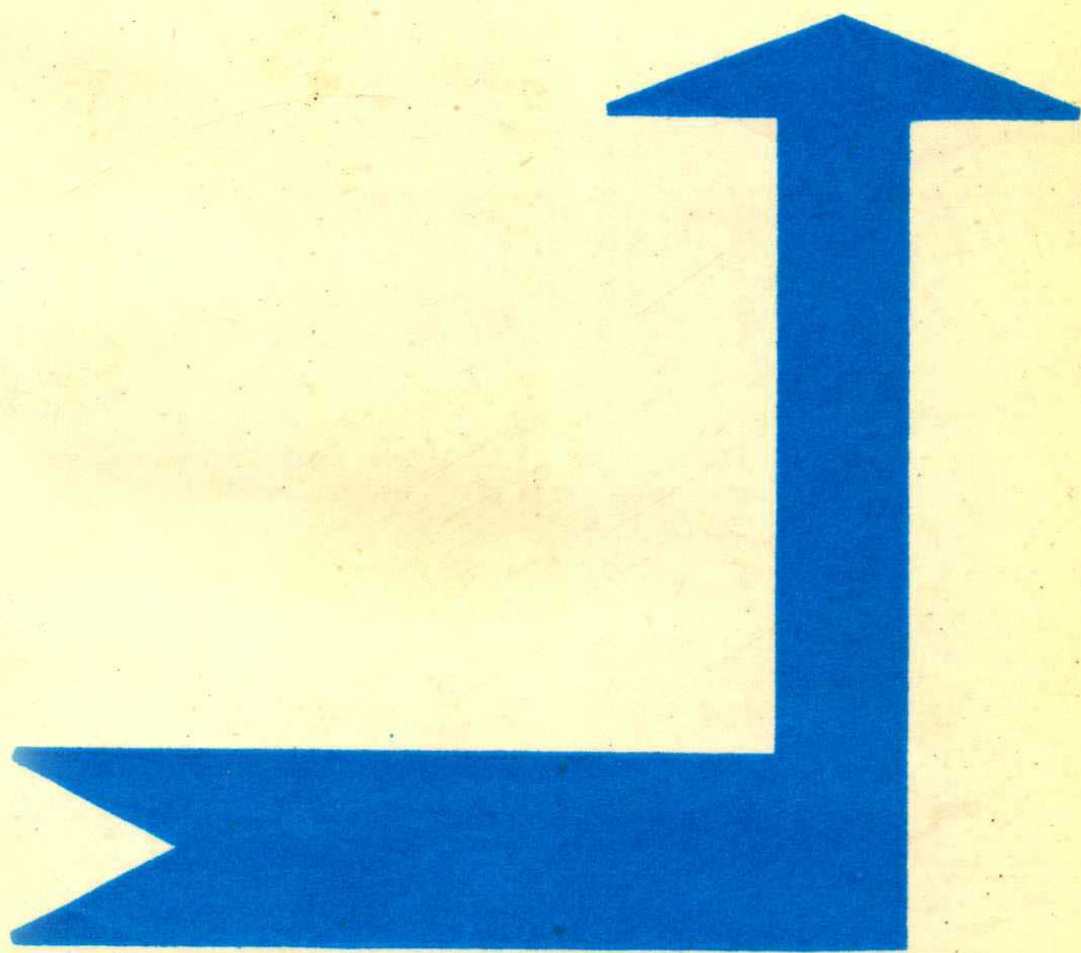


计算机情报检索导论

宋余庆 编



中国医药科技出版社

计算机情报检索导论

宋余庆 编

中国医药科技出版社

内 容 提 要

本书旨在为从事科学技术工作者学习使用计算机情报检索这一先进情报收集方法而作。全书共分六章,力求融计算机基本知识与情报检索于一体,系统阐述计算机情报检索基本知识、计算机情报数据库结构及其组织方法、计算机情报检索原理及过程和中文计算机情报检索系统、国际联机情报检索系统以及国外光盘情报检索系统的使用。

本书适用于从事科学技术工作者学习计算机情报检索;也可以用作高等院校“文献检索课”补充教材;还适用于图书情报学师生参考。

计 算 机 情 报 检 索 导 论

宋 余 庆 编

*

中国医药科技出版社 出版

(北京西直门外北礼士路甲 38 号)

(邮政编码 100810) (物化探所实验工厂印刷厂 印刷)

全国各地新华书店经销

*

开本 787×1092/mm 1/32 印张

字数 80 千字 印数 1—3100

1993 年 8 月第 1 版 1993 年 8 月第 1 次印刷

ISBN 7-5067-0969-4/Z. 0064

定价:5.40 元

主编

宋余庆

编者

宋余庆

周毅华

陈路平

邵 军

刘 骥

陈森森

钱 俊

许 祥

袁晓红

汤永祥

序

计算机情报检索,既是历史悠久手工检索系统的发展与深化的必然结果,也是在手工检索的基础上,运用计算机及其他相应辅助设备高速、准确、自动地功能与手段,使其完成资料的存贮、修订、增补、排序、组合、检索等实现过程。这里值得一提的是有关“Information”一词的译法,几十年来它都被译成中文的“情报”或“信息”等。然而,在中文里的“情报”与“信息”(台湾学术界译作“资讯”)它们的外延与内涵均有颇大的差异,二者不是“相同”“相等”、“平行并列”关系,而应属于从属关系,则“信息”的范畴大于“情报”的范畴。据新华社 1992 年 9 月 20 日北京电,国家科委决定,采用“科技信息”的称呼取代原来的“科技情报”。并决定将该系统内,从事科技信息收集、整理和服务的工作及机构,同时随之改名。如将“国家科委科技情报司”改为“国家科委科技信息司”,“中国科技情报研究所”改为“中国科技信息研究所”。将“情报”改为“信息”考其原因,理由之一是:随着科技对外交往的增多,在东南亚和我国港台地区,因“情报”常与“谍报”相混淆,造成某些不必要的误解。目前,尚有一些机构也正在陆续更名,理由之二是:借助于“信息”含盖面大等特点,试图拓宽其研究领域与业务范畴,扩大人才培养的口径。但是,这决不意味着“信息”可在任何意义上任何场合中都能取代“情报”,更不意味着“情报”一词将从图坛学苑乃至字典辞典一类字书中悄然消逝。辞汇的更新、消失、复出、兼并、转换,需要经过历史的筛选甚至是漫长的衍变过程。作为“情报学”、“情报检索”等的名称,我认为还会继

续存在并发展下去。应该把某些业务工作及其相应业务机构名称的更替与科学的名称区别开来,因为后者更具有相对的稳定性、兼容性与独立性。也许正因为如此,尽管一些高校有关系名已经作了或今后还有可能作出相应的改变,但其相关专业设置,学科、学位、课程名称等还依然继续采用“情报”这一名称。

计算机情报检索若视为一种独立学科,约使于本世纪 50 年代后期与 60 年代初期。纵观其技术发展轨迹,大体历经着:脱机批处理检索阶段,联机检索阶段,网络化检索阶段等。自 70 年代以来,微电子技术、现代通讯技术与情报事业的发展 and 相互交叉渗透,使情报检索工作得以实现自动化、网络化。近十多年来,计算机情报检索事业发展很快,仅以古特拉 (Cuadra) 的《联机数据库名录》计录统计:

一、联机系统检索的增长

1979/1980 年	59 个
1980/1981 年	93 个
1981/1982 年	170 个
1982/1983 年	213 个
1983/1984 年	272 个
1984/1985 年	362 个
1986 年	454 个
1987 年	528 个
1988 年	555 个
1990 年	644 个

二、数据库生产者的增长

1979/1980 年	221 个
1980/1981 年	340 个
1981/1982 年	512 个
1982/1983 年	718 个
1983/1984 年	927 个
1984/1985 年	1189 个
1986 年	1379 个
1987 年	1568 个
1988 年	1685 个

三、联机情报数据库的增长

1979/1980 年	400 个
1980/1981 年	600 个
1981/1982 年	965 个
1982/1983 年	1350 个
1983/1984 年	1879 个
1984/1985 年	2453 个
1986 年	2901 个
1987 年	3369 个
1988 年	3699 个
1990 年	4465 个

我国计算机情报检索工作,近年来有了很大的发展,正向一个新台阶迈进。截至 1990 年,我国自建数据库约 400 种,引进国外数据库约 70 种。我国约有 50 个城市已开通与国外 12 个大型情报检索系统联机检索服务。在我国计算机情报检索正在迅速发展,其使用方法也亟待推广。

《计算机情报检索导论》的成书与出版十分必要,亦非常及时。宋余庆同志,既攻读过医学又读过文献情报学,更为可贵的是他在这一领域内作了大量工作并在实践中总结整理出一套切实可行的学习方法。本书是著者集多年实践与教学之结晶,用通俗、生动的语言并结合实例系统介绍了利用各种计算机情报检索系统进行医学情报检索的全过程。这部书稿与同类或近似的著作相比,有其许多独到之处。著者力求将计算机基本知识和情报检索知识融于一书,让初学者能独立进行计算机情报检索;著者十分注重理论联系实际,让读者既能知其然更能知其所以然,以致可以举一反三;著者非常关注手检与机检之间的密切联系与区别,始终瞄准计算机情报检索领域内的学科前沿课题。我深信这本书稿,确实不失为一部佳作,值得向广大读者推荐。

倪 波

1992年11月26日于南京

前 言

当今科学技术的发展,促成了科学研究成果和科技情报量的激增。逐渐增大的情报信息量使人们难以用原始落后的情报搜集方法,满足自己的情报需求。在当代社会中,若不借助于现代情报搜集手段,任何一个科学工作者,要想搜集全本课题、本学科内的所有情报资料;掌握学科最新进展,是绝对不可能的。

计算机情报检索是现代情报搜集手段之一。它是随计算机事业的发展而产生和发展的。本世纪 70 年代以来,微电子技术、通讯事业以及情报事业的共同发展,使情报检索工作实现了自动化、网络化。仅 30 多年的时间,全球计算机情报检索系统已达 300 多个;数据库已逾二千。

个人电脑的发展,使发达国家利用电脑检索情报资料较为普及,许多科学工作者在家里或工作间,都能随时将自己的电脑与大型计算机情报检索系统进行联机,进而方便地查得自己所需的情报资料。

我国自 1983 年以来,陆续建成了许多国际网络化计算机情报检索终端站/点,现已遍布各大城市。国内自己的情报数据库亦已开发,并不断完善。我国的计算机情报检索硬环境已经形成。如何应用这种良好的环境为科技人员的工作服务,摆在了我们面前。1992 年国家教委要求将计算机情报检索列入高校《文献检索》课教学的基本要求。这说明现代科技工作者掌握计算机情报检索之重要。据此,我们写成此书,试图从理

论与实践相结合的角度,让读者真正学会计算机情报检索这一现代化情报收集手段。

南京大学倪波教授对本书进行了认真审校,并热枕为之作序。这里,我们深表谢意。镇江医学院图书馆、中国教育图书进出口公司光盘科为本书作了大量工作,镇江医学院王兰芳小姐承担了全部书稿的打印,我们一并表示感谢。

宋余庆

1993 年 1 月

目 录

第一章 导论	(1)
第一节 基本知识	(1)
一、概况	(1)
二、基本概念	(2)
第二节 计算机情报检索系统的基本构成	(3)
一、硬件	(3)
二、软件	(4)
三、情报数据库	(6)
第三节 计算机情报检索系统类型	(6)
一、脱机情报检索系统	(6)
二、联机情报检索系统	(6)
三、网络化情报检索系统	(7)
第二章 数据库结构与检索原理	(10)
第一节 计算机数据库系统	(10)
一、数据库及其管理系统.....	(10)
二、数据库数据组织.....	(11)
三、数据库数据文件结构.....	(13)
第二节 情报数据库系统	(14)
一、概念.....	(14)
二、情报数据库的基本类型.....	(15)
三、情报数据库的结构.....	(16)
四、情报数据组织的标准化.....	(17)

第三节 检索原理	(19)
一、计算机数据检索原理	(19)
二、计算机查找方法	(19)
三、计算机情报检索过程	(20)
第三章 检索策略	(23)
第一节 检索途径和检索词	(23)
一、分类途径和分类号	(23)
二、主题途径与主题词	(24)
三、著者途径与著者姓名	(25)
四、序号途径与序号	(26)
五、来源途径与刊名	(27)
第二节 算符	(27)
一、逻辑算符	(27)
二、位置算符	(28)
三、字段限定符	(29)
四、算符的优先权	(30)
第三节 制定检索策略	(30)
一、选择数据库	(31)
二、选择检索词	(31)
三、拟定检索表达式	(33)
四、准备修改方案	(33)
第四节 检索技巧	(34)
第四章 国内中文情报检索系统	(37)
第一节 概况	(37)
第二节 汉字及其输入法	(38)
第三节 中文科技期刊篇名数据库	(41)

一、简介·····	(41)
二、操作步骤和方法·····	(42)
三、检索·····	(43)
四、显示/打印结果·····	(47)
五、退出系统·····	(47)
第五章 国际网络化联机情报检索系统 ·····	(49)
第一节 概况 ·····	(49)
第二节 字段代码及命令集 ·····	(50)
一、字段代码·····	(50)
二、命令集·····	(51)
第三节 检索步骤和方法 ·····	(53)
一、进入数据通信网络·····	(53)
二、接通联机情报检索系统·····	(54)
三、检索·····	(54)
四、退出系统·····	(55)
第六章 国外光盘情报检索系统 ·····	(57)
第一节 概况 ·····	(57)
第二节 字段代码与命令集 ·····	(58)
一、命令集·····	(58)
二、字段代码·····	(58)
第三节 检索方式 ·····	(59)
一、普通检索·····	(59)
二、标题词库自动检索·····	(65)
三、索引词自动检索·····	(70)
第四节 操作步骤和方法 ·····	(78)
一、启动计算机·····	(78)

二、进入检索系统.....	(78)
三、检索.....	(78)
四、显示检中内容.....	(78)
五、打印检出内容.....	(81)
六、转存检索结果.....	(83)
七、退出系统.....	(86)
附录一 DIALOG 医学及其相关文档表	(88)
附录二 禁用词表	(91)
附录三 常见出错的处理	(93)
主要参考文献	(93)

第一章 导论

情报检索是一门新兴学科。自创立以来发展很快,尤其是本世纪 40 年代计算机问世以来,情报检索科学进入了崭新的历史时期。自此,计算机与情报检索两门学科间便结下了不解之缘,在发展中彼此相互促进。计算机科学为情报科学的飞速发展奠定了基础,情报科学为计算机科学提供了十分广阔的应用前景。

计算机情报检索是计算机与情报检索结合而成的一个新应用领域。是现代科技工作者必须掌握的一项技能。对一般科技人员而言,要掌握这项技能,首先必须了解计算机与情报科学的基本知识。本章从应用的角度对与计算机情报检索相关的一些基本知识作一介绍。

第一节 基本知识

计算机情报检索的有关基本知识是我们首先要掌握的内容。它涉及到计算机和情报学的一般知识,这里仅对与计算机情报检索直接相关的最基本的概念作一简介。

一、概况

计算机情报检索是继电子计算机问世之后发展起来的一项现代情报检索领域。30 多年以来,随着全球微电子技术和通讯事业的飞速发展,计算机情报检索事业亦随之高速发展,现已实现了情报检索的自动化和网络化。目前,全球计算机情报检索系统已达 300 多个,各种情报数据库已逾二千。例如:

西欧的 ESA—IRS(总部设在意大利的费拉斯卡蒂(Frascati))拥有 70 多个情报库,美国的 DIALOG(总部设在帕洛阿尔托(Palo Alto))拥有 300 多个情报库。计算机情报检索事业的发展为情报检索科学的进步奠定了强有力的基础。

在科技高速发展的今天,信息对科技工作者而言,就象水在人的生命活动中的作用一样重要。现代社会中的信息数量之大,亟需一种能够快速准确寻求对己有用信息的技术,计算机情报检索,正符合了这种需求。当今,在我们这个地球上,具有良好的网络化计算机情报检索手段。用之,我们就可以在世界的各个角落,借助通讯、电子设备,查找到任何地区的科技信息。尤其在欧美等发达国家,用计算机情报检索系统查找信息就象打电话一样方便。许多科学工作者在自己家里或工作室就可以随时将自己的计算机与世界各地的大型计算机情报检索系统相联接,进行各种信息资料的查询。

我国自 1983 年以来,陆续建成了许多国际网络化情报检索终端站/点,现已遍布除青海、西藏、新疆之外的 40 多个大中城市。国内亦已建成数十个情报数据库(如:中文科技期刊篇名数据库、中国专利公报、商情数据库等)。与此同时,国家已把通讯信息事业的发展列入国家重点发展规划,不久的将来,我国的网络化情报检索事业会有更大发展。

二、基本概念

1. 计算机情报检索

情报检索是指一种查找过程,即在汇集的情报信息源中找出自己所需情报的过程。要在众多散在的情报信息源中,有效地找到自己特定需要的情报信息,首先必须使得众多、散在、杂乱的情报信息源达到有序。这项工作是由情报人员完成的,他们用人们共知的情报检索语言作为标准,将情报信息整

理成有序的情报信息集合。这样,情报需求者就可以方便有效地从有序的情报信息集合中找出自己所需的情报信息资料。

情报检索语言是情报人员专门创制的,用来描述情报信息外表和内容特征以及情报查找需求概念的规范化符号体系。情报人员据此整理情报信息,使之达到有序;情报需求者据此在有序的情报信息集合中找到自己所需的情报信息资料。情报检索语言是沟通情报人员与情报用户的一种特殊人类语言。

计算机情报检索是指利用计算机技术设备,在汇集的情报数据库中找出自己所需情报资料的过程。其中,情报数据库是指存放由情报人员加工整理好的有序情报信息的仓库,这里所指的仓库是存放于计算机存储介质中的数据仓库。

2. 计算机情报检索系统

情报检索系统是指在一定检索设备支持下,由情报人员将整理好的有序情报信息存贮在某种载体上,供检索者方便查找的系统。

计算机情报检索系统是指借助于计算机技术设备,用计算机存储技术和手段,根据人们共知的情报检索语言,将各种情报信息源输入、存贮在计算机数据库中并使之有序,以会话方式提供情报检索的自动化查询系统。

第二节 计算机情报检索系统的基本构成

计算机情报检索系统由计算机硬件、软件和情报数据库三个部分组成。

一、硬件