

联机检索基础

[美]查尔斯·T·梅多 波林·A·科克伦 著

陶涵或 任月娟 译

祝希龄 校

科学技术文献出版社重庆分社

联机检索基础

[美]查尔斯·T·梅多 波林·A·科克伦 著

陶涵或 任月娟 译 祝希龄 校

责任编辑: 林云梯

科学技术文献出版社重庆分社 出版

重庆市市中区胜利路132号

新华书店重庆发行所 发行

科学技术文献出版社重庆分社印刷厂 印刷

开本: 787×1092毫米1/32 印张: 11.125 字数: 24万

1987年10月第1版

1987年10月第1次印刷

科技新书目: 155—346

印数: 5000

ISBN7-5023-0131-3/TP·14

统一书号: 15176·743

定价: 2.60元

目 次

第一章 绪论	(1)
第一节 几个基本定义.....	(1)
第二节 检索服务机构.....	(3)
第三节 联机检索服务的利用.....	(4)
第四节 某些词汇.....	(5)
第二章 交互检索原理	(13)
第一节 数据库的组织.....	(13)
第二节 用户和数据库的连接.....	(18)
第三节 选词.....	(21)
第四节 组配逻辑.....	(26)
第五节 浏览和打印.....	(31)
第六节 命令语言的不同形式.....	(32)
第七节 联机检索的费用.....	(33)
第三章 情报专家和情报需求者的预检商谈	(36)
第一节 先求助于情报专家.....	(36)
第二节 预检商谈的步骤.....	(38)
第三节 检索的概念化.....	(43)
第四节 预检商谈中的人际接触.....	(46)
第五节 预检商谈用于联机检索的实例.....	(48)
第六节 小结.....	(51)
第四章 终端与网络	(57)
第一节 通信与计算机.....	(57)
一、数据通信网络.....	(58)
二、终端.....	(61)

三、调制解调器·····	(63)
第二节 终端的使用·····	(63)
一、ASCII键盘和APL 键盘·····	(65)
二、专门功能·····	(65)
三、其他专用开关·····	(66)
第三节 通信故障·····	(68)
一、计算机停机·····	(68)
二、网络故障·····	(68)
三、终端故障·····	(69)
四、用户错误·····	(69)
第五章 检索命令语言简介 ·····	(71)
第一节 命令语法·····	(71)
第二节 隐含和缺席命令与值·····	(72)
第三节 检索方式·····	(73)
第四节 机器翻译命令·····	(74)
第六章 数据库 ·····	(77)
第一节 可用数据库的范围·····	(77)
一、政府机构建立的数据库·····	(78)
二、专业协会建立的数据库·····	(78)
三、商业机构建立的数据库·····	(79)
第二节 记录结构·····	(80)
第三节 文档的组织·····	(89)
一、文档的类型·····	(89)
二、数据元的类型·····	(89)
三、词典文档·····	(92)
四、文档的同时使用·····	(92)
第四节 文档维护·····	(94)

第五节 利用文档卖主提供的文档介绍·····	(96)
第七章 基本命令 ·····	(99)
第一节 基本命令的功能·····	(99)
第二节 选词命令·····	(99)
一、词典和叙词表检索·····	(100)
二、集的形成·····	(103)
三、集形成命令的变化·····	(105)
第三节 集的组配·····	(107)
一、集组配的基本方法·····	(108)
二、集组配逻辑·····	(109)
三、DIALOG系统和BRS系统的变化·····	(111)
四、组配算符的使用说明·····	(111)
五、概念组或面·····	(112)
六、概念集的组配·····	(114)
七、集的大小·····	(114)
第四节 集合中某些含义的限制·····	(115)
第五节 显示与浏览·····	(117)
第六节 检索实例三则·····	(119)
一、DIALOG系统·····	(119)
二、ORBIT系统·····	(126)
三、BRS系统·····	(129)
第八章 文本检索 ·····	(132)
第一节 文本检索的起源·····	(132)
第二节 规范词表和非规范词表·····	(133)
第三节 文本检索的使用·····	(138)
第四节 词段落检索·····	(139)
第五节 词串检索·····	(141)

第六节	何时使用文本检索	(144)
第七节	全文本联机检索：一项新的实验	(147)
第九章	检索的开始与结束	(150)
第一节	与计算机连接	(150)
一、	与网络连接	(152)
二、	与终端相匹配	(152)
三、	终端类型的识别	(152)
四、	目的计算机的识别	(153)
五、	网络与检索服务的连接	(153)
六、	用户识别符号	(153)
第二节	口令	(153)
第三节	消息	(154)
第四节	数据库的选择	(155)
第五节	检索的重新开始	(159)
第六节	文献的订购	(160)
第七节	时间和费用	(160)
第八节	请求脱机	(160)
第十章	检索结果的存储和定题服务	(164)
第一节	检索结果的存储	(164)
一、	DIALOG系统	(164)
二、	ORBIT系统	(165)
三、	BRS系统	(167)
第二节	定题服务	(170)
一、	DIALOG系统	(170)
二、	ORBIT系统	(170)
三、	BRS系统	(171)
第三节	交叉文档检索	(172)

第四节	小结	(174)
第十一章	辅助系统	(175)
第一节	评定检索集合	(175)
第二节	向计算机提问	(178)
第三节	电话和其他辅助系统	(179)
第十二章	检索策略	(185)
第一节	什么是检索策略	(185)
第二节	概念组和词的检索策略	(189)
一、	先用最专指面	(189)
二、	先用最少记录面	(190)
三、	积木式概念组	(191)
四、	引文珠形增长	(193)
五、	连续分段	(196)
第三节	小结	(197)
附录A	检索语言概述	(200)
附录B	市售的联机数据库	(228)
附录C	数据库介绍和检索指南实例	(251)
译后记		(339)

第一章 绪 论

第一节 几个基本定义

本书的目的是向您介绍用计算机检索数据库。我们着重阐述文献数据库和联机交互检索的计算机有关问题。

文献数据库(*Bibliographic data base*)包括文献著录(*descriptions*), 即记录的文档。这些记录可供您用来判定是否需要检索原文献。我们所讲的这些文献可能是期刊论文、报告、专利、图书或其它类似的资料。这些记录中著录的内容包括标题、著者、出版者和出版日期以及主题。主题可以包括主题分类词(*term*)或关键词, 也可以包括文摘。我们使用词这一概念表示标题或文摘外的任何一个著录项目。叙词(*descriptor*)是一个由标引人员用来描述文献主题内容的专用词。

联机检索(*Online searching*)是指检索者通过检索程序与计算机直接通信。用更为技术性的话来说, 它指的是直接或通过电话把终端与计算机连接起来, 终端(*terminal*)是用来把信息传递给计算机并从计算机接受信息的一种设备。交互(*interactive*)检索是检索人员和计算机互相对话、交换信息和提出问题的一种检索。如果不进行联机, 计算机检索系统实际上就不可能是交互的。因为, 如果不进行联机, 信息交换就会费时太久——正如用电报而不是用电话交换信息一样。由于这两个特性经常互相关联, 联机一词在通常使用中也就赋有交互的含义。在后面, 我们还将更详尽地探讨这些术语。

联机检索的兴起至今只有十年左右，但发展极为迅速。我们现在论及的数据库，有些不限于文献资料，而且也可能包括统计资料。有朝一日，我们能期望检索文献的全文，而不仅仅是摘要。

如果您不熟悉该领域中使用的术语，或者该领域中涉及的公司、服务机构、产品和主要用户的名称，那么，您可以查阅这样一些出版物：《联机》(Online)、《联机评论》(Online Review)、《数据库》(Data Base)或《情报管理者》(The Information Manager)。大多数开展联机文献检索业务的公司在这些期刊中登载有广告。它们免费赠送的小册子(和类似的书籍)有助于您大胆地进入正在飞快发展的该领域。有许多这类公司在诸如美国图书馆协会年会这样的专业会议上举办展览。

例如，打开《情报管理者》，会有鳞次栉比的广告展现在您面前，使您了解许多情报检索系统和服务机构，数据总库和数据库提供者，甚至有精于“标引系统词表”的情报专家公司。该领域内的职业市场在日益扩大。

有一则广告是这样描述联机检索的：

“计算机文献检索——几分钟内能找到有效情报，以易读的形式显示检索结果；迅速交付完整的原始文献，只支付使用费，……使用设在贵机构中的终端和电话能查找全世界广泛主题领域的文献。”*

同一期杂志的另一则广告并没有提到计算机，但它画了一幅附有终端的图书馆书库图，而且把“情报检索系统”称为“世界首位的联机检索服务，提供70多个数据库和18,000,000篇以上的参考文献”。“我们对图书馆提供许多方便。”

* 《情报管理者》(Information Manager), 1978年8月号。

象洛克希德的DIALOG检索服务系统,系统开发公司(SDC)的ORBIT检索系统、环境科学情报联机检索(Environline),国会情报服务处(CIS)和纽约时报的情报库(Information Bank)等名称都是通过这一商业性刊物反映出来的。该商业性的刊物是“为那些从事于把许多不同情报处理技术综合成一个总系统的人们所用的。”

我们希望,本书将有助于您阅读和理解《联机》或《情报管理者》这类期刊中的广告和论文,并帮助您成为一名能独立进行检索的联机检索人员。

第二节 检索服务机构

提供有助于联机检索开展的计算机服务和辅助服务之组织机构称为检索服务机构(search service)。这些机构往往是私人公司(卖主),有时候是政府机构和大学的中心。数据库通常由另一组织私人公司、政府机构或专业文摘社所建立,它们以租借的方式和/或收使用费的方式向检索服务机构提供数据库。例如:美国化学学会所属的化学文摘社(CAS)、美国石油学会中心文摘和索引服务社(CAIS)和美国教育总署的教育资源情报中心(ERIC)等均为数据库的提供者。在第六章和附录C中将详细介绍其它的数据库及其文献工作。

检索服务机构提供能同时为许多用户共用的计算机。它们从许多数据库提供者那里(洛克希德的DIALOG情报检索系统提供100多个数据库)汇集数据库,供用户使用,并提供辅助服务,如人员培训、文献加工和业务咨询工作。检索服务机构还作为安排用户和执行检索的计算机之间通信的机

构。图书馆和其它情报中心使用的主要联机服务系统如下：

- (1) DIALOG (洛克希德情报系统)；
- (2) ORBIT (系统开发公司——SDC)；
- (3) BRS (书目检索服务公司)；
- (4) MEDLINE (国家医学图书馆)；
- (5) New York Times Information Bank (纽约时报情报库)；
- (6) LEXIS (米德技术实验室数据中心)。

在本书中，讲述检索方法时力图使描述的内容具有普遍性，但是，我们将从DIALOG、ORBIT和BRS等系统中选取具体例子加以说明。

今天，检索服务的主要买主是图书馆——院校图书馆、专业图书馆、公共图书馆，甚至学校图书馆。这些图书馆可以向其它顾客转售，也可以免费或收取部分费用的形式向用户提供服务。有一些（但并不是所有的）图书馆鼓励用户学会独立进行检索，但是绝大多数的图书馆则雇用受过高级培训的专家，他们可称为检索专家、情报专家或数据系统馆员。这些中介人帮助用户检索或替用户进行检索。本书着重叙述联机文献检索的过程，其中包括利用检索服务机构和用户间职业中介人的问题。

第三节 联机检索服务的利用

图书馆参考咨询台周围的推荐性小册子和告示牌常常向用户指明负责联机文献检索的工作人员。这些告示牌向用户表示“机助研究”、“联机数据库检索”、“书目咨询服务”、“定向书目服务”或“情报检索”等服务项目。这种服务通常被视为

传统参考咨询服务的扩展，而且还可以供其它地方，例如化学图书馆、医学中心图书馆和参考咨询部等处使用。这些小册子除了介绍可利用的检索服务项目外，还介绍可资检索的数据库。

由于检索费用的高低是取决于这样一些条件：如图书馆是否想全部或部分收回费用，对用户检索中所用的数据库、检索输出的范围和联机的时间有何限制等。因此，如果检索服务要收取费用，那就得预先向用户说明。

图书馆偶尔会向用户提供几种文献指南。当用户确定要对这些类目中的项目进行联机检索时，图书馆就会介绍这种服务，并详细加以解释。

除图书馆之外，在多数大城市中还有情报服务机构或“情报经纪商”代为进行收费联机检索。他们可能在电话簿的黄页*中列出。他们的工作性质颇似上述的图书馆服务。

第四节 某些词汇

如同所有新兴领域一样，特定意义的联机检索和一般意义的情报管理（或者如通常称谓的情报存储和检索）都有一些术语方面的问题，许多词汇交替使用。有些概念各有几个附属的名称。许多“同义词”从不同的方面或不同的发展过程中衍化而来。如果不能适当地给所有的概念下定义，那么我们提供的准同义词表和同义词表（表1.1和表1.2）也许有助于您阅读其它作者的著作。

这些词表也许会让您了解到各种词如何成为同义的。严

*国外有些参考工具书中的某些插页用黄色纸印刷，以示区别。查阅时一目了然。

格说来,例如计算机,交互和联机不是同义词,因为后两个词描述整个计算机系统,但在联机文献检索(即本书的重点)范畴内,这些词作为形容词交替地用来描述该领域内计算机系统或服务。

表1.1第一栏中这些词的含义是:服务或系统(第二栏)是以计算机为基础的。采用了计算机系统中最现代化的形式,即允许用户进行人机对话(交互),或者甚至在用机时使计算机暂停并仍然保持联系(联机),而不致完全停止一项工作,只能等以后收到脱机完成的检索结果。所以,一些作者使用联机这个形容词,另一些作者使用交互这个词,还有些作者则称自动化或计算机化。

表1.1的各栏包括“同义”词。一栏中的一个词能与相邻一栏的词或短语组配,构成一个含义更广泛的概念。在一栏中找出一个您能定义的词,它能帮助您浏览该栏目,以了解其它涵义相近的“同义词”。例如,一篇关于该主题的论文可能包括下述句子:带有用(1)叙词表和其他(2)可检索字段中(3)叙词来评价(4)文献记录的(5)数据库的(6)以计算机为基础的(7)文献检索设备。(8)检索语言允许(9)中介人来帮助(10)终端用户查找(11)原始文献。

这两个句子可用这11栏中各栏的其它词或词组来表示,对某些读者有同样含义,对另一些读者略有不同。(见表1.2。)各栏括号内的词是计算机出现以前图书馆界沿用的传统术语。

如果这点不会使你迷惑,那倒是有某些帮助,因为实际上,情报一词六百年左右并未有过合适的定义。(根据《牛津英语辞典》该词于1387年第一次被使用。)当时拼写为“informacioun”!现在这单词有许多定义和有关的概念。

表 1.1 描述联机文献检索概念所使用的准同义词表

(1)	(2)	(3)	(4)	(5)	(6)
机 器	系 统	数据库	文 献	数据元	叙 词
计算机	服务(机构)	数据总库	参考文献	字 段	词
以计算机为基础的	检 索	文 档	记 录	信息字段	文 本
自动化的	检索服务	索 引	书目记录	引文元	检索词
计算机化的	检 索	数据 集	书刊种类	存取标识	单元标号
联 机	情报系统	(图书馆目录)	单元数据	可检索字段	关键词
人机对话(交互)	检索系统	(文摘服务)	单元记录	记录结构	规范语言
机械化的	检索服务	(图书索引)	(目录卡片)	标识符	概 念
远程终端	题录检索	(文献)		(目录款目)	自由文本/自然语言
	文献检索			(目录卡片的	(主题标目)
	文献服务			项目)	
	卖 主				
	情报提供者				
	情报检索(参				
	考服务)				

(7)	(8)	(9)	(10)	(11)	(12)
词典	指令	中介者	用户	检索	输出
词表	对话	联机检索员	终端用户	需求题目	检索结果
基本索引	功能	检索分析员	检索员		原始文献
倒排文档	检索语言	情报专家	需求者		(图书馆藏书)
叙词表	计算机软件	(参考馆员)	读者		(图书)
目录			(图书馆读者)		(期刊论文)
检索词	检索逻辑				(报告)
(图书馆目录款目)	集合组配				
(标引系统)	检索语句集				
	系统特性				
	检索策略				
	(目录检索)				

注: A. 各栏(1—12)列出该领域中各作者最经常交替使用的词或词组。
 B. 在该领域为文献中, 一栏中的词与相邻一栏的词组合, 以形成更广泛的概念, 或提供上述总系统的一些详情。例如第1—2栏中的词表明整个领域; 第3—4栏中的词表示系统中的内容, 第5—7栏表示文档中内容和文档来源; 第8—11栏表示某人为某人做什么事; 第12栏的词表示输出。
 C. 括号中的词把本“现行”目录与图书馆学文献中的旧术语联系起来, 这些术语是准同义词, 例如数据库或文档(第3栏)相当于图书馆目录和图书索引; 叙词(第8栏)相当于主题标目; 可检索字段(第5栏)相当于目录款目; 中介者(第9栏)相当于参考馆员等等。

表 1.1 的原文对照表

(1)	(2)	(3)
Machine	System	Data base
Computer	Service	Data bank
Computer-based	Search	File
Automated	Search service	Indexes
Computerized	Retrieval	Data set
Online	Information system	(Library catalog)
Interactive	Retrieval system	(Abstracting service)
Mechanized	Retrieval service	(Book index)
Remote terminal	Bibliographic retrieval Literature searching Literature service Vendor Supplier Information retrieval (Reference service)	(Literature)
(4)	(5)	(6)
Documents	Data elements	Descriptors
References	Fields	Words
Records	Information fields	Text
Bibliographic records	Citation elements	Search terms
Items	Access points	Element labels
Data	Searchable fields	Keywords
Unit records	Record structure	Controlled language
(Catalog cards)	Tags	Concepts

	(Catalog entries (Elements of a catalog card)	Free text/natural language (Subject headings)
(7)	(8)	(9)
Dictionary	Commands	Intermediary
Vocabulary	Dialogues	Online searcher
Basic Index	Functions	Search analyst
Inverted file	Search language	Information specialist
Thesaurus	Computer software	(Reference librarian)
List	Protocols	
Search words	Search logic	
(Library catalog entries)	Set combination	
(Indexing system)	Set of search statements	
	System features	
	Search strategy	
	(Catalog search)	

(10)	(11)	(12)
User	Search	Output
End-user	Request topic	Search results
Searcher		Source documents
Requester		(Library collection)
Client		(Books)
(Library patron)		(Journal articles)
		(Reports)