

发挥 CDS/ISIS 软件功能 用好《中国石油文献》数据库

乔和建* 相秀芳

(四川石油管理局科技处)

摘 要 介绍了利用 CDS/ISIS 2.3 版软件对《中国石油文献》数据库建立单汉字索引,进行单汉字检索;发挥该软件的字段级无倒排检索功能并采用二次检索的方式进行全字段查找;以及利用该软件的 Pascal 高级语言接口开发编程,进行数据加工和文献统计等方面的情况。通过分析上述数据库的有关统计结果,提出了收录文献时既要重点摘录学术性强、理论水平高的论文,也要兼顾实用性强、篇幅短的文章;采用多库互补检索,提高查全率的建议。

主题词 CDS/ISIS 软件 文献检索 单汉字全文检索 文献统计 中国石油文献 数据库

数据库是知识的成果,是控制和处理信息的现代化手段。而机读数据是现代化信息社会中组织与交流信息的主要工具,是构成计算机信息检索系统的重要数据源。随着国民经济建设对石油资源需求的不断增长,加之石油科技的迅速发展,使得人们对文献的需求更加广泛且要求获得的信息更准、更新。《中国石油文献》数据库的开发和建立正好为人

们提供了一种交流和获取信息的现代化检索手段。《中国石油文摘》是国内石油行业的权威性检索刊物,该数据库的建立,实现了对中文石油文献的集中加工处理和建库工作。随着微机的普及,中国石油天然气总公司信息所又将数据库移植在微机上,利用 CDS/ISIS (以下简称 CDS 软件) 软件开发了微机检索系统。Micro CDS/ISIS 软件是一种功能很强

⑦扩大二级单位使用固定资产承包交纳占用费方案。为加强企业使用固定资产的经济责任,提高使用效益,1993 年局对六大机厂、三大运输公司在“八五”后三年采取按国有资产帐面余额(净值)实行资产交纳占用费承包,分三年逐步到位方案,效果是好的。各个承包单位都有盈利,新增国有资产规模得到自我控制。

⑧明晰产权关系,对分离出来的专业性公司或全民所有制面向社会服务的多种经营实体,以及生活服务公司、劳动服务公司,凡使用国有资产,都应当按银行的同期利率交纳占用费,在享有权益的同时,还要承担国有资产在使用过程中的义务。

⑨对申报“封存”和申请“启封”的资产,应由专业部门提出意见,归口由国有资产管理部门审批。对批准封存的资产,可视为未使用的资产,停止计提折旧。对同意启封的资产,恢复折旧。

⑩加强对“热点”国有资产管理,如:凡使用生产指挥车的单位和使用无绳电话的单位或个人,要分别明细登记,要单独按银行的同期利率向所在单位的国有资产管理部交纳占用费。

⑪加强对已报废资产的管理。利用已报废资产要建立利用台帐。处理残值只能统一由国有资产管理部门进行。

⑫对不需资产应首先在我局范围内平衡后积极设法进行对外价让。对外价让资产,只能由国有资产管理部门按规定进行。

⑬尽快在局内建立内部租赁市场,建议挂靠在国有资产管理处。依托二级单位的资产库及资产,试行工作。租赁公司在局内应形成一个信息网络体系兼执行体系。

(本文收稿 1994—01—28 编辑 赵 勤)

* 乔和建,1954 年生,助理馆员;1991 年毕业于四川电大图书馆专业;从事图书馆现代管理工作及研究,与人合编有《汉英石油技术词汇》一书,擅长对国内外数据库的加工处理。地址:(610051)成都市府青路一段 3 号。电话:(028) 3324911 转 211977。

的信息检索软件, 该软件的检索功能尤其突出, 这已是国内上千家用户的共识。本文就如何发挥 CDS 软件功能, 用好《中国石油文献》数据库谈谈我们的肤浅体会, 以求同仁指教。

1. 增设检索入口, 拓宽检索途径

(1) 建立题名单汉字检索入口

我国历来重视题名检索点, 就连过去习惯以责任者作为主要检索点的西方国家, 现在也都越来越重视题名检索点了。文献题名本身属于自然标识的范畴。按数据库索引结构划分, 题名字段又属于基本索引, 即主索引之一。一般来讲, 从信息检索角度而言, 利用题名字词进行检索其查准率较高。由于该检索方法完全从作者用词与用户用词出发, 较好地保证了每个标引词必为有检索价值的词汇。据《数据库通讯》1993 年第 3 期报导,《中文期刊篇名库》(简称《中篇库》) 已入库的 95 万条数据中, 每篇文献的作者一般都是通过题名来表达文献中心内容, 可以说 85% 左右的科技文献题名非常符合正文所反映的内容。对题名不能完整表达正文内容的约 15% 文献, 显而易见, 增设题名检索入口对主题词检索是一种有益的补充; 尤其是采用单汉字检索方式, 其检索效果更为突出。

首先, 单汉字检索是把中文数据库中的信息, 用单个汉字作为标引和检索的基本单元而建立起来的一种检索方法。换言之, 单汉字检索是以单个汉字为单位的“主题词”。采用单汉字模式建立的检索数据库, 不同于以词为单位进行标引和检索的数据库, 在检索方式、方法上有其独到之处。就单汉字而言, 汉字字典的体积最小 (在计算机中标准化的汉字字库中, 一级和二级国标汉字共计 6 763 个。从这点上分析得知, 汉字关键词索引文件比西文关键词索引要小得多); 就其自动切分方法而言, 汉字单字的切分方法最为简单 (以双字节字符自动抽取), 最接近自然语言。尤其是单汉字在存储处理新词和处理任意汉字中的天然单元时, 具有其他方法不可比拟的优点。

第二, 单汉字检索系统是最彻底的后组式索引模式, 它适用于一切专业数据库。检索时不用查找词表, 组配灵活, 检索速度快, 尤其能实现检索词在左、中、右截词的全部功能。而一般的检索系统虽都能进行右截词检索, 但左截词检索实现较困难。对检索中文信息来讲, 左截词有相对重要的意义。因为汉语的构成词法具有“中心词后移”的现象, 即把各种修饰性成分放在前面, 实质性的部分放在后面。例如: “陆相盆地” “含油气盆地” “定向钻井”

“丛式钻井”等, 最关键的部分是“盆地”和“钻井”, 如要全面检索“盆地”和“钻井”方面的文献, 使用左截词检索效果最佳。另一方面, 利用单汉字检索还可以任意调整检索的专指度和泛指度。就专指度而言, 检索时组配的汉字越多, 就越能达到专指的水平。它不受控制词汇系统的限制, 也不受新词滞后的影响。尤其对一些缩写、约定俗成词语等, 用一般的检索方法是无法查全的, 而这正是单汉字检索的突出特点, 同时也符合用户的使用习惯。例如: 我们检索“双燃料发动机”课题时, 由于库中无“双燃料”主题词选择, 改用单汉字检索在题名中查到。就泛指度而言, 单汉字检索具有极强的字面成族能力, 这又是单汉字所特有的突破性功能。例如: 在数据库题名中以“盆”字检索, 则可以“带”出国内、国外的各种盆地以及盆地特征等方面的信息。笔者用“盆”字共检中 1 959 条记录, 而采用主题词“盆地+右截词符”只检出 479 条记录。在 1 959 条中除极少数记录中的“盆”字不太切题外 (如: “盆山省”等), 其余的 1 400 多条题名中的“盆”字均属于中心词后移情况。另外, 用 CDS 软件做右截词检索也有一些缺憾, 不可避免的要带出一些不相关的检索词。如用主题词“采油\$” “海洋\$” “石油\$”和“钻井\$”等检索, 尽管可以通过字段操作限定符限定在某一字段范围内命中, 但检索查找过程中象《采油工艺》《海洋地质》《石油物探》和《钻井工艺》等期刊名称也被当作查找对象一起带出, 逐个被显示在屏幕上, 只是最终不作为命中记录处理罢了。无疑这将大大降低检索速度, 尤其用分类号+截词符检索时, 检索速度更慢, 这一点 CDS 用户的体会都颇深。

第三, 在单汉字检索中笔者发现, 编制检索提问式时, 可以采用特征度大的字或者称为辨识度大、网络度小的字进行组配检索。这样可以缩短检索时间, 提高查准率。这也是单汉字检索有效性的一个体现。所谓特征度大的字, 就是指出现频率少、分辨力强的字。例如: “油”字, 在《中国石油文献》库中分辨力很低, 而在图书专业数据中却很高; 反之“图”字在《中国石油文献》库中分辨力高, 在图书专业库中却很低。就字的分辨力大小来说, 它是与数据库的具体内容主题相联系的, 是相对而言的。

第四, 在单汉字检索操作中, 根据 CDS 软件提供的资料信息, 在编制检索提问式中该软件提供了两种检索运算符及其他一些限定符号。一种是布尔逻辑算符“或”“与”“非”, 另一种是位置逻辑算符

“.”“\$”等。这些运算符比“逻辑与”算符有更大的限制。另据 CDS/ISIS 资料介绍,算符的优先级最高的是位置逻辑算符,其次是布尔逻辑算符。在单汉字检索实践中,我们主要采用了位置逻辑算符,理由是:利用布尔逻辑算符对检索词进行组配检索时,由于无法限定检索词之间的位置关系,易造成误检,严重影响查准率。例如检索式:“油*气*田”,检到的题名是:

①对高升油田热采试验区蒸汽吞吐试验的初步认识

②克拉玛依油田的非背斜油气藏

而采用位置逻辑算符检索正好弥补了布尔逻辑算符的这一缺憾。

我们利用单汉字这一突出特点,用 CDS2.3 版软件将《中国石油文献》库中的题名字段(包括正题名和副题名)处理成单汉字索引方式(原该字段未作为检索点),作为新增的检索途径。利用单汉字在题名中检索某方面的“新型”“新技术”“新工艺”“新理论”“新进展”“新发现”等文献,非常容易实现。这对于了解石油天然气行业的动态和进展信息是很有帮助的;而用主题词检索效果很差。又如利用单汉字方式还可以从题名中检索出“产品型号”“以英文命名的方法名称”“英文缩写符号”“英文油气田名称”“英文会议名称”“地名”和“特种文献号”等等;而用主题词检索虽能检到一些,但结果很不理想。

采用单汉字检索优点较多,但也并非尽善尽美。如单汉字揭示隐含主题功能欠缺,单汉字倒排文件占空比较大等。但是随着计算机外存容量的增大和光盘存贮系统的出现等,不惜以多开销空间来换取检索上的快速度这一做法已被人们所接受。因此,在检索方面,将主题词与单汉字相结合,采用布尔算符和位置算符将多个检索词组配起来反映多维内在关系,提高检索深度和广度,仍不失为一种行之有效的办法。

(2) 完善文献类型检索入口

从《中国石油文献》数据库收录的文献品种来看,除了公开和内部发行的科技期刊和科技图书外,还有一些特种文献资料,包括会议文献、科研成果报告、学术论文、考察报告、专利文献和技术标准等。从文献类型字段的数据分析得知,该字段的数据内容视文献类型的不同而分为文献类型标识、译文标识和文种标识。如:“刊”“会译英”等。原该库对这一字段建立索引时采取整字段抽词技术,如文献类型为“中译文”著录用双标法时,那么用整

字段抽词技术则无法将三种标识分别抽出。通过实践,笔者利用 CDS2.3 版软件提供的抽取字段或子字段中部分内容的命令符,对这三种标识进行了分别索引。检索证明,这种索引方式是可行的。完善这一检索入口可以对所检索的文献内容进行类型范围限制,任意选择文献类型、译文类型和文种,以便尽可能的就地索取到一次文献。

2. 采用无倒排检索和二次检索方式,实现全字段查找

使用 CDS 软件进行快速检索,通常是通过建立倒排索引来实现的。就一个文献数据库而言,它是若干字段的集合,建立倒排索引时既要考虑实用性,又要考虑占空比。一般应抽取文献中具有检索意义的描述性内容。如分类号、题名、作者、主题词字段等等,这样其它的字段将无法检索到。CDS2.3 版软件提供了一种备用检索方法,即字段级无倒排全文检索方法。使用它对那些没有建立倒排的字段也能实现检索查找(尽管检索速度要慢一些)。可以讲,CDS2.3 版新增这一检索功能为该软件锦上添花,同时也减少了倒排文件的占空比。字段级无倒排检索式由无倒排检索识别符,字段助记符和字段、子字段标识符,包含运算符、单引号、字串、单引号等构成。如: ? V35: '钻柱稳定技术'。在编制无倒排检索提问式的实践中笔者发现,除了可以使用上述常规的检索式外,在编制无倒排检索提问式中还可以使用布尔算符、添加数值函数、字符串函数、布尔函数和抽取字段部分内容命令符等,组配成各种检索提问式,进一步增强和拓宽检索功能,达到全方位查找之目的。下面例举几个无倒排检索提问式:

? S (V5, V6, V34, V35): '深源天然气'

? VAL (V51 * 2.4) = 1435 AND VAL (V51 * 2.4) <= 3188

? S (V35): 'IFP 法' OR S (V35): 'Fe⁺⁺⁺'

? NOT S (V5, V6): '倾角测井'

? NOT P (V7) AND A (V35)

? S (V5, V6, V34, V35): '天然气' AND S (V5, V6, V34, V35): '流量计'

在以上检索提问式中分别用到了 S (字符串函数), VAL (数字函数), A、P (布尔函数), *, (抽取字段部分命令符, AND、OR、NOT (布尔算符) 等。这里有两个值得一提的函数: 一个是 VAL 函数, 另一个是 S 函数。VAL 函数是把字符型数字转换成数值型数字, 这解决了检索数据库中数值类型的问题; S 函数是字符串函数, 它隐含着布尔算符“或”运算功能, 比“或”运算速度快得多, 且检索

提问式简短。例：

①? V5:/'气田'/OR V34:/'气田'/OR V35:/'气田'/

②? S (V5, V34, V35):/'气田'/

①与②相比较,查找时间都为 39 s,前者查中 1 322 条记录,后者查中 1 445 条记录。

二次检索。二次检索亦称第二级检索。二次检索是在一次检索(即用倒排文档直接检索)的基础上,再对其检中的结果用顺序扫描的方法直接查找文中的任意字串或单字,实现文中任何字的检索

功能。采用二次检索机制可以大大提高检索速度,最大优点是节省盘空间。二次检索时,先进行布尔检索或位置检索,将一次检索命中的检索式号再与无倒排检索提问式组配起来进行二次检索。例如,一位读者需索取 SPE 论文的一次文献,共计 15 篇。虽然 SPE 原文我们都有,但我们先查《中国石油文献》库看有无译文,根据读者提供的 SPE 文献号,利用二次检索获得结果。先通过题名检索出 SPE,即一次检索式为 SPE,二次检索提问式为:

序号 数据库名 命中数 检索提问式

1	CPA	511	SPE
2	CPA	4	? #1 VAL (V5) =12558 OR VAL (V5) =14249 OR VAL (V5) =14795 OR VAL (V5) =17616 OR VAL (V5) =17924 OR VAL (V5) =18033 OR VAL (V5) =18055 OR VAL (V5) =19945 OR VAL (V5) =19553 OR VAL (V5) =19956 OR VAL (V5) =21613 OR VAL (V5) =22355 OR VAL (V5) =22533 OR VAL (V5) =223438

经过查找检出 4 篇,即 SPE17616、SPE18633、I-ADC/SPE 19945 和 SPE 22355。另一方面,利用无倒排检索方法和二次检索机制,尤其对检索文摘字段特别有用。就文摘而言,文摘能同时起到报导和检索的双重作用,能充分满足用户的需要。从信息角度而言,文摘较之主题词、分类号来讲,对信息内容有更深入地揭示;从检索角度而言,文摘的引得深度大,检索途径多而丰富。除了能从文摘中检索到词、词组外,还可以从文摘中检索出各种缩写符号、公司名称、地层名称、计算公式、分子式、温度数值、深度数值、重量数值和长度数值等等。而这些正好是叙词、关键词不反映的或者说是难以表达的。

3. 利用 CDS/ISIS Pascal 高级语言接口开发编程,对数据进行加工处理和统计

CDS/ISIS Pascal 是一种基于通用信息管理系统 Micro CDS/ISIS 的程序设计语言,是为了进一步开发 CDS 软件而设计的。CDS2.3 版增设高级语言接口,为该软件增添了活力,使用户开发自己所需要的程序成为可能。

利用数据库进行文献检索服务是实现数据库价值的关键所在,但不是全部。笔者认为,可以充分利用数据库开展一些文献计量学统计方面的工作,对文献进行更深层次的计量分析和加工,这样可以大大拓宽数据库的应用范围,增加使用价值。基于

此目的,利用 Pascal 语言接口开发编程,对《中国石油文献》库中的多责任者(包括著作字段、译者字段和发明人字段)记录、组配分类号记录、核心期刊记录和一些录错了的数进行了加工处理,并标注了检索标识符号,基本满足了快速检索和统计的需求。首先,据统计,第一著者发表文章在 10 篇以上的作者有 34 人,其中陈元千单独发表文章 162 篇,与他人合著 3 篇;译者翻译文章在 20 篇以上的有 12 人,其中陈文华 113 篇。从这点分析得知,这些作者是石油行业的核心作者群。掌握他们目前的研究选题,对于寻找各相应学科的研究热点、预测近期的研究发展趋势等,无疑是有很大帮助的。第二,该库中多责任者共计 9 227 条记录,占库总数的 30%左右。从中可以看出,随着现代科学不断向纵深发展,科研的难度也在不断加大,许多研究项目如仅靠一人独自完成,不仅耗时长,且困难很多。加强合作研究,充分发挥集体的智慧和力量,必将促进科研进程,有利于加快我国石油科技发展的步伐。第三,该库中有 6 026 篇论文的类目涉及两个类目,占库总数的 20%左右。我们发现对这部分记录的分类标引采用组配方法,能较好地揭示复合主题概念,能揭示不断出现的新概念和专深化的内容。掌握这一情况检索时采用“与”算符可以提高查准率。第四,统计表明各核心期刊共计有 14 784 条记录,约占库总数的 48%,占期刊类型总数 21 776 条的

67%左右。由此可知,有关石油科学的刊物论文也遵循了大量论文相对地集中在一定数量的刊物上,而剩余部分则依次分散在其它大量的相关刊物上这一普遍规律,即符合布拉德福定律。笔者认为,了解核心期刊所占比例情况,可以指导我国石油科技人员重点查阅文献,提高专业信息利用率,同时也为各油田信息、图书部门搞好中文石油专业刊物的采访、收藏工作提供了客观依据。第五,就文献类型而言,该库摘录的科技期刊论文数量最多,其次是特种文献,科技图书的比例最少(表1)。

表1 《中国石油文献》库文献类型统计表

序号	文献类型	记录数量/条
1	期刊论文	21 776
2	汇编文献	3 882
3	会议论文	1 910
4	专利文献	1 549
5	科研成果报告和考察报告	517
6	科技图书	369
7	技术标准	165

在统计中笔者还得知,数据库中摘录了相当数量的译文文章。据统计共摘录译文8 412条,占库总数的27.8%。所摘录译文中包括的语种有:英文6 786条、俄文1 202条,日文302条、法文19条、德文13条等。由此可知,在吸收国外石油工业最新研究成果和信息时,文献来源主要是英文,其次是俄文、日文、法文和德文等。译文作为一种重要的文献源,扫除了阅读语言障碍问题,在检索时可以利用译文标识来限制检索,这样可以了解和掌握原文被翻译的情况,避免重复劳动工作。第六,从使用引用刊物的数量方面看,《中国石油文献》库从1986年的197种刊物增加到1992年的332种,7年间共增加145种刊物。这说明该库每年都在不断扩大收录石油专业的新刊和相关刊物的品种数量(表2)。

从表2中不难看出,引用刊物的品种数量呈上升趋势,而实际引用刊物数量则逐年减少,最少的1992年实际引用刊物数量仅占该年引用刊物表数量的30%左右。笔者认为数据库建设应注重“全”,不全也就无所谓新。在收录文献时既要重点摘录学

表2 《中国石油文献》库各年引用刊物数量统计表

年份	引用刊物表数量种	实际引用刊物数量种	未用到的刊物数量种	占引用刊物表数量的百分比%
1986	197	218	+21	110
1987	233	183	50	79
1988	257	181	76	70
1989	272	169	103	62
1991	309	124	185	40
1991	315	137	128	59
1992	332	101	231	30

术性强、理论水平高的文章,同时也要兼顾实用性、篇幅短的文章。如笔者在为四川大学检索“油罐液位自动检测”课题时,除了在《中国石油文献》库查到几篇外,还利用《中篇库》检索到20多篇文章。从《中篇库》检出的信息中发现,有的文章刊登在《光纤光缆传输技术》《军用航油(国外版)》《国际经贸探索》《江苏工学院学报》《核工业自动化》《山东工业大学学报》和《电工技术》等刊物上。在检索中笔者体会到由于现代科学的交叉渗透和文献的离散分布性,很难讲哪一个数据库能够全部包罗。为了提高查全率,笔者认为采用多库互补检索,不失为一种有益的补充。

尽管如此,瑕不掩瑜,《中国石油文献》数据库仍得到了广大用户的好评。我们相信,通过大家的共同努力,并不断探索、实践、总结、提高,一定会将CDS软件和《中国石油文献》数据库应用得更好。

参 考 文 献

- 1 陈光祚. 论单汉字检索系统. 情报学报, 1992; 11 (1): 11~18
- 2 刘继昌. 全标引汉字情报检索系统研究. 情报科学, 1992; 13 (1): 19~22
- 3 陈光祚, 王军. 利用全文检索技术制作电子出版物. 情报学报, 1993; 12 (1): 1~7
- 4 中国科技情报所重庆分所维普资讯公司. 《中刊篇名库》题录加文摘. 数据库通讯, 1993; (3)
- 5 曹希平. 数据库建设要注重功能的开发. 情报理论与实践, 1992; (1): 40~41

(本文收稿 1993—11—08 编辑 居维清)

type of tipburn reborn. The material diluted with steam, the noble metal Pt on globular alumina-zinc spinel carrier is used as a catalyst. The single progress convert ratio of material is 50%~55% (mol). The selectivity of products is 94.5%~98% (mol), a suite of installation ran in Oct. 1992, two suite of installations are designing and construction and the contract of three suite of installations were signed. The FBD-4 technique adopts the ebullience fluidized bed reactor-regenerator system. Reacting and regenerating process ran in two independent vessels, the fuel supplement is supplied by regenerating vessel, and the reaction needed of heat which is supplied by sensible heat burning deposit carbon overlay the catalyst to heat cycling particle of catalyst, the oxide chromium and auxiliary agent on globular Al_2O_3 carrier is used as a catalyst. The single progress convert ratio of isobutane is 50% (mol), the selectivity of isobutane is 91% (mol), and the old technique of thirty-seven suite of installations ran in 60' s. The improved new technique of the first suite of installation has industrialized. In addition, six suite of installations are designing and construction.

SUBJECT HEADINGS: alkane, catalyzed dehydrogenation, catalyst.

Guo Sheng-qiu (*National Property Managing Department of Sichuan Petroleum Administration*); DISCUSSION ON STRENGTHENING NATIONAL PROPERTY MANAGEMENT, NGI 14(4), 1994: 77~79

ABSTRACT: Stressing national property management is an important work which penetratingly implements document spirit of strengthening national property management, brings national property benefit into full play to keep a-boriginal value or rise value, and builds up a system of socialist market economy. Apart from strengthening institutional framework, establishing and perfecting rules and regulations, strengthening national property management should lay stress on reforms, try the share system; Meanwhile, national property will play a full role in national economic construction by doing well the research job and performance of national property management ways with economic profit as center, changing ideas, improving business level of competence, bringing organization effect into play, strengthening effect analysis and expanding lease contract system.

SUBJECT HEADINGS: national property, management, present situation, suggestion.

Guo Sheng-qiu, engineer, graduated in mineral mechanics from Southwest Petroleum Institute. Now he is a vice section chief of National Property Managing Department of Sichuan Petroleum Administration. Add: (610051) No. 3, Sec. 1, Fuqing Rd, Chengdu, Sichuan.

Qiao He-jian (*Science-Technology Department of Sichuan Petroleum Administration*), Xiang Xiu-fang; GIVING REIN TO CDS/ISIS FUNCTION AND MAKING GOOD USE OF THE DATA BASE OF CHINA PETROLEUM DOCUMENTS, NGI 14(4), 1994: 79~83

ABSTRACT: The conditions such as building up single Chinese character index of China Petroleum Documents to retrieve by using CDS/ISIS 2.3 version software, giving rein to the software's function of free text searching retrieving by secondary level searching, developing and compiling programs by using the software's Pascal advanced programming services, processing data and document statistics etc are introduced. By analysing the above statistical result, some suggestions that not only the papers being good academic contents and high theory level are laid stress on extracting, but also those being good application are given consideration are proposed.

SUBJECT HEADINGS: CDS/ISIS, document retrieval, single Chinese character, free text searching, document statistics, China Petroleum Documents, data base.

Qiao He-jian, assistant librarian, graduated in library speciality from Sichuan TV University in 1991; He has written a book titled English-Chinese Petroleum Technology Vocabulary, and he specializes in processing the data bases at home and abroad. Add: (610051) No. 3, Sec. 1, Fuqing Rd, Chengdu, Sichuan. Tel: (028) 3324911—211977.