

Implementation of a Federated Search System: Resource Accessibility Issues

Fei Xu

During 2009, the Hilton M. Briggs Library, South Dakota State University, went live with MetaLib, a federated search engine from Ex Libris. This paper presents the implementation process, addresses resource accessibility issues, and suggests a number of resources that can be consulted about these issues. Readers can apply similar strategies to projects involving federated search systems, as many implementation projects involve the same basic considerations presented here. The author did not find any studies specifically focused on configurations of MetaLib resources, so this content should be singularly useful. *Serials Review* 2009; 35:235–241.

© 2009 Elsevier Inc. All rights reserved.

Keywords: MetaLib; library systems implementation; federated searching; Metasearching; Information Resource Description (IRD); resource accessibility; Yaz client

Introduction

Hilton M. Briggs Library at South Dakota State University (SDSU) is a member of the South Dakota Library Network (SDLN), a multitype library consortium. In June 2003, SDLN selected the Ex Libris suite of products of ALEPH, MetaLib and SFX to replace its existing PALS system. The Briggs Library, however, did not bring its MetaLib instance live to its users until 2009.

Many academic libraries have introduced federated search solutions to their communities to meet the library users' desire to streamline their searches across search engines, library catalogs, and subscribed databases. The Briggs Library is no exception in seeking to maximize usage of its resources. Immediately following the implementation of SFX in 2008, the library implemented MetaLib in 2009.

Literature Review

Today “federated searching” and “metasearching” are very popular terms. According to Fryer, while the US National Information Standards Organization (NISO) and many libraries equate federated searching with metasearching, system vendors prefer using different terminology to prevent confusion with the metasearch engines on the free Web.¹ Metasearch engines crawl open Web sites, while federated search systems focus on subscribed databases processed through a professional indexing mechanism to enable structured, in-depth, content-oriented retrieval.² While use of the terms “federated searching” and “metasearching” vary, they are treated interchangeably in this article to describe federated searching systems that simultaneously search multiple resources.

Research on federated searching has concentrated on three major themes: system performance and technical development, system implementation, and usability testing. Daniel Dorner and AnneMarie Curtis developed a set of evaluation criteria for the performance of federated search systems. The evaluation criteria included search functionality, user interaction, interface customization, authentication, design, database communication protocol, vendor support, and system platform. Their study identified the top five products: MuseSearch, EnCompass, MetaLib, SingleSearch, and WebFeat.^{3,4}

Bob Gerrity, Theresa Lyman, and Ed Tallent detailed the installation and configuration of MetaLib and SFX at Boston College.⁵ Charles Lockwood and Patricia MacDonald addressed the lessons learned from implementing a federated searching system and concluded that more teamwork between system and reference staff would better address issues, such as technical capabilities, limitations of the system, and reconciliation between user preferences and librarian ideals.⁶ Georgia Library Learning Online (GALILEO) identified a hybrid approach to federated search products. Their experience suggested that consortial implementations should reflect the unique technical and administrative environment that requires solutions to be scalable and flexible.⁷ The Five Colleges Libraries in Massachusetts hosted SFX and MetaLib installations with Ex Libris. Lori Mestre and others presented how the Consortia planned, made decisions, and took actions regarding the implementation.⁸ Deborah Becker mentioned that the most difficult part of the project was determining the list of resource categories and allocating resources to each category.⁹ Marybeth Grimes elaborated the methodology and process of allocating library resources to MetaSearch Categories and Subcategories.¹⁰ Michael Nelson, Mary Ann Harlow, and Cassandra Kvenild recommended that having a good vendor–customer relationship is key to the success of a project.¹¹

As for usability testing, Susan Avery, David Ward, and Lisa Hinchliffe explained in detail three layers of technical, functional,

Xu is Integrated Systems Librarian and Assistant Professor, Hilton M. Briggs Library, South Dakota State University Libraries, Hilton M. Briggs Library, SBL 2115, South Dakota State University, Brookings, SD 57007-1098, USA; e-mail: fei.xu@sdstate.edu.

and usability testing during implementation of a federated product.¹² Karen Calhoun conducted interviews of forty-five Cornell University Librarians, many of whom were unsatisfied with search performance, including limitations in database availability, search speed, retrieval precision, and result comprehensiveness.¹³ Anne Highsmith and Bennett Ponsford described search configuration changes, and user interface modifications and detailed the method for pre-release usability testing.¹⁴ Usability studies at Boston College showed that students preferred simple interfaces and had little interest in advanced searching techniques, such as subject searching or combining indexes. Most of their students just left the search interface as the default setting.¹⁵ At Northwestern University, studies indicated that students preferred a “best bets” group of three major databases in each subject area, with secondary groupings for further research, rather than long lists.¹⁶ Both of the studies at Boston College and Northwestern University showed that keyword searching is the most popular and that a single search box is preferred.

As the literature indicates, many implementation projects concentrated on selection criteria, implementation decisions, database categorization, authentication, user interface customization, and employee training techniques. However, no studies were located that concentrated on configurations of MetaLib resources, which is the focus of this article.

Implementation Process

For the Hilton M. Briggs Library, the implementation process was very straightforward since MetaLib had been chosen and the appropriate user authentication mechanism had already been implemented. The system implementation process consisted of three major phases: configuring and testing resources, allocating resources to categories, and developing a MetaSearch Quick Search Box.

Configuring and Testing Resources

In order to add library resources, such as databases, OPACs, Web sites, and subject gateways to MetaLib, subscription and technical descriptive data called Information Resource Description (IRD) must be created. IRD lies in the MetaLib KnowledgeBase. Ex Libris maintains a Central KnowledgeBase (CKB). Every MetaLib installation includes a copy of the CKB that serves as the basis for creating a Local KnowledgeBase. Each institution must localize its own IRD records by providing information that is specific to a subscription, such as authentication data. This process is called configuring resources.

The Briggs Library has approximately two hundred subscribed databases, all of which were planned to be included in MetaLib. Almost all vendors of those databases were contacted for database configuration. MetaLib CKB Updates and Ex Libris were also consulted. Right after configuring each resource, the resource accessibility would be tested by performing actual searches in the User Interface of MetaLib. Normally resources can successfully pass those tests. If an error message occurs or 0 hit is obtained, further tests with Yaz client and MLWP need to be conducted to identify if the problem is with the resource itself or with MetaLib. This specific error message will be discussed later.

Allocating Resources to Categories

MetaLib allows libraries to assign their resources to categories. Categories contain subcategories, and subcategories contain resources. To ensure the resources within a subcategory are searchable, the number of resources assigned to the subcategory

cannot exceed a certain value, which is determined by the global parameter in MetaLib, `www_metalib_search_limit`. The maximum value of the parameter is 500. However, increasing the value could severely diminish MetaLib's functionalities. For example, search time would greatly increase so that some databases would time out. So the maximum value of the parameter has been set to 15 for our MetaLib instance, which was decided by Ex Libris and us for the maximum number of databases that may be concurrently searched.

Categorizing resources is an important step as it greatly impacts patrons' use of metasearching. The Briggs Library has forty-four subject areas/disciplines in the library subject guides, which reflect the academic programs offered at SDSU. We hoped the way of organizing the library resources in MetaLib could meet the needs of academic programs and reflect the library's collections. Therefore, forty-five categories, including a general category, were created in MetaLib via the Categories Admin module in the Management Interface. The librarians had some difficulty deciding which databases would go into which categories as some databases fit into several categories while other databases did not seem to fit any of them. The systems staff solved this problem by asking the library bibliographers to furnish a list of core and additional databases in their subject areas. Accordingly, categories in our MetaLib instance were usually divided into two subcategories: core databases and additional databases. Each subcategory was assigned up to ten databases. If more than twenty databases were included in one subject area by the library bibliographer, new subcategories would be created in MetaLib to which those databases would be assigned.

Developing a MetaSearch Quick Search Box

Prior to the Briggs Library going live with MetaLib, the librarians extensively discussed where MetaLib should be placed for user access. The statistics indicate that the “quick search” is an extremely heavily used feature.^{2,15} In seeking to provide a better search experience for the patrons, we decided to place both a link to MetaLib and a MetaSearch Quick Search Box on the library's home page. The former is the full functionality of MetaLib and the latter is the “lite” version.

Based on MetaLib deep linking, the MetaSearch Quick Search Box was developed with JavaScript, allowing librarians to build tailored resources for users to search without having to locate MetaLib.¹⁷ It offers access to MetaLib through IP recognition. So off-campus access will need to log in first, while on-campus access is relatively seamless without a login. The MetaSearch Quick Search Box is a “Google-like” search box, which can search the library catalog, databases, and journals. The default value for the drop-down menu in the Quick Search Box was set to the “General Search” category (see [Figure 1](#)).

The “Google audience” favors simplicity and intuitiveness, so no instructions or explanation were provided at this point as those details are included elsewhere on the library Web site. Patrons using the Quick Search Box will be taken immediately into the library gateway and then search results will be presented to users. By providing the Quick Search Box, the library dismissed the confusion encountered when MetaLib was first introduced to the users.

Resource Accessibility Issues

The configuration of MetaLib resources itself seemed simple and straightforward, but it proved to be a very complicated and prolonged process, whether modifying an existing resource in the CKB or adding a new one to the CKB. The Briggs Library began the implementation of MetaLib several years earlier, but it had never

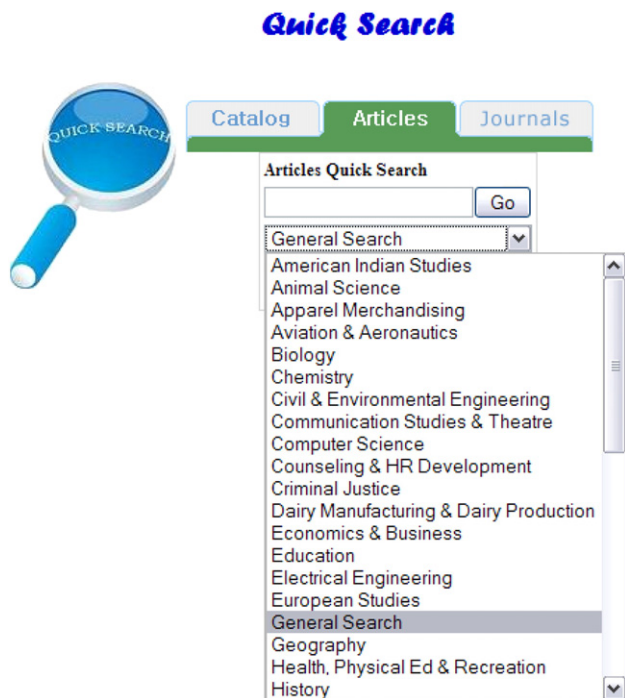


Figure 1. The MetaSearch quick search box.

brought the MetaLib instance live to the users until 2009. When the author started to implement MetaLib, some library resources had already been added to MetaLib. To begin this project, the first step was to check a resource in MetaLib against the list of the library resources on a one-by-one basis. Then the resource accessibility was tested if the resource was found in MetaLib, otherwise the resource was added to MetaLib, followed by the testing. If a resource could not pass the tests, the configuration for the resource would be consulted with the database vendor and Ex Libris. During the testing, the configurations for many resources in MetaLib were inaccurate. This testing, then, revealed resource accessibility issues.

This project presented itself as an opportunity to identify and solve resource accessibility issues for implementing MetaLib. There are many factors contributing to these issues. When the issues occur, the important configurations for resources need to be checked, such as hostname:port, database code, authentication, registration of MetaLib server IP with vendors, access method, and configuration code. To solve resource accessibility issues, database vendors, MetaLib CKB Updates, and server-side testing tools need to be consulted.

Database Vendors

The first step for configuring databases was to contact database vendors, which could involve long interactions back and forth. Initial persons contacted at the vendors might not be familiar with their databases. This unfamiliarity can complicate and prolong



Ongoing Menu

<h4>MetaLib KnowledgeBase</h4> <ul style="list-style-type: none"> • Resource Cataloging and Configuration <ul style="list-style-type: none"> ◦ Find a Resource ◦ Browse List of Resources ◦ Add a New Resource ◦ View Resources Added in CKB Update ◦ View Resources Changed in CKB Update ◦ Update or Activate Resources • Export-Import Procedures • Restart Z39 Gate 	<h4>MetaLib Administration</h4> <ul style="list-style-type: none"> • Set-Up <ul style="list-style-type: none"> ◦ Institutional Settings ◦ Portal Administration ◦ Define Default Profiles ◦ Manage QuickSets ◦ Categories Administration • Staff Administration • User Administration <ul style="list-style-type: none"> ◦ Register Users ◦ List of Users • Portal Look and Feel • Generate Deep Link • Reports & Statistics • MetaIndex Management
---	---

[GO TO INITIAL MENU](#)

Figure 2. The MetaLib management interface.

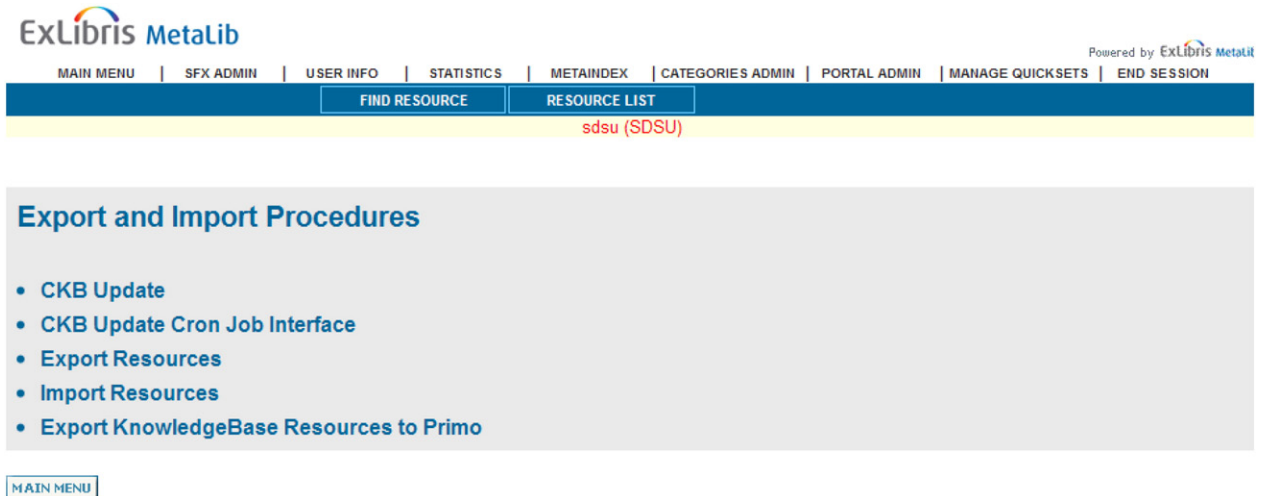


Figure 3. The export–import procedures.

communications. In difficult cases, asking Ex Libris for assistance during interactions with vendors was helpful. It should be noted, however, that Ex Libris usually does not negotiate with vendors on behalf of customers. The Briggs Library subscribed to SpringerLink through MetaPress. When configuring the SpringerLink database, many interactions with MetaPress took place. The MetaPress support staff seemed unfamiliar with the Metalib federated search setup. The configuration for SpringerLink eventually was solved with the assistance of the Metalib KnowledgeBase Team.

A special situation regarding resource accessibility issues was encountered when implementing our Metalib instance. Often a library will subscribe to multiple databases from one vendor. The configurations for those databases need to be considered as a unit instead of individually – a great time saver. In general, except for database code, other configurations such as hostname:port and

authentication are the same. The Briggs Library subscribed to many databases from EBSCO Publishing, such as Academic Search Premier, Business Source Premier, CINAHL, EBSCO MegaFile, and Electronic Journals Service (EJS). All these databases worked with Metalib except for EJS. The configurations for the EBSCO databases were checked, and there seemed nothing wrong with them. This particular situation was reported to EBSCO. During investigation, EBSCO found EJS was missing from the Z39.50 access profile (the Z profile) on their server. After EJS was re-added to the Z profile, the problem was solved.

Some vendors initially claimed that their databases are not compatible with Metalib, but subsequent conversations with different individuals proved otherwise. Consulting Metalib CKB Updates and Ex Libris also proved to be a way to acquire correct configurations for some databases.

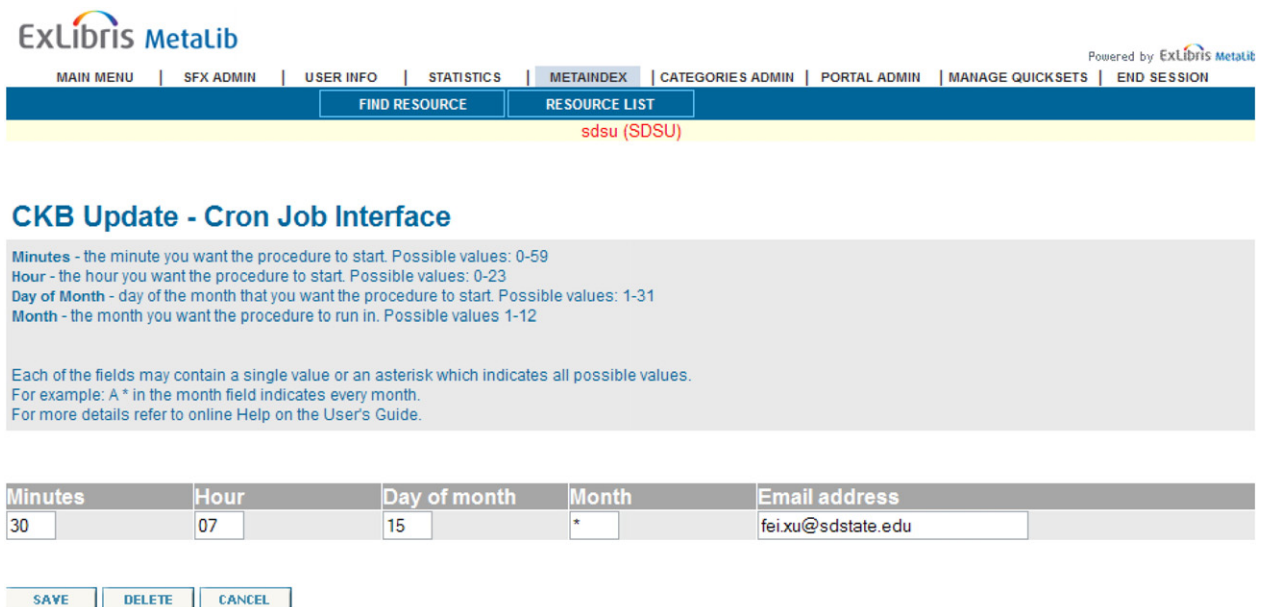


Figure 4. CKB update – cron job interface.

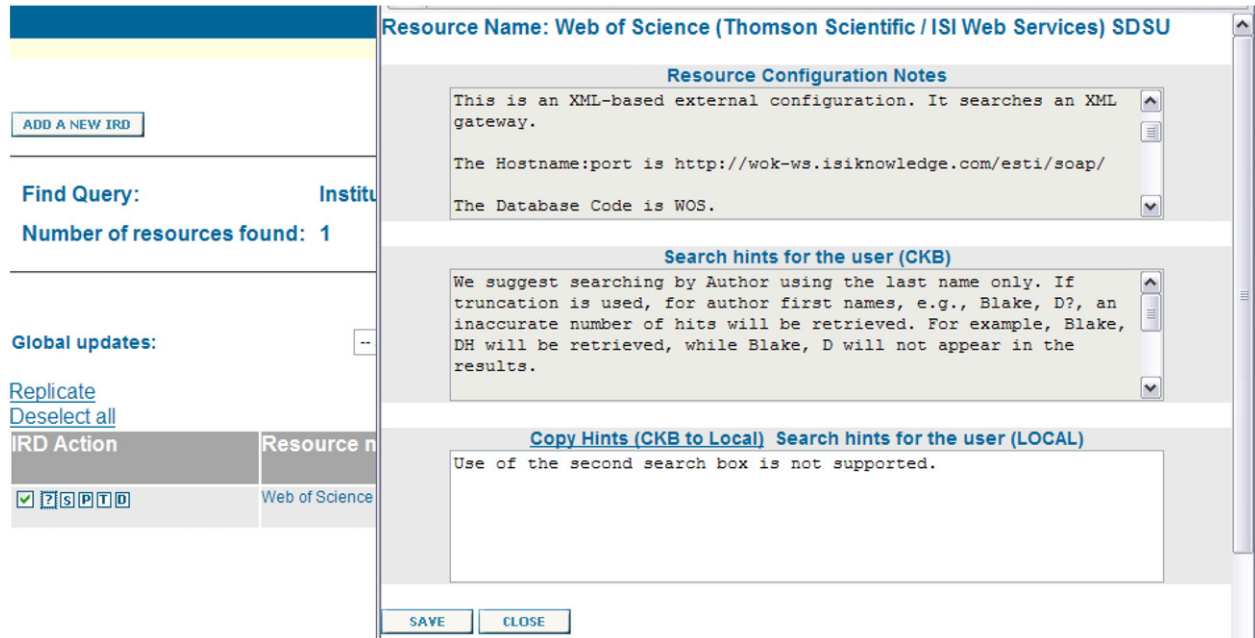


Figure 5. The configuration notes in the CKB.

MetaLib CKB Updates

Each month Ex Libris releases MetaLib CKB Updates, which include newly added and enhanced existing resources. When implementing MetaLib, if a resource cannot be found, the chances are that the resource may have been recently added to the MetaLib CKB; if an existing resource does not work with MetaLib, the configuration for the resource might have been updated or enhanced by Ex Libris. So consulting with MetaLib CKB Updates might solve those problems. MetaLib CKB Updates can be obtained in two ways: run CKB Updates in MetaLib and subscribe to the MetaLib-Discuss Listserv for Resource Configuration Release Notes.

To run CKB Updates in MetaLib, select "Export-Import Procedures" in MetaLib management interface and then click on "CKB

Update" (see Figures 2 and 3). Actually CKB Updates can be run automatically by arranging a cron job (a scheduled job). To arrange the cron job, click on "CKB Update Cron Job Interface." The time when the cron job starts can be specified. The cron job for this MetaLib instance was arranged to start on the 15th of every month (see Figure 4). After obtaining the most recent CKB by running CKB Updates, click on "View Resources Added in CKB Update" and "View Resources Changed in CKB Update" (see Figure 2) to view resources that were added and changed in the latest updates. Configuration notes for those resources can be found under the "?" icon of the IRD (see Figure 5). Based on the configuration notes, changes need to be made accordingly. Finally, test accessibility for those resources by performing actual searches in the User Interface of MetaLib.

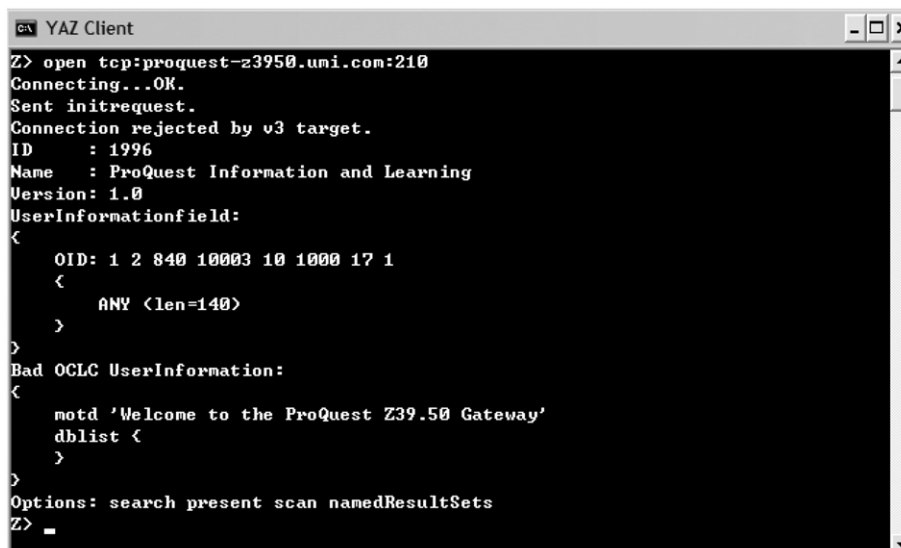


Figure 6. The Yaz test rejected by Z39.50 Server.

```

YAZ Client
Z> open tcp:proquest-z3950.uni.com:210
Connecting...OK.
Sent initrequest.
Connection accepted by v3 target.
ID      : 1996
Name    : ProQuest Information and Learning
Version: 1.0
UserInformationfield:
<
  OID: 1 2 840 10003 10 1000 17 1
  <
    ANY <len=49>
  >
>
OCLC UserInformation:
<
  motd 'Welcome to the ProQuest Z39.50 Gateway'
  dblist <
    'PQD'
  >
>
Options: search present scan namedResultSets
Z> base PQ-current
Z> find @attr 1=4 hnl
Sent searchRequest.
Received SearchResponse.
Search was a bloomin' failure.
Number of hits: 0, setno 1
Result Set Status: none
records returned: 0
Diagnostic message(s) from database:
  [109] Database unavailable -- v2 addinfo 'PQ-current'
Z>

```

Figure 7. The Yaz test with incorrect database code.

Resource Configuration Release Notes are distributed by Ex Libris to MetaLib Discussion List every month. The release notes usually only include the current month's updates and additions.

The full release notes that dated back to 2004, including more than 5,000 resource titles, were posted on the EL Commons, which is the Ex Libris customer wiki and Developer Zone.¹⁸ MetaLib CKB

```

YAZ Client
Z> open tcp:proquest-z3950.uni.com:210
Connecting...OK.
Sent initrequest.
Connection accepted by v3 target.
ID      : 1996
Name    : ProQuest Information and Learning
Version: 1.0
UserInformationfield:
<
  OID: 1 2 840 10003 10 1000 17 1
  <
    ANY <len=49>
  >
>
OCLC UserInformation:
<
  motd 'Welcome to the ProQuest Z39.50 Gateway'
  dblist <
    'PQD'
  >
>
Options: search present scan namedResultSets
Z> base PQ_current
Z> find @attr 1=4 hnl
Sent searchRequest.
Received SearchResponse.
Search was a success.
Number of hits: 1035, setno 1
records returned: 0
Z>

```

Figure 8. The Yaz test passed.

Updates are valuable for configuring resources. However, consultation with MetaLib CKB Updates requires discretion. For instance, according to the Resource Configuration Release Notes April 2008, SpringerLink is a free database. But this information was an error. SpringerLink requires a subscription and the IP of the MetaLib server needs to be registered with MetaPress.

If the configuration for a resource cannot be found either in MetaLib or MetaLib CKB Updates, recommendations could be made to add a resource to the CKB. Such recommendations can be submitted to the SFX/MetaLib Users Group (SMUG), Ex Libris, or database vendors to develop a standard gateway, such as Z39.50 and XML. Ex Libris has a CKB "acquisitions policy." If resources are used globally, have a standard gateway, or have been requested for a configuration by SMUG, Ex Libris would give preferential consideration to design a corresponding configuration for those resources.

Server-side Testing Tools: Yaz Client and MLWP

If a resource does not work with MetaLib, the resource accessibility can be tested using Yaz client and MLWP.¹⁹ The former is used to test Z39.50 resources, and the latter is used to test all other resources. If tests pass using Yaz or MLWP, the problem lies somewhere in MetaLib, such as IRD, search configurations, etc. If those tests fail, the problem is not in MetaLib. Common reasons that lead to this failure include: specific ports are blocked by firewalls; incorrect hostname:port, database code, authentication, access method, etc.; and/or, the IP of the MetaLib server is not registered with vendors.

Accessibility tests for ProQuest Academic Research Library using Yaz client are illustrated as in the following figures. The hostname:port for Academic Research Library is proquest-z3950.umi.com:210, and the database code is PQ_current. Figures 6 and 7 show the typical symptoms that Academic Research Library is not working with Yaz. Figure 6 indicates that connection was rejected by the ProQuest Z39.50 server due to incorrect authentication information.

In Figure 7, the ProQuest Z39.50 server was connected successfully and Yaz performed a search, but an error message was given because the database code is wrong. While the correct database code is PQ_current for Academic Research Library, PQ-current was used in this test as the correct database code intentionally. Figure 8 illustrates that Academic Research Library successfully passed the Yaz test, and Yaz performed a successful search for "H1N1" and the number of hits is 1035.

Server-side testing tools can help identify if the accessibility problem is with a resource itself or with MetaLib. If the resource can be accessed using Yaz or MLWP, the problem is in MetaLib, and the resource should work with MetaLib. The IRD or search configurations for the resource needs to be checked with MetaLib CKB Updates or Ex Libris. If the testing tools are not able to access the resource, the problem is not in MetaLib. Whenever there is an error message during testing a resource using server-side tools, the network administrator or the database vendor needs to be consulted to fix the underlying problem before the resource will work with MetaLib.

Conclusion

The Briggs Library implemented MetaLib following the implementation of SFX. To streamline searches across search engines, library catalogs, and subscribed databases, the library incorporated the resources in MetaLib based on what the library owns and the categories to which the resources were assigned by the library bibliographers according to the library subject guides.

The implementation of MetaLib was a difficult, prolonged and complicated process. There are many factors to consider when

resource accessibility issues occur. This article suggested a number of resources that can be consulted about these issues and also discussed server-side testing tools to diagnose resource accessibility. The experience at South Dakota University's Briggs Library will shed light on problem solving in similar projects.

Acknowledgment

The author thanks Professor Mary Caspers-Graper at South Dakota State University for appropriate suggestions in the preparation of this paper.

Notes

1. Donna Fryer, "Federated Search," *Online* 28, no. 2 (2004): 16–19.
2. Rong Tang, Ingrid Hsieh-Yee and Shanyun Zhang, "User Perceptions of MetaLib Combined Search: An Investigation of How Users Make Sense of Federated Searching," *Internet Reference Services Quarterly* 12, no. 1/2 (2007): 211–236.
3. Daniel G. Dorner and AnneMarie Curtis, "A Comparative Review of Common User Interface Products," *Library Hi Tech* 22, no. 2 (2004): 182–197.
4. Daniel G. Dorner and AnneMarie Curtis (2003), *A Comparative Review of Common User Interface Software Products for Libraries: Report Commissioned by the National Library of New Zealand*, Wellington, NZ: School of Information Management, Victoria University of Wellington.
5. Bob Gerrity, Theresa Lyman and Ed Tallent, "Blurring Services and Resources: Boston College's Implementation of MetaLib and SFX," *Reference Services Review* 30, no. 3 (2002): 229–241.
6. Charles Lockwood and Patricia MacDonald, "Implementation of a Federated Search System in the Academic Library: Lessons Learned," *Internet Reference Services Quarterly* 12, no. 1/2 (2007): 73–91.
7. Lauren Fancher, "Wanted, Dead or Alive: Federated Searching for a Statewide Virtual Library," *Internet Reference Services Quarterly* 12, no. 1/2 (2007): 133–158.
8. Lori S. Mestre, Christine Turner, Beth Lang and Barbara Morgan, "Do We Step Together, in the Same Direction, at the Same Time? How a Consortium Approached a Federated Search Implementation," *Internet Reference Services Quarterly* 12, no. 1/2 (2007): 111–132.
9. Deborah Becker, Rolf Proske, Nikki Crowster and Luinda Lilley, "A First in Africa: Implementation of MetaLib and SFX at the Cape Technikon Library," *Electronic Journal of Academic and Special Librarianship* 5, no.1 (2004).
10. Marybeth F. Grimes, "MSU Libraries' Implementation of Federated Search Software," *Mississippi Libraries* 71, no. 1 (2007): 8–10.
11. Michael L. Nelson, Mary Ann Harlow and Cassandra Kvenild, "Implementing Federated Search at the University of Wyoming," *Online* 33, no. 2 (2009): 26–31.
12. Susan Avery, David Ward and Lisa Janicke Hinchliffe, "Planning and Implementing a Federated Searching System: An Examination of the Crucial Roles of Technical, Functional, and Usability Testing," *Internet Reference Services Quarterly* 12, no. 1/2 (2007): 179–194.
13. Karen Calhoun, "An Integrated Framework for Discovering Digital Library Collections," *Journal of Zhejiang University Science* 6A, no. 11 (2005): 1318–1326.
14. Anne L. Highsmith and Bennett Claire Ponsford, "Notes on MetaLib Implementation at Texas A&M University," *Serials Review* 32, no. 3 (2006): 190–194.
15. Ed Tallent, "Metasearching in Boston College Libraries — a Case Study of User Reactions," *New Library World* 105, no. 1/2 (2004): 69–75.
16. Frank Cervone, "What We've Learned from Doing Usability Testing on Open URL Resolvers and Federated Search Engines," *Computers in Libraries* 25, no. 9 (2005): 10–14.
17. Ex Libris, *Metlib Version 4.0X Deep Linking Guide* (Jerusalem, Israel: Ex Libris, 2007), 13.
18. Ex Libris, 2009, EL Commons Collaborative Website, <http://www.exlibrisgroup.org> (accessed February 28, 2009).
19. Ex Libris, 2007, Troubleshooting Non-Searching Resources in MetaLib: Yaz and MLWP::Simple, <http://docs.exlibrisgroup.com/docportal> (accessed February 28, 2009).

Copyright of Serials Review is the property of Routledge and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.