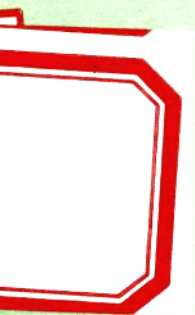


国外科技文献与检索工具

计算机检索 用户指南



中国国防科技信息中心

国外科技文献与检索工具简介之十三

计算机检索用户指南

吴成芳 编

中国国防科技信息中心

一九九一年十二月

前 言

本书介绍了我中心提供服务的五种国外文献数据库：WPI、GRA(NTIS)、INSPEC、MIL、GCK。内容包括这些数据库的专业范围，检索特点等，着重论述了各库的检索策略。书中例题均为真实题目，为了保密，隐去或修改了关键技术。

本书可供科技人员，情报人员，大学师生参考。

在编写过程中，参考了吕能扬同志有关检索经验和一些作者的研究成果(见参考文献目录)。张复华、张杏忠同志对全文作了审核，修改。在此，向他们一并致谢。

因水平有限，不妥之处，敬请读者批评指正。

吴成芳

1991.12

计算机检索用户指南

一、计算机检索系统发展概况及用户对策

(一) 计算机检索发展史

1946年，当世界上第一代计算机问世不久，美国就开始研究用计算机进行科技文献检索问题。首先把计算机用于情报、图书馆部门的是美国海军兵器中心。1954年该中心图书馆首先用IBM-701型计算机实现了情报检索。1958年，经过改进，用IBM-704型计算机，可以检索题目和文摘。到60年代，一些工业发达国家相继在情报、图书馆使用计算机。

1972年，著名的美国“DIALOG”系统建成并向全世界提供情报检索服务。把计算机检索推向“国际联机”情报检索新阶段。到1988年1月，“DIALOG”系统已发展到拥有300多个数据库，约360个文档，文献总量达1.5亿篇。其中包括商业经济类数据库约50个，化学化工类数据库近30个，医学数据库26个、其余数据库为基础科学，工程技术，人文科学等。“DIALOG”系统数据库平均以每年递增20%的速度增加

(二) 国内联机情报检索

1975年，当时的国防科委情报所和机械工业部情报所，先后在北京航空学院的FELIX256计算机上进行了我国最初的国外文献磁带计算机检索试验。1980年，北京文献服务处从美国引进UNIVAC 1100/10型计算机及检索软件UNIDAS。

此后，用引进的GRA(又称NTIS)磁带建成GRA库。这是我国第一个“实用型”计算机联机情报检索系统。它的建成投入运行，标志着我国计算机检索从试验进入实用阶段。

1986年，北京文献服务处联机情报检索系统完成了更新换代。主要标志是：①用UNIVAC 1100/70型计算机取代老机型。②用自行研制的BDSIRS联机情报检索软件代替进口的旧软件。该软件具有中西文兼容，检索功能强，响应速度快，输出格式多样，建库开销少，存贮效率高，造库及更新内容容易，使用灵活方便等特点。③数据库从一个增加到5个：GRA(又称NTIS)、(美国政府研究报告索引)，WPI(世界专利索引)，MIL(美国军用标准目录)GWQK(国外期刊文献库)，GCK(中国国防科技信息中心馆藏库)。此外，还建成《中国国防科技报告索引》、《战略文献索引》等中文数据库。文献总量达800多万篇，每年递增约60万篇。这些数据库检索功能和技术水平与国外同类型数据库相当。④用户检索终端从几台发展到50多台，分布在北京、上海、南京、长沙、西安、成都、洛阳、哈尔滨等约20个城市，成为我国目前最大的国内联机情报检索系统。该系统累计检索课题达4万个以上。

与此同时，其他单位如中国科技情报所、机电部机械情报所、中国专利局、化工部情报所、上海科技情报所等亦先后建成或即将建成国内联机情报检索系统。现在全国拥有大、中、小型计算机100多台，微机上千台，引进数据库约70种，累计文献量达3000万篇，已装入系统约1000万篇。

中国科技情报所联机情报检索系统采用美国IBM4381型计算机及CDS/ISIS联机情报检索软件。到1989年已建成

7 个数据库, 文献总量达220万篇。其中包括①INSPEC(英国科学文摘), 文献155万篇, 每年递增约16万篇。②COMPENDEX(美国工程索引), 86—88年, 35万篇, 每年递增约10万篇。③EI—Meeting, 来源于《工程索引》公司出版的EI—Meeting 磁带, 每年约2 000个会议, 8万篇论文。④西文馆藏库, 文献约3万篇, 年递增1.5万篇。⑤西文馆藏期刊库, 含1986~1988年西文期刊11 470种, 每年更新一次。⑥中国学报数据库1987~1988年, 文献量6 000篇。⑦学位论文库, 包括博士、硕士论文, 每年约7 000篇。此外, 该所还建有《中国重大科技成果库》、《中国适用技术成果库》、《星火计划适用技术信息库》、《中国学术会议论文库》等。

中国科技情报所通过多路转换器转接国内外通讯线路, 已实现与国内20多个城市的联机, 允许终端利用同一规程同时检索国际联机数据库和该所数据库。

机电部机械情报所联机情报检索系统, 采用加拿大《国际发展研究中心》(IDRC)HP-3000计算机, MINISIS 联机情报检索软件。该所从1985年到1987年9月用3年时间对MINISIS 软件进行二次开发成功, 使中西文兼容。从此, 该软件有英、法、西班牙、阿拉伯、中文5种版本, 能处理10种文字。由于该软件功能强, 在我国很快推广。先后采用该软件的单位有北京大学, 中国人口情报中心, 国家体委, 解放军总医院等约20个单位。

机械情报所联机情报检索系统建成数据库: ①INSPEC (英国科学文摘), 1974年开始, 约248万篇; ②COMPENDEX (美国工程索引), 1974年开始, 140万篇; ③METADEx

(美国金属文摘), 1974年开始, 54万篇; ④ISMEC (美国机械文摘), 1973年开始, 20万篇; ⑤EMA (美国工程材料文摘), 1988年开始, 3万篇。

目前, 机械情报所在武汉、合肥、西安、南昌、长春、南京等地建有10余家远程终端用户。通过直拨电话联通北京, 可以直接检索该所文献数据库。

上海地区科技情报检索网是我国第一个自行设计的分布式数据库联机网络, 通过数据分组交换把建立在不同机型计算机系统上的数据库联成一体, 实现数据库查询, 事务并发控制, 恢复和目录管理等分布式数据库的基本功能。目前上海及附近地区, 通过电话拨号, 利用普通的IBM PC终端和一套公共检索命令, 可访问网上各数据库。上海地区科技情报检索网的建成, 反映了我国情报检索联机网络最新发展。

参加上海地区科技情报检索网的主要单位有:

① 上海科技情报所, IBM4381计算机, 文献库有WPI (世界专利索引), GRA (美国政府研究报告索引), STDS (国际工业标准) 等。

② 华东化工学院, Burroughs B6935计算机, 文献库CA (美国化学文摘)。

③ 同济大学, Siemens 7536计算机, 文献库COMP-ENDEX (美国工程索引)。

④ 复旦大学, Micro VAX II计算机, 文献库INSPEC (英国科学文摘)。

⑤ 上海交通大学, Honeywell ZPS8计算机, 文献库INSPEC (英国科学文摘)。

⑥ 中国科学院上海测试计算中心, Burroughs B6935

计算机, 文献库BTA (生物技术文摘)。

(三) 国际联机情报检索

机械电子工业部 210 所于 1981 年 12 月与美国“DIALOG”系统第一个实现了国际联机情报检索。1982 年 9 月中国科技情报所与欧洲空间组织情报检索中心联机成功。此后, 国际联机情报检索在我国迅速发展。现在已发展到在北京、上海、南京、杭州、福州、厦门、广州、石家庄、武汉、长沙、南昌、宁波、嘉兴、济南、青岛、西安、成都、重庆、绵阳、包头、沈阳、长春、哈尔滨等 30 多个城市约 100 多个终端。通过国际卫星通讯线路, 我国与 10 多个大型国际情报检索系统实现了国际联机情报检索, 主要有美国的 DIALOG、BRS、ORBIT 系统; 欧洲航天局的 ESA-IRS 系统; 西德、日本、美国联合建立的 STN 系统; 西德的 FIZ-Technik 系统等。

(四) 光盘数据库

70 年代出现了光盘新技术。不久, 把光盘技术移植到情报存贮和检索获得成功。1980 年 飞利浦/索尼公司推出首片小型数字式光盘。1984 年末, 几种牌号的 CD-ROM 光盘驱动器样机问世。1985 年 5 月, 美国数字设备公司 (DEC) 抢先把 CD-ROM 作为 Micro VAX II 型计算机系统的组件, 当年掀起了 CD-ROM 驱动器和光盘数据库发行的第一个浪潮。

1985 年世界第一个光盘数据库 BIBLIOFILE (书目文档) 问世以来, 发展非常迅猛。到 1988 年底世界上已出现 CD-ROM 数据库 300 多种, 其发展速度大大高于磁带数据库初期。

1986 年 4 月, 国家海洋局科技情报所首次在我国引进 CD-ROM 数据库 ASFA (水科学和渔业文摘)。现在我国已

有近60个单位引进CD-ROM产品40多种,而且还继续增加。

MEDLINE (美国医学索引) 是我国引进最多的光盘数据库, 主要引进单位有: 北京图书馆、中国医科大学、北京肿瘤研究所、军事医学科学院情报所、中国医学科学院、上海第二医科大学、上海医学情报所、上海第二军医大学、西安第四军医大学、河北医学科学院、山东医学情报所、广东医学情报所、河南医学情报所、新疆地方病防治所等20多家。

GRA(NTIS) (美国政府研究报告索引) 也是引进较多的光盘数据库之一, 仅北京地区就有中国科技情报所、机械电子工业部210所、航空航天部628所等单位。

引进的其他光盘数据库有: 中国农业科学院情报所的AGRICOLA (农业文献索引); 北京大学、清华大学、复旦大学等约10所高等院校图书馆的BIBLIOFILE (书目文档); 北京师范大学图书馆的ERIC (教育文摘); 北京图书馆的LISA (图书情报文摘)。

除文献型数据库外, 中国科技情报所还引进了《联邦德国公司名录》、《美国托马斯公司名录》等事实型光盘数据库。

(五) 我国文献数据库建设

文献数据库建设是我国计算机情报检索事业发展最快, 见效明显的领域之一。到1989年, 全国已建造文献数据库近300种, 已投入使用的200多种。主要特点是: ①国外文献数据库利用加强。包括CD-ROM在内, 我国引进国外文献数据库约70种, 目前有 $\frac{1}{3}$ 可全套进行计算机检索。②计算机化

联合目录和馆藏目录有明显发展。影响较大的有：中国科学院文献情报中心的西文连续出版物联合目录，北京大学图书馆的西文图书联合目录，中国科技情报所的中文期刊联合目录，国防科技情报单位的西文期刊联合目录，北京图书馆的西文期刊联合目录等。大约20个部委情报所的计算机馆藏目录已初具规模。③自建文献数据库与文摘索引刊物出版相配套。我国书本型文摘索引约129种，其中《中国机械工程文摘》、《中国国防科技报告通报与索引》、《中国电子科技文摘》、《中国化学工程文摘》、《中国海洋文摘》、《中国航空文摘》、《通信科技文献索引》、《兵工文摘》、《中国地质文摘》、《中国核科技文摘》、《中国林业文摘》、《中国医药文摘》等已实现数据库化。④数据库应用领域不断扩大。事实型数据库发展迅速。例如中国科学院的计算机市场信息库。中国国际信托投资公司的世界金融、经济、贸易机构数据库。轻工业部的轻工产品国际贸易数据库。北京科技情报所牵头的全国科技成果交易信息数据库。机电部210所牵头的兵工数据库等。⑤国产数据库开始国际交流。据不完全统计，近10个国际或国外文献数据库与我国有关部门建立了合作关系。把我国相关专业文献摘要译成英语，纳入他们的数据库。例如中国专利文献已纳入英国德温特公司的世界专利索引(WPI)。中国大学学报论文英文版数据库将进入DIALOG系统。中国科技情报所的《中国企业公司》事实型数据库西文版1990年进入西德、日本、美国合营的STN系统。

(六)用户对情报检索系统选择对策

前面介绍了我国目前有三种计算机情报检索系统：①以中国科技情报所为代表的国际联机情报检索系统；②以北京

文献服务处为代表的国内联机情报检索系统；③以机电部210所为代表的光盘情报检索系统。

上述三种情报检索系统各有特色。国际联机情报检索系统有数据库种类多，情报内容更新快的优点，缺点是检索费用昂贵。光盘检索系统最大优点是检索费用低廉，不需要通讯线路，缺点是数据库单一，情报内容更新慢，只能单人检索。国内联机情报检索系统的特点介于国际联机和光盘检索两者之间。例如检索费用，国内联机约为30~100元/题，低于国际联机150~300元/题，高于光盘检索10~30元/题。情报内容更新情况，国内联机检索系统内容更新周期为一月左右；国际联机为一周左右；光盘检索系统为一季度至一年。数据库种类，国际联机系统有几百种，国内联机系统有几种至十几种，光盘检索系统每单位有一种或几种。

用户要根据各情报检索系统特点及对情报内容检索要求优化选择。选择检索系统基本原则是：“需求、保密、费用”6个字。例如，国内联机或光盘检索能满足课题检索要求，不必用国际联机检索。检索有关国防或国家安全的课题，应尽量避免国际联机，以防泄密。查商业经济最新国际情报，最好利用国际联机检索。对于文献检索，最好先在国内联机或光盘上检索，获得最佳检索策略，然后在国际联机上补查有关内容。这样作能达到降低检索费用和满足检索要求的目的。

很明显，三种情报检索系统各有优缺点。它们互相补充，取长补短，共同发展，以满足不同用户的要求，用户根据检索课题特点，在三种情报检索系统中优化选择。

二、数据库及其检索

计算机联机情报检索系统一般由四部分组成：①计算机硬件，包括主机和外部设备（终端操作显示设备，输出打印设备等）；②通讯设备；③计算机联机情报检索软件；④数据库，包括文献数据库（二次文献数据库和全文数据库）、事实数据库和数值数据库。

以上是计算机联机情报检索系统必要条件。为了情报检索，还应具备使用条件——操作指令和检索策略，对用户来说，只要熟悉有关数据库及其检索策略就可以实施情报检索。为此，本书着重介绍有关数据库特点及其检索策略。帮助用户熟悉和掌握计算机检索这一现代化手段。这是撰写本书主要目的。数据库种类很多，本书只介绍用户常用的五种西文文献数据库。它们是：WPI(世界专利索引)；GRA(N TIS)(美国政府研究报告索引)；INSPEC(英国科学文摘)；MIL(美国军用标准)；GCK(中国国防科技信息中心馆藏库)。

(一) WPI数据库及其检索

WPI(World Patents Index)《世界专利索引》数据库是引进英国德温特公司磁带建成的。德温特公司创建于1951年，已有40年的历史。它是国际上公认的收集、整理、发行国际专利文献(专利说明书)的权威组织，具有体系完整，专业面广，数量大，报导速度快，多途径检索，全部采用英文报导等特点。它同时出版发行书本型和磁带两种专利文献目录索引。

WPI(世界专利索引)包括美、苏、英、法、德、日、中在内的30个国家，两个国际专利组织，两种期刊发表的专利文献。主要内容包括农业、轻工业、医学、纺织、造纸、

运输、化学、冶金、建筑、采矿、机械、物理、电学、通讯、仪表、光学、摄影、能源等各技术领域，小到钮扣、别针大到飞机、火箭内容极为广泛。WPI是可以同时检索多国专利文献的有效工具。

目前我国已建成WPI库的单位有：北京文献服务处，中国专利局，上海科技情报所等。北京文献服务处的WPI文献库录入1964~1990年文献总计约500万篇。按年代分为4个文档：1*文档1964~1980年；2*文档1981~1983年；3*文档1984~1987年；4*文档1988~1990年。其中1*文档只有题录，没有文摘，其余文档都有文摘。

WPI库检索项或检索入口主要有：国际专利分类号、德温特分类号、规范化单元主题词、专利权所有者（包括公司代码）、专利发明人、专利号等。本书只介绍最常用的检索项：国际专利分类号和规范化单元主题词。

1. 国际专利分类号检索

国际专利分类号(以下简称分类号)既是手工检索世界专利最重要手段，也是计算机检索WPI库重要入口。分类号检索属范畴检索或族性检索，主要特点是，查全率高，不易漏检；缺点是查准分类号较困难。下面介绍三种查找分类号常用方法。

〈1〉用国际专利分类表查分类号

用国际专利分类表查分类号是最基本最常用的方法。分类号结构分5级：部、大类、小类、主组、小组。分类表按专业内容分为8个部：A-人类生活必须品(农、医、轻工)；B-作业、运输；C-化学、冶金；D-纺织、造纸；E-固定建筑物(建筑、采矿)；F-机械工程；G-物理；H-电学。

用分类表查分类号的基本方法是：从上至下或者说从大到小逐级查找，即所谓递推法或逐级查找法。查找顺序是：首先查部，再查大类，然后查小类，最后查主组和小组。例如查《火箭筒击发机构》课题的分类号：①F-机械工程；②F41-武器、爆破；③F41F-火炮、枪、枪架或其所用摇架，导弹发射架；④F41F3/00-自身推进的导弹发射装置；⑤F41F3/045-适于人员携带和使用的火箭或鱼雷发射装置，如火箭筒。F41F3/045即为检索上述课题用分类号。

〈2〉用《国际专利分类表技术用语索引》(简称索引)查分类号。

这本索引是使用《国际专利分类表》的辅助工具。当检索课题分类号用递推法无法查到时，借助索引的帮助，可顺利查到分类号。《索引》以类目名称首字汉语拼音字顺排列。使用方法是：①对检索课题首先进行主题分析，找出关键词；②像查汉语词典一样，查关键词在索引中的位置及对应的分类号；③把查得的分类号与《国际专利分类表》对照、核实、修正。这样就查到了切题的分类号。

例如检索《聚合物浸渍混凝土》的课题。该题关键词是“混凝土”。按混(Hun)字在索引中字顺位置，查到混凝土对应类目为C04B，再用分类表按递推法查到切题的分类号C04B15/02-含高分子化合物的有机粘结剂组成的砂浆、混凝土、人造石或其他类似的材料。

〈3〉用相关专利分类号为检索用分类号

众所周知，密切相关专利文献分类号一定相同。世界各国专利文献按规定都标有国际专利分类号。据此，可利用已知相关专利文献分类号进行检索。那么第一篇相关专利文献

怎样获得呢？基本作法是：用主题词检索WPI库，查得一篇或几篇相关专利文献。查看这些相关专利文献标记的分类号。这些分类号就是课题检索用分类号。

例如，检索《喷气发动机灭火装置》课题。先用主题词在WPI库中查到一篇西德专利《喷气发动机用于煤矿灭火》。这篇专利分类号是A62C3/02。用这个分类号再检索，又查到一篇苏联专利《涡轮喷气发动机灭火器》。这样经多次反复检索，最后一定能查全有关文献。

〈4〉用国际专利分类号检索注意事项

① 应熟悉国际专利分类表结构和特点，理解分类号代表的物理概念。国际专利分类表结构属层垒和数字混合编制。前4级用层垒制，第5级各类目间是族性制。各类目间从属关系不是以数字大小来区别，而是以分类号前面园点多少、来区别。园点表示：第一，小组细分时，上下位类之间从属关系。第二，用园点代替高一级小组类名，避免类名重复。这对正确理解各小组类号物理意义是很重要的。

② 分类表每五年修订一次。每次修订会随着科学技术发展而作相应改动。所以，要针对欲查文献年代范围，选相应分类表版本。

③ 选分类号时，要明确各上下位类包括的文献范围。或者说，哪些文献内容应放入上位类，哪些应放入下位类。上位类包括的文献，不等于各下位类文献总和。

④ 分类表以学科应用和功能两个体系为基础编制的。分类法对主题分类，有功能分类和应用分类的区别。分类规则规定，某发明技术主题，既可按功能分类，又可按应用分类。本表以功能分类优先。由于发明人、专利文献标引员思

维方法不一定和检索用户思维方法完全吻合。所以，用户选分类号时，既从功能，又从应用以及其他方面查全分类号，以防漏检。

为了查全，防止漏检，经常采用“过量”分类号检索。换句话说，每个课题不是选一个分类号，而是选多个分类号检索。

例如，《飞机应急救生用化学氧气发生器》，选4个分类号检索。

① A 62 B 7/08 (含有产生氧的化学品呼吸保护装置)。

② A 62 B 7/14 (高空飞机用呼吸保护装置)。

③ A 62 B 21/00 (呼吸保护装置用化学物质产生氧的设备)。

④ C 01 B 13/02 (氧的制备)

又如，《轻骨料旋转式机枪防御工事》，选6个分类号检索。

① E 04 H 9/10 (单独的掩蔽部，单独的防弹墙)

② E 04 H 9/04 (防空袭或其他战争行动的建筑群或掩蔽部)

③ F 41 H 5/08 (人员使用的防御物)

④ F 41 H 5/12 (适用于轻武器的防御物)

⑤ F 41 H 5/18 (旋转式防御物)

⑥ C 04 B 14/16 (多孔颗粒材料增强的砂浆、混凝土、人造石)

2. 用主题词检索

用主题词在 WPI 数据库中检索文献，必须使用《世界专利索引——规范化主题词表》。该表有以下特点：①主题词

由单元词构成，不像其他主题词表如INSPEC、NTIS等主题词表大量使用多元词。②对主题词进行了严格规范化处理。或者说，对一个词的名词单复数，动词及其分词，形容词，副词都规范成一个词作为正式主题词，用于实际检索。其他形式的词作为非正式主题词，不得用于检索。例如：

DETECT (15 491)

Detectability

Detectable

Detected

Detecting

Detection

Detections

Detective

Detector

Detectors

Detects

其中只有第一个用大写字母表示的DETECT是正式主题词，可用于检索。其余词都是非正式主题词，不得用于检索。数字15 491表示DETECT词的文献频率。这种规范化处理，对提高文献查全率效果明显。这是该词表最大优点。

WPI库除1*文档用词不严格外，其余文档(2*-4*文档)用词均很严格。换句话说，除1*文档外，其余文档不得使用非正式主题词或自由词检索。这是WPI库最大检索特点。

主题词检索主要优点是：①查词简便、迅速、方便；②用户容易掌握；③查准率高。缺点是：若选词不全，易漏检。