Goddard, Lisa. "Getting to the Source: a Survey of Quantitative Data Sources Available to the Everyday Librarian: Part II: Data Sources from Specific Library Applications" *Evidence Based Library and Information Practice* [Online], 2 14 Mar 2007
Final Version: http://ejournals.library.ualberta.ca/index.php/EBLIP/article/view/152/241

Getting to the Source: a Survey of Quantitative Data Sources Available to the Everyday Librarian: Part II: Data Sources from Specific Library Applications [Post Print Version]

Lisa Goddard
Emerging Services Librarian/Division Head for Systems
Memorial University of Newfoundland Libraries
St. John's, Newfoundland, Canada
Email: lgoddard@mun.ca

Received : 01 December 2006 Accepted : 07 February 2007

Introduction

This article is the second in a two-part guide intended to help practitioners identify and access Web server log data that may be useful for quantitative analysis of resource and service use. Part I introduced readers to the standard data elements that appear in Web server logs and provided practical pointers for aggregating and interpreting this log data. Part II explores logs and reports generated by library-specific applications such as proxy servers, electronic index and journal sites, link resolvers, federated search tools, virtual reference software, and integrated library systems. The reader should be aware that this information may change over time as applications evolve and mature. This article is intended as a practical guide to help researchers undertake quantitative analysis in the current environment.

Remote Patrons: Measuring Proxy Server Use

A proxy server is used to authenticate off-site users prior to their access to licensed electronic products. The proxy server logs hold a host of information about remote use of e-books, e-journals, and e-index sites. Because proxy servers are basically Web servers, they log much of the standard data available in Web server logs, including the URLs of

the remote patrons who make requests, date/time stamps, and the URLs of the licensed resources requested.

```
142.162.51.235 - - [10/Nov/2006:13:52:57 -0700] "GET
http://www.somedb.com:80/index.html HTTP/1.0" 200 1234
```

As in Web logs the query portion of a URL appears in the proxy log:

```
&title=Journal+of+Transcultural+Nursing&atitle=The+effect+of+social+sup
port+and+acculturation+on+postpartum+depression+in+Mexican+American+wom
en
```

Proxy server logs can also be configured to record more or less information, depending on the logging format chosen. Like standard Web logs, they can be run through the same analysis tools that are used to analyze library Web site logs. Proxy logs are mined for information about the frequency with which off-site patrons access licensed resources held on other servers.

Some proxy server packages provide summary reports of activity within their administration interfaces. At Memorial University of Newfoundland (MUN) libraries we have used EZproxy from Useful Utilities (<http://www.EZproxy.com>) since 2002. The 'Hosts' report within the administration interface provides a summary of use for each vendor server configured in the log from the time that entry was first made.

Host	Created	Accessed	Counts
www.ncbi.nlm.nih.gov:80	10/31/2002 14:39	11/4/2006 21:41	3029226
www.sciencedirect.com:80	10/31/2002 15:26	11/4/2006 21:47	2282226
www.blackwell-synergy.com:80	10/31/2002 15:12	11/4/2006 21:29	2279443
info.library.mun.ca:80	10/9/2001 13:20	11/4/2006 21:39	1798910
spWeb.silverplatter.com:80	10/31/2002 13:55	11/3/2006 19:41	1483264
134.153.184.164:80	10/31/2002 13:55	11/4/2006 8:42	1469768
www.jstor.org:80	10/31/2002 15:28	11/4/2006 21:48	803051
www3.interscience.wiley.com:80	10/31/2002 15:28	11/4/2006 21:48	515526
online.statref.com:80	10/31/2002 14:39	11/4/2006 21:47	425362
serials.abc-clio.com:80	10/31/2002 15:08	11/4/2006 20:01	387078
muse.jhu.edu:80	10/31/2002 15:28	11/4/2006 19:20	237479
www.csa.com:80	10/31/2002 13:55	11/4/2006 21:24	148247

Table 1. EZproxy Hosts Report. From EZproxy administration interface.

The date 'Created' represents a patron's first access of that server via the library proxy. 'Accessed' refers to the last time that a remote patron accessed that resource.

When reading the EZproxy hosts report, it is important to remember that it is arranged according to server name, rather than by specific services offered by vendors. Because major indexes or e-journal packages may be hosted across multiple servers, all of the host entries for a given domain (e.g., <ejournals.ebsco.com:80> and <www-ca.ebsco.com:80> will need to be identified. It is therefore necessary to know which server names and domains are registered to a given service provider. Aggregating all hits on a particular domain may give you more useful data about comparative use of resources according to the vendor or provider.

If URLs to restricted resources have been wrapped with the EZproxy starting point URL, then it is possible to use proxy logs to generate 'gate-count' information about both local and off-site database access. "Wrapping" a URL simply means to put the EZprozy login link at the beginning of the URL. EZproxy starting point URLs take the following form:

```
http://qe2a-proxy.mun.ca/Login?url=http://restricted.resource.com/
```

Useful Utilities, the manufacturer of the Ezproxy software, offers a proxy directive called 'LogSPU' which creates a separate log file for starting point URLs. Both local and remote users must pass through the proxy URL in order to access the destination URL, *http://restricted.resource.com*. When the server administrator has configured LogSPU, EZproxy creates a separate log file that records the URL of the requested resource, a session identifier, and an indicator of whether the requesting user IP was local or remote. Unlike the main proxy log, the SPU log does not record every subsequent hit on the remote server, but provides a simplified dataset which reveals how often each entry point to a restricted resource was chosen by a user, and indicates whether that user was on-site or off-site. This is a very useful mechanism for measuring the number of times each index or aggregated e-journal package was accessed by a patron. It does not, however, provide information about the searches executed from within that index or the full-text resources selected from within the vendor gateway.

Use Data Provided by Vendors

Most database and e-journal vendors make usage reports available to their clients. In order to access these reports it is necessary to know the administrative URL and login for each separate vendor interface. This can be a major data collection project, if the library subscribes to many different resources, and if several different departments administer them.

Typically vendors will offer reports that summarize all site activity for a selected period of time. Table 2 is a summary report from ebook vendor Ebrary that provides title level activity for a particular e-book package.

Title	User Sessions	Pages Viewed	Pages Copied	Pages Printed
Survival Analysis	**157**	**3992**	**16**	**1337**
Antigone	**33**	**440**	**0**	**0**
Zen & the Brain: Toward an Understand of Meditation & Consciousness	**30**	**315**	**25**	**27**
Making Social Science Matter : Why Social Inquiry Fails and How It Can Succeed Again	**28**	**467**	**3**	**38**
Complete Critical Guide to Geoffrey Chaucer	**20**	**322**	**0**	**8**
Microsoft Excel VBA Programming for the Absolute Beginner	**19**	**613**	**10**	**0**
Witchcraft Continued : Popular Magic in Modern Europe	**16**	**89**	**0**	**0**

Fundamentals of Cellular Network Planning and Optimisation : 2G/2. 5G/3G- Evolution To 4G	15	138	11	0
Racial Theories in Fascist Italy	15	36	0	0
Visual Basic 2005 Express : Now Playing	15	686	1	0
7 Hidden Reasons Employees Leave : How to Recognize the Subtle Signs and Act Before It's Too Lat	14	131	5	0

Table 2. Ebrary Title Level Activity Statistics Report. Generated from Ebrary administrative interface.

Summaries for database activity are usually also available. The example in Table 3 provides a monthly summary of the total number of searches performed compared against each database from a particular provider (in this case, ProQuest) and the formats provided.

Searches	Database	Cit / Abstract	FT Format	Total
2883	CBCA Business	344	930	1274
4372	CBCA Current Events	598	1965	2563
3569	CBCA Education	519	1561	2080
3899	CBCA Reference	850	2700	3550
1116	Canadian Research Index	8	0	8
2098	Dissertations and Theses	237	0	237
2098	Dissertations & Theses: A&I	237	0	237
17937	Grand Total	2556	7156	9712
10860	Total Unique Searches (Search Button Pressed)			

Table 3. Database Usage Report. Generated from ProQuest administrative interface.

Metasearch engines provide some additional challenges for database use analysis, because they initiate searches against many products at once, driving up the number of searches reported by database vendors. In order to interpret the statistics in a meaningful way, libraries will need a way to distinguish the number of search sessions initiated from the native interface from those initiated from metasearch interfaces.

More granular database usage reports are available to provide information on the number of times specific journal titles are viewed in each database. The following report from ProQuest (Table 4) also provides information on the kinds of document formats provided.

Database	Journal Name	Citation	Abstract	Text	Total
CBCA Education	Physical & Health Education Journal	33	8	274	315
ProQuest Historical Newspapers NYT	New York Times (1857-Current file)	196	185	0	381
CBCA Current Events	Canadian Press NewsWire	7	2	420	429
Dissertations &	ProQuest Dissertations and	14	315	0	329

Theses: A&I	**Theses**				
CBCA Reference	**The Canadian Nurse**	**11**	**10**	**106**	**127**
CBCA Reference	**Canadian Journal of Public Health**	**25**	**4**	**33**	**62**
CBCA Education	**Education Canada**	**8**	**18**	**134**	**160**
CBCA Current Events	**CanWest News**	**2**	**5**	**202**	**209**
CBCA Education	**Canadian Journal of Education**	**15**	**10**	**132**	**157**
CBCA Current Events	**The Globe and Mail**	**175**	**0**	**0**	**175**
CBCA Business	**Canadian HR Reporter**	**13**	**20**	**58**	**91**
CBCA Education	**Alberta Journal of Educational Research**	**2**	**9**	**119**	**130**

Table 4. Derived from Document Usage by Database/Journal Report. ProQuest administrative interface.

Because librarians will not have access to the vendor's raw log data, and because one doesn't necessarily know the criteria upon which reports are based, it can be a challenge to interpret and compare the data. In recent years libraries and vendors have worked to develop a standard means to record and exchange online usage data.

One of the most successful initiatives has been project COUNTER (Counting Online Usage of Networked Electronic Resources). COUNTER has developed and published Codes of Practice for the dissemination of statistics related to electronic journal, database, book, and reference resource use. Vendors who have implemented COUNTER-compliant reports are required to clean, sort, and organize usage data according to the methods prescribed in the relevant Code of Practice. Libraries using COUNTER-compliant vendor reports can be assured that data is generated in a reliable and consistent way across different vendor services and products. There are a number of COUNTER-compliant reports, and a vendor may implement one or many, depending on the services offered. The available usage reports include:

- Journal Report 1: Number of Successful Full-Text Article Requests by Month and Journal
- Journal Report 2: Turnaways by Month and Journal
- Book Report 1: Number of Successful Title Requests by Month and Title
- Book Report 2: Number of Successful Section Requests by Month and Title
- Database Report 1: Total Searches and Sessions by Month and Database
- Database Report 2: Turnaways by Month and Database
- Database Report 3: Total Searches and Sessions by Month and Service
- Total Searches Run by Month and Service

A full list of COUNTER Usage Reports, descriptions, and examples are available in the COUNTER Code of Practice for Journals and Databases: <http://www.projectcounter.org/r2/COUNTER_COP_Release_2.pdf> and in the COUNTER Code of Practice for Books and Reference Works: <http://www.projectcounter.org/cop/books/cop_books_ref.pdf>.

COUNTER provides a standard set of methods and reports to measure the use of electronic resources, but it does not help to solve the problem that vendor-delivered reports are scattered all over the Web in a variety of password-protected interfaces. It is time-consuming to collect and aggregate reports from different vendors, even when comparable data is available. Enter SUSHI, NISO's Standardized Usage Statistics Harvesting Initiative.

SUSHI is a protocol that automates the retrieval of vendor-generated use reports. The protocol is designed to support the transfer of COUNTER-compliant reports. It allows libraries to schedule programs that automatically harvest statistical reports from vendor Web sites on a regular basis. SUSHI delivers usage reports as XML documents structured according to the COUNTER XML schema. An application that understands how to interpret and parse COUNTER XML is needed to display these reports in a user-friendly fashion. This could be a custom XSL style sheet written by a programmer, or it could be part of a commercial Electronic Resource Management System (ERMS).

Many libraries have implemented ERMSs to keep track of electronic subscriptions, terms of use, vendor contact information, administration interface URLs, passwords, and usage data. Verde, from Ex Libris, (<http://www.exlibrisgroup.com/verde.htm>) is an example of an ERMS that includes support for the implementation of SUSHI to automatically gather usage reports from many vendors. This is an ideal model, as it not only streamlines the process of gathering data, but it also integrates electronic resource usage information into the larger management system without a need for staff intervention or re-keying.

Link Resolver Data

Link resolvers are poised to provide a critical piece of the data puzzle for libraries. Like all online resources, link resolvers run on Web server technology and generate standard Web logs of transaction information. Link resolver logs have several advantages over other library data sources. Firstly, the link resolver provides usage information that cuts across many different resources and services offered within the library's digital environment. Secondly, link resolver hits indicate that a patron has found a record for a resource that interests him sufficiently to request the full text. Thirdly, link resolvers provide usage information at the article- or book-chapter-level because of the granularity of the data contained within OpenURL requests. OpenURL provides a standard way to encode a full article-level citation into the query portion of a URL. Web logs are capable of recording the query section from each URL (the portion following the ?). Following is an example of an OpenURL query:

```
http://mun-resolver.asin-
risa.ca:8080/?genre=article&isbn=&issn=0278193X&title=Working+Mother&vo
lume=29&issue=8&date=20061001&atitle=Building+the+new+stereotype.&aulast
=Gunn%2c+Eileen+P.&spage=92&sid=EBSCO:Academic+Search+Premier
```

It encodes the full article citation:

Author: Gunn, Eileen P.,
Title: Building the new stereotype.
Journal Title: Working Mother **ISSN:** 0278-193X

Volume: 29 **Issue:** 8 **Date:** 2006-10-01 **Pages:** 92

Citation information is valuable for serials collection analysis. It can identify heavily used electronic journal titles as well as those rarely accessed. The citation information can also help to identify journal titles requested from indexes, but not owned by the library. This is possible because the resolver records every request it receives, regardless of whether or not the resource is available in the library. Most link resolvers will provide online reporting tools for commonly sought metrics, including the number of requests for a specific journal title (Table 5) or ISSN (Table 6) within a given time period.

Title	Count
Science	465
Journal of advanced nursing	383
Nursing Times	332
Nature	262
Nursing Science Quarterly	208
Nursing	204
Pediatrics	202
Journal of nursing education	197
New England Journal of Medicine	187
Journal of college student development	184
Canadian Journal of Zoology	180

Table 5. Top *n* Requests by Journal Title Report. Generated from SirsiDynix Single Search Administration Interface.

ISSN	Title	Count
NO ISSN		65898
0954-7762	Nursing Times	323
0730-3084	Journal of Physical Education, Recreation and Dance	277
0309-2402	Journal of Advanced Nursing	272
0002-936X	American Journal of Nursing	206
0008-4263	Canadian Journal of Public Health	195
0036-8075	Science (Washington D C)	192
0362-4331	New York Times	188
0419-4209	Dissertation Abstracts International Section A: Humanities and Social Sciences	182
0148-4834	Journal of Nursing Education	173
0008-4301	Canadian Journal of Zoology	172
0028-0836	Nature	170

Table 6. Top *n* Requests by ISSN Report. Generated from SirsiDynix Resolver Administration Interface.

OpenURL links also provide information about the referring index. In the OpenURL example above, the user used EBSCO's Academic Search Premier database to locate the referenced article citation. Referrer identities specify which indexes generate frequent requests for full-text material. In addition to title-level reports, standard reports are also usually available that tally requests according to the index used.

rfr_ids Request Counter: Memorial University of Newfoundland	
Referrer	**Count**
epnet.com:CINAHL	9560
ncbi.nlm.nih.gov:	4598
epnet.com:Academic Search Premier	4093
epnet.com:PsycINFO	2865
hwwilson.com:OMNIFT	2711
silverplatter.com:BXCD	1962
abc-clio.com:XDB	1893
pqil:mla	1667
csa.com:socioabs-set-c	1438
epnet.com:ERIC	1328
abc-clio.com:AHL	913
abc-clio.com:HA	820
Google.com:	806

Table 7. Top *n* databases referring users to link resolvers. Generated from SirsiDynix Resolver Administration Interface.

The data collected in link resolver logs has obvious implications for local collection development, as it contains a great deal of information about e-index, e-book, and e-journal usage. Some have explored the use of link resolver data to create new indicators of scholarly quality at a global level. The bX project, sponsored jointly by Ex Libris (developers of the widely used SFX link resolver) and the Los Alamos National Laboratory, seeks to aggregate link resolver data from many institutions and subject it to data-mining techniques. The goals of the project are to reveal large-scale trends in scholarly communication, to provide new measures of the impact and prestige of publications, and to provide insight into use patterns by diverse scholarly communities (Bollen et. al, 2005).

Link Resolvers and Overlap Analysis

The back-end of every link resolver is a large vendor-maintained knowledge base that maps periodical holdings to each of the full-text indexes and e-journal packages containing that particular journal title. We can see, for instance, all of the full-text journals available within Academic Search Premiere, and the years for which those titles are available. As well, we can see every package that offers full or partial access to a specific journal title. This information can be leveraged to compare the holdings and coverage of two or more databases or e-journal packages being considered for purchase. CUFTS, a link resolver developed at Simon Fraser University Library for the Council of Prairie and Pacific University Libraries (COPPUL), makes a holdings comparison tool

freely available over the Internet. The CUFTS product allows comparisons of up to four aggregated resource packages and identifies those journal holdings that may be duplicated across the packages. A development version of CUFTS can be found at <http://cufts2.lib.sfu.ca/MaintTool/public/compare>.

It is also possible to perform collection overlap analysis through the use of the library's link resolver. This is because link resolvers not only track the partial or full title runs available within each aggregated resource, but they also contain information about each of the packages to which the library subscribes. The SFX resolver from Ex Libris - <http://www.exlibrisgroup.com/sfx.htm> - offers overlap analysis tools for subscribers, as does Serial Solutions' Article Linker - <http://www.serialssolutions.com/overlap.asp.>.

The library can also take advantage of Serials Solutions' overlap analysis tools if it subscribes to the e-journal Full MARC Records service, whether or not Article Linker has also been implemented. Overlap analysis reports provide data on the duplication of titles in subscribed packages and help determine which packages may be de-selected.

Another important feature of these tools is the ability to evaluate new aggregated packages being considered for purchase. The overlap analysis tool makes it possible to compare the holdings of any available package with those to which the library has current subscriptions. It is able to generate reports that indicate the percentage of partial and full overlap with current electronic subscriptions, as well as the percentage of unique holdings within the package under consideration. It is possible to drill down to the title-level to see which titles are completely unique and which will be duplicated in full or in part.

Federated Search

Federated search (or metasearch) tools have been adopted by many libraries as part of their resource discovery arsenals. Some popular metasearch tools for libraries include Endeavor's ENCompass, WebFeat's Prism, and Ex Libris' MetaLib. All of these tools send simultaneous searches to multiple distributed database targets and return results to the user in a single, consistent results interface. These tools provide single search access to groups of databases clustered around specific subjects or resource types, and they may also offer sorting and de-duping functions.

Metasearch applications generate two types of logs: Web server logs and metasearch application logs. Both may hold valuable information about the ways in which the tools are used. Web server logs, as we know, can provide information about the times and dates of searches, user IPs, the number and length of sessions, the types of queries, and the items that are viewed. The metasearch application's internal logs may also contain some or all of this information, but are more likely to organize data collection around metrics specific to the metasearch process:

- number of targets searched simultaneously
- number of simultaneous sessions

- speed with which results are returned from different targets
- user profile information from the authentication database.

Most metasearch engines provide some kind of reporting interface, although the types of reports available vary from product to product and depend on the way the application logs have been configured. WebFeat's SMART reporting interface, <http://www.Webfeat.org/products/smart.htm>, offers one of the most full-featured reporting tools available for metasearch applications. SMART is COUNTER-compliant, and provides database activity reports that include searches conducted from the federated search interface, as well as those executed in the native interfaces of the databases. Information available from metasearch logs includes information on user sessions, search activity, and database activity:

User session information:

- number of successful and failed logins
- session length
- user IP address
- browser and operating system
- average number of simultaneous active sessions (by date and time)
- number of session timeouts

Search activity:

- total number of searches over a period
- number of targets searched simultaneously per session
- search queries

Database activity:

- number of searches conducted against each specific target
- speed with which results are returned from each database
- number of database turnaways or failed searches
- number of hits returned from each target database
- number of times full record requests were made from each target database
- number of times a 'next set' request was made from a database (Database result sets are usually truncated in metasearch result sets, so users can request 'next set of *n* results' from a specific database.)
- number of times users clicked through to the native database interface
- number of full record requests made for a specific journal title, article title, author, database, or ISSN.

Institutional Repositories and Use Data

Most of the data sources discussed so far are those that help evaluate the use of electronic resources hosted remotely and delivered by third-party vendors. Many libraries have electronic collections created and stored locally. Digital archives and institutional repositories (IRs) are two ways that libraries have adopted an expanded role as the custodians and gatekeepers of local electronic collections with unique or rare materials.

These collections may be the result of large and time-consuming digitization efforts within the library. They may contain theses, dissertations, pre-prints, presentations, and other scholarly materials produced within the larger institution. Some of these collections provide a means for institutions to organize, search, and access administrative documents, annual reports, policy papers, and the results of internal studies. All of these local collections represent a significant investment of time and money, and librarians must determine the value of these investments in terms of increased visibility and use.

Some commonly used digital repository platforms are:

- DSpace, an open source platform originally developed by MIT Libraries and Hewlett-Packard <http://www.dspace.org>
- GNU ePrints, open source repository software <http://www.eprints.org>
- ContentDM, a commercial software package offered by OCLC. <http://www.oclc.org/contentdm/>

Activity data on these systems is held in both Web server logs (for end-user access data) and system logs (to record transaction information about staff workflows).

Reporting tools are still under development for each of these systems, as are metrics to assess the use and effectiveness of digital repositories. Because search and access information can be captured in standard Web logs, Web analysis tools can be implemented to uncover standard measures of end-user activity. This includes data about the most heavily used collections, frequently accessed objects, number of items viewed during a typical session, activity by time and date, search strategies, referring URLs and search engines, geographic location of user-based IP addresses, and file downloads.

The current version of DSpace uses a Java-based logging utility called log4j to generate system activity logs as plain text files, although add-on tools to generate SQL and XML versions of the logs are under discussion and development by members of the DSpace community. CONTENTdm administrative activity is logged in a SQL database, and custom administrative activity reports can be created through the use of Crystal Reports or other SQL query software. The CONTENTdm 4.0 administrative interface provides canned reports for some of the more common usage measures, such as those in Tables 2-6 and 2-7, identifying which hits on collections / objects and user search strategies respectively.

CONTENTdm Server Report Collections Accessed from 2006-09-01 to 2006-10-31		
Collection	**Image**	**Times Viewed**
Centre for Newfoundland Studies - Digitized Maps	A general chart of the island of Newfoundland with the rocks...	36
Centre for Newfoundland Studies - Digitized Maps	Tierra nueva (B) [cartographic material].	26
Centre for Newfoundland Studies - Digitized Maps	A chart of the banks and harbours of Newfoundland. --	21
Maritime History Archive - Forbes Family fonds	Unidentified fraternal group standing in front of a large …	19
Maritime History Archive - Forbes Family fonds	"August 1925"	19
Maritime History Archive - Grenfell Photograph Collection	A Grenfell Mission nurse standing outside in a garden	18
Maritime History Archive - Grenfell Photograph Collection	A nurse and three injured children sitting on the front step...	18
Maritime History Archive - International Grenfell Association...	Grenfell mission building with the sign "Faith Hope and L...	17
Centre for Newfoundland Studies - Digitized Maps	A new map of the only useful and frequented part of New Foun...	16

Table 8. Collections and objects accessed report. Generated from CONTENTdm 4.0 Administration Interface

CONTENTdm Server Report Top Search Terms from 2006-06-01 to 2006-08-30	
Search Terms	**Times Searched**
grenfell mission st. anthony	57
bell island	43
st anthony husky	29
tilley	24
st anthony	17
harbour grace	15
eilley	14
avila	12

Table 9. Top Search Terms Report. Generated from CONTENTdm 4.0 administrative interface.

Collection metadata is held in the repository database and can generate statistics about the character of the library's digital collections, such as the total number of objects, the size of each individual collection, the number and extent of compound objects, dates when each of those objects was created, the number of null items (items with no associated file), and breakdowns of the numbers of items by file type (e.g., JPG, PDF, TIFF, MP3, or WMV).

Some libraries have created custom tools in order to perform more detailed analysis. An example is Oregon State University's CONTENTdm Controlled Vocabulary Analyzer, <http://oregonstate.edu/~reeset/contentdm/downloads.html>. The tool generates reports

on controlled vocabulary usage across all of a library's CONTENTdm collections. Resulting data can be used to identify conflicting terms within a controlled vocabulary or to understand the scope of different collections by indicating the frequency with which certain terms have been applied.

End-user searches and views are contained in Web server log data, but other logs exist that describe staff-side activity on the system. This data quantifies the amount of time documents spend at each phase of the submission and staging process, the number of times each document is touched by a staff member during the process, staff login counts by date/time and session duration, and other workflow information that can be analysed to identify bottlenecks affecting document turn-around time.

One of the interesting features of IRs is that libraries can attach administrative and statistical information about each digital object as a linked part of the record for that object. In this way statistical and processing information can become part of the metadata for a specific object. A good example of this can be found in the ePrints repositories of the University of Melbourne (<http://eprints.unimelb.edu.au/>) and the University of Tasmania (<http://eprints.utas.edu.au/>). These institutions have developed a joint program to provide links to statistical information from each ePrints record, allowing users to view the download statistics for that particular item.

A second interesting attribute of IRs is the ability to store entire reports within the digital repository system as discrete objects with their own metadata. The digital repository can act as the repository for log and report information. If these logs have a common XML format, they can also be exposed for harvesting by other IR systems using a protocol developed by the Open Archives Initiative (OAI). By harvesting log data from various IR systems, libraries will be able to compare IR content and usage data across different institutions. The Interoperable Repository Statistics (IRS) Project, <http://irs.eprints.org/about.html>, is an initiative that seeks to promote data sharing through the use of OAI harvesting.

Electronic Reference Services

There are several different approaches to electronic reference services:

- e-mail/Web-based asynchronous reference
- chat/synchronous reference through the use of instant messaging clients, e.g., AIM, MSN Messenger, and Meebo
- full-featured co-browsing systems, e.g., OCLC's QuestionPoint or Ask A Librarian from Tutor.com

The data available from virtual reference software depends on the type of service and the data the library requests from users before allowing them to initiate a virtual reference session. Identifiers for patron and responder are available in transcripts, logs, or e-mail headers. These may take the form of a login name, an e-mail address, or a chat display name. Because reference chat logs and transcripts often contain personally identifying data, librarians need to strip this kind of information before beginning data analysis.

Randomizing programs can be written to change meaningful strings such as e-mail addresses to meaningless series of characters, even while preserving a unique identifier for each string. In this way repeat visits can be tracked without exposing a patron's identity. Through the collection of other information, (e. g., patron-level, location, and institutional affiliation), patron types can be correlated with query types and subjects.

Date/time information is also usually available. For e-mail-based reference, this data would be included in the message header; in chat services the date, hour, minute, and second of each line in the correspondence is logged in the session transcript. Librarians can use this data to determine the frequency of electronic reference queries, and the distribution of queries over the days of the week and hours of the day. Date/time stamps can also be used to evaluate turn-around times for electronic reference services.

All types of electronic reference products have the capability of saving a complete transcript of the conversation between the librarian and the patron. While email and virtual reference software keep a transcript automatically, most instant messenger (IM) clients have to be specifically configured to log session transcripts. Librarians may use transcripts to evaluate the type of questions received (e.g., hours, services, known-item queries, technical help, or research help), the types of sources to which users have been referred (e.g., Web pages, indexes, guides, or reference sources), as well as a subject breakdown of research questions received (e.g., business, engineering, or medicine).

Because these transcripts are manifest as unstructured natural language questions and answers, human mediation is usually required to evaluate and characterize the interactions they record. Qualitative analysis software such as ATLAS.ti, <http://www.atlasti.com/>, or QSR NVivo, <http://www.qsrinternational.com/products/productoverview/NVivo_7.htm>, is designed to analyse 'soft' data (i.e., unstructured text files) that lack a controlled vocabulary. This type of software can be useful to librarians who wish to undertake in-depth analyses of large numbers of virtual reference transcripts.

Full-featured virtual reference applications also offer the option for patron exit surveys at the end of each session, where user feedback can be solicited regarding the usefulness of the answer received and the software's ease of use.

The Digital Reference Electronic Warehouse (DREW), under development by Scott Nicholson and R. David Lankes from the University of Syracuse, is a project that aims to create an archive of cross-institutional digital reference transactions. DREW architects are developing an XML schema to represent transaction information from e-mail, IM, and commercial virtual reference products. By bringing many similarly structured transactions together into a single space, Nicholson and Lankes hope to produce a rich dataset to which Online Analytical Processing (OLAP) and data-mining tools can be applied to "discover patterns in users, experts, question metadata, and works referred to by experts in the answer" (Lankes and Nicholson, 2005). The purpose of DREW is to provide a collaborative research space, management information system, and toolset, which can be used to understand and improve digital reference services. By amalgamating virtual reference data from many different libraries, DREW's developers

aim to provide a tool that allows us to better understand the character of electronic reference and to tailor our services more closely to patron needs.

ILS Data Sources

The integrated library system (ILS) holds a great deal of rich and detailed data about many of the library's key operations. Information about funds, vendors, spending trends, number of items catalogued within specific call-number ranges, number of items added to the collection each month, turn-around times for item processing, and a great deal of other administrative data is held in the ILS database. ILS logs are even more useful, as they contain details about every transaction that has ever transpired in the system, whether or not that data exists in the current production database. A patron or vendor record may have been removed from the database, but the existence of that account and all of the activity attached to it will still be contained in the server history logs.

The ILS circulation module links to patron records that may contain:

- ZIP or postal code
- primary branch affiliation
- departmental affiliation
- major area of study
- status as undergraduate, graduate, faculty, or alumni
- date of birth

This data from the ILS circulation records allows the study of patron activity in a much more granular fashion than might be possible in Web server logs, which frequently record no identifying information about a user.

ILS logs, like most other types of logs, are held in dense, largely inaccessible text files on the ILS server. The raw log data is encoded through the use of proprietary vendor data codes, unlike Web logs that follow internationally recognized standards for encoding data. This means that a vendor-provided translation tool or datacode manual will be necessary in order to make sense of the log data. The following is a sample of log data from the MUN libraries' SirsiDynix Unicorn ILS:

```
E200510280830240058R^S36IVFFLGODDARD^FcNONE^FEQEII^IQXX(1727366.1)^NI31
^Nz33^daLT^ND32^NH245^NG10^NEArchâeologie du vin et de l'huile dans
l'Empire romain /|cJean-Pierre Brun.^OM^^O
E200510280830250058R^S37IVFFLGODDARD^FcNONE^FEQEII^IQXX(1727366.1)^NI32
^Nz34^daLT^ND33^NH260^NG^NEParis :|bErrance,|cc2004.^OM^^O
E200312121523590020I ^@38hDIQT 174.5 C64 1983^I4T 174.5 C64
1983^IS1^NQ31162004594801^NSQEII^ILCHECKEDOUT^ltnqCHARGED^^Z
E200312121523590020F ^5nrN^jxN^CKLISA^CLLisa's test
account^CA12/12/2003,15:23^CI10/4/2004,23:59^CX4MUR$2^CGCHECKEDOUT^IACo
llingridge, David.^IBTechnology in the policy process : controlling
nuclear power / David Collingridge. --^^O
E200312121523590020R^S39ZnFFSIRSI^FcNONE^FEQEII^MI1^MO2^MLitem_checked_
out,due^^O
E200312121523590020F ^@39ZnMI1^MQENGLISH^MPY^MM002C~0
item_checked_out||item checked out|due||due|^^O
```

There is an ethical concern about accessing raw ILS log files. As evident in the example above, the circulation record in the raw log file contains elements that disclose the identity of individuals and the particular items they borrow. Because librarians are concerned about privacy issues, many ILS vendors provide tools to strip logs of patron barcodes and other data that might allow the identification of specific users while retaining broader profile information for the purpose of statistical analysis. Ethical concerns dictate that librarians undertaking ILS log analyses should ensure that user IDs, barcodes, and personal information are removed from the raw data before making logs available for study of any kind. This is even more critical if the logs are to be duplicated in more than one location or shared with several people for the purpose of analysis.

Beyond ILS history logs there are Web logs, a second-level of logging for the Web-based OPAC interface. Web logs, like all of the other Web server logs discussed here, record the query portion of searches executed and may also indicate the frequency with which patrons choose to browse or keyword search a particular field.

```
198.165.141.237 - - [31/May/2003:10:23:19 -0230] "GET
/uhtbin/cgisirsi/x/0/5?user_id=REMOTEWEB&searchdata1=Remote+Sensing+of+
Sea+Ice HTTP/1.1" 200 69189

134.153.164.143 - - [31/May/2003:10:46:19 -0230] "GET
/uhtbin/cgisirsi/x/0/5?user_id=REMOTEWEB&searchdata1=lazarus+and+folkma
n HTTP/1.1" 200 28582

142.163.11.5 - - [31/May/2003:11:10:30 -0230] "GET
/uhtbin/cgisirsi/tyQtIrCYGa/127540009/8/104295/AMERICAN+JOURNAL+OF+DISE
ASES+OF+CHILDREN HTTP/1.1" 200 29928
```

This information can be used to analyse the catalogue search behaviour of patrons. Because these are standard Web server logs, commercial or freely-available Web log analysis software can draw out information about dates and times of heavy OPAC usage and help to isolate and cluster search strategies.

The ILS has a number of canned reports that can be used to gather data about the number of classroom reserves that are created, items catalogued, items circulated, and other baseline activity. ILS reporting systems are useful to gather aggregated data, i.e., the number of AV material circulations that occurred in a given month, but they are not optimised to allow detailed examination of that data. The limitations of ILS reporting tools have made the extraction and manipulation of the huge data stores contained in ILS logs very difficult.

A number of strategies have emerged in recent years to help librarians access and use ILS data in new ways. One approach has been the development of new reporting tools, modelled on the kinds of business intelligence systems widely adopted in the corporate world. OLAP reporting tools, such as Director’s Station from SirsiDynix, provide multi-dimensional views of library activity. These tools allow librarians to find relationships between activities that occur in different library units such as selection, acquisitions, cataloguing, and circulation. These tools are intended to help librarians identify use

patterns and trends over time. OLAP reporting tools allow librarians to view aggregated activity data and then drill down into that data for more specific information, adding and removing different dimensions to create custom views of the data. Other features include the ability to perform custom calculations on the fly, create natural language queries ("Which hour of the day and day of the week had the highest circulation?"), sort data within and across hierarchies, export report data to spreadsheets for further manipulation, and create alerts for certain thresholds of activity, e.g. create an alert when the total amount of fines owed at my branch reaches $1000.

The kinds of questions that can be answered easily using an OLAP reporting tool includes a wide range of topics:

- What is the average age of my collection of chemistry books?
- What is the distribution of publication dates in that collection?
- What is the average cost per-item in my collection of chemistry books?
- What is the distribution of chemistry book purchasing across vendors and funds?
- How many chemistry books with publication dates before 1995 circulated this year?
- What has been the spending trend over the last ten years for chemistry books?
- Which are the top one hundred circulating chemistry books over the past three years?
- What percentage of my chemistry collection circulated to undergraduates, graduates, or faculty members over the past five years?
- What is the distribution of those borrowers across university departments?
- How does circulation per-item in the chemistry collection compare with that of the general collection?

ILS Data & Cross-Library Comparison Tools

The tools identified above are those that help evaluate collection use in a single library or library system. Vendors are also developing tools, such as the Normative Data Project (NDP) or OCLC's WorldMap that allow ILS data to be integrated with GIS programs for demographic analysis. OCLC has several data mining projects that attempt to leverage its WorldCat data for cross-library comparison. These include the Systemwide Print Book Collection study, the Comparative Collection Assessment project, and the OCLC Mining for Digital Resources study

The Normative Data Project (NDP), <http://www.libraryndp.info/>, combines GIS-based demographic information with ILS data to aid library planners and decision-makers. This project combines transaction and bibliographic data extracted from the ILS logs of a number of different libraries, with census data, and U.S. library data from the National Center for Education Statistics. The NDP project seeks to extract and normalize ILS data from many different vendor platforms to provide a standard view of data for cross-library data sharing and comparison. Understanding which titles, formats, and languages are in heavy demand by different categories of patrons in comparable libraries can help predict which resources will be used by your own patrons.

OCLC's GIS-based information system, WorldMap, seeks to gather key statistical information from libraries around the world, including holdings and titles by place of publication, types of libraries, number of libraries, librarians, users, volumes, and annual expenditures. WorldMap provides an intuitive graphical entry point for comparing titles and holdings at national and regional levels. It also permits regional comparisons on staffing, expenditures, and accessibility of library services.

OCLC's WorldCat database contains more than 65 million bibliographic records representing over 950 million holdings from more than 20 million libraries around the world. This creates an invaluable data source by which libraries can evaluate their own collections in a larger context, and opens up opportunities for cross-library collection analysis, and cooperative regional collection development. OCLC's WorldCat Collection Analysis service allows libraries to compare their own holdings against those of peer-group libraries to identify subject-matter strengths, gaps, and overlaps.

OCLC has several other current ILS data-mining projects. The Systemwide Print Book Collection study analyses aggregate print book holdings from WorldCat to identify rare or unique materials in individual library collections, as well as titles that are widely available within a particular region. This information can help libraries target titles that are widely available in other collections and therefore easily procured through inter-library loan. OCLC's Comparative Collection Assessment project incorporates holdings information with ILL usage information to help libraries identify collection strategies for print and e-book collections. The OCLC Mining for Digital Resources study aims to analyse holdings patterns, material types, and cataloguing activity for digital resources across library collections to identify trends and best practices for collecting and organizing digital materials.

Conclusion

This article has focused on data sources that are likely to be immediately available to the typical practitioner who wishes to engage in statistical analysis of collections and services within his or her own library. There is an enormous amount of quantitative data available to support evidence based studies, and a surprising amount of that data is either collected within the library, or readily accessible through standard reporting tools offered by service and resource providers. This data can help to identify problems with existing interfaces and tools, to identify and analyze search strategies across different types of resources, to evaluate the frequency and type of use of existing services and tools, to evaluate service turn-around times, to identify high-demand resources, to identify infrequently used products and services, to identify new services and resources which may be in demand, to identify the ebb and flow of demand by month, day or hour, to discover more about the behaviour of specific patron groups, to analyze information requests according to subject groupings, and to make cross-branch or cross-library comparisons on all of those measures.

Beyond the need to locate and access data sources, a number of further challenges face libraries in the area of quantitative analysis. Libraries require well-developed standards for reporting use data to allow for more accurate comparisons across similar vendor

products, services and resources. Libraries need to aggregate data that currently tends to be available in silos according to specific applications or providers. We must combine different kinds of use data into common spaces that will allow us to find correlations between evidence from a number of sources, and to identify trends that cut across different services and applications. It is fairly easy to extract data to track a single patron's behaviour during a single session and across a single type of resource or service. It is extremely difficult to get a holistic view of how a single patron interacted with all of the possible resources during a single information-seeking session. That patron may have started at the library's Web site, conducted an OPAC search, linked to the metasearch interface from within the OPAC, linked to a record in a vendor's native interface via the proxy server, from there clicked the resolver button to find local holdings for an article of interest, and finally viewed the full-text of the resource on the site of an e-journal vendor. Ideally we will develop tools to track, analyse and correlate all of this activity in a single data space (a data farm or data warehouse) that contains views of log data from all of the available products and services. The scope of work is even more daunting when one imagines trying to create a standardised warehouse for data sources across a number of different libraries. Lastly, of course, metrics and key performance indicators have not been clearly defined for many of our electronic resources, so even with accessible data, it is difficult to know exactly how to best measure certain activities, and what inferences confidently be made.

Resources for Log Analysis in Libraries

Proxy Servers:

EZproxy Discussion List. 2000-2006. State University of New York. 14 Nov. 2006. <http://ls.suny.edu/read/search/results?forum=EZproxy&words=LogSPU>.

Zagar, Chris. "EZproxy Support: Log Format." 2006. Useful Utilities. 14 Nov. 2006. <http://www.EZproxy.com/support/cfg/logformat/>.

Use Data from Vendors

COUNTER. Counting Online Usage of Networked Electronic Resources. 2006. COUNTER. 17 Nov. 2006 <http://www.projectcounter.org/>

Ex Libris. "Ex Libris Verde ERM System Successfully Tests SUSHI to Deliver Raw Usage Statistics," Library Technology Guides. 23 Nov. 2005. 27 Feb. 2007 <http://www.librarytechnology.org/ltg-displaytext.pl?RC=11684>.

Hiott, Judith. "Collecting and Using Networked Statistics: Current Status, Future Goals," Library Quarterly 74.4 (Oct. 2004): 441-54.

Nicholas, David, Paul Huntington, Hamid R. Jamali, and Anthony Watkinson. "The Information Seeking Behaviour of the Users of Digital Scholarly Journals." Information Processing & Management 42.5 (2006): 1345-65.

NISO Standardized Usage Statistics Harvesting Initiative (SUSHI). 2005. NISO. 17 Nov. 2006 <http://www.niso.org/committees/SUSHI/SUSHI_comm.html>.

Link Resolvers:

Bollen, Johan, Oren Beit-Arie, and Herbert Van de Sompel. "The bX Project: Federating and Mining Usage Logs from Linking Servers." CNI Fall 2005, 5-6 Dec. 2005, Phoenix, AZ. 27 Feb. 2007 <http://library.lanl.gov/cgi-bin/getfile?LA-UR-05-9439.pdf>.

CUFTS Overlap Analysis Tool. 27 Feb. 2007 <http://cufts2.lib.sfu.ca/MaintTool/public/compare>.

Federated Search:

Chandler, Sarah, and Nan Hyland. "Federated Searching and Academic Libraries: One Size Fits All?" 226th American Chemical Society (ACS) National Meeting, New York City. 8 Sept. 2003. 27 Feb. 2007 <http://hdl.handle.net/1813/3040>.

WebFeat. SMART. 27 Feb. 2007 <http://www.webfeat.org/products/smart.htm>.

WebFeat. SMART: Statistical Measures Available Real Time. April 2005. 27 Feb. 2007 <http://www.carli.illinois.edu/uploads/File/wf/WebFeat-SMART-Detail.pdf>.

Institutional Repositories:

DiMeMa. CONTENTdm Online Help, Version 4. "Reports." 2006. 27 Feb. 2007 <http://www.dimema.com/help4/server-admin/reports.html>.

DSpace Federation. DSpace Wiki: DSpaceStatisics. 16 Nov. 2006. 27 Feb. 2007 <http://wiki.dspace.org/index.php/DspaceStatistics>.

Frost, Jim, and Henry Pisciotta. Penn State University Libraries. "Visual Image User Study (VIUS): VIUS REPORTS 8.7- Authentication Logs." Oct. 2003. 27 Feb. 2007 <http://www.libraries.psu.edu/vius/8.7.pdf>.

Organ, Michael. "Download Statistics - What Do They Tell Us? The Example of Research Online, the Open Access Institutional Repository at the University of Wollongong, Australia," D-Lib Magazine 12.11 (2006). 27 Feb. 2007 <http://www.dlib.org/dlib/november06/organ/11organ.html>.

Reese, Terry. Oregon State University CONTENTdm Development Projects/Resource Page: Oregon State University CONTENTdm Controlled Vocabulary Analyzer (Aug. 2006). 27 Feb. 2007 <http://oregonstate.edu/~reeset/contentdm/downloads.html>.

Sale, Arthur, and Christian McGee. Tasmania Statistics Software Download (2006). 27 Feb. 2007 <http://eprints.utas.edu.au/262/>.

University of Southampton, UK, and Key Perspectives Ltd., UK. “IRS: Interoperable Repository Statistics.” 3 Mar. 2005. 27 Feb. 2007 <http://irs.eprints.org/about.html>.

Electronic Reference Services :

Lankes, David L. and Nicholson, Scott. “Creating the Infrastructure for Collaboration between Digital Reference Services and Researchers: The Digital Reference Electronic Warehouse (DREW) Project,” JCDL ’05, 7-11 June 2005, Denver, CO. <http://portal.acm.org/citation.cfm?id=1065497&dl=acm&coll=&CFID=15151515&CFTOKEN=6184618>

Marsteller, Matt, and Paul Neuhaus. “The Chat Reference Experience at Carnegie Mellon University.” Poster Presentation. Annual Meeting, American Library Association. 16 June 2001. San Franciso, CA. 27 Feb. 2007 <http://www.contrib.andrew.cmu.edu/~matthewm/ALA_2001_chat.html>.

ILS Data Sources :

Blecic, Deborah D., Nirmala S. Bangalore, Josephine L. Dorsch, Cynthia L. Henderson, Melissa H. Koenig, and Ann C. Weller. “Using Transaction Log Analysis to Improve OPAC Retrieval Results,” College & Research Libraries 59.1(Jan.1998): 39-50.

Cooper, Michael D. “Usage Patterns of a Web-Based Library Catalog,” Journal of the American Society for Information Science and Technology 52.2 (Jan. 2001): 137-48.

Cummins, Caroline. “Below the Surface,” Library Journal Net Connect 15 Jan. 2006. 27 Feb. 2007 <http://www.libraryjournal.com/article/CA6298559.html>.

Hamilton, Scott, and Helen Thurlow. “Transaction Log Analysis @ State Library of Queensland,” 3rd International Evidence Based Librarianship Conference. Brisbane, Queensland, 17 Oct. 2005. 27 Feb. 2007 <http://conferences.alia.org.au/ebl2005/Hamilton.pdf>.

Lavoie, Brian, and Roger Schonfeld. “A Systemwide View of Library Collections.” OCLC Online Computer Library Center and Ithaka. CNI Spring Task Force Meeting. 5 April 2005. 27 Feb. 2007 <http://www.oclc.org/research/presentations/lavoie/cni2005.ppt>.

Lavoie, Brian, Lorcan Dempsey, and Lynn Silipigni Connaway. “Making Data Work Harder,” Library Journal Net Connect. 15 Jan. 2006. 27 Feb. 2007 <http://www.libraryjournal.com/article/CA6298444.html>.

Nicholson, Scott. “Proof in the Pattern,” Library Journal Net Connect 131(Jan. 2006): 2-6. 27 Feb. 2007 <http://www.libraryjournal.com/article/CA6298562.html>.

Normative Data Project for Libraries (NDP). 27 Feb. 2007 <http://www.libraryndp.info/>.

OCLC. Data Mining Research Area. 2007. 27 Feb. 2007 <http://www.oclc.org/research/projects/mining/>.

TLC. The Library Corporation. "Press Release: TLC Partners with Civic Technologies, Inc." 4 May 2004. 27 Feb. 2007 <http://www.tlcdelivers.com/tlc/press/pr050404.asp>.

General Resources for Quantitative Analysis:

Association of Research Libraries. Association of Research Libraries Statistics and Measurement Department. MINES for Libraries: Measuring the Impact of Networked Electronic Services. 2006. http://www.minesforlibraries.org/mines/index.cfm>.

Borgman, Christine L., and Sandra G. Hirsh. "Rethinking Online Monitoring Methods for Information Retrieval Systems: From Search Product to Search Process," Journal of the American Society for Information Science 47.7 (Jul. 1996): 568-83.

Jansen, Bernard J. "Search Log Analysis: What Is It, What's Been Done, How to Do It," Library & Information Science Research 28.3 (June 2006): 407-32.

Nicholson, Scott. "Approaching Librarianship from the Data: Using Bibliomining for Evidence-Based Librarianship," Library Hi-Tech 24.3(2006): 369-75.

Nicholson, Scott. "The Basis for Bibliomining: Frameworks for Bringing Together Usage-Based Data Mining and Bibliometrics Through Data Warehousing in Digital Library Services," Information Processing and Management 42.3 (May 2006): 785–804.

Zucca, Joe. "Traces in the Clickstream: Early Work on a Management Information Repository at the University of Pennsylvania." Information Technology and Libraries 22 .4(Dec. 2003): 175-9.

Software

Article Linker c. 2003. Serial Solutions. Seattle, WA. Feb 28, 2007. <http://www.serialssolutions.com/overlap.asp.>

Ask A Librarian 2.2 c. 2007 Tutor.com. New York, NY. Feb 28, 2007. <http://www.tutor.com/products/aal.aspx>

ATLAS.ti 6 c.2007. Berlin, Germany. Feb. 28, 2007 <http://www.atlasti.com/>

ContentDM 4 c. 2006. OCLC. Dublin, OH. Feb. 28, 2007.
<http://www.oclc.org/contentdm/>

CUFTS 2 c. 2007. Council of Prairie and Pacific University Libraries (COPPUL). Victoria, BC. Feb. 28, 2007. <http://cufts2.lib.sfu.ca/MaintTool/public/compare>.

Director's Station 3.8 c. 2006. SirsiDynix. Hunstville, AL. Feb. 28, 2007.
<http://www.sirsidynix.com/Solutions/Products/analytical.php>

DSpace 1.4 c. 2006. Originally developed by MIT Libraries and Hewlett-Packard. Released under the GNU General Public License. Feb. 28, 2007.
<http://www.dspace.org>

EZproxy 4.0 c. 2006. Useful Utilities. Peoria, AZ. Feb 28, 2007.
<http://www.EZproxy.com>

EPrints 3 c. 2006. Lead Developer, Christopher Gutteridge. Released under the GNU General Public License.Feb. 28, 2007. <http://www.eprints.org>

MetaLib 4 c. 2007. Ex Libris Group. Des Plaines, IL. Feb. 28, 2007.
<http://www.exlibrisgroup.com/metalib.htm>

WebFeat 3 c. 2005. Old Brookville, NY. Feb 28, 2007.
<http://www.webfeat.org/products/prism.htm>

NVivo 7 c. 2006. QSR International Pty Ltd. Australia. Fe. 23, 2007.
<http://www.qsrinternational.com/products/productoverview/NVivo_7.htm>,

QuestionPoint c. 2006. OCLC. Dublin, OH. Feb. 28, 2007.
<http://www.oclc.org/questionpoint/default.htm>

SFX 3 c. 2005. Ex Libris Group. Des Plaines, IL. Feb. 28, 2007.
<http://www.exlibrisgroup.com/sfx.htm>

Verde 2.0 c. 2006. Ex Libris Group. Des Plaines, IL. Feb. 28, 2007.
<http://www.exlibrisgroup.com/verde.htm>