

· 专业教育 ·

将 Dialog 联机检索系统引入文献检索课程

程爱平 战玉华 钱俊雯 黄晓玲

(清华大学图书馆 北京 100084)

文 摘 介绍了 Dialog 联机检索系统的信息资源分类、数据库扫描以及逻辑算符、通配符、位置算符等指令的使用技巧,并将其引入文献检索课程中,以帮助学生构建合理的检索策略、根据 Dialog 的扫描结果充分利用馆藏资源或免费网上资源获取相关的文献信息,提高学生的信息素质。

关键词 文献检索课 数据库 教学 Dialog 联机检索

Introduce the Dialog Searching System to Literature Retrieval Course

Cheng Aiping Zhan Yuhua Qian Junwen Huang Xiaoling

(Tsinghua University Library, Beijing 100084)

Abstract: In order to improve students' information literacy, this paper introduced the database features, scanning process and the use of Boolean logic operators, wildcard characters and proximity operators in Dialog system, and introduced the Dialog into literature retrieval course to help students construct suitable searching strategies and acquire literatures from library resources or free web-based document resources according to Dialog scanning results.

Key words: Literature retrieval course, Database, Teaching, Dialog

为提高学生的信息素质,很多高校都开设了文献检索课。图书馆拥有大量电子资源,网上也有许多免费资源,但 Dialog 作为世界最大的国际联机检索系统,具有数据库数量多且可以进行多库扫描及检索等诸多优点,将其引入文献检索课程中,可以更好地提高学生的信息能力。本文通过介绍 Dialog 联机检索系统的特点及检索技巧,说明将其引入文献检索课程的意义及实用性。

1 文献检索课程的目标及引入 Dialog 的意义

当前,互联网和搜索引擎已经高度普及,但对学来说,仍然存在着或多或少的信息获取障碍。而且随着图书馆电子资源数量的大幅度增长,如何在众多电子资源中选择适合自己课题的数据库等信息资源,对学生来说也有一定的困难。

为帮助广大读者很好地利用各类资源为学习和科研服务,各图书馆在资源的整合及利用方面做了

大量的工作。以清华大学图书馆为例,面对大量以“孤岛”形式存在的电子资源,开发了“学术信息资源门户”,提供信息查找的“统一界面”、提供多数据库的整合检索和获取目标信息的一站式服务。但即便这样,当具体到某一研究课题尤其是跨学科的研究课题时,还是有可能遇到一些问题:一方面,学生可能会因为难以确定其专业方向而难以选择合适的检索工具,或者仅选择自己熟悉的少量数据库而漏掉一些比较好的其他数据库;另一方面,有些学科的相关数据库很多,学生也难以选择哪些是更适合自己的课题的数据库;而且不同的检索系统或数据库,其检索方式不尽相同,检索式一般不能通用。这些都会对检索效率及检索结果带来一定的影响。

为使学生能更好地利用信息资源,很多高校开设了文献检索课,《文献检索与利用》即是清华大学图书馆为全校本科生及研究生开设的课程之一^[1]。文献检索课的开设目的是使学生了解信息检索的基

础理论和基本知识,以及检索工具的利用方法和技术;了解相关资源,学会选择、评价信息资源,制定合理的检索策略,分析处理和利用检索结果,进而能完成一些具体的应用^[2]。在文献检索课的教学过程中,文献信息的获取是非常重要的环节。

Dialog 联机检索系统拥有几百个数据库,具有明确的学科分类、多库的灵活选择、运算符的灵活搭配、使用同一检索式可实现跨库检索、检索结果可以剔重等多个特点。在高校图书馆和科技情报所等机构,Dialog 检索系统一般用于科技查新过程中,它是保证文献查全率的必需资源。而在文献检索课中引入 Dialog 检索系统,可以帮助学生制定和优化检索策略,并根据扫描命中的数据库查找相应的馆藏资源和免费资源,为学生方便快捷、准确全面地获取各类科技信息提供有利的帮助。

2 Dialog 联机检索系统及其利用

Dialog 是目前世界上最大的国际联机检索系统,收录了全球 600 多个数据库,文献量超过 3 亿多篇^[3]。利用 Dialog 系统,可为课题立项、文献调研、项目查新、专利申请、新产品开发,以及了解市场动态和竞争对手进而进行经济预测等多种需求提供服务。

2.1 Dialog 信息资源概况

Dialog 的信息资源按数据类型分为:书目文摘型数据(Bibliographic,主要是科技、专利类信息),名录手册型数据(Directories,主要是各种工商企业名录、百科全书、手册药典等),全文型数据(Complete-text,主要是专利、市场行业报告、分析报告和工业报告、新闻报道、期刊报纸等),数值型信息(Numeric,主要是价格、进出口数据、生产、销售数据等各种统计信息)。

Dialog 的信息资源按主题分为商业、新闻、专利商标、科学技术、医药、政府和法律、社会科学和人文、参考资料等共八大类。其中科学技术类数据库的使用率最高,这在高校尤为突出。科学技术类数据库按学科可以细分为农业和营养学、化学、能源与环境、医学与生物学、药理学、工程与技术等六大类,共 280 多个数据库,主要为书目文摘型数据。利用科学技术类数据库获取书目信息后,可以再通过其他途径获取全文。

2.2 Dialog 检索界面及选择利用

Dialog 有多种信息检索界面,如:

DialogWeb 综合性智能傻瓜界面: <http://www.dialogweb.com>;

DialogSelect 简易菜单式检索界面: <http://www.dialogselect.com>;

DialogClassic Web 专业指令界面: <http://www.dialogclassic.com>;

DialogLink 为 Dialog 的检索软件平台,可直接采用命令检索。

利用综合性检索界面 DialogWeb 进行数据库扫描即预检索,可以判定检索词是否规范、检索式是否合理、检索结果(从数量上看)是否合适。利用 DialogWeb 的特点之一,是可以在确定的学科范围内利用同一检索式对多个数据库进行扫描检索,获得命中数据库的名称及其命中文献的数量(不能看到具体文献信息)。根据命中文献结果,可以调整检索式,再进行扫描,直至检索式合适为止。确定最终的检索式后,可以利用 DialogClassic Web 或 DialogLink 进行正式联机检索,获取文献(该过程需要正式账号)。而仅用 DialogWeb 进行数据库扫描,则可以使用免费试用账号,且扫描过程几乎不产生费用。

2.3 免费账号的获取及数据库扫描

利用 DialogWeb 进行数据库扫描之前,需要进入 Dialog 主页获取免费试用账号和密码,步骤如下:

2.3 免费账号的获取及数据库扫描

利用 DialogWeb 进行数据库扫描之前,需要进入 Dialog 主页获取免费试用账号和密码,步骤如下:

① 进入 Dialog 的主页(<http://www.dialog.com>);

② 点击 Support 菜单下的 Training,进入 Training 页面;

③ 点击 Free Practice Searching,获得 ONTAP User Numbers and Passwords for Dialog。

获取试用账号后,就可进入 DialogWeb 综合检索界面进行数据库扫描,步骤如下:

① 进入 DialogWeb 界面(<http://www.DialogWeb.com>) 输入账号、密码,点击“logon”;

② 点击“Databases”进入数据库选择页面;

③ 按照信息资源类型选择数据库,如选择科技类数据库(Science and Technology)。进入选中的数据库组后,可直接输入检索式进行扫描,也可以进一步选择数据库,如点击“computer, electronics, and telecommunications”则进入计算机、电子和通讯类数据库组;

④ 选择好数据库类别后,可以对该类别的所有数据库进行扫描,也可以从中选择某些(至少两个)具体数据库文档进行扫描(见图 1)。注意:至少选择两个以上数据库才能免费扫描。

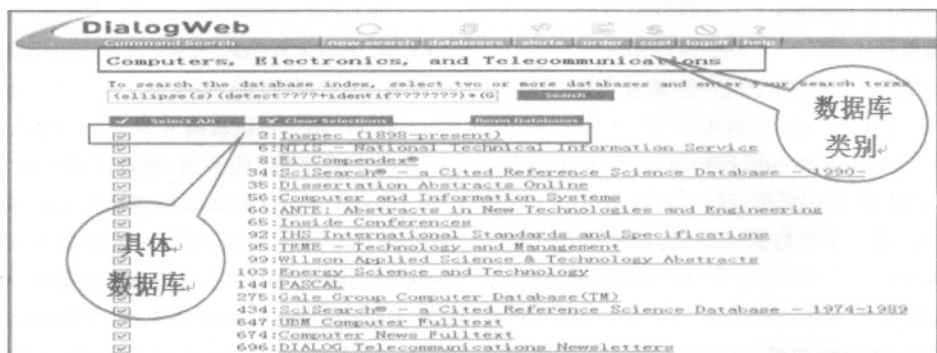


图1 选择数据库进行扫描

⑤ 扫描结果

利用检索式“ellipse(s) (detect???? + iden-

tif?????) * (GPU + graphic?? () processing () u-nit??)”进行数据库扫描,结果见图2。

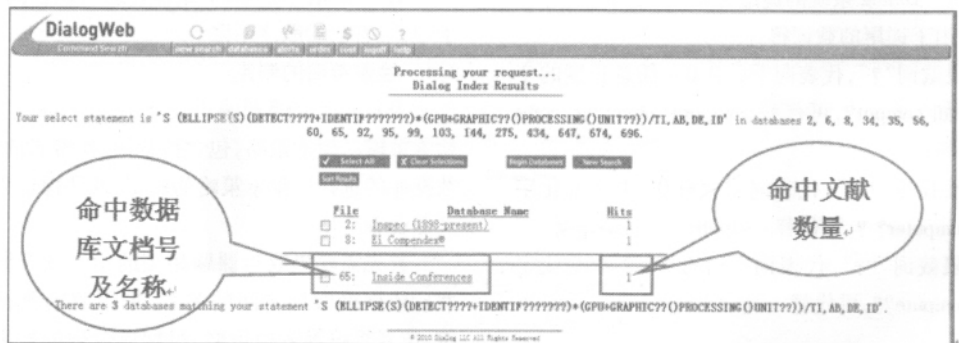


图2 扫描结果显示

2.4 利用馆藏资源或网上免费资源获取文献信息

利用 DialogWeb 进行数据库扫描后,可以根据命中数据库的名称,在图书馆购买的电子资源中查找相应的数据库,或者查找网上的免费资源,以获取相关文献的信息。

2.4.1 馆藏资源的查找

很多高校的图书馆购买了大量电子资源,Dialog系统中的许多数据库都能从中找到。以笔者所在的清华大学图书馆为例,Dialog中的2号库Inspec(科学文摘)、8号库Ei Compendex(工程索引)和6号库NTIS(美国政府报告)这三个数据库都在本馆订购的Engineering Village 2平台上,而Dialog中的34号库SciSearch则包含在馆购电子资源Science Citation Index Expanded中,此外Dialog中的14、23、56、64号等多个数据库也都在本馆订购的剑桥科学文摘(CSA)中。

对应馆藏资源的查找方式:在本馆主页的数据库导航中,可以输入扫描命中的数据库名称“按数据库名称检索”,或者按字顺浏览数据库,或者从常用电子资源平台/数据库列表中查找,以确定本馆是否购买了该数据库。找到并进入馆藏数据库后,采用

与扫描式对应的检索策略(各数据库可能有所差异)进行检索,就可以获取相关文献的详细信息。

2.4.2 免费资源的查找

对于多数专利数据库,如Dialog中的340号库(CLAIMS(R)/U.S. Patents)、347号库(JAPIO)、348号库(European Patents Fulltext),分别可以从网上免费资源美国专利数据库(<http://patft.uspto.gov/>)、日本专利数据库(<http://www.ipdl.inpit.go.jp/homepg.ipdl>)、欧洲专利数据库(<http://ep.es-pacenet.com>)中找到。

通过馆藏资源或网上免费资源,可以获取相关文献的书目信息甚至全文,为所进行的课题研究提供参考。

3 利用 Dialog 运算符制定优化检索策略

在课题的研究过程中,检索策略的制定非常重要^[4]。采用不同的检索策略,检索结果不同。因此如何制定和优化检索策略,是文献检索课中学生信息能力培养的关键环节。

Dialog系统具有统一的检索算符,其中除“与”、“或”、“非”逻辑算符外,还有位置算符和通配符等。利用这些算符,可以灵活地制定和优化检索策略,提

高文献的查全率和查准率^[5]。

3.1 逻辑算符

Dialog 的“与”、“或”、“非”逻辑算符, 可以用“and”、“or”、“not”表示, 也可以用“*”、“+”、“-”表示。用 DialogWeb 进行扫描时, 检索式的字符数量没有限制, 采用哪种逻辑算符都可以。正式联机检索时, 由于字符数量有限制, 如果检索式较长, 最好采用“*”、“+”、“-”, 以减少字符数量。

3.2 通配符

通配符检索是指利用检索词的词干或不完整的词形进行检索。通配符的使用, 既可以减少检索词的输入量, 又可以提高查全率。通配符分为截词符和屏蔽符。Dialog 系统的通配符以“?”表示。

3.2.1 用于词尾的截词符

无限截词“?”代表词干后有 0 - 任意长度的变化字符: 如 comput?, 可代替 compute、computer、computation 等;

有限截词“??”代表词干后有 0 - 1 个变化字符: 如 computer??, 可代替 computer、computers 等;

有限截词“???”代表词干后有 0 - 2 个变化字符: 如 compute???, 可代替 compute、computer、computers 等;

有限截词“??...?”(n 个?, $n \geq 2$ 的自然数), 代表词干后有 0 - n 个变化字符。

3.2.2 用于词中间的屏蔽符

在一个词中间插入一个或多个“?”, “?”的个数为可代替的字符数。例如:

fib?? board, 可代替“fiberboard”、“fibreboard”等。

3.3 位置算符(W)/(nW)、(N)/(nN)和(S)

位置算符用于限定检索词之间的位置关系, 即通过限定检索词之间的顺序和位置, 提高文献的查准率。尤其是当检索词为固定词组, 或两个检索词之间有禁用词或被标点符号隔开时, 使用位置算符更有必要。

检索词之间的位置关系有前后顺序固定、前后顺序可以互换两种情况, 利用数字 n 还可以限定检索词之间插入词语或字符的数量。

3.3.1 位置算符(W)/()、(nW)

A(nW) B: 表示检索词 A、B 之间可插入 0 ~ n 个任意词语或字符, 且 A、B 的前后顺序不能颠倒。如: optical(1w) network, 检索结果包括 optical network、optical communication network、optical transmission network 等。

(W) 常用来替代词组中间的空格或“-”, 可以简写为()。如: x() ray 检索结果包括 x ray、x-ray。

3.3.2 位置算符(N)/(nN)

A(nN) B: 表示检索词 A、B 之间可插入 0 ~ n 个任意词语或字符, 且 A、B 的前后顺序可以互换。如: distributed(3N) system, 检索结果包括 distributed system、distributed management system、distributed workflow management system、management system of low-voltage distribution network 等; streaming(N) media 检索结果包括 streaming media、media streaming。

3.3.3 位置算符(S)

A(S) B: 表示检索词 A、B 必须出现在同一个句子中, 二者前后顺序不限。

3.4 检索策略的制定

为获取研究课题的相关文献, 需要制定相应的检索策略。检索策略, 包含检索词、检索式的确定及数据库的选择。检索策略的制定, 既要保证查准率, 又保证查全率。

首先需要提取与课题相关的检索词及同义词; 然后依据检索词之间的关系, 采用逻辑算符、通配符、位置算符等多种策略, 组配成合理的检索式; 再在相应的数据库中进行检索, 获取相关文献信息。

Dialog 检索系统虽然有几百个数据库, 但其运算符是相同的, 因此可以采用这些运算符组配成同一检索式进行多数据库的跨库扫描和检索。如课题“基于 GPU 的实时椭圆检测”, 制定如下检索式“ellipse(s) (detect???? + identif???????) * (GPU + graphic??) processing() unit??”。

4 将 Dialog 引入文献检索课程的教学实例

文献检索课程的目标是提高学生的信息素质, 而信息素质中最重要的是信息能力的培养, 即培养学生获取、筛选、分析、加工处理和吸收信息, 并创造出新信息的能力^[6]。将 Dialog 系统引入文献检索课程中, 可以更好地培养和提高学生的信息素质。

在学生的学习和科研过程中, 获取尽可能全面且具有重要参考价值的文献信息并加以分析利用, 是保证这些工作顺利进行的基础。下面以一具体研究课题“超灵敏度质谱检漏系统”为例, 介绍如何利用 Dialog 系统获取相关的文献信息。

4.1 分析研究课题, 提取检索词, 制定检索式

课题分析: 该课题的研究对象为具有超灵敏度的质谱检漏系统, 在质谱检漏系统中, 检漏灵敏度是最重要的参数。由此可见, 在该课题中, 质谱检漏、

超灵敏度是主要的概念。

检索词的提取: 检索词要保证其准确性和全面性, 很多检索词除了主题词、叙词等规范性词语外, 还有自由词等其它同义词, 而且可能存在多种词性表达形式。该课题中, 提取检索词: 检漏 (leak???? - detect?????)、超灵敏度 (ultra - sensitiv???? / ultrasensitiv????)。说明“检测”、“灵敏度”英文有多种词性表达形式, 如“检测”有 detection、detecting 等, “灵敏度”有 sensitivity、sensitivities 等, 所以分别在其词干后面采用通配符“?”。

检索式的制定: 根据检索词之间的逻辑关系组配检索式。该课题中检索词之间的逻辑关系为: 检漏 和 超灵敏度。确定了检索词之间的逻辑关系

后, 采用检索词的英文形式, 依据 Dialog 系统的逻辑算符、位置算符和通配符等指令, 组配 Dialog 检索式如下: (leak???? () detect????? * (ultra () sensitiv????? + ultrasensitiv?????) /ti ,ab ,de ,id

4.2 数据库选择及扫描

制定好检索式后, 要根据课题所属学科类别选择数据库进行扫描, 流程见 2.3 所述。

对于该课题, 其研究的质谱检漏系统按照研究主题分类属于仪器仪表类, 而仪器仪表还隶属于工程技术大类。分别选择“Engineering Research”和“Technology Patents”两个数据库组进行扫描, 结果如下(见图3):



图3 扫描结果

说明: 如果扫描得到的文献量过多或过少, 则需要调整检索词和检索式, 对检索条件加以限制或放宽, 再次扫描检索, 直至文献量合适为止。

4.3 根据扫描结果查找馆藏资源或免费资源

由图3可知, 利用该检索式在 Dialog 系统中扫描命中了 13 个数据库。其中 2、6、8、14、23、32、34、57、60、108、351 共 11 个数据库在清华大学购买的电子资源中能够找到, 而 340 号库则对应于免费网上资源 - 美国专利数据库。

4.4 从馆藏资源或免费资源中获取文献信息

进入与 Dialog 相对应的馆藏资源中, 利用相应

于该数据库的检索式进行检索, 可以获取相关文献的信息, 其中的某些文献, 还可以通过多种途径获取全文。

由图3可见, 34号库命中了5篇文献。对应于 Dialog 的 34 号库, 在馆购资源 ISI Web of Knowledge 平台上, 选择 Web of Science 数据库中的 SCI - EXPANDED, 采用检索式“leak* - detect* and (ultra - sensitiv* or ultrasensitiv*)”在主题字段进行检索, 得到这 5 篇文献的书目信息(见图4)。对于每一篇文献, 还可以点击 SFX 查找原文。

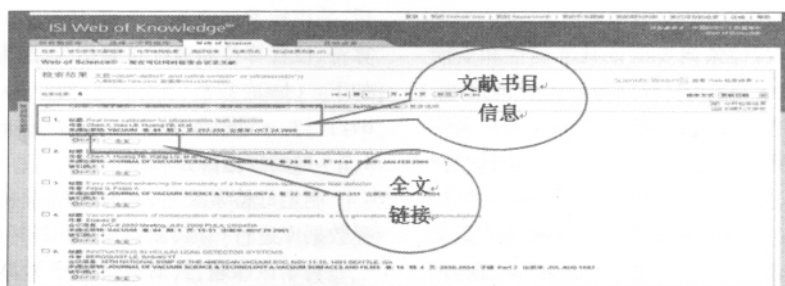


图4 SCI-EXPANDED 检索结果

5 结语

在文献检索课程中引入 Dialog 联机检索系统,可以帮助学生利用逻辑算符、通配符等指令组配合理的检索式,并根据 DialogWeb 的扫描结果,查找与命中数据库相对应的馆藏资源和免费资源,获取相关文献信息。Dialog 的引入,提高了学生分析课题、构建检索策略、获取文献信息的能力,为学生信息素质的培养及以后的学习和科研打下了坚实的基础。通过笔者在所授文献检索课中引入 Dialog 系统的尝试,取得了较好的效果。但如何更好地发挥 Dialog 系统在文献检索课中的作用,还需要进一步研究和探索。

参考文献

- 1 花芳,程爱平,曾晓牧.以“学”为中心:文献检索与利用课程的教学设计理念[J].教育探索,2008(11):63-64

(上接第81页)

- 10 岳珂著,王煊蕃校对.刊正九经三传沿革例[A].王云五主编.丛书集成初编[Z].上海:上海商务印书馆,1936:1-2
- 11 张政烺.读相台书塾刊正九经三传沿革例[A].张政烺.张政烺文史论集[M].北京:中华书局,2004
- 12 杨伯岳著,钱侗考证.九经补韵附录[M].北京:中华书局,1985
- 13 周密著,吴企明点校.癸辛杂识[M].北京:中华书局,1988:84-85
- 14 王国维.两浙古刊本考[A].王国维.王国维遗书:第7册[Z].上海:上海书店,1983:361-362,432
- 15 乔衍管.宋代书目考·绪言[M].台北:文史哲出版社,1987:2
- 16 朱彝尊著,黄君坦校点.静志居诗话[M].北京:人民文学出版社,1990:227,219
- 17 孙庆增.藏书纪要[M].上海:古典文学出版社,1957:38-359
- 18 卢文弨.抱经堂文集[M].北京:中华书局,1990:23,121
- 19 瞿冕良.中国古籍版刻辞典[M].济南:齐鲁书社,1999:123
- 20 金蟠辑.十三经古注[Z].同治八年浙江书局修补本.

- 2 http://www.lib.tsinghua.edu.cn
- 3 http://www.dialog.com
- 4 张岚,张柏秋,于非等.探讨科技查新中检索策略式的制定[J].现代情报,2008(10):151-152
- 5 孙君,陈陶.提高文献查全率和查准率的有效途径——逻辑运算符、位置算符和通配符的灵活运用[J].现代情报,2006(10):167-169
- 6 李晓新,陈孝林.信息技术的演进与信息能力的形成[J].图书馆工作与研究,2005(3):5-8

程爱平 女 1966- 清华大学图书馆,副研究馆员。

战玉华 女 1970- 清华大学图书馆,馆员。

钱俊雯 女 1965- 清华大学图书馆,馆员。

黄晓玲 女 1960- 清华大学图书馆,副研究馆员。

(收稿日期:2011-05-27 编发:刘 娴)

- 21 黄永年.古籍整理概论[M].上海:上海书店出版社,2001:31-32
- 22 陈鳢.经籍跋文[A].中国书店编.海王邨古籍书目题跋丛刊:第3册[Z].北京:中国书店,2008:19-20
- 23 余萧客.古经解诂沉[A].永瑒,纪昀.文渊阁四库全书:194册[Z].上海:上海古籍出版社,2005:358-359
- 24 蔡锦芳.清代藏书家朱文游与乾嘉学术[A].上海大学.文衡[M].上海:上海大学出版社,2009:169-179
- 25 叶德辉.郎园读书志[A].中国书店编.海王邨古籍书目题跋丛刊:第5册[Z].北京:中国书店,2000:205,205
- 26 徐珂.清稗类钞[M].北京:中华书局,1984:3846
- 27 丁丙.善本书室藏书志[M].北京:中华书局,1990:444
- 28 黄丕烈著,屠友祥校注.莒圃藏书题识[M].上海:上海远东出版社,1999:54
- 29 陈惠美.东海图书馆藏孟子书目[A].台中:东海大学图书馆馆讯,2004.4(31)

高正伟 1974- 华东师范大学中文系诸子研究中心博士。

(收稿日期:2011-01-05 编发:梁颖)