

# 对 4 种国际著名跨库检索系统的评价

吕瑞花, 崔宇红, 崔红伟, 吴克友

北京理工大学图书馆参考咨询部, 北京 100081

**[摘要]** 用界面设计、核心技术、检索功能、管理功能、个性化设计和其他 6 个指标, 对国际上著名的跨库检索系统 Metalib, Multisearch, Resaerch PRO 和 WebfeatExpress 进行综合评价, 结果显示 Metalib 和 WebfeatExpress 技术比较成熟, 而 Metalib 的使用需要投入较大的人力和物力。

**[关键词]** 跨库检索; 整合检索; 统一检索; 评价

**[中图分类号]** G354.4

**[文献标识码]** A

**[文章编号]** 1000-7857(2007)16-0038-04

## Evaluation of Four International Metasearch Systems

LV Ruihua, CUI Yuhong, CUI Hongwei, WU Keyou

Reference Department, Library of Beijing Institute of Technology, Beijing 100081, China

**Abstract:** With six aspects used as indexes, including interface design, core technique, searching function, management, personalized design, this paper gives an evaluation of the Metalib, Multisearch, Resaerch PRO and WebfeatExpress. The result shows that the Metalib and WebfeatExpress are better than the others in technical aspects, but the use of Metalib needs more personal and material resources

**Key Words:** metaserching; integrated searching; simultaneous searching; evaluation

**CLC Number:** G354.4

**Document Code:** A

**Article ID:** 1000-7857(2007)16-0038-04

### 1 跨库检索的背景

近几年来, 图书馆通过采购或者自建的方式, 积累了大量来自不同信息供应商、不同数据格式、不同访问协议、不同认证方式、不同检索界面的信息资源。例如图书馆 OPAC、全文数据库、文摘/索引数据库、引文数据库、电子期刊、电子图书、字词典/名录/指南/百科全书等。在 Internet 浏览器中, 这些数据库以各自专有的客户端检索界面进行数据检索和读取, 也就是说不同的数据库需要不同的检索方法, 这使很多读者感到非常困惑。为找到需要的文献信息, 读者要在不同的数据库中进行重复检索, 造成时间和精力浪费。因此尽快地改变这种资源分散性、缺乏统一管理局面, 实现数字图书馆各类数据资源跨库检索, 是摆在我们面前的迫切任务。

跨库检索又称为“整合检索”、“跨平台检索”、“统一检索”、“联邦检索”和“平行检索”等<sup>[1-2]</sup>, 它是指用户通过统一的检索界面, 可以一次并发地对多个网络数

据库(全文、文摘、OPAC)同时进行检索, 并将各个网络数据库的检索结果去重、排序后归并, 一次提交用户, 是在实体资源分散的情况下实现“虚拟的资源整合”<sup>[3]</sup>。理想中跨库检索系统应该为用户提供一个简单、快速、全面的导航及检索系统, 同时满足不同层次用户的需求。然而事实是否如此呢? 为了解决北京理工大学的数字资源整合, 并为广大的图书馆选购跨库系统提供依据, 本文对国外 4 个著名的跨库检索系统进行了综合评价, 评价结果与大家共享。

### 2 评价对象及指标

目前国内外跨库检索系统较多, 国外有 Metalib, Endeavor, Webfeat Prism 等产品, 国内有 Calis 的统一资源检索、TRS 的资源整合等。本文将目标集中在那些致力于大型图书馆和信息中心提供跨库检索的系统, 经过滤选, 本文选择了在国内具有应用市场的国外系统 Ex Libris 公司的 Metalib, CSA 公司的 Multisearch, 以

收稿日期: 2007-06-28

作者简介: 吕瑞花, 北京市中关村南大街 5 号北京理工大学图书馆, 副教授, 研究方向为信息素养教育和参考咨询服务;

E-mail: lvrh@bit.edu.cn

及在国内没有应用市场的著名国际系统 Innovative 公司的 Resaerch PRO, Webfeat 公司的 WebfeatExpress, 进行系统测试和评价。

本文选取界面设计、核心(或后台)技术、检索功能、管理功能、个性化设计和其他 6 个指标对上述 4 种国际著名跨库检索系统进行综合评价。

### 3 评价结果

#### 3.1 界面设计

界面是用户的直接接触点,界面的好坏直接关系到最终用户的接受程度。对于界面的评价我们主要选取了友好程度、界面汉化、导航功能、资源的揭示程度以及界面的可选性 5 个指标。4 种跨库检索系统的界面评价列于表 1。

从表 1 可以看出, Metalib 的界面友好, 汉化完全符合中国人的思路 and 习惯 (Metalib 界面汉化由清华大学图书馆完成), 资源导航功能强大、全面, 资源查找方便。Metalib 的界面管理功能灵活, 图书馆完全可以自己开发和选择界面, 因此 Metalib 界面功能整体很好。Webfeat Express 的界面设计基本符合用户需求, 而 Multisearch 和 Resaerch PRO 界面的汉化和友好程度都不够, 不能满足一般用户的需求。

#### 3.2 核心技术

核心技术决定资源整合的质量和数量。Metalib 和 WebfeatExpress 均采用自己公司开发的跨库技术, 发展潜力大。而 CAS 的 Multisearch 和 Research PRO 的核心技术均采用第三方技术——MUSE 的跨库技术, 其发展受控于第三方。4 种跨库检索系统基本都能整合目前市场上较大的商业化的英文数据库, 但是与中文数据库的整合存在很大差异。目前为止, 清华大学图书馆已经利用 Metalib 整合其馆藏 90% 以上的电子资源, 其中包

括 60 多个中文数据库, 而 WebfeatExpress, Multisearch 和 Research PRO 中文资源整合较少。Multisearch 已经成功整合了中国最大的期刊数据库——中国期刊全文数据库, Research PRO 虽然已经整合了万方的学位论文数据库, 但还存在很多问题。WebfeatExpress 目前还没有整合任何中文资源。由于 WebfeatExpress, Multisearch 和 Research PRO 对中文资源的资源整合刚刚起步, 还需要进一步地验证其核心技术。

#### 3.3 检索功能

理想的跨库检索功能应不丢失单库检索的功能, 检索的结果应同于单库检索。而实际上, 在实现异构数据库的统一检索时, 只能保留住数据库中的共性检索功能, 一些特殊的字段或功能则可能在跨库检索中损失, 因此检索结果也会出现差异。4 种跨库检索系统的检索功能列于表 2。

从表 2 可以看出, Metalib, Multisearch 和 WebfeatExpress 都可以支持全字段检索, 支持去重, 而 Research PRO 则不支持全字段检索, 不支持去重; 相关度排序是检索系统的重要功能, 它可以将系统认为最相匹配的结果最先提供给用户, 单库检索基本都具备此项功能, 而在跨库检索中, 需要等所有结果返回, 重新核定相关度而排序, 所以跨库检索的相关度排序难度较大, 难以实现。Metalib 和 WebfeatExpress 都支持总结果的相关度排序, 而 Multisearch 和 Research PRO 则不能支持总结果的相关度排序。Metalib 不支持结果批量输出, 检索速度尚可。而 WebfeatExpress 的跨库检索功能毫不逊色于单库检索, 给读者的感觉是如同在原库中检索, 而且速度非常快, 因此 WebfeatExpress 在检索功能方面具有独特的优势, 既实现了跨库检索, 又不丢失原库的检索功能。

表 1 4 种跨库检索系统界面评价

Table 1 Evaluations based on the interface of four metasearch systems

| 界面性能  | 产品                        |                               |                |               |
|-------|---------------------------|-------------------------------|----------------|---------------|
|       | Metalib                   | Multisearch                   | WebfeatExpress | Research PRO  |
| 友好程度  | 很好                        | 较差                            | 一般             | 较差            |
| 界面汉化  | 已经完全汉化并符合中国人的思路 and 习惯    | 初步汉化                          | 基本汉化           | 已经汉化但不符合中国人习惯 |
| 导航功能  | 导航功能强大, 提供分类、查找、列表等多种导航方式 | 有分类和列表, 但数据库列表置于第二页面, 不利于资源发现 | 有分类和列表导航       | 有分类, 但分类乱     |
| 资源揭示  | 简单揭示, 有详细说明链接             | 没有揭示                          | 详细揭示           | 没有揭示          |
| 界面可选性 | 用户可以修改                    | 用户不能选择                        | 有 3 个备选界面      | 用户不能选择        |

注: 界面友好程度的评价结果来自于对用户调查的结果。

表 2 4 种跨库检索系统的检索功能评价  
Table 2 Evaluations based on the retrieval functions of four metasearch systems

| 功能       | Metalib                                   | Multisearch               | WebfeatExpress            | Research PRO              |
|----------|-------------------------------------------|---------------------------|---------------------------|---------------------------|
| 简单检索     | 字段缺省                                      | 字段缺省(题名/文摘/主题词/作者)        | 全部/关键词/题名/作者/摘要/主题        | 关键词/标题/作者/主题              |
| 高级检索     | 主题词/题名/作者/ISSN/ISBN/年份/任意自段               | 任何字段/作者/题名/主题/关键词         | 全部/关键词/题名/作者/摘要/主题        | 关键词/标题/作者/主题              |
| 逻辑算符     | 支持                                        | 支持                        | 支持                        | 支持                        |
| 位置算符     | 支持                                        | 不                         | 不                         | 不                         |
| 截词算符     | 支持                                        | 支持                        | 不                         | 支持                        |
| 词组检索     | 不                                         | 支持                        | 支持                        | 不                         |
| 限制检索     | 不                                         | 支持                        | 支持                        | 不                         |
| 二次检索     | 支持                                        | 支持                        | 不                         | 不                         |
| 排序       | 相关度、作者、标题、年份、数据库                          | 日期、篇名、作者、作者/题名、相关性排序且局部排序 | 相关度、作者、标题、日期、仅限选择         | 无                         |
| 去重       | 支持                                        | 支持                        | 支持                        | 不支持                       |
| 结果输出     | 只有完整格式才能保存/E-mail/添加到我的电子书架/查看原始格式,不能批量输出 | 打印、下载、E-mail              | 打印、保存、E-mail              | 不支持批量下载,需要到原库下载           |
| 结果查看     | 用户可以在所有结果都传回之前,开始查看部分检索结果                 | 不能                        | 用户可以在所有结果都传回之前,开始查看部分检索结果 | 用户可以在所有结果都传回之前,开始查看部分检索结果 |
| 兼容文献管理软件 | 支持                                        | 支持                        | 支持                        | 不支持                       |
| 检索速度     | 尚可                                        | 慢                         | 快                         | 慢                         |

### 3.4 管理功能

Metalib 的服务器设在本地,图书馆馆员自行对跨库检索系统进行配置、更新、认证、统计等管理,可以及时处理跨库检索中的问题,例如某数据库链接不通、检索结果出错等。而 WebfeatExpress, Multisearch 和 Research PRO 的数据库配置都是由提供商进行远程维护,如出现问题,图书馆不能自行修改,需及时通知跨库检索系统提供商进行排除和修改,因此解决问题的时效性难以保证,但是系统的提供商往往拥有很强的技术力量,而且他们与数据库的提供商联系紧密,可以阻止因链接改变而产生的死链等问题。WebfeatExpress 虽不能自行进行数据库配置,但可实现简单管理,系统管理员可很方便地设置和修改用户的访问权,进行数据库分类和使用统计等。Multisearch 的配置由提供商完成,图书馆只能进行简单分类、激活。Research PRO 的管理界面暂时没有提供,性能未知。由谁来进行数据库配置的问题,目前还有许多争议,决定权不但在于技术,还在于图书馆的发展理念。

### 3.5 个性化设计

个性化设计是近几年数据库商极力推出的功能,也受到很多图书馆用户的认可。目前 Metalib 支持个性化设计,可以让每个读者都能够创建自己的资源列表和电子期刊列表,建立个人的电子书架、保存检索历

史、设置定题资源提醒等,包括我的数据库、检索历史、定制 ALERT 等。WebfeatExpress 和 Multisearch 也支持部分个性化服务,而 Research PRO 则不具备。

### 3.6 其他指标

实际上在购买一个系统时,除了需要考虑技术因素外,还需要考虑一些非技术因素,例如产品的价格、需要投入的人力和物力、售后服务和本土化效应以及国际化效应等。表 3 中列出了购买跨库检索时需要考虑的其他因素。

从表 3 中可以看出, Metalib 已经实现了本土化,在服务支持上也有绝对优势。但是从人力、物力的投入上来看, Metalib 比其他 3 个系统的投入都要多得多。

## 4 结论

通过以上的对比评价,可以看出, Metalib 在界面设计、核心技术、管理功能、售后服务等方面都有绝对优势,可称为是一个技术程度成熟、服务有绝对保障的产品。WebfeatExpress 虽然在本土化的进程上稍慢于 Metalib,但是其检索功能比 Metalib 更容易为用户所接受。Multisearch 和 Research PRO 的开发较晚,因此在核心技术和界面功能、检索技术等方面都有待于提高。

然而,真正购买一个自动化系统时,除了考虑技术

表 3 4 种跨库检索系统的其他评价因素  
Table 3 Evaluations based on the other factors of four metasearch systems

| 评价因素 | Metalib                               | Multisearch                   | WebfeatExpress             | Research PRO                  |
|------|---------------------------------------|-------------------------------|----------------------------|-------------------------------|
| 人力   | 1~2 个人管理                              | 不需要                           | 少量投入                       | 不需要                           |
| 物力   | 系统购买费用+服务器费用+维护费                      | 维护费                           | 维护费                        | 维护费                           |
| 国内用户 | 包括清华、复旦、北师大、国家图书馆等著名图书馆 15 家          | 已有多家试用                        | 在中国还没有用户                   | 在中国还没有用户                      |
| 国际用户 | 在全球 38 个国家和地区拥有 800 多个图书馆用户,占有最大的全球市场 | 全球接近 40 家单位正购买使用              | 超过 1 500 家图书馆用户,包括康奈尔大学    | 新一代产品全球共有 5 家用户在试用            |
| 服务支持 | 公司在中国设有代表处,能够提供直接的、有效的技术支持与服务         | 公司在中国设有代表处,能够提供直接的、有效的技术支持与服务 | 在中国没有服务支持人员,无法提供直接的技术支持与服务 | 公司在中国设有代表处,能够提供直接的、有效的技术支持与服务 |

外,还要考虑人力和物力的投入,还要考虑谁来维护电子资源的更新更有利,是图书馆还是跨库检索提供商。康乃尔大学就放弃了自我管理的 Endearvor 系统,而选择了提供商远程管理的 Webfeat 产品。因此选择产品前要全面权衡。

跨库检索系统是一种新生的改变迅速的技术,它处于一个爆炸性的环境,有多个竞争者,最大的竞争对手是 Google。目前的跨库技术还都有它本身的局限性,在实现异构数据检索时,牺牲了原数据库的许多特性和结果。期待未来有 Google 这样的提供商将所有异构数据同构化,实现真正的统一检索。

### 参考文献 (References)

- [1] SADEH T. The challenge of metasearching [J]. New Library World, 2004, 105(1198/1199): 104-112.
- [2] TALLENT E D. Metasearching in Boston college libraries—a case study of user reactions [J]. New Library World, 2004, 105 (1196/1197): 69-75.
- [3] 陈冰云. 标准化跨库检索的设想 [J]. 科技情报开发与经济, 2005, 15(6): 231-232.  
CHEN Bingyun. The presumption of normalized trans-database retrieval system[J]. Sci/tech Information Development & Economy, 2005, 15(6): 231-232.

(责任编辑 王 芷)

### · 举报信箱 ·



欢迎举报

E-mail: hjy@cast.org.cn

## 学术不端行为举报反馈

编辑同志:

您好!

我是河北工业大学青年教师赵晓艳, 贵刊 2007 年第 2 期“学术不端行为举报”栏目中,有读者反映我和其他合作者不同刊物上发表的 2 篇文章内容完全相同。经过本人核实,发现 2 篇文章内容确实基本相同,但随着课题的不断深入,试验内容不断丰富,其中一个刊物上的文章增加了计算程序流程图和缺陷桩速度响应曲线,并且结论也有些许不同。该读者反映的情况确实是作者造成的。借此机会,我谨向广大读者致

以深深的歉意,同时感谢科技导报社严谨的工作作风,以及为净化科学共同体学术环境所做出的不懈努力。

本人及课题组、协作单位必将以此为鉴,本着严谨求实的科学作风,坚决杜绝此类事件的再次发生,请贵刊及广大同仁予以监督。

此致

敬礼!

赵晓艳

编辑部回应:

本刊自 2006 年第 3 期开设“学术不端行为举报”栏目以来,已刊登举报信息 21 条,并给每位当事人及其所在单位寄送了当期刊物。赵晓艳同志是第一位主动给本刊回复并认真做出反省的当事人,我们对该同志勇于承认错误、知错就改的态度谨表示敬意,并希望通过我们大家的共同努力,进一步净化我国的学术环境和科研氛围。

本刊编辑部