Joint Information Systems Committee Project Report

Digital Repository Management Practices, User Needs and Potential Users: An Integrated Analysis

Alesia Zuccala[1], Mike Thelwall[1], Charles Oppenheim[2], and Rajveen Dhiensa[2]

[1]School of Computing & Information Technology
University of Wolverhampton
MI Building, City Campus-South
Wulfruna Street, Wolverhampton, WV1 1SB
United Kingdom

[2]Department of Information Science,
Loughborough University,
Loughborough, Leicestershire, LE11 3TU
United Kingdom

31 August 2006

Table of Contents

1. Acknowledgements

The authors would like to thank the JISC Digital Repositories programme for funding this project and Dr. Neil Jacobs for his professional guidance and support. We would also like to express our gratitude to the repository managers who gave us their valuable time when invited to participate in an interview.

2. Executive Summary

This project was designed to evaluate five different types of public repositories in the UK, a digital library (The National electronic Library for Health), a subject repository (CogPrints), an institutional repository (e-Prints SOTON), a data repository (UK Data Archive), and an e-learning repository (Jorum). Our project team's approach was to carry out the evaluation using a set of qualitative and quantitative research techniques, including webometric analyses of links to the selected repository Web sites, interviews with repository managers, and an online survey of repository users. Unique Web intelligence reports were given to the repository managers to demonstrate how links to and co-links with their Website can be used to monitor and/or assess their current Web network positioning. The overall aim was to help repository managers to better understand and address user community needs, and work towards building stronger links (i.e., Web and real-world links) between themselves and other relevant organisations or activities, at both a national and international level.

The webometric reports, generated from our LexiURL software tool, were of interest to the interviewed managers. Positive feedback was given regarding the colink data network maps since this was not something the managers had seen before or considered for assessment purposes. Some managers were surprised to see a significant number of international links to their resource. Most links retrieved for the individual resources were expected. For instance, we expected the academic resources to be situated within an academic context on the Web and the link analysis results indicated that this was true for the Jorum, The UK Data Archive and E-prints SOTON sites. The results of the management interviews were also positive in that the chief managers and development team members were open to discussing the value of collaborative work, and conveying significant details concerning their publicity strategies and methods of reaching and understanding users. Our user survey results indicated that word-of-mouth networking (i.e., colleagues or friends) generates a significant amount of awareness concerning a new library/repository. Digital library /repository users who use a resource frequently (i.e., a few times per week) tend to bookmark the site on their Web interface, while less frequent users search for the resource by name. A good percentage of our surveyed users deemed the digital resources (i.e., NeLH, CogPrints, E-Prints SOTON, UK Data Archive, Jorum) easy to use and satisfactory for meeting their information needs, although a number of them put forth suggestions concerning the availability of other types of information/materials/resources. We collected some qualitative information concerning the benefits users have experienced as a result of using the digital library/repositories, and the results indicated that many benefits have been work-related, but some were unexpectedly personal. The information that a manager obtains from users, by way of interviews or surveys, is useful for understanding users and user needs; however, what users say about a resource and its benefits (i.e., success stories) can also help to improve public relations.

The most important outcome or achievement associated with this project was the introduction of LexiURL to the digital library/repository community. Since LexiURL was new at the time of this study, we hope that further results will be achieved through the use of this webometric tool in the future. Link statistics collected using LexiURL can be used to supplement a log file analyses (e.g., using *WebTrends* or another software), and based on our project findings we suggest that digital library/repository managers use the software every six months or so to uncover new links to their site, to detect colink network changes, and to search for lists of links to 'competitor' sites (i.e., other similar international libraries or repositories). With a comparative LexiURL link analysis it is possible to acquire insight into new or potential user groups: if more links or different types of links are found to be directed to the site of a similar international resource, then perhaps these links represent previously unrecognized users, or areas for further outreach or cooperation.

3. Background

In October 2004, Mike Thelwall and Alesia Zuccala, initiated a SCIT Cybermetrics consultancy unit at the School of Computing and Information Technology, University of Wolverhampton and received a consultancy contract from the UK National Archives. The members of the National Archives' Online Content and Partnership Development (OCPD) group had called upon SCIT Cybermetrics to retrieve a list of links to their website for the purpose of developing a "Web intelligence" report. A set of link statistics were presented to the OCPD group, which included a colink "map" to show them the archive's Web network positioning and help the group identify interesting sites for potential organisational partnerships.

Following this consultancy work, Prof. Thelwall and Dr. Zuccala recognized that more webometric analyses of this type needed to be carried out with digital libraries, repositories and public archives. In March 2004 they responded to a call for thematic projects concerning digital repositories, issued by the Joint Information Systems Committee (JISC). The Wolverhampton team's project proposal focused on creating a Web link analysis software tool and Web intelligence reporting system for repository managers, and referred to the definition of *repository*, which appeared in a previous JISC report written by Heery and Anderson (2005):

- Content is deposited in a repository, whether by a creator or third party
- The repository architecture manages content as well as metadata
- The repository offers a minimum set of basic services, e.g., put, get, search, access control
- The repository must be sustainable, trusted, well supported and well managed (p.2).

Prof. Thelwall and Dr. Zuccala invited Prof. Charles Oppenheim, a recognized digital repository expert, to join the project, as well as his Masters Degree student, Rajveen Dhiensa. In collaboration, the Wolverhampton-Loughborough group decided to develop a mixed-methods research project, using Web link data, interview data and survey data to carry out a full evaluation of a selection of different digital repositories.

4. Aims and Objectives

The general aim of the project was to conduct a comparative analysis of five different types of digital repositories in the UK:

- A Digital Library
  The National Electronic Library for Health (http://www.nelh.nhs.uk/)
- A Subject Specialty Repository
  Cognitive Sciences E-Print Archive: "CogPrints" (http://CogPrints.org/)
- An Institutional e-Prints Repository
  The University of Southampton E-prints Archive (http://eprints.soton.ac.uk)
- A Research Data Repository
  The UK Data Archive (http://www.data-archive.ac.uk/)
- An e-Learning Repository
  The Jorum Learning Repository (http://www.jorum.ac.uk/)

All five of the repositories were evaluated from a *webometric* perspective (described in section 5.1), and further assessed with in-depth management interviews. A user survey was also implemented on the Web to learn more about the habits of individuals who use the different repositories, including individuals who do not use them because they do not need to or might not be aware of the fact that the repositories exist.

5. Methodology

A mixed methods approach was designed for this study, using a special set of quantitative and qualitative research techniques. We chose a mixed-methods approach due to the type of issues addressed – i.e., Web network structures, repository management/service goals; user perceptions and needs. Essentially, each research technique was selected for its power to complement the other, and expose specific types of information.

The first research technique, referred to as *webometric link analysis* is described below, including the software tool, *LexiURL*, which was created in order to retrieve and organize link data for analytic purposes. We also describe our semi-structured interview approach with the repository managers, and the primary elements of our user questionnaire.

5.1 LexiURL and Web link analysis

LexiURL (see Figure 1) is a specific software programme designed to retrieve link data from search engines, like Yahoo!, Google, or AltaVista and calculate summary statistics for lists of links or URLs. Its output is a series of standard reports. Although LexiURL is a flexible, generic program, we describe its functions that are most useful for digital repository/library link analysis.

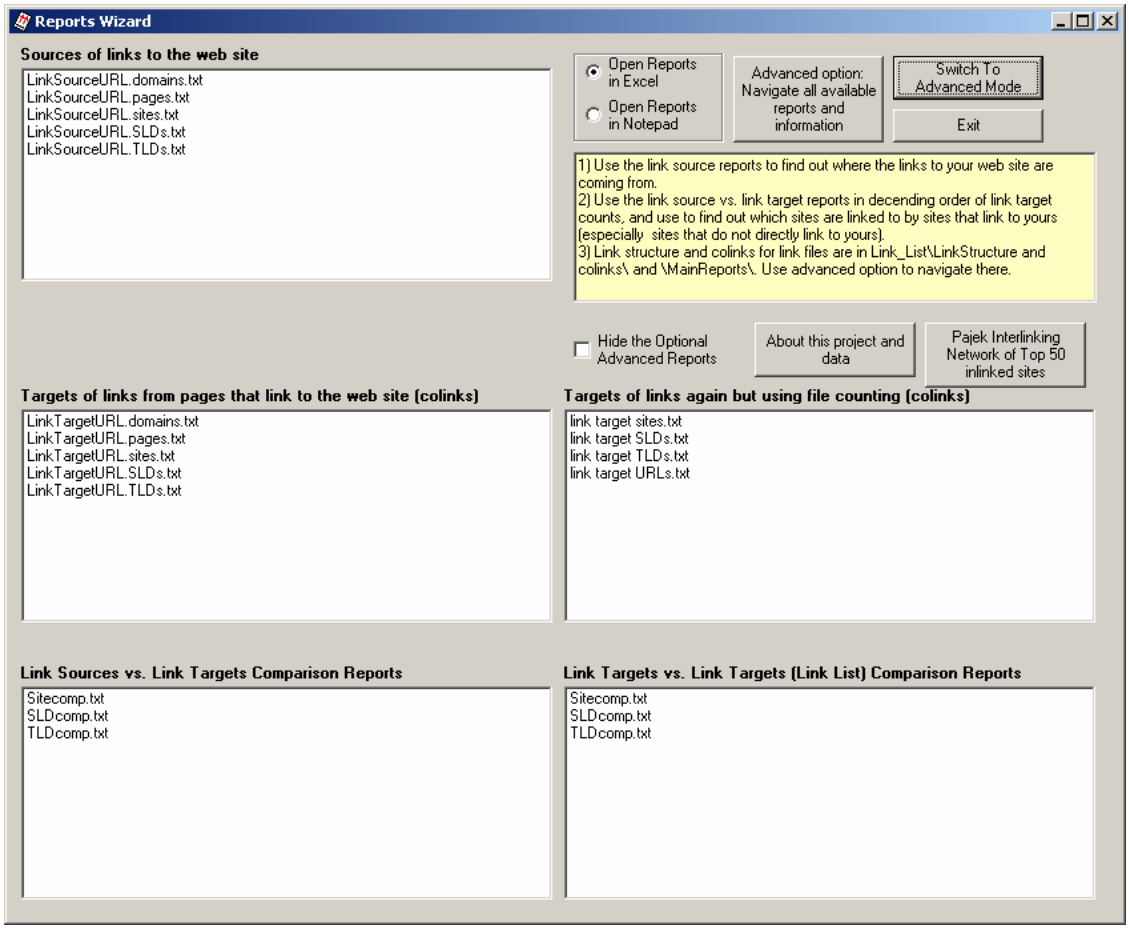


Figure 1. LexiURL Reports Wizard.

A *webometric link analysis* involves collecting link data and organizing or sorting the links for content review or context review purposes. Two types of link data can be retrieved from the Web: 1) *directed links* and 2) *colinks*. A *directed link* is either an *in-link* (a link to a web page), an *out-link* (a link from a web page), or a *self-link* (a link from a web page to the same page or to any page in the same site). Co-linked websites occur "when two pages both have inlinks from a third page" (Thelwall, 2004, p. 5).

In order to analyse all of the links to a digital repository/library web site, actual links must first be found. Although the web is clearly too large to manually browse for links, they can be identified through advanced commercial search engine searches. For example, in December, 2005 the following queries could be used for the National electronic Library for Health:

- **link:www.nelh.nhs.uk** in **Google** would list up to 1,000 pages containing a link to the page http://www.nelh.nhs.uk/ (page in-links)

- **linkdomain:nelh.nhs.uk NOT host:nelh.nhs.uk** in **Yahoo!** or **AltaVista** advanced search would list up to 1000 pages containing a link to any page in the nelh.nhs.uk web site, excluding pages within the same site (site links).

- **link:http://www.nelh.nhs.uk -site:www.nelh.nhs.uk** in **MSN** would give up to 250 hits for pages that link to the home page from outside the www.nelh.nhs.uk web site.

Note that all three of the above give different results. In particular, Google returns only links to the home page whereas Yahoo! and MSN returns links to any page in the site. Moreover, Yahoo! and MSN, unlike Google, support queries that exclude links from the same site (site self-links). This latter point is important: site self-links are normally not used in link analysis (Thelwall, 2004) because they are widely used for navigation rather than to point to resources (Bar Ilan, 2004).

In the past researchers sent queries to search engines, copied down the URLs of matching pages for all of the pages of result. This was a highly time-consuming process, but recently search engines have provided a facility for automatically extracting this data, an Applications programming Interface (API) or a Web Service.

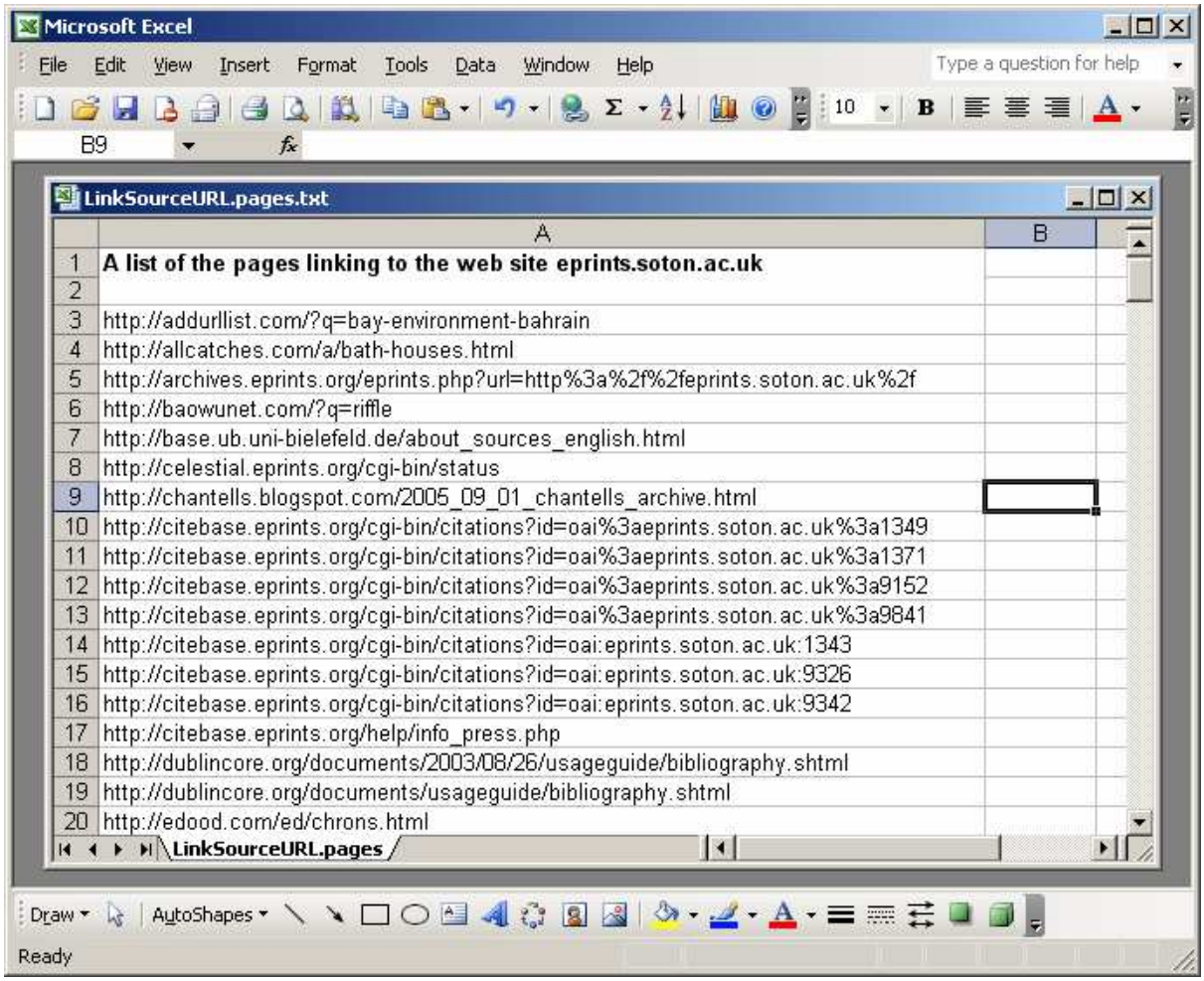
- The Google API (www.google.com/apis/) allows programmers to submit up to 1,000 queries per day.

- Yahoo! Search Web Services (developer.yahoo.net/search/web/V1/webSearch.html) allows computers to submit up to 5,000 queries per day (per IP address).

- MSN Search (msdn.microsoft.com/msn/msnsearch/) allows computers to submit 10,000 queries per day (per IP address).

LexiURL uses the Yahoo! API so that site self-links can be excluded. Researchers wishing to use a different search engine can manually construct a list URLs and feed it directly to LexiURL. If using the default Yahoo! API method then the researcher only needs to enter the home page URL or their digital library and it will form appropriate Yahoo! queries, submit them to the Yahoo! Search Web Service, and analyse the results.

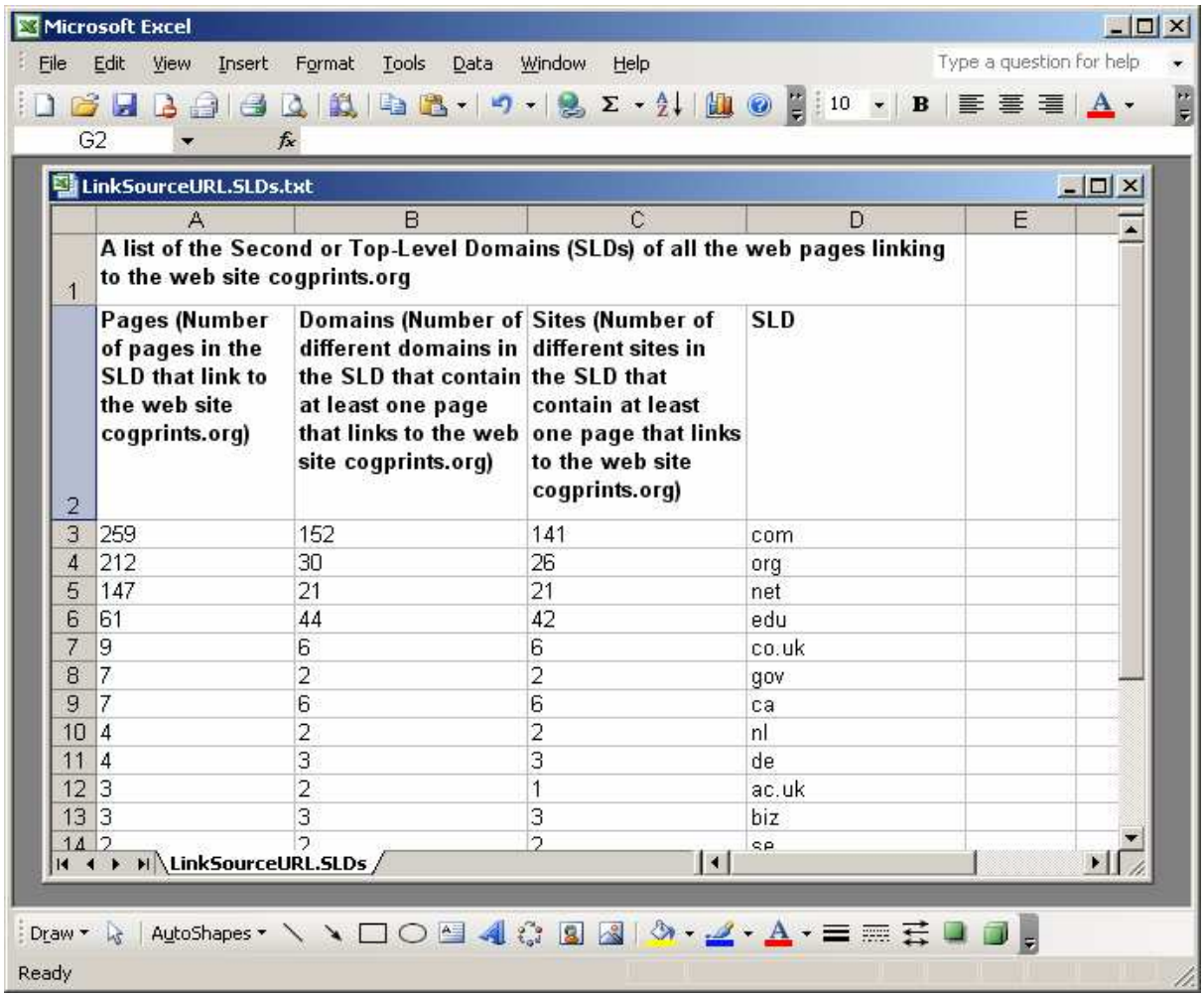
Table 1. Sample LexiURL report of pages linking to the web site www.eprints.soton.ac.uk. Links were retrieved on October 20th, 2005.



Commercial search engine data has several limitations. First, no search engine covers the entire web and so all are likely to return incomplete results. This problem is exacerbated by the typical limitation of 1,000 results per query and for Google and Yahoo! their automatic search services report only a fraction of the results known by the parent search engine. Hence for a large digital repository/library, LexiURL could be expected to find perhaps only 10% or less of the links to the site.

A second limitation is that the method by which each search engine finds pages is unknown, as is the method for ranking results, except in broad detail (Chakrabarti, 2003). This is undesirable for research since either method may introduce unwanted biases, e.g. for or against older sites, more popular pages or certain national or generic Internet domains (Vaughan & Thelwall, 2004). As a result of these issues, LexiURL's link data must be seen as a 'convenience sample' and a (possibly small) fraction of the complete set of links.

Table 2. Sample LexiURL report of Second or Top-Level Domains of all pages linking to the website www.CogPrints.org. Links were collected on October 18th, 2005.



Microsoft Excel

File Edit View Insert Format Tools Data Window Help

LinkSourceURL.SLDs.txt

	A	B	C	D	E
1	**A list of the Second or Top-Level Domains (SLDs) of all the web pages linking to the web site cogprints.org**				
2	**Pages (Number of pages in the SLD that link to the web site cogprints.org)**	**Domains (Number of different domains in the SLD that contain at least one page that links to the web site cogprints.org)**	**Sites (Number of different sites in the SLD that contain at least one page that links to the web site cogprints.org)**	**SLD**	
3	259	152	141	com	
4	212	30	26	org	
5	147	21	21	net	
6	61	44	42	edu	
7	9	6	6	co.uk	
8	7	2	2	gov	
9	7	6	6	ca	
10	4	2	2	nl	
11	4	3	3	de	
12	3	2	1	ac.uk	
13	3	3	3	biz	
14	2	2	2	se	

LinkSourceURL.SLDs

Draw AutoShapes

Ready

For any digital repository or digital library site in question LexiURL reports the following types of statistics:

- The URLs of all found inlinking pages (normally the first 999 retrieved from a search engine)
- The domains and the sites of all found inlinking pages, together with a count of the number of inlinking pages from each domain/site
- The Second Level Domain (SLD) and Top Level Domain (TLD) of all found inlinking pages together with a count of the number of inlinking pages/sites from each STLD/TLD
- A list of the top 49 websites colinked with the site, as well as matrix data for creating a multidimensional representation of the colink structure – i.e., a colink map

Tables 1 and 2, present sample reports for a typical LexiURL link extraction (see Tables 1 and 2 above). Each report can be viewed as a Microsoft Excel table, so that different sorting options can be used. For instance, it is often useful to sort the extracted list of inlinking pages alphabetically, or sort the sites linked to the digital repository according to their page count. Note that page counts must be treated with suspicion because sites sometimes include links in a standard navigation bar replicating it across the site and swamping the counts. In such cases site counts are often more robust measurements, in addition to their intrinsic reporting value.

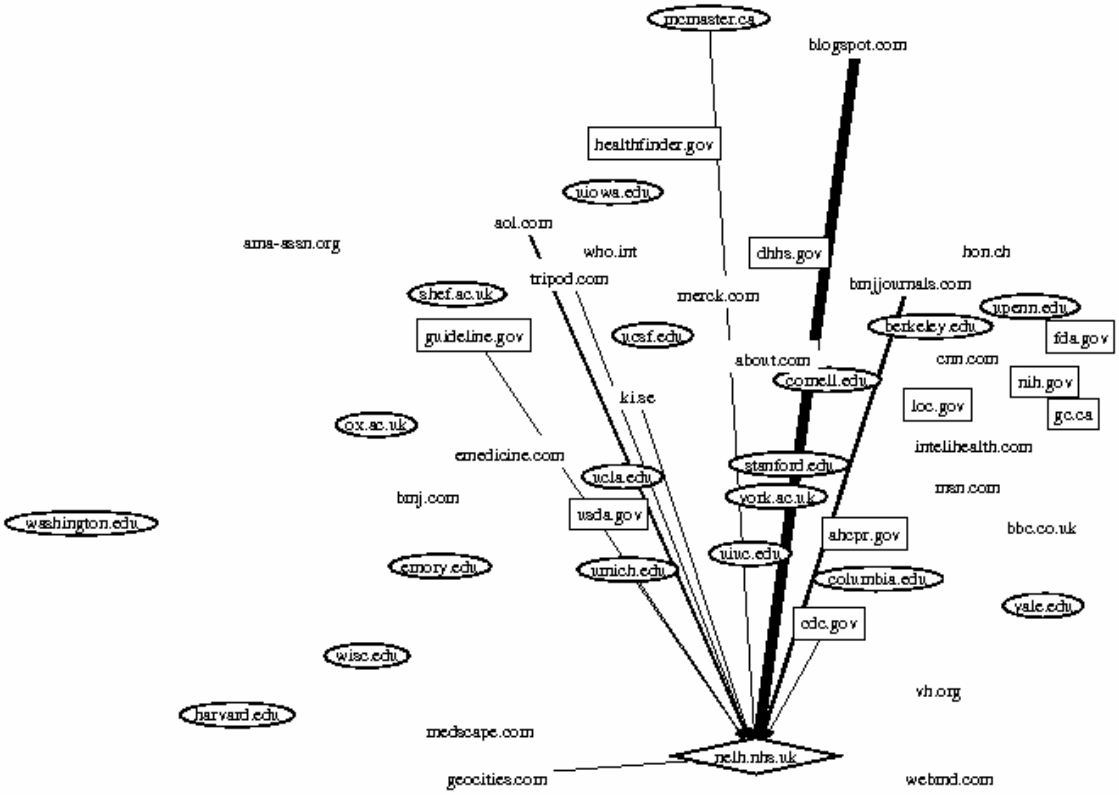
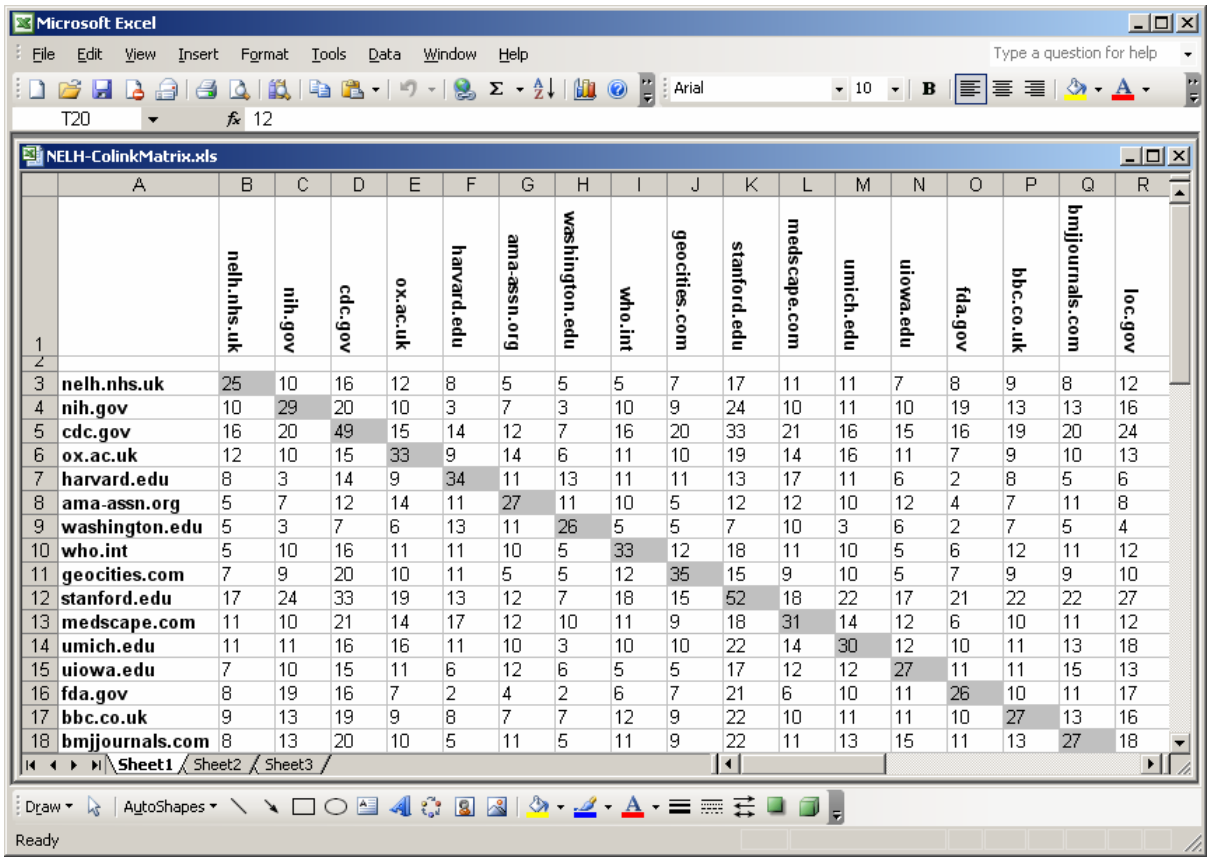


Figure 2. Top 49 sites colinked with the nelh.nhs.uk, including directed links to the NeLH (diamond) and link nodes for all educational sites (ellipses) and government sites (boxes). Links for this configuration were collected on September 20th, 2005.

Figure 2 presents a sample colink structure for the National electronic Library for Health. All educational domains (e.g., ac.uk; edu; .ca) are highlighted with ellipses and all government domains (.gov) are highlighted with boxes. Directed links to the nelh.nhs.uk site are also visible, with the strength of the link indicated by the thickness of the line. The Excel matrix data file used to create the NeLH colink map is shown in Table 3. *LexiURL* has the capacity to generate two original "reports wizard" files in plan text format – a list of the top 49 colink sites and raw data matrix file. Both files can easily be cut and pasted into Excel for reformatting, and used as input into any type of data analysis/ visualisation program (e.g., SPSS-X or the Pajek network visualization program).

Table 3. Sample colink matrix for the National electronic Library for Health (NeLH). Link data were collected on September 20th, 2005.



Microsoft Excel

File Edit View Insert Format Tools Data Window Help — Type a question for help

Arial 10 B

T20 fx 12

NELH-ColinkMatrix.xls

1	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
		nelh.nhs.uk	nih.gov	cdc.gov	ox.ac.uk	harvard.edu	ama-assn.org	washington.edu	who.int	geocities.com	stanford.edu	medscape.com	umich.edu	uiowa.edu	fda.gov	bbc.co.uk	bmjjournals.com	loc.gov
2																		
3	**nelh.nhs.uk**	25	10	16	12	8	5	5	5	7	17	11	11	7	8	9	8	12
4	**nih.gov**	10	29	20	10	3	7	3	10	9	24	10	11	10	19	13	13	16
5	**cdc.gov**	16	20	49	15	14	12	7	16	20	33	21	16	15	16	19	20	24
6	**ox.ac.uk**	12	10	15	33	9	14	6	11	10	19	14	16	11	7	9	10	13
7	**harvard.edu**	8	3	14	9	34	11	13	11	11	13	17	11	6	2	8	5	6
8	**ama-assn.org**	5	7	12	14	11	27	11	10	5	12	12	10	12	4	7	11	8
9	**washington.edu**	5	3	7	6	13	11	26	5	5	7	10	3	6	2	7	5	4
10	**who.int**	5	10	16	11	11	10	5	33	12	18	11	10	5	6	12	11	12
11	**geocities.com**	7	9	20	10	11	5	5	12	35	15	9	10	5	7	9	9	10
12	**stanford.edu**	17	24	33	19	13	12	7	18	15	52	18	22	17	21	22	22	27
13	**medscape.com**	11	10	21	14	17	12	10	11	9	18	31	14	12	6	10	11	12
14	**umich.edu**	11	11	16	16	11	10	3	10	10	22	14	30	12	10	11	13	18
15	**uiowa.edu**	7	10	15	11	6	12	6	5	5	17	12	12	27	11	11	15	13
16	**fda.gov**	8	19	16	7	2	4	2	6	7	21	6	10	11	26	10	11	17
17	**bbc.co.uk**	9	13	19	9	8	7	7	12	9	22	10	11	11	10	27	13	16
18	**bmjjournals.com**	8	13	20	10	5	11	5	11	9	22	11	13	15	11	13	27	18

Sheet1 Sheet2 Sheet3

Draw AutoShapes

Ready

5.2 Semi-structured interview schedule

A semi-structured interview schedule was prepared for each interview held with the repository service managers. Five interviews were carried out at separate times, based on the same set of questions. One advantage of a semi-structured interview was that it served as a guideline, yet allowed us some flexibility at certain moments when there was a need for further lines of questioning. Sometimes it was necessary for us to probe the interviewee for details or expand upon a particular topic for clarification.

The questions for each interview – 22 in total – focused on five thematic areas:

1) The managers' rationale for creating the repository/digital library;
2) Procedures associated with the repository's development;
3) The identification of users and the repository team's marketing or publicity approach;
4) Known benefits of the repository (to users); and
5) The service director's response to the Web link analysis or "Web intelligence" report.

A copy of the full interview schedule is shown in Appendix A. Since the interviews were audio taped for transcription, a consent form was also written for the managers' signatures (see Appendix B.). We do not identify the service managers in this report, but we have been given consent to use their interview comments.

5.3 Online questionnaire

For the survey stage of the study, we prepared five online user questionnaires – one for each repository type. In order to reach as many users as possible, we forwarded an e-mail announcement concerning the surveys, including their online links to relevant JISCMAIL news lists (e.g., academic lists; health practitioner lists; scholarly research lists, etc). The questionnaires were constructed using the my3q Web system (http://www.my3q.com/) and designed to obtain the following information: the user's gender, age, occupation group (e.g., lecturer/professor/academic researcher) country of residence, usage experiences (e.g., frequency of use), level of satisfaction with the repository's design and services, and usage experience other types of digital repositories. In the final section of the questionnaire, respondents were asked if they would be willing to participate in an online (e-mail) interview. A copy of the questionnaire template, comprised of 23 questions, is located in Appendix C.

6. Implementation

6.1 Study plan

To achieve the best results from a mixed-methods research design, we prepared a project plan comprised of five stages:

Stage 1): August to September, 2005
Select the repositories and inform managers of the nature of our research

Stage 2): August to September, 2005
Create LexiURL software tool for link analyses and "Web intelligence" reports

Stage 3): September to December 2005
Schedule management interviews and prepare "Web intelligence" reports

Stage 4): January to February, 2006
Develop and instigate online user questionnaires

Stage 5): February to August, 2006
Analyse the link data, interview data, and survey data, and integrate all findings

Although the five research stages were scheduled for implementation at different time periods, some of the periods overlapped. For instance, one of our management interviews (i.e., with the Jorum e-learning repository) was carried out at the same time as the digital repository user survey.

6.2 Interviews with repository managers

A few challenges arose throughout our study, and when they did, we were able to find a sufficient resolution. Our first challenge was to explain the purpose of our research to the repository managers and achieve their cooperation as interview participants, but by mid-September (2005), one of the repository managers declined to participate. We began a search for another repository to add to our comparison list, and fortunately the manager of CogPrints agreed to be our alternative interviewee. All of our interview requests were sent to the repository managers by e-mail, including an attached formal invitation letter.

Dear Repository Manager:

Attached to this e-mail message is a formal letter of invitation to participate in a digital repository research initiative, which is currently being sponsored by JISC (The Joint Information Systems Committee). This research will be undertaken cooperatively by researchers from the University of Wolverhampton and Loughborough University. We have selected your repository – [repository name] - for this study because it is an important and interesting resource and because we believe that you will benefit greatly from our free webometric/user needs analysis.

We hope that you will take up this valuable opportunity. Please do not hesitate to contact me for further information and let me know if you would like to make arrangements for an interview. During the interview we hope to learn more about why you have set up the repository and how you have ascertained its primary users.

Yours Sincerely,
Alesia Zuccala

Statistical Cybermetrics Research Group
School of Computing & Information Technology
University of Wolverhampton MI Building,
City Campus South, Wulfruna Street,
Wolverhampton, WV1 1SB
United Kingdom

The individual management interviews were scheduled for the dates of October 4th, 2005 (*National electronic Library for Health*), October 25, 2005 (*CogPrints*), November 4th, 2005 (*e-Prints SOTON*), December 1st, 2005 (*UK Data Archive*) and April 19th, 2006 (*Jorum*). During three of the interviews, more than one person was present to answer questions. The Jorum interview was the last to occur because the last to occur due to the very (or relatively) recent release of the repository to registered users in January 2006.

6.3 "Web intelligence" reports

Throughout the month of September (2005) Mike Thelwall and Alesia Zuccala, worked together to develop a unique Web intelligence report based on retrieved links (page URLs) to the selected repositories. Each report was given to the repository managers at the time of their scheduled interviews. The purpose of the report was to present retrieved link data in its raw, aggregated form, including a brief demonstration of how the data could be analysed for insights, for example, by following interesting or surprising links for context reviews. The reports were similar and based on a question and answer format (e.g., *What is a colink? Why do I want to know about the sites that are colinked with the CogPrints site?),* so that all of the managers could quickly and easily receive the webometric information.

6.4 Repository user survey

This stage of the study was the most difficult to implement due to the problem of locating repository-specific users and non-users. To obtain our survey participants we compiled a set of relevant of jiscmail news-lists on the internet and obtain permission from the list managers to circulate an e-mail announcement concerning our questionnaires. In some cases, the repository managers that we interviewed were also helpful in advertising our survey, and allowed us to send them a copy of the survey announcement to post on their private mailing lists.

The text of our e-mail announcement included a note about the Wolverhampton-Loughborough research team and our research homepage (http://cybermetrics.wlv.ac.uk/DigitalRepositories/index.html), as well as a link to the homepage of the digital repository selected for evaluation (e.g., http://www.jorum.ac.uk) and the exact location of the my3q questionnaire (e.g., http://www.my3q.com/home2/75/repository/32237.phtml).

During the initial stages of designing our user questionnaires, consideration was given to the length of the survey and time that it would take to complete – i.e., approximately 10 minutes. We also focused on creating a survey implementation strategy that would be fairly non-intrusive, yet still encourage participation, with no incentive other than the individual's willingness to cooperate. All of the individuals who characterized themselves as users were given a specific set of questions to answer, as well as those who considered themselves to be non-users (see questionnaire template, Appendix C).

7. Outputs and Results

7.1 Web statistics

With LexiURL as a research tool, Web statistics are much easier to download than if they are collected manually using a commercial search engine, such as Google or Yahoo! Digital repository and digital library managers can now use LexiURL to collect webometric data on a regular basis; however, the meaning of the link statistics will depend on the individuals who collect them and their interpretive focus.

LexiURL was new at the time that our first "Web intelligence" reports were created; hence we were very excited about learning as much as we could from the repository managers about the impact of the data. We could not anticipate what the managers would say in response to our reports; however our intention was to provide them with as much useful "Web intelligence" information as possible. What we discovered was that the managers were very interested and willing to provide us with some positive feedback (see section 7.2).

The tables and figures in the next section summarize the key elements of the link analysis reports, excluding all directed inlink lists (page URLs). Directed links to the main site in question (site identified by a diamond shape) are visible however in the colink network maps and the strength of this link is indicated by line thickness. The full webometric reports are available by request from the members of the Statistical Cybermetrics Research Group (http://cybermetrics.wlv.ac.uk/DigitalRepositories/index.html).

7.1.1 National electronic Library for Health (September 20th, 2005)

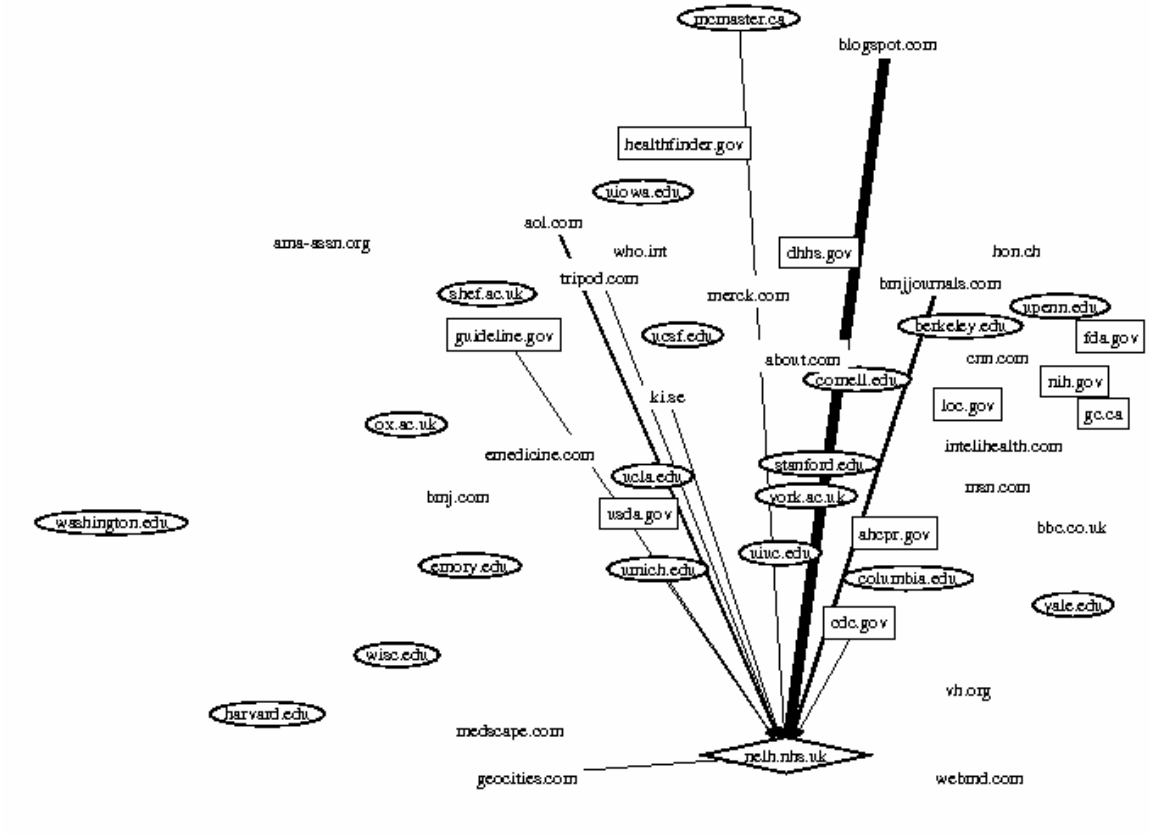


Figure 2. Top 49 sites colinked with the nelh.nhs.uk, including directed links to the NeLH (diamond) and link nodes for all educational sites (ellipses) and government sites (boxes).

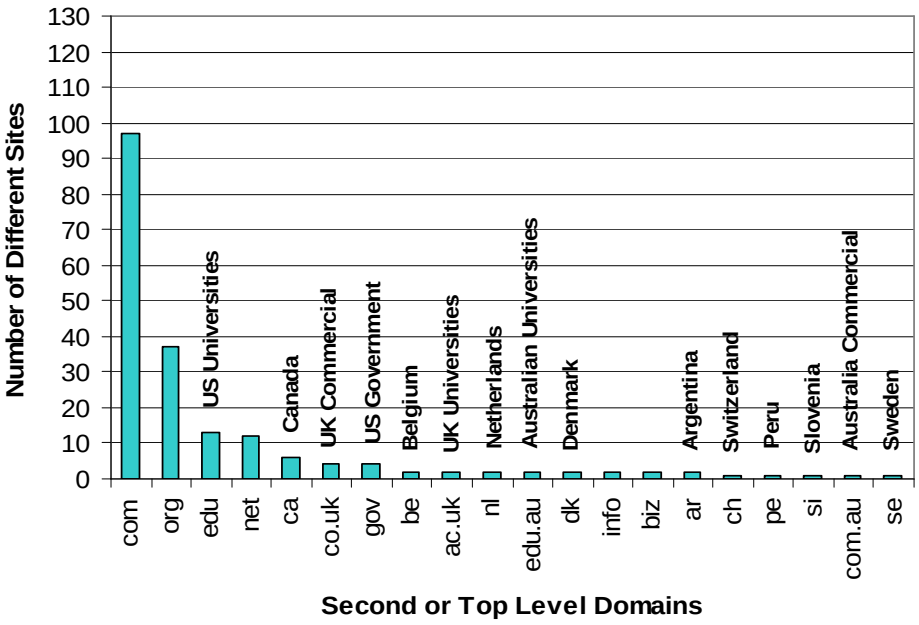


Figure 3. Number of different sites in the second or top level domains that contain at least one page that links to the NeLH website: www.nelh.nhs.uk.

7.1.2 CogPrints: Cognitive Sciences E-prints Archive (October 18, 2005)

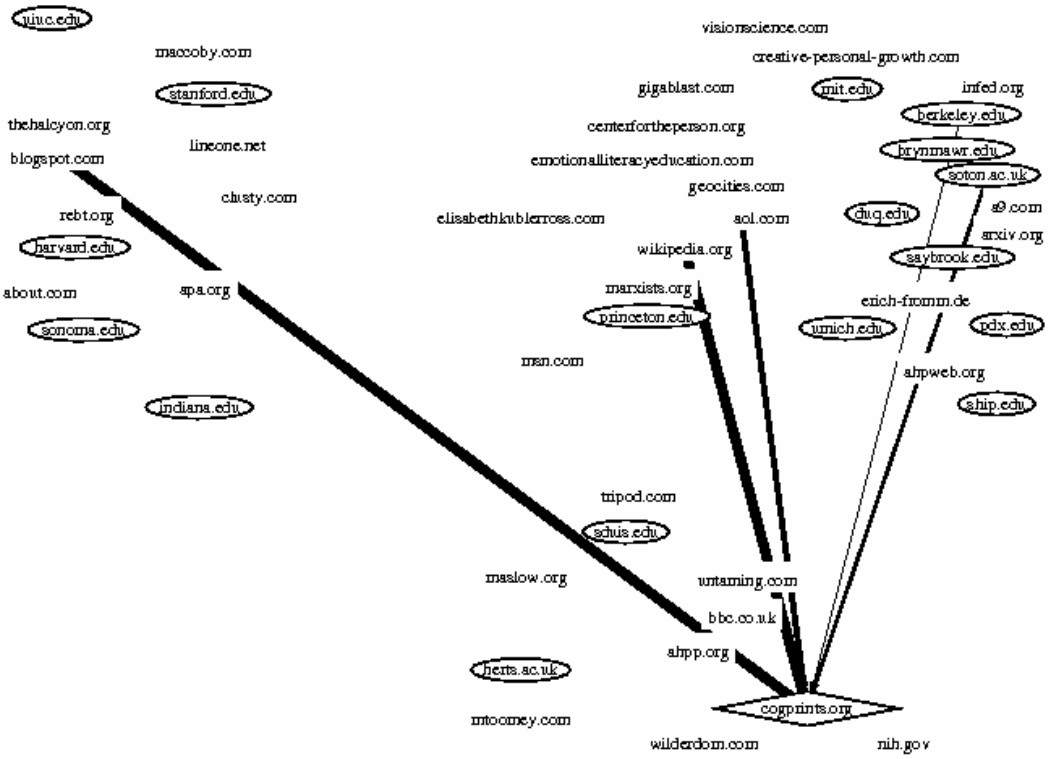


Figure 4. Top 49 sites colinked with the CogPrints.org, including directed links to CogPrints (diamond) and link nodes for all educational sites (ellipses).

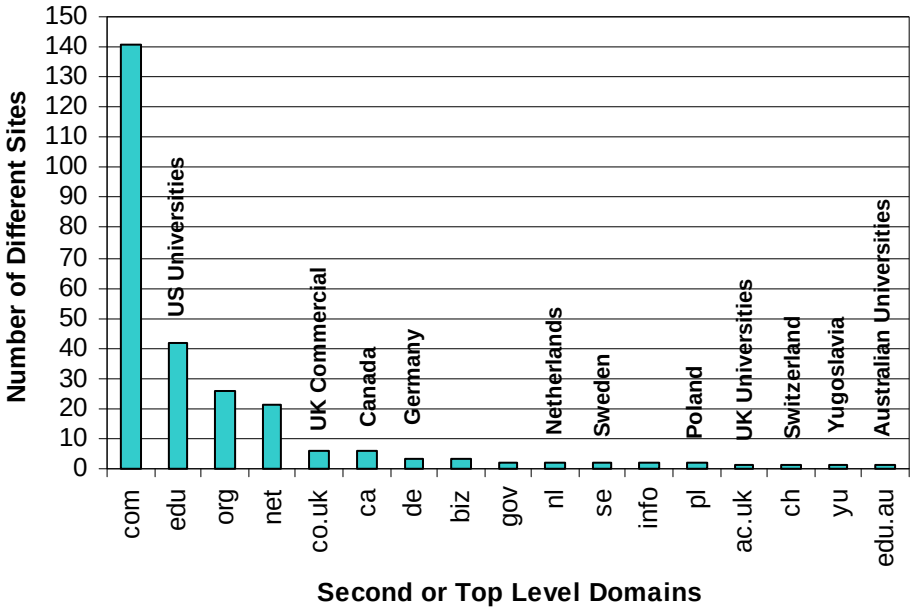


Figure 5. Number of different sites in the second or top level domains that contain at least one page that links to the CogPrints website: CogPrints.org.

7.1.3 University of Southampton E-Prints (October 20th, 2005)

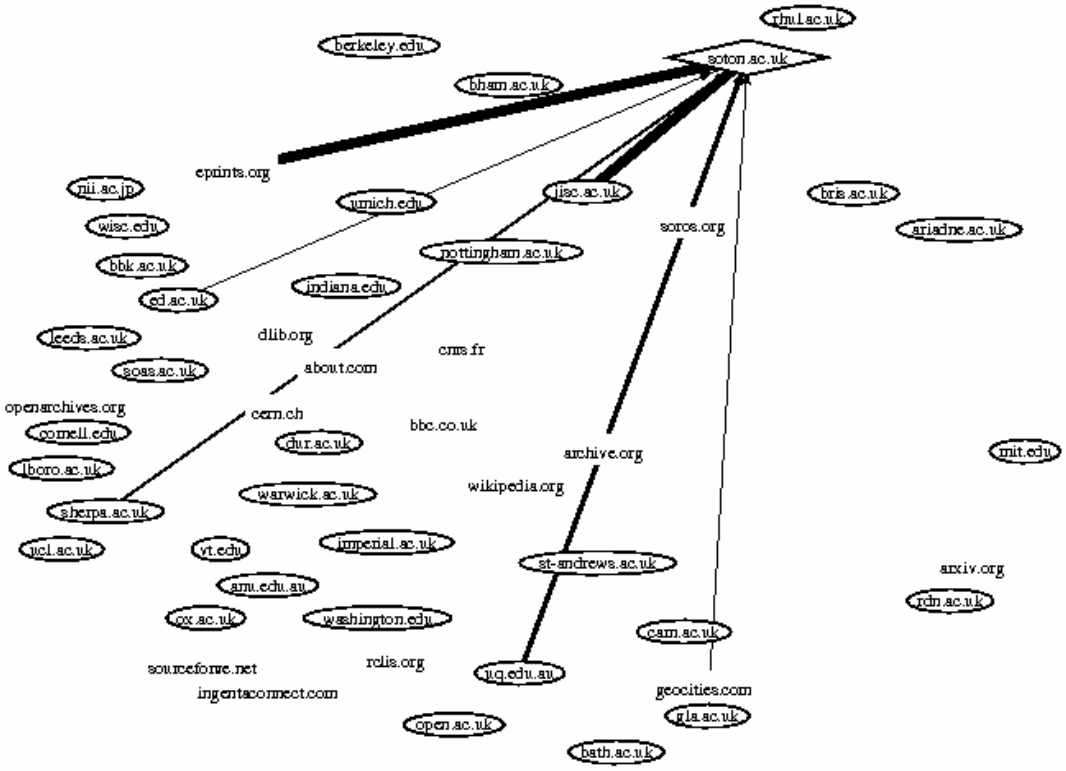


Figure 7. Top 49 sites colinked with e-prints.soton.ac.uk, including directed links to the Southampton e-prints site (diamond) and link nodes for all educational sites (ellipses).

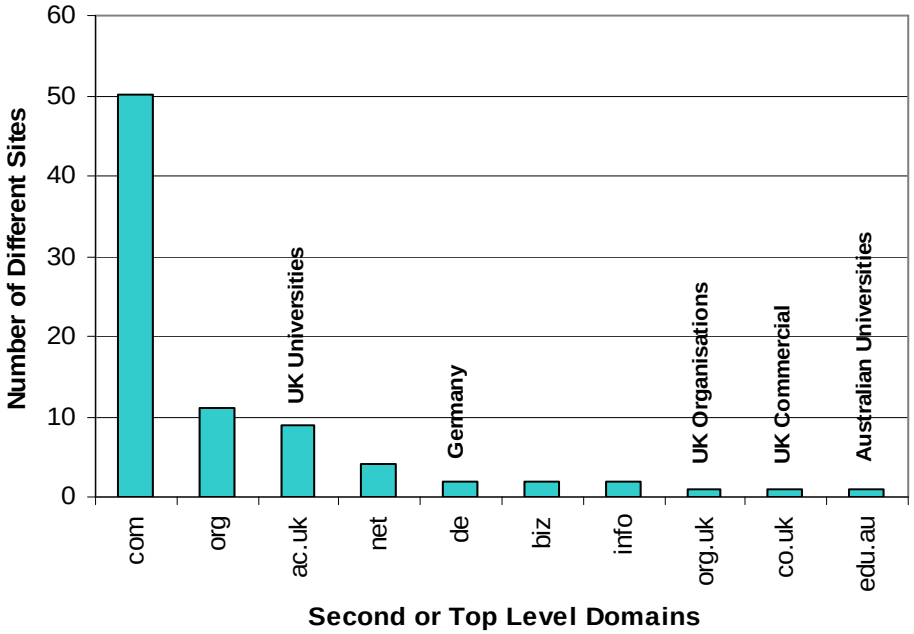


Figure 8. Number of different sites in the second or top level domains that contain at least one page that links to the University of Southampton e-prints website: eprints.soton.ac.uk.

7.1.4 UK Data Archives (December 12, 2005)

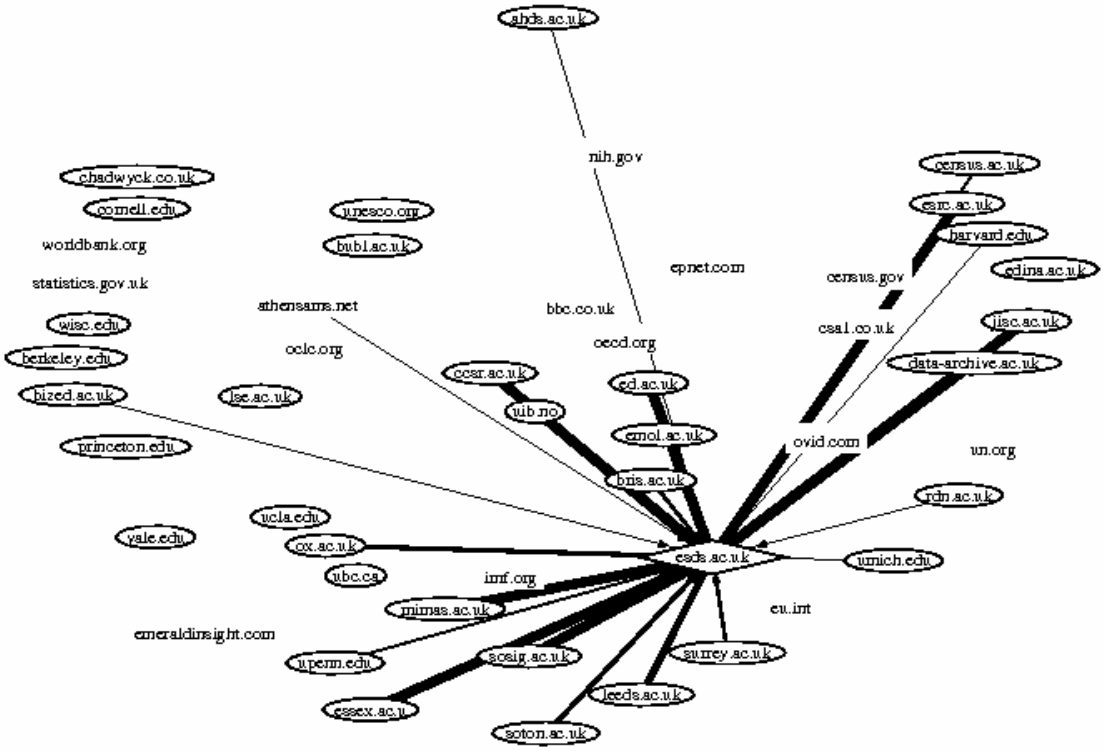


Figure 9. Top 49 sites colinked with esds.ac.uk, including directed links to the UK Data Archives (diamond) and link nodes for all educational sites (ellipses).

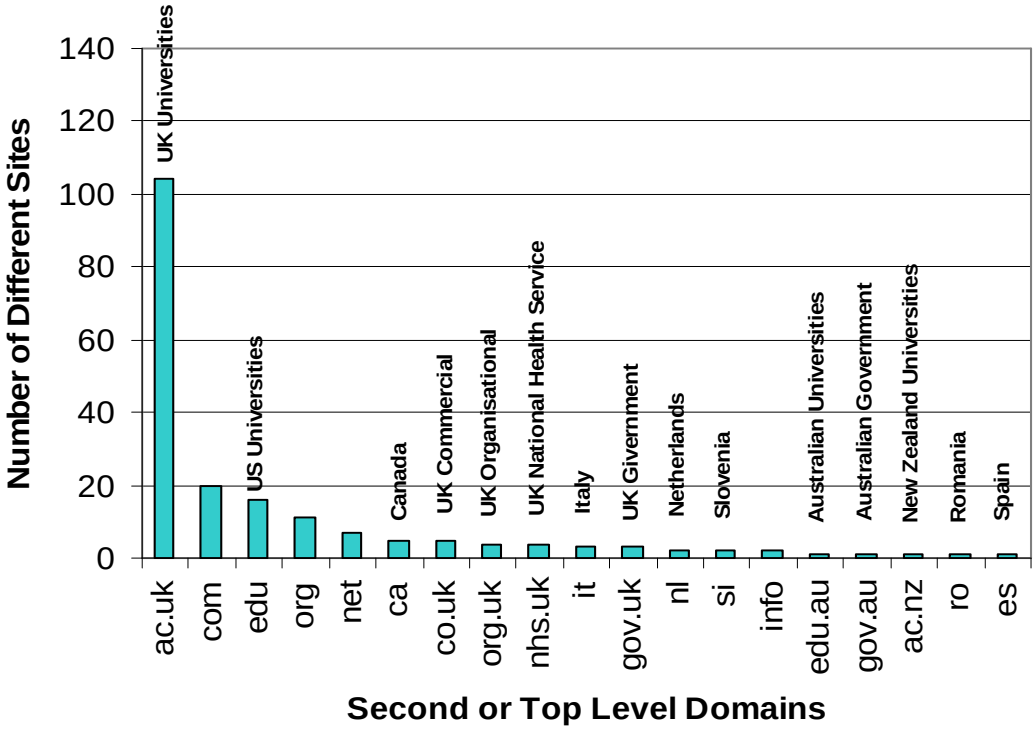


Figure 10. Number of different sites in the second or top level domains that contain at least one page that links to the UK Data Archives website: esds.ac.uk.

7.1.5 Jorum E-learning Repository (April 3, 2006)

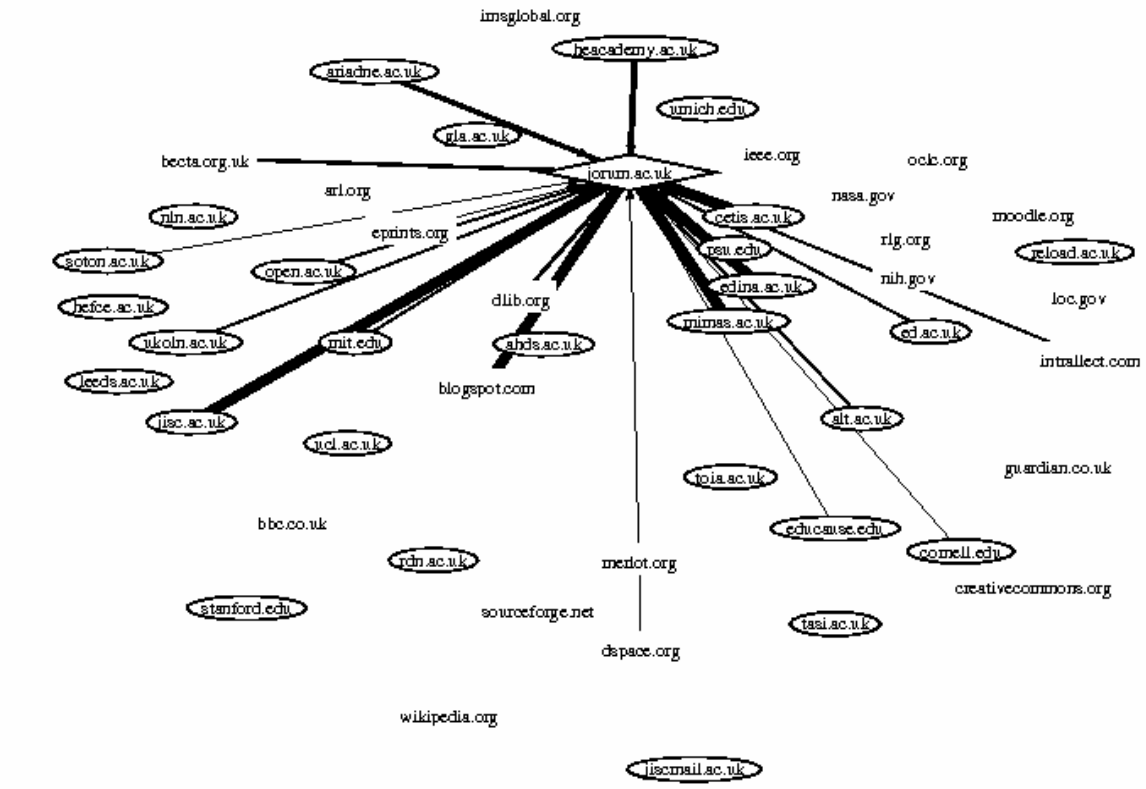


Figure 11. Top 49 sites colinked with jorum.ac.uk, including directed links to the Jorum E-learning site (diamond) and link nodes for all educational sites (ellipses).

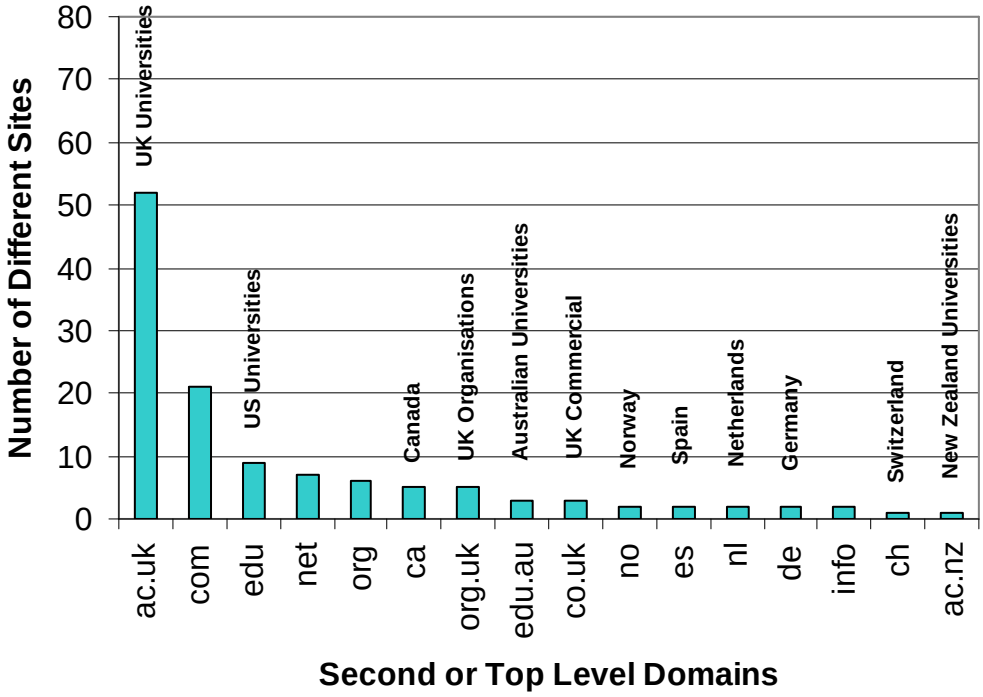


Figure 12. Number of different sites in the second or top level domains that contain at least one page that links to the Jorum website: jorum.ac.uk.

Table 4. Web link statistics summary.

	Expected inlinks	Surprising links or links of interest	Colink network at the time of the analysis
NeLH	• University of Oxford • Sheffield University • BMJ Journals	Several Blogspot links (e.g., *Confessions of a Quackbuster* and *HyScience*) produced repeated (sidebar) acknowledgement links to the NeLH. A link from a doctor's personal home page (Dr. Abhishek Arora) and another link from the personal website of a now deceased researcher in Library and Information Science from the University of Iceland (Professor Ann Clyde) The Karolinski Institute of Sweden McGill University Health Centre-Nursing Research Ball Memorial Hospital in Muncie, Indiana, U.S. Several missing university links from the United Kingdom as well as hospital website links.	The NeLH was firmly embedded in an academic and medical publishing context on the web, but not necessarily a UK-based context, given the significant number of US colinks. It also seemed to be situated in a significant commercial and governmental context.
CogPrints	• University of Southampton • Eprints.org • Citebase.eprints.org	A few blogspot and wikipedia links were retrieved. Many inlinks from Robots.net were directed to specific articles in CogPrints. Links to CogPrints were found at www.furl.net – a free service that saves a personal copy of any page found on the Web, with recommendations for new pages of interest. Links from Canadian Peter D. Turney's Website (www.apperceptual.com) were directed to his papers in CogPrints.	CogPrints was situated in a predominantly commercial link context, including a few academic sites (mostly American), but not as many academic sites as one would expect.

E-Prints SOTON	• The Joint Information Systems Committee • The National Oceanography Centre library • E-prints.org	The Science Oxygen Website for mining links to scientific knowledge contained quite a few link references to papers in e-prints SOTON (e.g., from researchers in the fields of botany, physics, chemistry and optoelectronics). The University of Queensland Australia had listed e-prints SOTON as a key resource on their Databases for Social Sciences page. The Web site of Tim Kelf (a PhD student in nano-optics, plasmonics and SRES, at The University of Southampton) links to e-prints SOTON page from his "Links" page under "Useful Science Web Sites." Aside from Tim's site, there are very few direct links to the repository from personal home pages	The University of Southampton e-prints repository site was situated in a predominantly academic Web context.
UKDataArchive (Economic and Social Data Service)	• The Economic and Social Research Council UK • Census registration service • BUBL Link Catalogue of Internet Resources • Edinburgh University Data Library • University of Kent Library • University of Essex • University of Central • England Birmingham • Social Science Information Gateway • Aston University Birmingham • Cardiff University • University of Bristol • University of Manchester (Cathie Marsh Centre for Census and Survey Research) • City University London • The Joint Information Systems Committee • University of Oxford • Napier University Edinburgh • MIMAS – Manchester Information & Associated Services	Personal website of Daniel L. Millimet (Southern Methodist University, Dept. of Economics) links to various data sources at the ESDS (e.g., British Household Panel Survey; 1970 British Cohort Study; UK National Child Development Study). The FQS (qualitative-research.net) Forum for Qualitative Research – e.g., one article links to the ESDS access and preservation guidelines for identifiers and anonymisation. Personal website of William Robinson (School of Information, University of Tennessee) links to the ESDS at his international organisations page. A link from the University of Toronto Centre for Industrial Relations and Human Resources "Statistics and Data Sets" page. The www.flickr.com website for storing, searching and sharing photos contains a page with Edwardian culture photos taken from the original study of Prof. Paul Thomson, "Edwardians: Family Life and Work Experience Before 1918"	The UK Data Archive's ESDS site was situated within an academic colink environment, and many of the colinked sites were directly linked to the ESDS.

Jorum	• The Joint Information Systems Committee and JISC Regional Support Centre • MIMAS – Manchester Information & Associated Services • EDINA JISC funded National Data Centre • Ariadne Web magazine • University of Bath • CETIS – Centre for Educational Technology Interoperability Standards • E-learning gateway at the University of Southampton • The e-learning page at SURF [Netherlands].	The FURL Personal Web (bookmarking) File provides a link to Jorum, including some repository rating sites eg., Jorum was rated 5 in Learning Objects on Jan. 21, 2005. A few blogs link to Jorum (eg., David Mattison's *The Ten Thousand Year Blog*, *David Davies' Weblog*, the *Moodle Journal* and the *Scholarly Electronic Publishing Weblog*). The personal Webpage of Jane Stevenson, an employee of MIMAS provides a link to Jorum, includes a report that she has written titled "Jorum Research & Development Report: Preservation Watch Report." An explanation of the Jorum resource and link to the site appears at the Spanish Centre d'Educació i Noves Tecnologies (CENT) de la Universitat Jaume I de Castelló	Jorum was situated in a predominantly academic colink context. Most of the other sites colinked with Jorum were related to e-learning (e.g. the MERLOT e-learning resource) or they were university sites.

7.2 Managers' responses to the Web statistics

NeLH:

--Comments regarding the lists of URLs (pages) that link to the NeLH:

"It's quite heartening that a lot of US sites appear to be [linked to] us, when in the past we've often looked at the US [for] examples of good practice. [Some] of the [other] major organisations are not surprising, because there are organisations [with which] we have quite significant relationships...."""

"...the only surprise that I do see is that there are fewer British organisations than I thought there would be. I mean, the BBC, I'm not surprised, and I would have thought more UK universities, possibly in the UK many health organisations that are not NHS, would have had links to us. That surprises me. Worries me a little as well, maybe, that our message isn't getting across to some other, other organisations..."

"Blogs are increasing in interest, certainly in the health area, people are communicating with a great, kind of diversity about health [since it is] one of the biggest topics, for, for discussion. So I'm not surprised about that either."

--Are the web statistics useful?

"Well....we've not, we've not looked at it in this way ourselves, we've looked at it in a very different way. [For instance we look at] who comes to us, and what they look for – [we look

at things in] a kind of user behaviour way. Co-relationship linking is something that I hadn't even thought about before I came here today, so, yes, it's very useful, actually...."

"...our focus at the moment, since our launch, obviously, is to improving usage in terms of the search interface. We've spent a lot of money on the search technology, because we're so often told now, that users don't want to browse, they want to search, and they want their results quickly.....

"...in the private sector we are often told it's all about building relationships with the customers long term, so [in terms of] building a relationship, this [colink information] certainly feeds into that idea. We haven't looked at the co-linking idea, or why people are linking to us, and what's their motivation. This is another source of information..."

CogPrints:

--Comments regarding the lists of URLs (pages) that link to CogPrints:

"…it's transitory and this is bad news in the sense….because it means instead of having the entire literature harvested… they're linking to an archive as if it were a place that has content; specific content that they want to go to directly. There is no point; they shouldn't be going to CogPrints directly at all, they should only be going in to harvest CogPrints metadata integrated with the metadata from all of the other institutional repositories, you see? As long as they're linking to CogPrints, we're not there [*i.e., not successful with open access to e-prints].* It's paradoxical, right? Where the name of the game is some sort of a local service, that's good news that they're linking to you; where the name of the game is a distributed universal [*open access e-prints*] service, its bad news that they're linking to you locally, because they shouldn't be…"

--Are links to CogPrints or colinks useful for determining new user bases?

"…Well, as I said, we're not really interested in developing user bases……[T]he links will become interesting but only at the point where we have all of the literature there, because then all of the links will in fact be, virtually all of them will be citation links and not just a link to an archive."

E-prints SOTON:

--Comments regarding the lists of URLs (pages) that link to E-prints SOTON:

"…it's not something we've done with the repository because we haven't yet tied the repository into, if you like, any marketing function within the University, but we could. You have the curation of the institution's knowledge, but this [i.e., referring to the webometrics report] is an issue of raising awareness, raising profile, is a key thing and it's another, it's another metric……"

"…it's interesting and it's an analysis that I've done in different contexts for things like a conference website: who's taking note of us? Who's referring, who's actually coming in to our call for papers? Does this [report] tell us something about the, how many, you know, how well our marketing's going?"

"...our concentration has been on getting the repository up and running and our users as authors with the assumption that the users of repositories generally will be searching via either general search engines like Google…..the concentration has not been on how well we

are linked within the general environment but how are the users who are searching going to find us as a first resort…..and that, we would assume, is the first resort, but nevertheless if we're looking at the repository both within the University and externally, there is certainly plenty of potential for looking into the wider front, I think…."

"...your report is very interesting because it makes me think well I can see what that and that comes in because that's just a service and so I can ignore that, but it's nice to see that we've got some of these scientific sites linking to us, but why haven't we got more? "

--Are directed links or colinks useful for determining new user bases?

"…I think what we're interested in is the context, the research context: the fact that this journal, this highly-respected magazine or this university has chosen to make some reference to some of our research and that's the, I suppose rather than trying to track down customers or users, it's more the issue that you said, what's the context that your repository's appearing in? Is it a, is it mainly, is it a research one? Is it a high-quality research one?"

--How often would you like to receive a link analysis report like this?

"…Well, it would be interesting if we looked in six months time, for example, I mean if there are lots and lots of good links here it's important if we're looking at preservation issues, which we are in our other project, it's really important to know that these links are there. If these links are significant then we need to care about our repository and that those links are maintained, and that we don't do something very simple that changes all the links, because we'll lose all the credit from those links. So it's actually worthwhile from a preservation point of view, understanding how important it is to other people, and that's something we don't know yet."

--Is there anything in this report that you would like to see added?

"…if we have people who have a good reason to link to us we'll be very happy. But, you know, how often are they going to link to the repository itself and how much are they likely to link to specific papers, is a good question really."

UK Data Archives:

--Comments regarding the lists of URLs (pages) that link to the UK Data Archive:

Note: The manager of this digital repository said that he would need to look at the report in-depth and speak to people from the information development group to respond to the questions. He also mentioned that the UK Data Archive is an umbrella organisation with lots of individual parts, with the Economic and Social Data Service (ESDS) being one of them. He said that people may be accessing the individual parts directly without visiting the UK Data Archive site because the "UKDA is the vessel not the product, ESDS is the product."

--Is there anything in this report that is surprising to you?

"…I don't really know, I'd need to have some time to, you know, digest it and obviously also ask, you know, people who work more closely in the website area in our information development section."

--Do you think a "Web Intelligence" report like the one we present would be valuable to you as part of a regular service?

"For the ESDS it would be particularly valuable."

--Is there any information in this report that you would like to see added?

"What I would like to see added is you almost do exactly the same thing for ESDS."

Jorum:

--Comments regarding the lists of URLs (pages) that link to Jorum:

"There's certainly quite a lot of international linkage."

"It's also surprising to see who does link and who does not. On the 'who does' side I'm surprised that it is sort of spread internationally, and regarding who doesn't, it is some of the organisations that we work with quite closely."

--Do you regard the information provided in this report to be useful to your repository programme?

"I would say it's very useful and its made me think about how we should possibly be locating this on a more regular basis, and I'd be interested to know how we could do that."

--Do you think a "Web Intelligence" report like the one we present would be valuable to you as part of a regular service?

"Absolutely"

--How often would you like to receive a report like this? (e.g. monthly, quarterly?)

"I would say four times a year, or every two months"

7.3 Repository management practices

The repositories that we selected for our study, including one digital library, were different in type; thus interesting to evaluate from a comparative perspective. In this section we present the qualitative information that was collected during our management interviews. The main purpose of the interviews was to learn more about the individual repositories' management practices, including the background work required to develop them, and rationales behind their creation.

Repository rationale: *Before a repository or digital library is created, a good development team, including the primary manager needs to have a clear understanding of its rationale or purpose.*

With the U.K's National Health Service, a new digital library – the National electronic Library for Health (NeLH) – was created due to "a realisation that clinicians, doctors, nurses, speech therapists, dieticians, and all the kind of therapeutic professionals, needed access to information quickly." According to the NeLH service manager, healthcare care professionals across the UK sometimes find it difficult to achieve quick and easy access to medical information when they need it. Normally this is related to the fact that the library of a hospital

or medical centre is located in a separate wing and is not always a convenient place to get to in order to do an information search. Often, medical professionals are also called to work outside a traditional medical setting; thus find themselves in a position where it is too time consuming to get to a medical research library. When using a Web-based digital library, a healthcare practitioner will only need to be in a place where he or she has access to the internet; hence the NeLH was developed as a Web-based service in order to "provide clinicians with access to the best current evidence on conditions and treatments to improve patient care." At the time it became available it was not "aimed at the public" but "it has a sister service called the NHS Direct that is." Both were developed "more or less at the same and "there [has been] a lot of cross usage….something like 10-15%....."

CogPrints, the subject specialty repository, was created for the cognitive science research community, due to the success of the Los Alamos physics e-prints arXiv. For several years the Los Alamos archive or "arXiv" has been a well-used, well-respected e-prints repository within the physics research community. The arXiv was "originally developed by Paul Ginsparg and started in 1991 as a [repository] for preprints in physics, [but] later expanded to include mathematics, computer science, nonlinear science and, most recently, quantitative biology….The term e-print was adopted to describe the articles" (Wikipedia, 2006). At our interview with the manager of CogPrints, we learned that there was a background interest "in demonstrating that [subject specialty repositories] were not just for physicists" and that they could "work for other disciplines." The CogPrints manager was convinced that if the new subject specialty repository "flew", or grew to be successful, it would show that "archives with self-archive papers [were] not just a special quirk of physics."

In conversation with the e-Prints SOTON management team, we learned that there were several reasons why they were interested in creating an institutional repository, including the fact that "the JISC Repositories Programme and FAIR programme were coming out" in the United Kingdom [Note: The JISC FAIR programme was instrumental in funding the E-Prints SOTON project]. E-prints SOTON was created shortly after the ECS database (http://eprints.ecs.soton.ac.uk/) at Electronics and Computer Science department was built, and close to that time (in 2002/2003), there was a growing interest from the National Oceanography Centre at the University regarding the development of a similar publications database. The development of e-Prints SOTON was, according to one of the interviewees, closely associated with "the issue of push and pull of the open access movement." This university-based repository team felt that it was a "natural progression in the publishing debate as a whole" and that the creation of E-Prints SOTON "would enable the university to organise its institutional research output in a way that would allow better analysis of where the research is going."

The UK Data Archive was originally created as a specialty research space in 1967/1968, because "the UK research council thought it would be a good idea [to create] a one-stop shop [for researchers] so that …rather than having to individually go to the data providers, mainly government departments and commercial data providers, [they] would be able to go to a central location and obtain all their data." The notion was to get one "organisation brokering access agreements and licensing arrangements and copyright arrangements rather than individuals having to do that on a one to one basis." Research councils have been major sources of financial support for "data collection exercises; therefore in order to maximise secondary use of the data, [sponsored researchers have been] required to offer data to the archive…." Essentially, "the data archive" was first created as an "archive of investments, made by the research councils themselves." According to the UK Data Archive manager, this particular repository "has become more important over the last few years… because of its change in status." Not only has it gradually become a digitized resource, it is also "a legal place to deposit; the only digital repository in the country that has legal place as status, [which] means that members of the public can come to [it in order to] acquire digital materials."

The newest service that we evaluated in this study, the Jorum e-learning repository, was created and funded by JISC (the Joint Information Systems Committee) for the purpose of "hosting content created for the [Higher and Further Education] community [as well as] to stimulate a community of users for teaching resources." Outside the United Kingdom "other teaching and learning repositories" have been created, "but none that were doing quite the same as Jorum." MERLOT (www.merlot.org) is the name of another "international repository, [which is] essentially a library catalogue system where people can come and search for content but the content isn't contained within the repository. This is not the case with Jorum, since it does both; it houses metadata records that describe the content that can be found elsewhere, but it can also be held in the repository itself."

Collaboration and teamwork: *When we asked our interviewees to provide us with a brief explanation of who was or who is currently involved in their individual repository/library projects, all gave us an answer confirming the importance of collaboration and teamwork*.

NeLH: "Several people were involved, including Sir Muir Gray who is very senior in the NHS. He works for the National Programme of IT (associated with IT developments in the NHS such as electronic booking of appointments, electronic health records) as a member of the senior management team. His interest in Web technologies and their application within hospitals made him the visionary. It was his vision initially that got the thing off the ground. He linked up with Dr. Ben Toff, director of Knowledge Management within the National Programme of IT. Dr. Toff was the sort of architect who designed an initial plan of how we might operate. So the two key drivers of NeLH were Sir Muir Gray and Dr Ben Toff, but now there is a team of about 15 people – people to run the service, develop, provide specialist help in certain areas, such as, search technology, web hosting, communications and Public relations.

CogPrints: "The first version of CogPrints was created by an Electronics and Computer Science PhD student at Southampton University. The second version, post Open Access Initiative (OAI), was rewritten by another PhD student to make CogPrints OAI compliant. The third version has been taken over by another PhD student at Southampton and is currently being maintained by him." Overall, this is a repository that was developed by a "very savvy population of computer scientists and pretty good hardware resources."

E-prints SOTON: "A project team was formed involving the oceanography centre, the computer science department and the Information systems services. It was a collaborative effort within the institution."

UK Data Archives: "A director (the interviewee), a senior management team made up of the heads of the various sections in the Archive. Underneath that, there is a Director's advisory group, comprising of both senior and middle management, these represent areas of activity within sections…. Back in the 60's the resources were fairly meagre. The resources were in the region of tens of thousands of pounds and basically about three people were employed. However, now the staff numbers have increased to sixty plus and there are multiple grants and sourcing…"

Jorum: "Jorum [is] a joint-funded venture between the two national data centres, MIMAS and EDINA. MIMAS is a JISC funded national data centre that provides services to the UK FE and HE community and EDINA is a comparable data centre based in Edinburgh. In terms of the management structure, there is a manager at each of the data centres, and senior managers above them that are the co-directors of the Jorum service. Above the co-directors there is the JISC programme manager who is quite closely involved with everything that is done." In the beginning, "we had a core team of about 8 people and we had to procure a system to underpin the Jorum Service. Half of the team were working on technical aspects of Jorum. There is now an Outreach and Promotions officer who supports the

outreach activities and dissemination work, and there is also a Training Officer, who works on the training side of things and a Metadata Officer, who plays an important role in looking after the metadata side of things."

Identification of users: *In order to populate a digital repository or digital library with useful materials, a professional development team needs to identify and sufficiently understand the needs of their service's primary users.*

At the NeLH, a number of user groups were identified from the digital library's initial pilot work: "we had panels of people reviewing material and coming to some sort of consensus. The user groups consisted of a wide range of people, including, doctors, nurses, various other allied health professionals and library and information workers as well." All of the users have now become "key stakeholders…because they act as advocates on [the NeLH's] behalf by getting people to use the library, and they also do a lot of training as well, in literature searching, for example, and use of databases…. [T]hrough these user groups [the management team at the NeLH] has identified a key quality resource that the clinicians feel they need to access quickly" [e.g., Hitting the Headlines]. Hitting the Headlines "is a review of the coverage of health issues in the press. Two or three times a week, a story is picked up from the press and examined, conclusions are then drawn as to the validity or otherwise of the newspaper reporting. Clinicians find this very useful as patients pick up on these stories from the TV or newspapers, and often clinicians are not aware of what the position is. So it helps clinicians to help patients."

The CogPrints manager said little about this repository's main user group; however, he told us that it was created as "a public archive for those who are cognitive scientists, as a place for them to put their material in all over the planet." During the interview he emphasized that "research itself is beneficial to the general public, that's why we fund it, we want it to grow, we want it to have impact, we want it to be used, we want it to be applied. So anything that increases the usage and application of research [i.e., CogPrints] is beneficial to the people that are funding the research."

At E-Prints SOTON, the repository development team "wanted to capture the whole output of the university, but "saw that [they] needed to start in a specific area. Research was [their] key focus, [and this included] conference papers, posters, project reports and all the different things that research encompasses." The "researcher" was the primary user that they had in mind; therefore one management interviewee said: "what the researcher thinks is important is what goes into the repository." Another interviewee added the following: "when we talk about users we mean users in terms of people who are depositing the work, such as authors. Users were wanted from a spread of areas across the university but we started with the ones that we knew were interested. Other groups were targeted which would set good examples for the rest of the university. For example, education was targeted as they wouldn't have a database that they could regularly deposit material into, so they would be encouraged to self archive via the repository and hence show other faculties that it is a good idea to be proud of their research and have it made visible."

The management team at the UK Data Archive had previously "undertaken user surveys" as part of their efforts to identify users and gave us the example of the Economic and Social Data Service (ESDS): when this was a major new facility and within the first quarter an in-depth survey was conducted "to try and understand what user needs were, both in terms of data types but also the support that they wanted relating to those data in terms of, training needs and so on." The chief manager who agreed to speak with us confirmed that certain employees at the UK Data Archive have gone "out into the community, [to hold] workshops both around particular data resources, based around themes and resource types." He admitted however that "the biggest problem [in identifying users or potential user groups] has

been to "understand the needs of those people who don't use [their service] but might use [it]."

When we asked the Jorum management team, "*How did you identify potential users or user groups during the early development stage of [your repository]?"* the following information was provided: "Jorum didn't really go through this process since it had been done [already]. Jorum had initially been set up to work with JISC funded content producing projects under the X4L Programme. JISC data centres have a history of providing services for institutions, therefore the target audience for Jorum was already defined, that being Higher Education and Further Education institutions."

Publicizing the repository: *After a repository or digital library is created, certain groups of people are expected to become users; however, new users will not necessarily recognize the service's value unless it is sufficiently publicized.*

At the National electronic Library for Health (NeLH) "the bulk of public relations and communications [has been] aimed at the health community and not the general public because NHS direct online [an alternative resource] serves their needs better." According to the NeLH's service manager this resource has been publicized "in several ways. The digital library [was] initially advertised in professional journals in the health service, and "in about a dozen major journals, [the management team] had advertise[d] periodically". A representative from the NeLH has also set up "stands at conferences regularly throughout the year" and "a lot of promotional literature [has also been] produced, such as, posters, leaflets, pens and badges [to distribute] at conferences and [to send] out to individual libraries so that they can distribute it to users to raise awareness. Sometimes "phone in sessions" are used to promote this digital library's use. During these sessions people can "phone in and be guided through [the NeLH] site and [its] key resources." Promotional work, we were told, "is an ongoing activity" for the NeLH management team. "One week of the year is called Health Libraries Week, which aims to promote the use of library information services to the health community. During this week, usually in November, [a number of] events are held in hospitals and other NHS organisations, such as a series of talks, lectures and live demonstrations."

When CogPrints was initially developed, "there were no problems associated with the actual setting up of the repository [i.e., its technical aspects], but people were unaware of what it was about and had worries and concerns about depositing material." According to the repository's key manager "a basic help [was made] available on the site in the form of instructions of how to use the repository and also an email system, whereby users [could] email the site administrators if they [were] experiencing any problems." Beyond this type of outreach to the public, CogPrints has not been advertised or marketed to a great extent, only "somewhat, through written articles etc. [It] was not meant to make any money, just meant to provide access to resources; therefore publicising it [has not been] a major concern."

At the time of the e-Prints SOTON management interview, we inquired about how the repository was introduced to faculty at the University of Southampton and other institutions across the U.K. One of the interviewees indicated that this repository was "initially a pilot (TARDIS project), which started in 2002. The TARDIS project officially finished in April 2005 and this was when university funding started and the phase of embedding the repository institutionally began." With respect to publicity work, another interviewee said that e-Prints SOTON was not publicized "as it was planned" although "a lot of internal publicising was done [within the University], working on many different levels, such as the head of research, individual researchers, the research administrator and secretaries. All these people needed to know about the repository not just the individual who deposits material. The traditional advocacy methods such as flooding people with posters and general meetings [were not]

used but it may still be possible…at a later stage. In the initial stages a personal approach was used, where a relationship was built with the people who were interested in the repository. Then it was often the case that these people sold it to other people. So it was a case of 'sowing seeds.'"

At the UK Data Archive leaflets are available and have been made available in the past to the general public concerning all branches of its service: "publicity takes on a variety of forms, including the distribution of hard copy documents, electronic documents, and specialist documents aimed at specialist audiences. The publicity materials inform users and potential users – e.g., the Archive's annual report – but they also but they serve another purpose of showing sponsors what [the Archive management team] is actually doing." Users of the UK Data Archive are invited to register and provide contact details so that they can access all materials on offer. The Archive's newsletter is available both as a hardcopy and pdf version on the Web. A newly registered user has the option of receiving a hard copy of the newsletter by ticking a box at the online registration form. According to the manager, the Archive has "20,000 registered users, but [approximately] only 4000 have asked for a hard copy." Mailing lists are also used to inform users of new releases of data, and a lot of promotional material such as paper brochures are produced and distributed at workshops and conferences.

The Jorum e-learning repository launched its resource for public use in two different stages. In the beginning, new depositors and contributors were given an opportunity to become familiar with the repository (in November 2005), then shortly after "the user service, [which allows] people to download content, went live in January 2006. The two separate services were staggered slightly to allow some content to build up." Throughout the two separate launches, articles were written, newsletters were produced and mailing lists were targeted. The Jorum e-learning repository is also promoted at events – some based on outside invitations, and others organised by the JORUM management team itself. One of the managers involved in the management interview provided the following information: "We promote the service to e-learning, ILT people, learning resource staff in institutions. We do this in a variety of ways in attempt to get at end users, so we promote to the right people in the right places to encourage uptake." With respect to user training, "a 'train the trainer' approach is used, whereby training and outreach events are held all over the country to give an overview of what Jorum is and showcase some of the [deposited] materials. These are typically half day events, for intermediaries who will in turn pass the information on to end users. The intermediaries are provided with the resources to deliver sessions to users at their institutions."

Measuring success: *There are different ways to measure the success of a new repository or digital library; however, in connection with the "webometrics" portion of the study, we were interested in understanding the way in which our interviewed managers had focused on obtaining information from the Web in order to better understand users.*

The NeLH interview service manager demonstrated a high degree of awareness regarding his digital library's users: "We know that our users, in many cases are overworked, exceptionally busy and have a number of competing priorities; therefore our strategy is really to try and sell NeLH to them, by telling them what's in it for them. The key messages are that we're always available, that we're easy to find and you can find the information within a few minutes of going onto our site." When we asked if usage statistics were collected for assessment purposes, the interviewee said: "Yes, we do, on a monthly basis. The statistical software used to track users is called *WebTrends* and it enables information such as what are the most visited pages, the averaged time spent on the site, entry and exit pages, so it enables, to a certain extent, the mapping of a users' journey through the site." We asked the manager if the NeLH traces where users come from and his response was: "Yes, Google was one of the highest entry points to [the NeLH Website]. Regular use of the NeLH "breaks

down something like 40% General Practioners (GP's), 30-35% nurses and 15% professions allied to medicine. The remainder is students and the general public. The students are from a variety of related areas, such as, life sciences."

--Have you come across any benefits to current or new users associated with the NeLH?

"Yes, we have and do. Success stories are a key part of NeLH's public relations. For example, we make a point of publicizing the fact that someone saw something on our site that directly benefited or contributed to patient care. Some individuals have said: "I've changed, or improved my practice through something I've read [on NeLH]". Testimonials of this nature demonstrate that people are finding NeLH useful. Some people volunteer this information through the feedback facilities available on the site. A large amount of positive feedback is received on what people found on the site, for example: 'I was able to do this because, because I found this.' Other information is sought out by asking clients in both the user community and the library information community and they relay the feedback that users have given."

The CogPrints manager mentioned that this repository's user base was located "worldwide" and that a majority could be identified as "almost certainly academics." He commented further that the subject-based repository was "not the kind that the layman would be particularly interested in." In terms of collecting 'actionable' information from the Web, the management team at CogPrints has implemented an online system for collecting Web-based statistics, but the manager did not express a concern about carrying out regular user assessments based on the statistics: "User satisfaction is not the relevant thing, the relevant thing is the provider satisfaction, which is the enhanced impact [i.e., more frequent citations for open access articles]..... [T]he relevant question is how do you get the 85% of the non-providers to be providers, so that they can get the enhanced impact. CogPrints should not be looked at.... in fact, open access [to published articles] should not even be looked at from a user standpoint; it should be looked at from a provider standpoint..."

The E-Prints Soton management team spoke about collecting some usage statistics for user assessment purposes, but admitted that this has not been a major part of their focus yet. One manager stated: "Yes, we have done a little bit of this, [but], our main focus has been to work on the [development of] the repository. We are very conscious about the fact that we need to see and show the vice chancellor some good statistics. At the moment statistics are modest but they will be much more sophisticated and will tie in with other statistics for other repositories around..." When we asked the question, "*Who have you identified recently as the main users of EPrints SOTON?"* the reply from one interviewee was as follows: "Academic users [i.e., faculty] within the University use it for their own reasons, whether it is to create a bibliography or see what other people are doing. There are also users from outside the university, internationally. We know this because we get emails from all over the world, particularly in nursing, so nurses from Australia, Italy and other places."

At the UK Data Archives, there is a "registration system which enables the identification of the exact user and the data set that has been requested. Everyone has an individual profile and [the management team] can look at user statistics either for a particular data resource, for a particular institution, for a particular individual and so on and so forth." The chief manager of the Archive stated that "usage is concentrated in certain key subject areas. Economists and sociologists are heavy users as well as users from other core social sciences and humanities." Some of the registration statistics have shown that "in particular subject areas usage isn't as heavy, for example, social anthropology and psychology," but the manager's rationale was that "there may be a very good reason for why social anthropologists don't use the UK Data Archive, because they might have their own mechanisms for data sharing, so it's about observing patterns of use and reacting to that."

He explained further that "what [they] do in terms of targeting new users is to look at groups that don't use [the Archive] very much pro-rata and see if there is more that [they] could do to promote use within that discipline or sub-discipline. This does not involve just going to specialised conferences but actually listening to user needs. So the strategy for recruiting new users is to listen and react to their needs."

The Jorum e-learning management team said that the collection of usage statistics was "one of the things that [they were] currently looking at." JISC services, including Jorum need to create a report under their SLDs (Service Level Definitions); however, because Jorum was (at the time of our interview) still becoming a fully fledged service they were at the stage where they had been looking at what JISC would require them to report on and indeed what they could report on. One of the interviewees stated: "currently we are collecting statistics on who is logging onto the service, the number of downloads etc." and with respect to numbers, this individual was able to tell us that they were up to 140 registered Higher Education /Further Education institutional members.

7.4 User survey results

After a period of 5 ½ months (February to mid July, 2006) our Web survey of repository/digital library users was completed and we obtained information from a total of 331 individuals. The graph below (Figure 13) shows the proportion of responses corresponding to each resource.

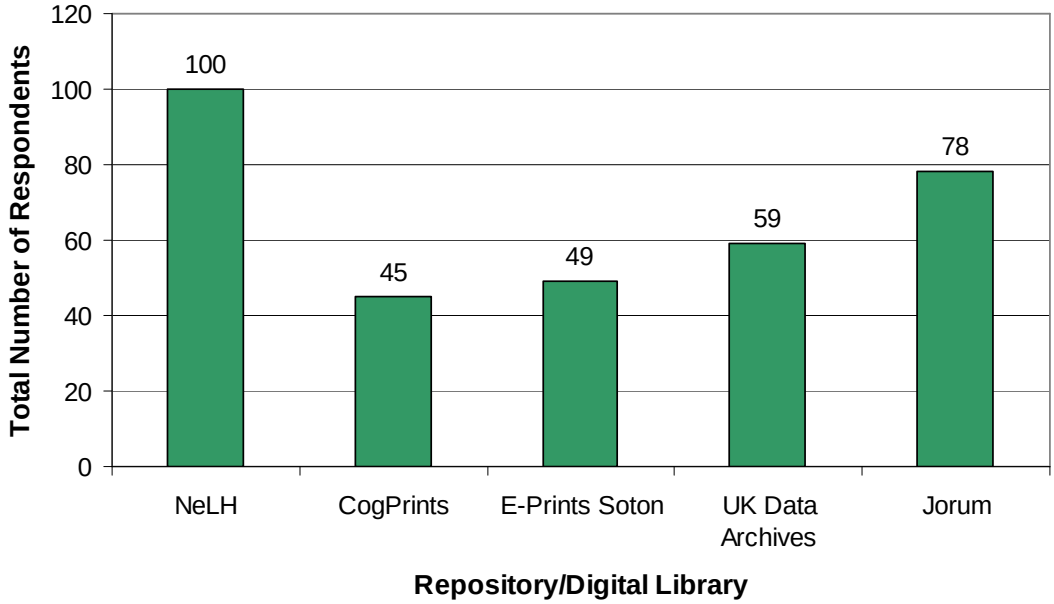


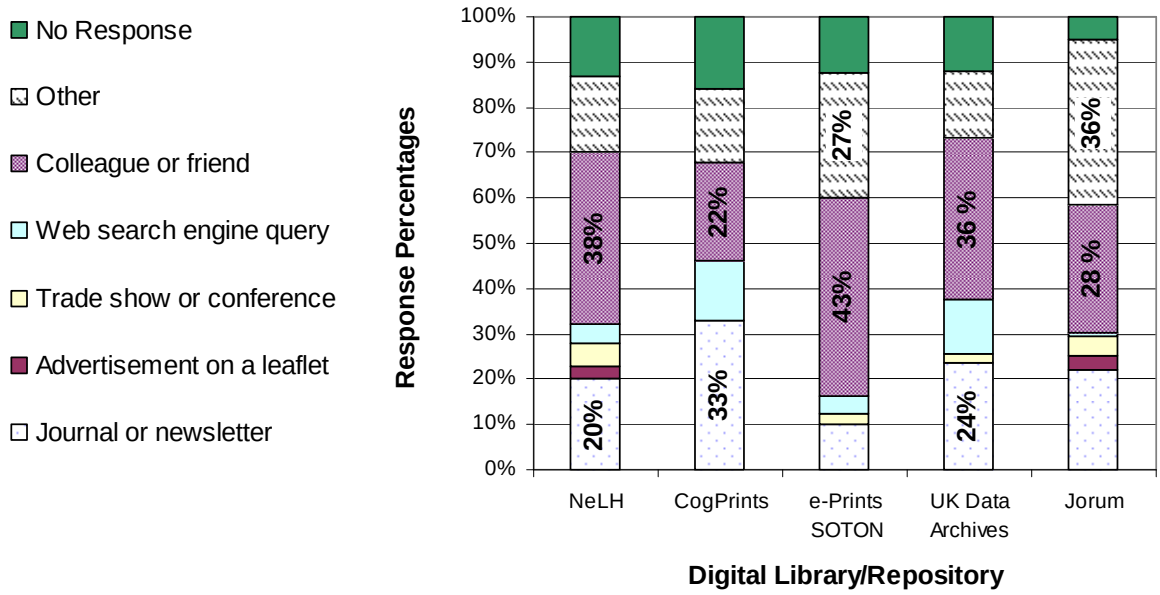
Figure 13. Total number of survey respondents corresponding to each service: NeLH (100), CogPrints (45), E-Prints Soton (49), the UK Data Archive (59) and Jorum (77).

7.4.1 Demographic Information

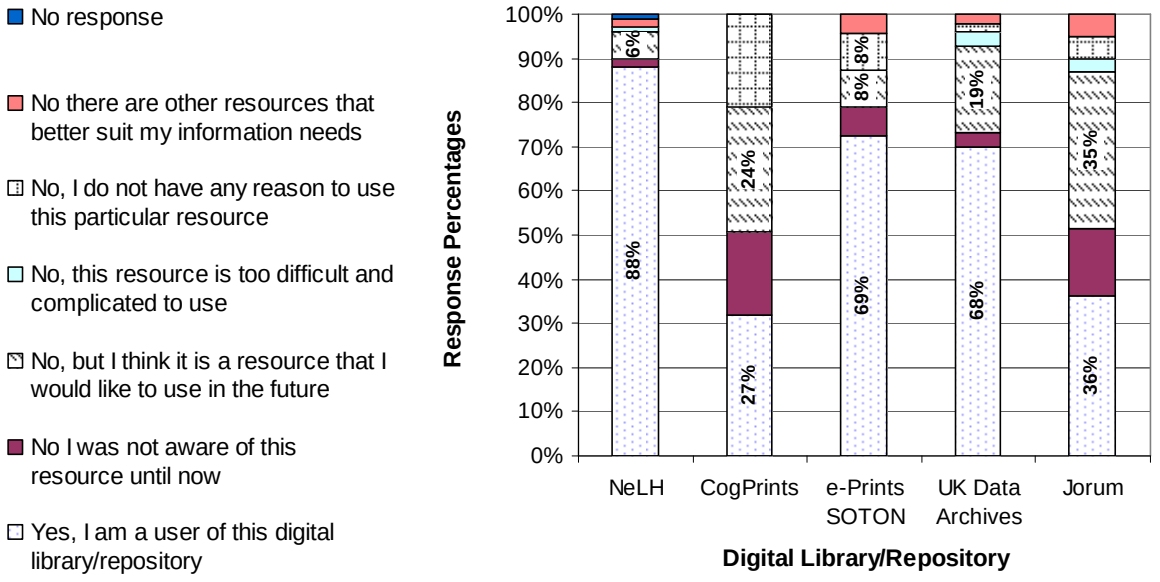
54% of the respondents were female; 44% were male (2% of the individuals surveyed did not respond to the gender question). As expected, the majority of individuals who completed the survey were between the ages of 25 to 65 (92 %) (2% did not respond to the question). 82% of our survey respondents noted that they were residents of the United Kingdom. A

proportion (16%) of individuals who responded from other countries were from the United States, Australia, Canada, New Zealand, Israel, China, India, Thailand, Haiti, Iran, Ecuador, Bolivia, Brazil, Mexico, Uruguay, Trinidad and Tobago, and parts of Europe (i.e., Italy, France, Germany, Turkey, Hungry, Finland). Most of the foreign survey respondents were associated with CogPrints, but this was expected since this subject specialty repository has a greater international focus than the other repositories/digital library evaluated in this study. In terms of profession or jobs, another expected finding was that most of our respondents were Librarians/Information Professionals or Lecturers/Professors/ Academic Researchers; however some Managers (i.e., IT or Project Managers), Nurses, Teachers, Students and Doctors, and Public Health Care Practitioners also completed the survey.

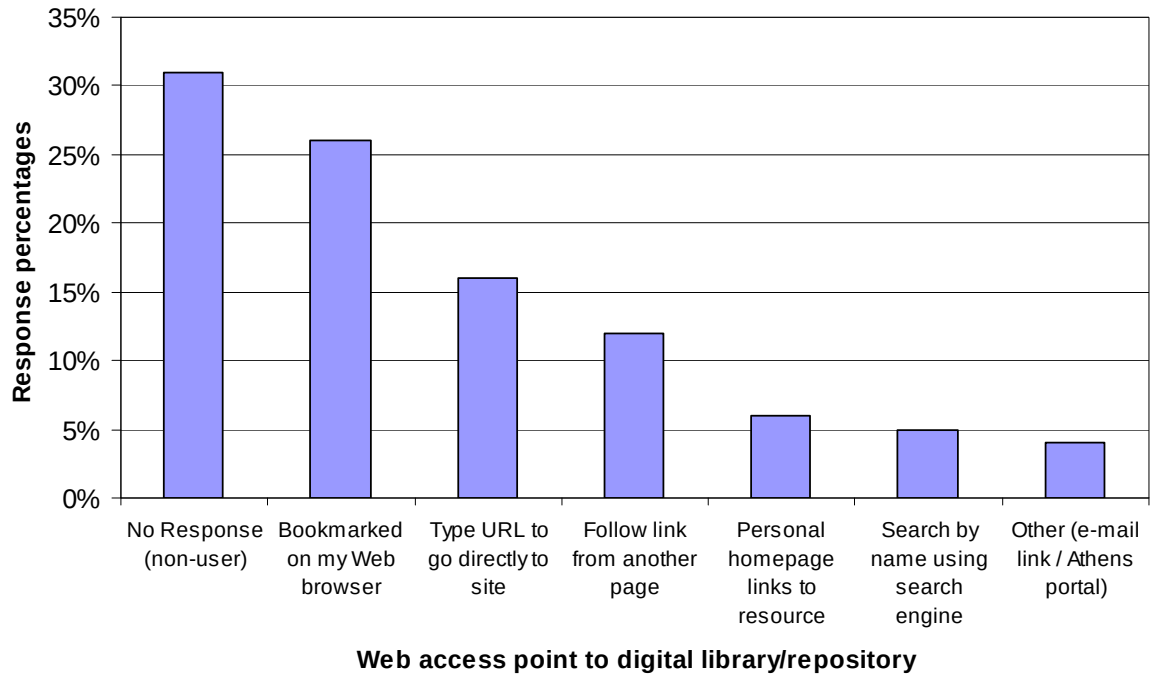
7.4.2 How did you first learn about the existence of the digital library/repository? (Figure 14).



7.4.3 Are you a user of the digital library/repository? (Figure 15).



7.4.4. What is your usual Web access point? (Figure 16).



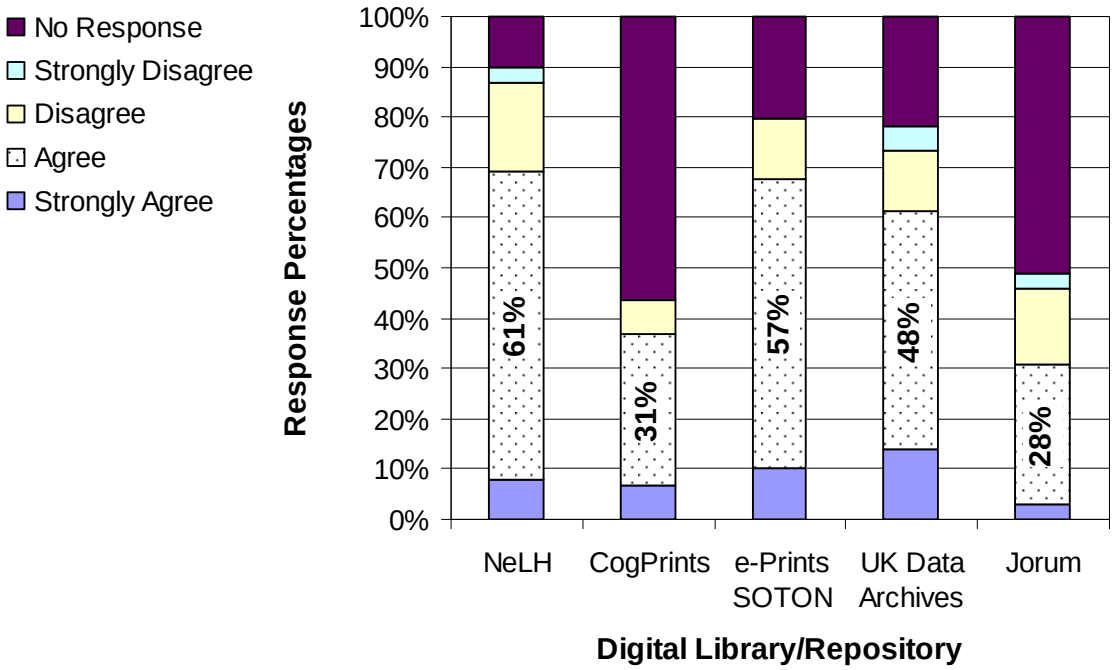
7.4.5 How often do you use the digital library/repository? (Figure 17).

We compared individual survey responses concerning Web access points to the digital library/repositories to the frequency of use responses and observed the following:

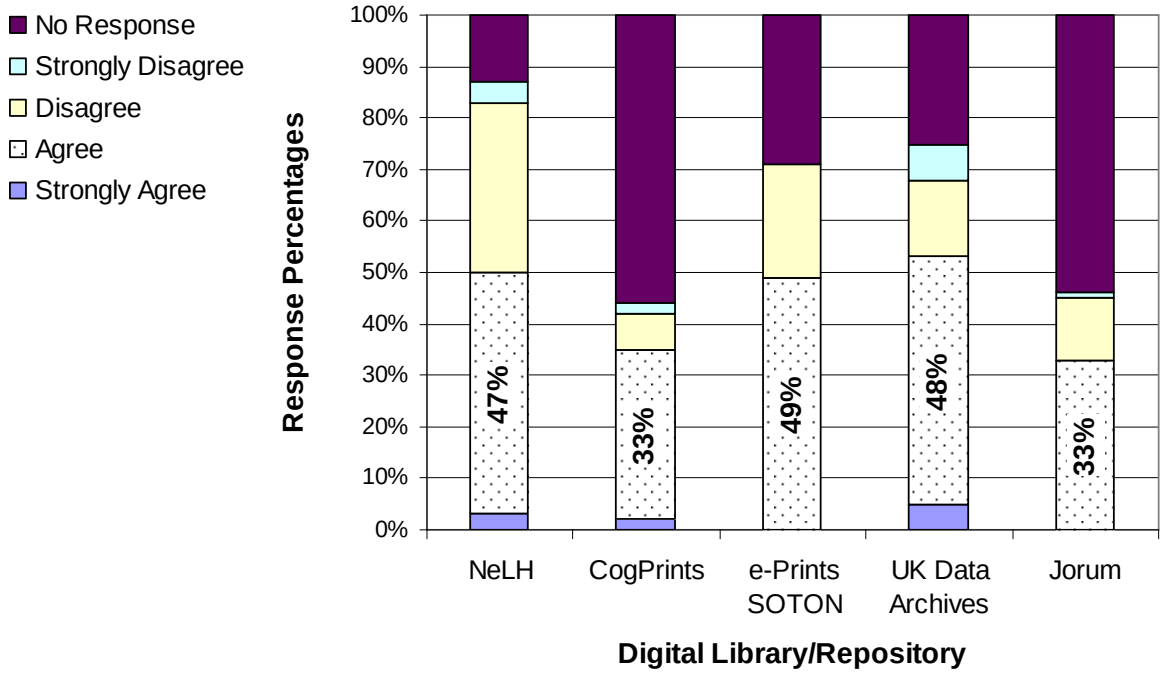
- Few users had created a personal homepage link to the digital repositories/library, although this practice could improve in the future.
- Users who created a bookmark or personal homepage link to one of the selected digital resources tended to use the resource a few times per week.
- Less frequent use of a digital library/repository was associated with searching for it by name on the Web.

Frequency of Use	Web Access Point to the Digital Library/Repository					
	Bookmarked (83)	**Type URL to reach site (50)**	**Follow link from other page (42)**	**Personal homepage link (20)**	**Search for it by name on Web (18)**	**Other (e.g., E-mail link / Athens portal / Desktop icon) (8)**
Everyday	16	2	3	3	0	0
2-3 times per week	**14**	12	6	**8**	1	1
Once per week	10	5	5	2	3	1
Approx. every 2 weeks	16	4	**6**	0	**3**	1
Once per month	4	**5**	6	2	2	0
Few times a year	23	22	16	5	9	**5**

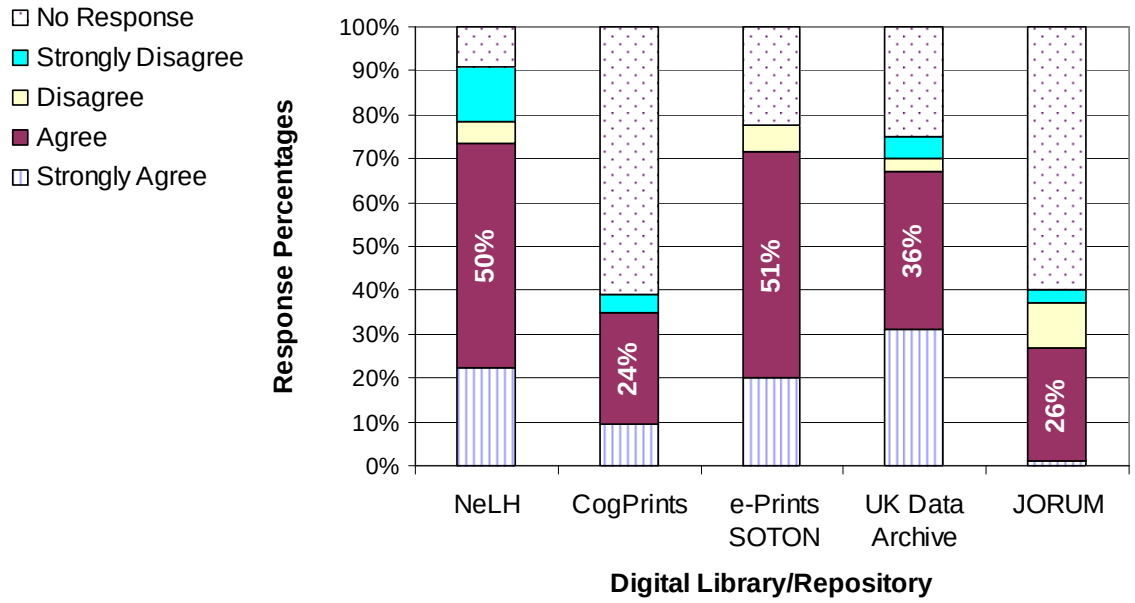
7.4.6 The digital library/repository is easy to use. (Figure 18).



7.4.7 Navigational assistance is available and helpful. (Figure 19).



7.4.8 The material on the digital library/repository is usually relevant to what I need. (Figure 20).



7.4.9 Is there any type of information that you would like to see available at the digital library/repository?

NeLH

- "It's pretty comprehensive already, I think, though I think the image resources could be better."
- "Access to a wider range of journals in full text"
- "A specialist library of substance misuse information covering drugs, alcohol, solvents and smoking."
- "Statistical information to be more obvious."
- "More statistics (national and local level)."
- "Easy access to more statistical information (e.g., number of children dying on surgical wards; number of sickle cell suffers in each PCT)."
- "I would like to see an About Us section, including staff names and contact information. I would like to be able to access policies for selection of content included in NeLH. I would like to be able to access a NeLH strategic plan. I would like to be able to access a NeLH newsletter."
- "Teaching packs available."
- "More nursing content"
- "Alerts to new copies of standard medical and nursing textbooks"
- "Public Health information similar to that which used to be available through HealthPromis and HDA online."

CogPrints

- "I like the preprints, postprints, archived presentations, etc. I've found and used. I'd be happy to see archived webcasts or podcasts."

- "Better subject delineation. I remain a bit confused about the exact subject coverage of the archive, which seems to have spread from its original aims (mind you, so has arXiv). I am also not sure about how up-to-date the archive is, though haven't investigated this in detail recently."
- "...a 'library' of the ID numbers of the publications."

E-Prints SOTON

- ".. more access to full texts."
- "Possible citation extraction or workflows."
- "Government and NGO reports; 'Grey' literature."
- "The ability to a) create reference lists in suitable formats for use in papers & theses; b) create custom reference repositories for personal research use"
- "The number of visits, the no. of downloads both with and without a later citation, citation for each paper."

UK Data Archive

- "More accessible machine readable metadata."
- "More qualitative data."
- "More business data."
- "There needs to be better links with other major information providers, such as Digimap at Edina (University of Edinburgh) towards e-social science."
- "More information re: social housing research"
- "Surveys that are comparable to UK sources from other countries (e.g. translated labour force surveys, household surveys etc.)"
- "Longer time series – i.e., integrated series that go back 50 + years"
- "The Question Bank is [quite difficult to use]. I also think that it is very confusing and overly complicated to have so many different places relating to deposited data - e.g. QB, ESDS, Data Archive. It would be so much easier for me (and I assume others) if these could be integrated. It would also be helpful to have information about private datasets with links to how to obtain these, where these are the most appropriate for a research project.
- "Retailing organisations (e.g. Tesco) have un-mined data on customers that would be good to access. Data collected by our national security services would also be interesting. Where data linkages are being made, this again should be archived annually and made available to the academic community."
- "Ethnographical Research, Discourse Analysis, Auto-ethnography."

Jorum

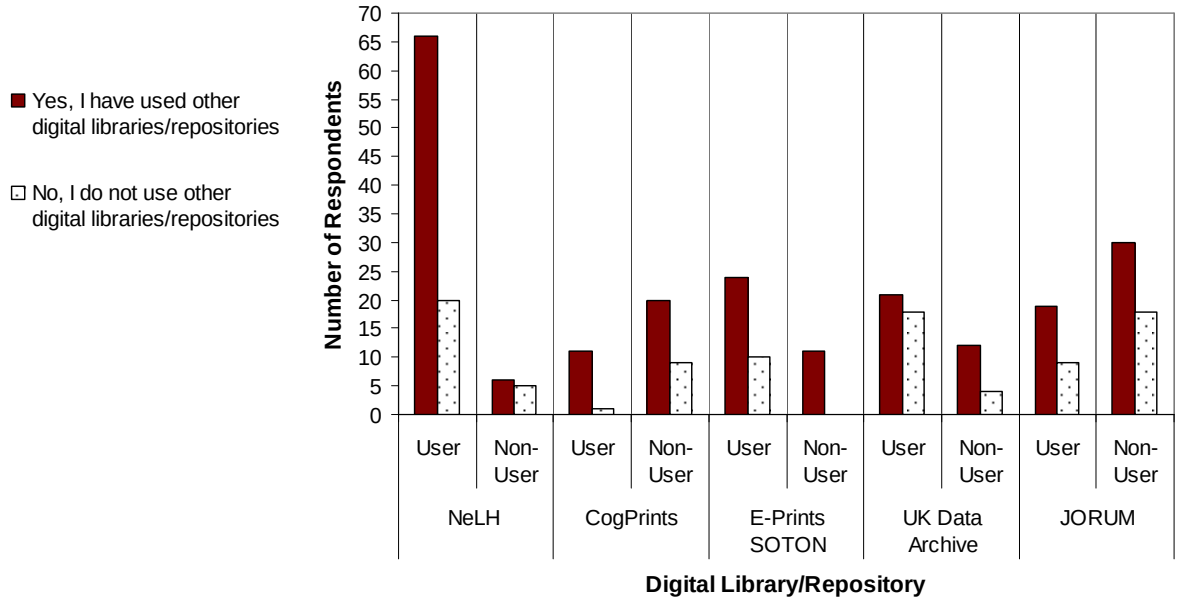
- "At the moment the resources/information on Jorum is very limited. More information for further education would be of great benefit."
- "Generic study skills; time management, research, essay writing etc."
- "[Information about] how people are actually making use of the resources they are downloading i.e. their pedagogic ideas about the use of a resource."
- "[More] relevant materials targeted at Further (rather than Higher) Educational requirements especially at the lower (Entry, Level 1 and Level 2) levels rather than the somewhat rarefied materials currently often available which need too much work for busy teachers to develop."
- "More general information on issues in learning and teaching."
- "Multi media for streaming."
- "More information literacy materials."
- "A greater range of resources e.g. research papers, project outcomes."

7.4.9 For what purpose(s) do you use the digital library/repository? Have there been any unexpected benefits as a result of your use? (Table 5).

	Reason for Use	Unexpected benefits?
NeLH	▪ for personal work/study purposes or literature searches ▪ to keep up to date on what is available ▪ for training purposes (e.g., to train library users) or for university lecturing ▪ to access electronic journals and access the Cochrane Library ▪ to access healthcare guidelines, hitting the headlines, document of the week ▪ to access health-care resources and documents relating to evidence-based policy ▪ occasionally to submit feedback ▪ to obtain references to patient information or clinical information	"I can obtain information that is clinically relevant much faster than I used to be able to" "Students particularly like the specialist libraries and the way they are laid out. Students find it easier to locate guidelines on their topics on the NELH than on other sites." "I have found information for some clinicians - an adjunct to information already retrieved." "... I accidentally stumbled onto some useful information concerning a condition I suffer from myself. This information was completely new to me and started me on the road to finding some more, which has offered me another treatment option and improved my own health."
CogPrints	▪ to access works publicized on lists or identified by links on other pages. ▪ to search for references and learn from an example of a subject-related open access repository ▪ to self-archive work and retrieve the works of others ▪ to access research papers archived with open access ▪ to stay in touch with the development of science and stay informed on what is happening in some interested fields ▪ for study purposes and to support research ▪ used for its specificity to the Cognitive Sciences	"I've found some good articles in computational neuroscience." "I get better in my work and I feel more comfortable being in touch with these great resources available for free!" "[CogPrints is a] good exemplar of what can be done in the open access arena. Useful backup resource when references cannot be sourced elsewhere (and I admit I still use such archives as a backup, not a primary source for searching)."

	Reason for Use	**Unexpected benefits?**
E-Prints SOTON	▪ to self archive journal articles, conference talks, and posters ▪ to enter new publications – to monitor the research output of research centre colleagues, and for Research Assessment Exercise simulation activities for the university ▪ to update a personal publication list (appearing also on a personal homepage) ▪ to look for important articles relating to fields of interest ▪ to gain access to full texts of articles not available in my library ▪ to monitor how the repository is developing – e.g., look at what types of material is hosted and whether it's available in full text.	"I am involved with managing an institutional repository at the Open University. I use Eprints Soton as an example of eprint practice. I tend not to use it to access research papers though I have done this a few times." "I have received contact from other researchers with similar interests." "Far more interest in my published work - especially in areas such as the Middle East and South America." "I am receiving requests for papers at a greater rate than previously." "[Repository link is] the best way to give colleagues/competitors quick access to my papers when I email them (rather than attaching things). They appreciate this." "It is quicker to review publication records."
UK Data Archive	▪ to retrieve UK comparative data for research ▪ to extract data from the British Household Panel Survey (BHPS) ▪ to look for relevant research carried out elsewhere in the UK ▪ to help library users find data for their own research ▪ to access datasets and information about those datasets (e.g., user guides, etc...) ▪ to locate information about questionnaires and methodology ▪ to download data for university users; in order to assist university users with their data requirements ▪ Research, Thesis (although I have not yet used it)	"The availability of data led to a new funded stream in my research program." "Made me aware of the existence of many more surveys relevant to health." "I have published one journal article directly based on UKDA data and have a couple of others in exploratory phases. Also, browsing through data bases is a way to keep up with various research interests. Conversely, an absence or scarcity of data may be an early indication that a contemplated line of research may be unfeasible." "Many contacts with fellow users have had tremendous unforeseen benefits."
Jorum	▪ so far have not used it but intend to discover its value and then promulgate this to others ▪ to search about the relevance of Learning Objects in the learning process ▪ Training; inspiration for development of materials for Virtual Learning Environment ▪ Evaluate learning resources for use by teaching staff. ▪ Browse to see what's new/current	"[Jorum] has encouraged me to adopt a more structured/rigorous approach to the management of learning resources within my own area." "Increased stimulation of teaching staff; motivation etc." "It is interesting to see how standards are applied / implemented e.g. metadata searching."

7.5.0 Have you used any other type of digital library or repository? (Figure 21).



7.5.1 Digital repository non-users

Amongst the individuals who completed our surveys (excluding the National electronic Library for Health – digital library) 50% percent identified themselves as repository users and 50% percent said that they were non-users. The positive outcome concerning non-users was that some acknowledged that they were still just learning about repositories (10%), and many claimed to be interested in using them in the future (22%). Only a few stated that they did not want to use them at all (8%).

Some of the survey respondents who indicated that they were not users of CogPrints, E-Prints SOTON, the UK Data Archive or Jorum, also said that they were users of other types of digital libraries/repositories; hence when we asked: "*For what purpose did you use another digital library or repository?"* the responses were as follows:

- To retrieve and use text-base resources for personal learning purposes (48%)
- To retrieve and use images, maps, diagrams, or other visual aids (17%)
- To retrieve teaching materials (27%)
- To contribute or deposit material into the repository for others to use (25%)

8. Outcomes

There are two major outcomes associated with this JISC project: 1) the descriptive statistics and written report generated from our complementary analyses, and 2) the LexiURL software program, which we have made available to all digital library and repository managers across the United Kingdom and worldwide. The first outcome – the present report – will benefit numerous digital library/repository managers because it will help them acquire new insight into users and recognize important best-management practices. It may also be used as a guide to carrying out future digital library/repository evaluations. The second outcome – LexiURL – is a valuable, new link assessment tool, which repository and digital library managers can exploit for further webometric evaluations. LexiURL, including a help

manual for downloading, installing and using this software tool, is available through the SCIT Cybermetrics Website at: http://lexiurl.wlv.ac.uk/.

9. Conclusions and Implications

This JISC-sponsored study was broad in terms of its breadth and scope, but successfully completed due to the cooperation we received from the digital repository managers and questionnaire participants. The analysis techniques selected for our research were appropriate; however, more time would have been valuable to expand upon our webometric applications. The introduction of LexiURL as a new assessment tool for digital repositories/libraries created an important challenge for our project team because the repositories selected for the study were embedded in a development or growth phase, i.e., at a fairly early stage for assessment. Our webometric reports were designed to be the innovative part of this project; thus we believe that more can be done in the future to educate repository managers on the benefits of LexiURL link analysis.

A link analysis using LexiURL can and should ideally be carried out for each of the individual repositories, including the NeLH digital library, approximately every four to six months. Over this period of time it is possible to detect changes in the co-link maps representing each resource's online link network or Web community. New links might appear, and a regular review of their context (i.e., where they are situated on a Web page, and what type of organisation is creating the link) would give managers an opportunity to think about places where new users might be "surfing" the Web and address the needs of potential user groups. From our initial analyses, we discovered that a growing proportion of links to the UK resources were coming from international websites (Example: The University of Queensland Australia directed a link to the e-prints SOTON site; listing it as a key resource on their *Databases for Social Sciences* page). As a result of this information, we are certain that managers will want to see that these links are preserved, and they will also want to know if such links are being followed as access points to their resource. A link analysis using LexiURL might therefore be used as a supplement to a log file analysis. Log files provide information about daily user activities on the Web, either in terms of the search engines used and phrases/words users' type to carry out a search, or the Web URLs (links) that are being followed. A LexiURL analysis is a complement to log file data because it extracts lists of links from the Web (via a well-known commercial search engine) that exist 'out in the wild,' which can be compared to log file (followed) URLs. Last but not least, we recommend that managers consider using LexiURL to perform comparative link analyses with 'competitor sites' or other international repositories similar in scope and purpose. If more links or different types of links are found to be directed to the site of another similar resource, then perhaps these links represent previously unrecognized users, or areas for further outreach and cooperation.

In terms of our overall research findings, we conclude this study with the following implications concerning repository management practices and repository users:

- Digital repositories/libraries are not static entities and require ongoing evaluations, not only to determine their quality consistency but to identify new directions for growth. Management teams of well-established and well-used repositories/libraries may need to become savvy with Web link statistics, download statistics or citation statistics in the future for a variety of analytic purposes, so that stakeholders will have an adequate measure of a repository's ongoing success.

- Repository uses can be as varied as users themselves; hence it is important for managers to communicate with users on a regular basis (e.g., an open forum) in order to share information with them and obtain feedback. Repository management teams who set up and maintain registration databases, listservs, or interactive newsgroups for users are engaging in a very important best-management practice.

- Based on the survey information generated from non-users, repository managers should not assume that non-use of their resource is due to an ignorance of or lack of familiarity with digital resources altogether. Potential users could be using other types of digital libraries and repositories, thus it is a good best management practice to try to find out more about what is attracting them to other repositories (online 'competitors' possibly), and develop publicity programs that will bring people up to date on what makes their resource especially valuable.

- Digital repository managers may need to give more consideration to the power of 'buzz' networking; that is, the importance of personal networking and information sharing among friends and work colleagues. A significant number of individuals surveyed for this project indicated that they had learned about The NeLH, CogPrints, JORUM, The UK Data Archive, or JORUM through a friend or colleague. Initial evidence was also found to suggest that repository user activity itself can and should contribute to collegial networking (Example: An E-Prints SOTON user on an unexpected benefit: "*I have received contact from other researchers with similar interests*.")

- Personal website links to online digital resources are normally not plentiful (e.g., Beaulieu, 2005); however based on this project's user survey we discovered that persons who frequently use CogPrints, NeLH, E-Prints SOTON, The UK Data Archive or Jorum sometimes do have a directed link from their personal website. A regular LexiURL link analysis should give the repository/digital library manager new insight into the number of personal pages linking to their resource over time, including some relevant and growing number of blogs. *What is the relationship between the source of the link and the link target? Does the source simply acknowledge the digital resource or provide descriptive information concerning parts that they appreciate, recommend to others, or have consulted to great benefit?*

10. References

Bar-Ilan, J. (2004). Self-linking and self-linked rates of academic institutions on the Web. *Scientometrics, 59(1),* 29-41.

Beaulieu, A. (2005). Sociable hyperlinks: an ethnographic approach to connectivity. In C. Hine (Ed.) Viritual Methods: Issues in Social Research on the Internet (pp. 183-197). Berg: New York.

Chakrabarti, S. (2003). *Mining the Web: Analysis of hypertext and semi structured data*. New York: Morgan Kaufmann.

Heery, R. & Anderson, S. (2005). *Digital repositories review*. University of Bath: UKOLN. Available at: http://www.jisc.ac.uk/uploaded_documents/digital-repositories-review-2005.pdf

Thelwall, M. (2004). *Link analysis: an information science approach*. San Diego, CA: Elsevier Academic Press.

Vaughan, L., & Thelwall, M. (2004). Search engine coverage bias: evidence and possible causes. *Information Processing & Management, 40(4),* 693-707.

11. Appendixes

Appendix A: Management Interview Schedule

Rationale for Creating the Digital Library or Repository

- What was the main reason for creating your digital library/repository?
- When did you first realise that a digital library/repository like the one you have created might be beneficial to registered users?
- Were there any examples of digital libraries/repositories that you followed? Why?

Development of the Digital Library or Repository

- Who was involved in the development of your digital library/repository? What was the senior management support like?
- What resources were required to set up the digital library/repository?
- How did you decide what materials would be included in digital library/repository?
- Were any training programmes used for those involved with its development? If so, what were these programmes and how were they implemented?
- Did you experience any problems/difficulties whilst you were setting up your digital library/repository? What were the difficulties and did they affect your progress?

Identification of Users and Publicizing the Repository

- How did you identify potential users or user groups during the early development stage of your digital library/repository?
- When did you introduce your digital library/repository to other institutions and members of the general public?
- Was your digital library/repository publicised? How was this done?
- Did you have any particular strategy for recruiting users?
- Do you have training programmes for users, and how have they been implemented?
- Do you collect any usage statistics of your digital library/repository? Do you track where your users come from?
- Who have you identified recently as the main users of your digital library/repository?

Benefits of the Digital Library or Repository

- Have you come across any benefits to users associated with the repository?

Web link Analysis Interview Schedule

- In what ways does the management team for your digital library/repository extract 'actionable' information from the Web in order to better serve users or potential users and their information needs?

[Present the LexiURL Web link analysis report to the interviewee and let him/her examine the information. Answer any questions and or explain the report details]

- Is there any information in this report that is surprising to you?
- Do you regard the information provided in this report to be useful to your digital library/repository programme? How?
- Do you think that a "Web Intelligence" report like the one we present would be valuable to you as part of a regular service?
- How often would you like to receive a report like this? (e.g., monthly; four times per year?)
- Is there any information in this report that you would like to see added?

Appendix B: Interviewee Consent Form

The purpose of this audio taped interview, held on the_____day of _____with Rajveen Dhiensa / Dr. Alesia Zuccala and the following participant(s):_____ _____ is to ensure all ethical standards and quality control aspects associated with this study.

As an Informant:

a) I understand that important ethical considerations concerning my rights, needs and values will be recognised throughout this interview and that there will be no known personal/organisational risks.

b) I understand that I have the option of withdrawing from this interview at any point in time, or refusing to answer any questions without any negative consequences.

c) I understand that if I participate in this interview that it will be tape-recorded, transcribed and used for future analysis.

d) I understand that the researchers assigned to this project may need to communicate with me by e-mail for follow-up questions after the interview, and that I have the option of continuing or not continuing my role as a research informant.

e) I understand that the project researchers may wish to use the information that I provide during this interview in a written research report, suitable for presentation at an academic conference or publication in an academic journal.

f) I understand that the information that I provide for any type of academic report will be associated with a pseudonym and not my real name.

g) I understand that I will have an opportunity to read and amend any or all facts included in a written report resulting from the information that I provide during the interview.

Informant's signature: E-mail address:

_____ _____

Appendix C: Digital Library/Repository User Questionnaire

The following questionnaire is designed to investigate your experience in using [name of digital library/repository (http://) and other types of digital libraries/repositories on the Web. Your participation will help to support a Joint Information Systems Committee (JISC) User Needs/Webometric Study of Public Repositories in the U.K.

A team of researchers from the University of Wolverhampton and Loughborough University is carrying out this study: http://cybermetrics.wlv.ac.uk/DigitalRepositories/index.html. If you have any questions or queries about the questionnaire please do not hesitate to contact Raj Dhiensa at Loughborough University: R.Dhiensa@lboro.ac.uk.

Before you begin, please note:

- We encourage you to complete the questionnaire regardless of whether or not you use [name of digital library/repository].
- The information gathered from this questionnaire will be kept in strict confidence, and analysed only in aggregated form.
- The questionnaire will take approximately 10 minutes to complete.

---

1. Gender

Male Female

2. Age

17 years and below 18 – 24 years
25 – 65 years 66 years and above

3. Occupation

Accountant Administrative Assistant / Secretary
Allied Health Professional Business Manager
Lawyer / Attorney / Barrister Lecturer / Professor / Academic Researcher
Librarian / Information Professional Nurse
Physician Pharmacist
Retired Student
Teacher Other (Please specify): _____

4. Country of Residence: _____

5. How did you first learn about the existence of [digital library/repository name]?

Through a journal / newsletter announcement
Through an advertisement on a leaflet or pamphlet
At a trade show / conference
Through a Web search engine query
From a colleague / friend
Other (Please specify): _____

6. Are you a user of [digital library/repository name]?

Yes, I am a user of this digital library
No, I was not aware of this digital library until now
No, but I think it is a resource that I would like to use in the future
No, this digital library is too difficult and complicated to use
No, I do not have any reason to use this particular digital library
No, there are other resources that better suit my information needs

[If your answer is YES, proceed to question number 7. If your answer is NO, proceed to question number 19]

7. What is your usual access point to [digital library/repository name]?

My personal home page links directly to the site
I follow a link to the site from another page on the Web
It is bookmarked on my Web browser as a favourite site
I search for it by name through a Web search engine
I type in the URL to go directly to the site
Other (Please specify): _____

8. How often do you use the [digital library/repository]?

Everyday
Once each week
2 – 3 times per week
Approximately every 2 weeks
Once per month
Only a few times throughout a year

9. For what purpose(s) do you use the [digital library/repository]? (If applicable, state more than one.)

10. The material on the [digital library/repository] is usually relevant to what I need?

Strongly Disagree Disagree
Agree Strongly Agree

11. Is there any type of information that you would like to see available through this digital library/repository?

12. Have there been any unexpected benefits as a result of your use of the digital library/repository?

Yes No

13. If your answer to question 12 is YES, please explain or comment on the benefits:

Read the following statements concerning the design of the digital library/repository and check one of the following responses:

	Strongly Disagree	Disagree	Agree	Strongly Agree
14. This digital resource is easy to use				
15. I tend to find the information that I need quite quickly				
16. Navigational assistance is available and helpful				
17.The files are quick to download				
18. The material is usually displayed in a format that is compatible with my PC				

19. Have you used any other digital library or repository?

Yes No

20. If your answer to question 19 is YES, please specify the name(s) of the digital libraries or repositories that you have used:

21. For what purpose did you use another digital library or repository? (You may select more than one option below).

To retrieve and use text-based resources for personal learning purposes
To retrieve and use images, maps, diagrams or other visual aids
To retrieve teaching materials for lesson plans
To contribute or deposit material into the repository for others to use

22. Would you be willing to be approached for an e-mail interview?

Yes No

If your response is YES, please provide your email address:
Note: An interview would provide us with an opportunity to ask a few in-depth questions related to your responses. It will not take up too much of your time, and you will remain anonymous; however some statements that you provide will be included in a report that can help us address important issues related to digital libraries/repositories in the U.K.

Thank you for your cooperation!