终端的作用。

1985年4月18日,中国科技情报研究所重庆分所国际联机情报检索终端开机。重庆分所的终端是通过北京的控制中心与通信卫星联系,从而实现对 ESA—IRS 和 DIALOG 的国际联机检索,使交通不便的西南地区亦实现了情报检索现代化。

二、国际联机情报检索系统的优点及费用

与普通手工检索方法相比,联机检索方法有许多优点,其主要点可归纳如下:

1. 打破国家距离、地理位置界限,共享情报资源。
2. 能获得最新信息,如 ESA—IRS 的46号文档,PRICEDATA 更新周期为170数据/天。
3. 检索速度快,利用联机终端查找一个课题,一般只要10分钟左右,而查找的文献范围可达几万,几十万甚至几百万篇;包括的年限从几年到几十年,而且包含多个国家,多种语言的文献。联机检索几分钟可以完成手工检索几十天甚至几个月的工作量,从而使科技人员摆脱大量的手检劳动,查全率方面还高于手检。
4. 检率途径多,计算机检索较之手工检索标引深度深,检索入口多,记录里的每一个项目都可作为检索入口。通过非规范词的标引,文章附带观点也输入作为检索入口。
5. 有很多启发性的功能及服务功能,例如通过文档的词库查询可展示某一文档词库所列的字顺表或相关词表,以启示用户选出合适的检索词。

通过重庆情报分所ESA—IRS终端,进行联机情报检索的服务功能有以下几方面:

- 1) 联机检索
- 2) 高速联机打印
- 3) 联机 SDI 服务
- 4) 联机订购原文
- 5) 联机数据登录:为用户创建自己的联机数据库。
- 6) 电子邮政(DDS)

即利用通信网络、在用户终端之间高速传递报文。

- 7) 用户帐目管理。

除了上述服务功能外,近来系统还增加了帮助用户选择相关文档的 QUESTINDEX 指令,以及对检索词进行自动频度统计的 ZOOM 指令。此外,ESA—IRS 还提供各种联机指导和各种训练文档,并且定期举办用户培训班和发行有关资料。

使用国际终端进行联机检索时，用户费用包括以下4项：

1. 数据库使用费

按使用时间计算，视不同数据库而异。ESA—IRS 数据库最高为每小时100美元，最低为每小时10美元，平均约每小时60美元。DIALOG 系统最高为每小时300美元。最低为每小时10美元，平均每小时约80美元。

2. 打印费

有脱机打印费（在国外打印）和联机打印费（在终端打印）两种。一般脱机打印一个记录（题录或文摘）为20美分左右。联机打印费因我国通信线路收费高，费用大大高于脱机打印费。

3. 通信线路费

按通信线路的连接时间和传输的信息量计算，即：通信线路费 = [连接时间 × 费用 (Au /单位时间)] + [传输信息量 (千字符) × 费用 (Au /千字符)]。 Au 为欧洲共同体货币结算单位， $1Au = 0.97797$ 美元。

邮电部规定，中意两国之间的直达业务：每千字符1.20元人民币，连接每分钟0.60元人民币。

4. 管理费

按上述费用的比例加收少量的管理费。

重庆情报分所ESA—IRS终端实际按下列公式收费：

$$P = (a \cdot b + e) c + d$$

式中：

P ——用户检索一个课题的用费。

a ——显示的机时费与联机打印费。

b ——该终端到两个系统的线路费系数（目前定为1.8）。

c ——美元与人民币的兑换比值（目前定为2.9）。

d ——管理费（目前定为每课题8元人民币）。

e ——脱机打印费（免收线路费）。

总的说来联机检索费用较高，据中国科技情报所1983年9月中旬到10月中旬一个月的试用结果表明，用户查找一个课题平均费用约合人民币60元，平均一个课题命中60篇文献，如果一个课题只查找一个数据库，平均费用在人民币30元左右。在进一步掌握了终端的使用技术后，费用还可能再降低。