



# University of HUDDERSFIELD

## University of Huddersfield Repository

Stone, Graham

Resource Discovery

### Original Citation

Stone, Graham (2009) Resource Discovery. In: Digital Information: Order or anarchy? Facet, London, pp. 133-164. ISBN 9781856046800

This version is available at <http://eprints.hud.ac.uk/5882/>

The University Repository is a digital collection of the research output of the University, available on Open Access. Copyright and Moral Rights for the items on this site are retained by the individual author and/or other copyright owners. Users may access full items free of charge; copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational or not-for-profit purposes without prior permission or charge, provided:

- The authors, title and full bibliographic details is credited in any copy;
- A hyperlink and/or URL is included for the original metadata page; and
- The content is not changed in any way.

For more information, including our policy and submission procedure, please contact the Repository Team at: [E.mailbox@hud.ac.uk](mailto:E.mailbox@hud.ac.uk).

<http://eprints.hud.ac.uk/>

# Digital Information: order or anarchy?

## Resource Discovery

Graham Stone

### Introduction

Today's researchers use a wide variety of tools to discover the information resources they require. These resources may be located within a physical library or available on the Web and if so, are available in a proliferation of formats and interfaces, which has often meant that users were directed to a number of different in-house or external systems to find the information they required. Over the past ten years libraries have, to a varying degree, sought to manage and expose this data, while attempting to dissuade users from flocking to the simplicity of Google (Scholar), Amazon, YouTube, etc. However, the provision of these different resource-discovery systems uses an increasing amount of the resources in today's academic library.

A review of the recent literature suggests that users prefer simple search interfaces such as Google<sup>1</sup>. The implication here is that libraries often fail to make their resources discoverable and that this may in turn affect the perceived value of the library.

In a recent report for vice-chancellors and senior institutional managers the Research Information Network (RIN) stated that:

"The usefulness of the content and collections provided by libraries and from other sources depends on how easily researchers can discover, locate and gain access to them. Institutions need to ensure that their researchers can readily make use, through the library and other providers, of services that enable them to discover, locate and gain access to information sources that may be relevant to their research<sup>2</sup>.

However, do libraries have a sufficient understanding of their users to provide this level of support in the way the libraries users expect and demand? Is there a very real danger of information overload from the plethora of different systems or lack of intuitive interfaces driving users towards Google? Is this actually a problem? In a time of severe fiscal hardship how can the libraries restore their importance and reclaim their position?

This chapter looks at both traditional methods of resource discovery and the next-generation systems entering the marketplace in 2009 and asks if these will be appropriate in the next five to ten years.

### Abstract and Indexing (A&I) Databases

"40% of university libraries plan cuts to book and journal purchases next year."<sup>3</sup>

The current financial climate has led many academic libraries to question their resource budgets. With a significant decrease in the amount of expenditure on books as a proportion of the library budget<sup>4</sup>, and evidence from CIBER (Centre for Information Behaviour and the evaluation of Research) suggesting that a further cut in journal budgets may have serious effects on universities' research capabilities<sup>5</sup>, libraries will have to look to other areas of the information fund to make the required cuts, and traditional A&I databases are an obvious choice.

## *Quality vs. Cost*

It is not necessarily suggested that the culling of all A&I databases is the answer. In fact, this would prove detrimental to resource discovery in the digital world. Librarians have a duty to give their users a wide variety of resource discovery options, which must include the potential to discover research that lies outside of the libraries' full-text holdings. However, gone are the days when a library could afford to subscribe to a wide variety of A&I databases on a similar subject; the library must choose the right amount of resources depending on the perceived quality of these resources, available budgets and the needs of the users.

Over the next few years every A&I database must show value for money in order to avoid cancellation. Quality is not just about accuracy of data. Any A&I database that does not include the following is at risk from cancellation:

*Article abstracts:* many A&I databases are still only indexes, e.g. 1-2 sentences of text describing the article. This is often frustrating to the user.

*Cover to cover indexing:* although many databases list a large number of journal titles in their coverage, further inspection often reveals that this is split between core content (cover to cover indexing); secondary content (where more than 50% of the material is indexed); and tertiary content (where less than 50% of the material is indexed). Any A&I database with a low proportion of core content or a high proportion of tertiary content is at risk.

*Full-text linking:* an A&I database that effectively sends users down a cul-de-sac, by not linking or providing the means through OpenURL linking (see below) to access external full text does not promote resource discovery.

*Date coverage:* unless specifically covering an archive period, A&I databases that purport to be current, but index a high number of ceased titles are arguably less relevant to users and provide little value for money.

*Geographical Coverage:* often research requires information from specific geographical areas; if these are not adequately covered then the resource is not valuable for research.

*Publisher coverage:* libraries want to use A&I databases to achieve a good spread of publishers; if this is not the case, then the resource is little better than searching a publisher's platform. Clearly this spread may be dependent on subject.

*Intuitive interface:* is the resource as easy to use as Google? If not, users might go elsewhere.

*Shibboleth authentication, EZProxy access as standard:* any A&I database that relies on individual usernames and passwords for access is creating a barrier to use.

*Unrestricted access:* analysis of turnaways and usage data at the University of Huddersfield shows that resources that restrict access by number of simultaneous users often leads to dramatic drop in usage over a period of time as users become frustrated by turnaway messages. In addition, restrictions by location, e.g. campus use/overseas, also result in potential low usage.

*COUNTER compliant usage data:* libraries are often suspicious of resources that do not reveal usage data, or become frustrated when lack of COUNTER compliance means that accurate comparisons cannot be made.

*Does not enable federated search as a minimum:* resources that cannot be added to the federated or harvested search (see below) are effectively making their resources invisible to today's user, who expects a 'just in time' approach to resource discovery.

Another issue is duplication of content. If funds do not permit the number of A&I databases we have been used to in recent years, then two resources with similar content are not economically justifiable. The JISC (Joint Information Systems Committee) Academic Database Assessment Tool (ADAT)<sup>6</sup> allows some of the above comparisons to be made with a limited number of A&I and full-text databases. However, simple manipulation of A&I database title lists in Excel (title lists are available on most vendor websites) can pay dividends, especially when looking at duplication of titles across a range of products.

### **Citation databases**

Citation databases are a different beast. As long as citation rates and impact factors are important in research assessment across the world – and early indication in the UK for the Research Excellence Framework (REF)<sup>7</sup> suggests that there will be an element of peer review informed by bibliometrics in the 2014 exercise – then citation databases will be required as standard by Universities with a research interest.

Indeed, with a number of enhancements over the past few years, citation databases may well have a significant impact on other smaller 'general' A&I databases that are already at risk of cancellation:

- Cross-referencing between ProQuest CSA (Cambridge Scientific Abstracts) and Scopus, available since 2007, allowing reciprocal searching, linking and de-duplicating of 4,500 CSA titles in Arts & Humanities and Social Sciences and 4,500 Scopus titles in Natural Sciences and Business & Economics<sup>8</sup>. This fills a noticeable gap for Scopus in Arts & Humanities and Social Sciences coverage.
- The announcement of Thomson Reuters Links Article Match Retrieval Service in 2009, which allows subscribers to Web of Science or Journal Citation Reports (JCR) to gain easier access to data, such as citation counts, in order to enhance institutional repositories etc. Subscribers to JCR can use this service to retrieve links to the record for a given journal<sup>9</sup> allowing a real-time lookup of bibliographic metadata such as DOI (Digital Object Identifier), author, source title, etc., against the Web of Science database. If a match is found, the service will return Times Cited information as well as links to view the full record, related records page, or citing articles page in Web of Science.
- The further announcement in 2009 that Serials Solutions Summon service will feature Web of Science citing references<sup>10</sup>.

However, in 1997, Cameron suggested an alternative to the subscription database<sup>11</sup>, where, “[o]ne approach towards the development of a universal (or semi-universal) citation database would be the establishment of a consortium of universities, academic societies and research library associations devoted to the purpose. ... A universal citation database would have considerable value as a tool for both literature research and the evaluation of scholarly work and hence could act as a strong catalyst for overall reform in scholarly communication. ... Indeed, a net savings may be achieved by rational and integrated reallocation of existing resources presently devoted to bibliography preparation, curriculum vitae maintenance and literature indexing.”

Canós, Nieto and Campos take this research a stage further by suggesting that citation data should be considered part of the scientific community's heritage, and that current technology is

available to allow the “existence of a global, community-maintained citation registry, generated via the early collection of citation data” and that this would be a viable replacement to commercial systems<sup>12</sup>. Any implementation of this technology could create a sea change over five to ten years.

### **Primary material**

So far this chapter has concentrated largely on secondary sources of information. Previously, primary sources of information, such as diaries, speeches, documents and raw data have been more difficult to trace; however, in a digital environment this information is becoming more retrievable.

The work of the JISC Digitisation Programme<sup>13</sup> has been an extensive source of primary material in the UK since 2004, with 22 projects already funded<sup>14</sup> and a further 25 projects nearing completion as part of the Enriching Digital Resources strand of funding<sup>15</sup>.

In addition, the UK Research Data Service Feasibility Study final report<sup>16</sup> highlights that over the past ten years research data has remained a “substantially untapped resource” and that it is “often unstructured and inaccessible to others”. The report goes on to note that similar views have been expressed internationally in the USA and Australia.

The challenge over the next five to ten years is for universities to collate and audit this data<sup>17</sup> and make it available either through a centralised facility or through University repositories.

### **Persistent Identifier Systems (PIs)**

*‘Error 404 - page not found’*

“The main frustration is not with the research discovery services themselves but with the problem of subsequently accessing identified sources and materials. The last mile of the process which actually delivers the document or other source that has been searched for is the focus of concern, with lack of access to journal articles because of a subscription barrier being the most frequently-expressed difficulty experienced. Librarians agree with researchers that the key problem is accessing online journals rather than problems with the discovery tools themselves.”<sup>18</sup>

Persistent identifiers (PIs) attempt to solve the problems of location-based identifiers such as the Uniform Resource Locator (URL), which can often lead to problems accessing resources caused by the resource being moved, relocated or renamed while external links to the resource remain unchanged. PIs give the resource a unique identifier that will not change over time, allowing for reliable referencing and access.

Emma Tonkin identifies a number of standards “at a mature stage in development”<sup>19</sup>

- The Uniform Resource Name (URN)
- The persistent URL (PURL); the Handle system
- The digital object identifier (DOI)
- National Bibliography Numbers (NBNs)
- The Archival Resource Key (ARK)
- The Open URL

The JISC Standards Catalogue<sup>20</sup> gives a detailed explanation of these standards, including those listed above and many others. For the purposes of resource discovery, this chapter will concentrate on the DOI and the OpenURL.

## *DOI (Digital Object Identifier)*

“The DOI System was developed as a cross-industry, cross-sector, not-for-profit effort managed by an open membership collaborative development body, the International DOI Foundation (IDF) founded in 1998”<sup>21</sup>. The DOI could be described as the telephone number of a journal article; however, like any telephone number it needs a directory to be discovered. In the case of scholarly content this directory service is provided by CrossRef, the official DOI link registration agency for scholarly and professional publications. Each record in the CrossRef database consists of a triplet: {metadata + URL + DOI} in order to facilitate resource discovery<sup>22</sup>.

Since the foundation of CrossRef, over 2800 publishers have participated, adding over 20,000 journals and e-books. This resource has proved crucial to resource discovery, in that citations listed in journal articles and institutional repositories have become instantly traceable. Many article citations now come from the author complete with a DOI, something that was brought to many researchers’ attention when DOIs were required by the 2008 UK Research Assessment Exercise (RAE)<sup>23</sup>.

However, an issue with DOIs is that, despite many publishers listing an article as having a DOI, not all of them are actually registered at CrossRef, which can lead to frustration. Essentially, if the item cannot be found in CrossRef the article becomes harder to retrieve. In addition, the DOI will typically link to the publisher site, which does not guarantee access at the point of use. For example, if the library has a subscription to the article via an intermediary then the DOI would be of no use for resource discovery; this is known as the appropriate copy problem<sup>24</sup>.

## *OpenURL*

“An OpenURL enables the transfer of metadata about an item (a journal article or book, for example) from a resource, where a citation is discovered (for example, an Abstracting & Indexing (A&I) database), to a link resolver. By providing a means to tell another system *what* something is, rather than *where* it is located on the Internet (the function of a normal URL), OpenURLs provide a means for link resolvers to take charge of directing users at particular institutions or organisations to appropriate, subscribed resources for the content, be they in electronic or print form”<sup>25</sup>.

Originally developed as part of a research project by Herbert Van de Sompel and Patrick Hochstenbach at Ghent University<sup>26</sup>, the OpenURL is now a NISO Standard<sup>27</sup>; however, Culling, in his 2007 report to the UKSG<sup>28</sup> identified a number of issues and barriers such as:

- Lack of understanding and close co-operation by stakeholders
- A significant number of stakeholders that do not make their collection details available
- Lack of open engagement and transparency regarding knowledge base requirements
- Lack of understanding as to the use which the data will be put, leading to inaccuracy
- Accuracy is not a major concern of full-text aggregators
- Competition between organisations hinders data sharing
- Lack of clarity and transparency regarding standards, frequency of data updates and linking syntaxes
- The absence of a code of practice and information standards to ensure optimal knowledge base compliance in addition to the OpenURL.

## *FRBRization*

If an A&I database or institutional repository does not have a direct link to DOI or OpenURL, or the ability to add this information via a link resolver product in order to link to full-text content (whether subscribed or not), that record is of little use to today's researcher. In addition, any publisher, and there are many, who does not register a DOI /OpenURL to all articles, including editorials, etc., will perpetuate the barriers to resource discovery. Those that register the DOI but do not display the DOI or OpenURL on the page effectively make themselves ex-directory as far as Repository managers are concerned, as trying to trace a DOI via CrossRef using citation details can prove frustrating.

Separately the DOI and OpenURL have gone a long way to satisfy different needs of the community; however, in their 2006 article Apps and MacIntyre ask whether, "incorporating a DOI within an OpenURL, where available, will take advantage of the exact resolution capabilities of DOI."<sup>29</sup> This could be crucial when looking to the future where the FRBRization<sup>30</sup> (the creation of a set of records that all represent the same work) of records in metasearch and other discovery services would enable an article that may be held in the repository, indexed in an A&I database, and held in full text through a journal publisher site or via an intermediary, to be displayed as one record.

Meyer reported that 59% of readers in a survey by Canada's National Research Council wanted linking from references to cited articles<sup>31</sup>. However, reference quality and accuracy have been weaknesses in scholarly publishing for many years. Meyer goes on to describe a number of areas of best practice for publishers, including checking references earlier in the submission process. In addition A&I databases can include incorrect citations; a quick search of one database in preparation for the opening section of this chapter revealed 18 records, one of which led the author to the wrong issue of the required journal.

In 1998, Tim Berners Lee suggested that, "Cool URIs don't Change"<sup>32</sup>; however, over ten years later, PIs are not routinely applied, and this is essential in the digital information world.

### *The key to resource discovery*

In January 2008, UKSG and NISO jointly launched KBART: Knowledge Bases And Related Tools working group<sup>33</sup> to:

1. Develop and publish guidelines for best practice to effect smoother interaction between members of the knowledge base supply chain. Knowledge base providers and their customers (primarily academic libraries) will benefit from provision of higher quality data by content providers. Publishers will benefit from accurate linking to their content and subsequently the possibility of increased usage.
2. Provide educational events that will address the roles of each (relevant) party in the supply chain, the functions each needs to carry out to improve supply of data to knowledge bases, and the value of doing so in each case. Content providers will benefit from a greater understanding of the needs and activities of those to whom they supply data; knowledge base providers and libraries will again benefit from improvements that can be expected when content providers are better educated.
3. Deliver a centralised information portal, to support educational activities and provide a comprehensive resource for further learning

The KBART Report from February 2009 indicated that progress is being made:

“We’ve also recently held discussions with CrossRef about how they can complement KBART’s work by providing related services to their members. CrossRef will likely proceed with developing a service that offers to collate its members’ KBART-related data for onward distribution to, or collection by, link resolver managers.”<sup>34</sup>

The success of this working group is essential if resource discovery is going to overcome the present barriers.

## **Federated (Meta) Searching**

Federated searching (or meta searching) is now a mature product, having been with us now for over ten years<sup>35</sup>, and since then a number of products have come and gone from the marketplace. However, is it the Holy Grail<sup>36</sup> that we were once led to believe it was?

“The jury is still out on federated search systems, even though more libraries now have them. There are murmurings that federated search has lower-than-expected use and may not be the magic search bullet we were led to believe”<sup>37</sup>.

Many librarians do not recommend federated search, despite actual usage increasing<sup>38</sup>. Anecdotal evidence from users’ comments at the University of Huddersfield concurs. Problems with federated search rank alongside lack of library text books as the perennial favourite in the students’ comments. So what are the arguments for and against?

### *Benefits*

*Simultaneous Search:* possibly the very reason many libraries purchased federated searches in the first place – an opportunity to take the fight to Google.

*Personalisation features:* the ability to personalise database and e-journal lists and to save searches are seen as benefits for the federated search.

*OpenURL:* integration with an OpenURL resolver is an essential part of any resource discovery platform.

### *Disadvantages*

*Cross-database Search:* users do not inherently know which databases to search when faced with having to search each individually. Although metasearch systems allow users to search multiple databases simultaneously, that does little in itself to resolve the question of why a user would choose to search, say, Project Muse and JSTOR simultaneously over WorldCat and Academic Search Elite<sup>39</sup>. Walker argues that in order to allow users to customise their searches, librarians are in fact reverting to old style lists.

*Non-searchable content:* no federated search can search all subscribed databases for a number of reasons; permission is withheld by the vendor, the resource is not compatible with federated search; or there is a restriction on access, e.g. simultaneous users, username/password access. In some cases, e.g. law, this can render federated searching of little use; this may result in the federated search being very difficult to market from the librarian’s point of view.

*Dumbing down:* the limited search capability means that controlled vocabulary, indexes or thesauri, or in some cases even date-range or peer-review filters, are lost to the federated searcher. This

may result in the possibility of inferior results; again, this often discourages librarians from marketing federated searching to certain users, e.g. researchers, faculty.

*Lack of direct links to full-text content:* although it is no more than a few clicks away using the OpenURL resolver, the federated search still brings no more full-text content back than searching A&I databases individually. This is one of the main reasons why librarians at the University of Huddersfield do not recommend federated searching for undergraduates who require full text immediately.

*Speed:* a federated search is only as fast as the slowest database to retrieve results, and therefore when searching across a range of databases, federated searches have a tendency for extreme slowness or even timing out, and this does little to promote this method of resource discovery.

*Interface too complex, not intuitive:* faculty and researchers have their favourite databases, and librarians will quickly be able to recommend a top three resources by subject; the federated search brings a new unfamiliar and perhaps unwieldy interface to the user for all the reasons stated above.

*Not Google-like enough:* although the prospect of a one-stop shop is tempting, federated searches do not find a variety of material that can be of use to the researcher, e.g. this chapter references a number of blogs and web pages, and how discoverable is this data? For almost as long as Google has been in existence, librarians have been fighting it, but it does have its uses! In addition, institutional repositories using Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) can be cross-searched, but this is very basic and is unavailable in federated searches. This could mean that open access via a university repository to a search result is not evident.

Are academic librarians correct in their assumptions<sup>40</sup>, “that the best way to lead our students to quality research resources... ..is to implement federated search interfaces that emulate the “best” features of the Internet” and that “ease and convenience are more important than providing in-depth searching via native interfaces...”. Of course, the real answer to this is usability testing. The historical answer to providing resource discovery for users was to duplicate lists in the library catalogue, web pages, metasearch products, A-Z lists etc.; no wonder then that federated searching has proved popular as users require a one-stop-shop approach<sup>41</sup>.

Tenopir stated that “[f]ederated systems remain controversial because they focus on what we think users want, at the expense of functionality, precision, and finesse. They are still a long way from providing a single, simple solution to information retrieval.”<sup>42</sup> Given that usability testing shows that the idea of a one-stop shop is popular in principle with students, but the concept is unpopular with librarians, is federated searching a transient technology like the CD-ROM and is the real Holy Grail just over the horizon?

### **Pre harvested search – the true one-stop shop?**

“The second major approach to federated search is to harvest all of the relevant sources of data, normalize them into a single metadata schema, and index all of them together in one large union index. This approach offers huge advantages in speed and in the logic that can be applied to the presentation and sorting of results.”<sup>43</sup>

Since January 2009 a number of products have been announced for beta or open beta testing, the big three are Serials Solutions Summon<sup>44</sup>; Ex-Libris Primo<sup>45</sup> and OCLC Worldcat Local<sup>46</sup>. These products intend to move a step beyond traditional federated-search products by creating a union

index solution by harvesting content. These products present a new generation of resource discovery by attempting to provide the best bits of federated search while eliminating the downside.

*Single search:* a Google-like search box.

*Relevancy ranking:* by using open-source algorithms such as Lucene, these products rank results so that relevant results appear at the top, rather than the results from the quickest database; this provides the quality of search results, rather than the quantity of a Google search.

*Hosting:* these services offer hosted support on centralised servers.

*Integration:* open APIs are used in order to integrate the products with existing library systems or next generation catalogues such as VuFind etc.

*Clustering and faceting:* these features are easier to implement with a hosted union index.

*Full text and peer review:* results can be limited to full-text and/or peer-reviewed content only, finally allowing fast full-text scholarly results to be displayed for the user.

*'Did you mean' suggestions:* guidance for alternative spellings, misspellings and results with low hits.

### *Union Index*

The harvested data does not only cover A&I databases; the library finally has the opportunity to add local holdings information, such as the library catalogue (including real-time loans information), the institutional repository, and locally and nationally digitised primary material. This in turn creates the potential to allow libraries to avoid duplicating their own lists; for example, if the majority of the library's e-journal and e-book holdings are directly searchable in the pre-harvested index, either through full-text availability via publisher web sites or aggregators, or via OpenURL from databases and institutional repositories, the need for this information through A-Z lists and library catalogues is minimal.

Can we really move away from A-Z lists completely? Can the library catalogue revert to a record of the print holdings in the library, essentially a facet of the union index? There are certainly benefits to the idea of a union index for resource discovery; however, there is still some debate as to whether this is the right choice.

"The danger with relying on any one service to provide you with access to its indexed content is that the service's criteria for source selection may not be yours."<sup>47</sup>

However, a counter-argument is that users of federated searching do not necessarily choose A&I databases either. Indeed as highlighted by Rochkind in his reply to the above blog post, "EVERY SINGLE content provider does NOT make their content available for federated search in the first place. Of the approximately 800 licensed databases we have listed in our collection, only about 300 are federated search-able. The remainder are largely not there because of lack of functionality on the content provider's end, not on our fed search vendor's end."<sup>48</sup>. In fact the recent press release from Serials Solutions seems to indicate that some content is more readily available to harvested search than federated search<sup>49</sup>.

Indeed, Google does not index everything on the web, but it indexes a lot. No tool for uncovering academic information is entirely comprehensive in its coverage. Research at Stockholm University

shows that students were not enthusiastic about Google Scholar or MetaLib; however, they agreed that Google Scholar was easy to use<sup>50</sup>.

### *Compare the market*

The success of the new breed of 'web discovery services' will be seen over the next five to ten years; however, the fact that three of the major players have all thrown their hats into the ring shows that they must believe there is significant benefit in developing such systems. It will be up to libraries to gauge whether the benefits of harvested search outweigh the benefits of federated search<sup>51</sup>.

Libraries will need to plan their e-resources strategy accordingly before assessing this new market, by developing a plan to see where they want to be in five years' time and mapping out the journey of how they want to get there. Libraries need to prioritise which of the benefits of harvested/federated search are most important:

- a first-class search engine for your users
  - increased user satisfaction
  - increased usage
- a single interface
  - 'one-stop shop' approach
- improved system management (interoperability and flexibility)
  - mix and match
  - future-proofing
- improved system management (management and administration)
  - less duplication of effort/more efficiencies
  - better control of subscriptions
- improved value for money
  - within existing budget

The order of these priorities should influence the terms of reference and will therefore assist in weighting any evaluation of products and services<sup>52</sup>.

With neither solution, as yet, providing a search of all A&I content, a switch from federated to harvested search may, in reality, mean a switch from one set of incomplete resources to another. However, in the long run, users will choose the method of resource discovery that they prefer.

### **Understanding our users**

The change in user behaviour as users move towards a fully digital library brings about a real possibility that librarians and libraries will become removed from their users<sup>53</sup>. However, this change in behaviour also creates an opportunity to collect user information and to monitor use; Walker likens this to being sat on a goldmine<sup>54</sup>.

### *Text and data mining*

Unlike the physical library, where little objective information could be gleaned from irregular usage checks, it is no longer satisfactory to think we know what is best for the user. We can utilise usage statistics to inform us of the potential value of a given resource, but we can also look much deeper into how users behave: which sites do they refer from, how long do they stay, what do they do when they are here, where do they go? This allows librarians to make informed decisions<sup>55</sup>. It does

not mean to say that we cannot select or recommend what the user may need, but we must do it in an informed way and in their language.

At the University of Huddersfield, Web 2.0 style recommender services and usage logs are being used to create additional features on the library catalogue to enhance resource discovery<sup>56</sup>. Although some of this functionality refers to both electronic and print resources, it should be noted that for many subjects print is (and will continue to be) a major source of much of the scholarly content and is therefore crucial to resource discovery, even in a digital world.

*Zero hits:* accounted for 23% of searches over a six-month period; this now automatically generate a 'did you mean,' '...or maybe spell checker and 'serendipity searches'. The serendipity search generates suggestions by running the search against a number of external reference web sites (including [www.answers.com](http://www.answers.com)). The results returned from these sites are then compared against the catalogue to generate a series of potentially relevant keyword searches. Users are then invited to use their own judgement to decide if the suggested searches (and also the results they generate) are relevant or not to the original search.

*Renew prompt:* during the six-month analysis of keyword searches, the word renew was entered 74 times. The following text now appears to aid the user: "To renew items you currently have on loan, please click on the "My Account/Renewals" tab."

*Tag cloud:* a visual depiction of the most popular keywords from the last two days; although the data often carries the same basic set of keywords, there are interesting anomalies throughout the year.

*People who borrowed this also borrowed:* analysis of borrower history has led to the creation of a suggestions list for additional material.

*Items with similar subject headings:* in order to give suggestions for similar items.

*Other editions:* using web services from OCLC<sup>57</sup> and LibraryThing<sup>58</sup> to locate other editions and related works within the holdings.

*Email alerts and RSS (Really Simple Syndication) Feeds:* to help users monitor new items added in their areas of research.

*ticTOCS:* The Journal Tables of Contents Service (ticTOCS) have made their data available, so it is now possible to add links to journal RSS feeds direct from the catalogue<sup>59</sup>.

*User rating and comments:* are available to add from the record.

*Loans per year:* are available for all print copies for users to gauge the popularity of the items.

*Google book search:* to allow users to sample a book online before they borrow/reserve it.

There is anecdotal evidence to show that the 'did you mean' spell checker follows the borrower usage pattern, the borrowing suggestions usage peaks one month after the peak in borrowing, perhaps indicating that users cannot locate the items they need as they are on loan, and so are using the suggestions option for related reading. The keyword tag cloud clicks peak during October, again suggesting that new users are unsure of the content they require. In addition the number of unique titles borrowed per calendar year rose significantly since borrower recommendations were soft-launched.

This is a largely print-based example; however, the implication for digital information is clear: these are the type of features that users are expecting from Google, Amazon, etc. The secret for libraries and resource-discovery vendors over the next five to ten years is how to go one stage further than the Huddersfield example, by using text and data mining to recommend both digital content at the user's fingertips and physical content, at the user's local library in one place. How to move away from the traditional A&I database experience – your result returns zero hits – by adding value, flexibility and personalisation.

### *Personalisation*

Following on from highly successful usage-based recommendations by commercial services such as Amazon.com, Ex-Libris use the bX recommender service, a collaboration of Ex Libris and the Los Alamos National Laboratory (LANL)<sup>60</sup>, to provide recommendations based on usage data as an add-on service for customers. This move prompted the University of Huddersfield to start logging data from its SFX menu requests to see if there was potential in a similar in-house system. Although the test is at a very early stage, and will never be able to utilize the sheer amount of data generated by Ex-Libris, there is enough data to make recommendations at journal level<sup>61</sup>. The quality of these recommendations is now being tested.

Low usage or non-usage of resources is a perennial problem in libraries. However, the use of anonymised personal data may be of use in encouraging others to use systems; e.g. a first-year psychology student would be able to log in to the University portal to check email, University virtual learning environment (VLE), etc., the portal would also be able to recommend which books and journal articles to use for the course based on previous undergraduates' experience without the student ever needing to enter the library or use traditional methods of resource discovery. This may be seen as dumbing down or spoon-feeding; however, if it addresses the needs of the non-user by introducing resource discovery through recommendation or "accidental discoveries", then there is a benefit of such as system<sup>62</sup>.

### *Privacy*

Is the cost of increased personalisation an invasion of privacy? Surely, to understand users' needs we need to research them? In general, we do not object to using supermarket store cards, but these are not for our benefit, they are so that the store can see what we buy in order to better market their products to us.

"Each time that your card is swiped, it registers not only the total amount spent, but exactly what you have spent it on. The supermarkets can therefore build up a customer profile of each card holder."<sup>63</sup>

A recent blog from Tony Hirst entitled "Why Private Browsing Isn't..." discusses private browsing in the latest browsers, but then goes on to discuss how Flash's 'Stored Objects', an equivalent to cookies, sits outside of private browsing and thus enables sites such as the BBC's iPlayer to resume a programme if you change browser<sup>64</sup>. Once again the use of this data is not considered an invasion of privacy.

Jenny Walker suggests that librarians have been very cautious in this area.

"To date, user information-seeking behaviour data has been largely overlooked for enhancing library services, but steps are being taken in this direction, such as the development of recommender services and new metrics for scholarly evaluation."<sup>65</sup>

This chapter does not intend to tackle the finer points of data protection and ethics, but rather to make the point that the use of *anonymised* data is the real issue. It is only through the use of anonymised personal data that resource discovery can be improved/ personalised in the ways described above. The only real concern is when a particular cohort is so small that usage can easily be mapped to individuals.

“Data Protection places significant demands on such an undertaking. Submitted records will not include individual user details and will be aggregated at the level of course /unit of study and item (e.g. book title). Furthermore, in the data used to derive activity patterns (‘users who did this also did that’), lone transactions in a given group will also be removed.”<sup>66</sup>

The JISC MOSAIC (Making Our Shared Activity Information Count) project plans investigate the technical feasibility and issues around exploiting data to assist resource discovery and evaluation in Higher Education. Data derived from

- The circulation module of Library Management Systems (the initial project focus)
- ERM (Environmental Resources Management ) system / Resolver journal article access
- VLE resource and learning object download
- In addition, reading lists (from a variety of institutional sources, without activity data) may provide key indicators

This work builds on the previous JISC TILE project<sup>67</sup> and the work of David Pattern at the University of Huddersfield<sup>68</sup> who released over 80,000 titles derived from a pool of just under 3 million circulation transactions spanning a 13 year period, a major portion of book circulation and recommendation data, under an Open Data Commons licence<sup>69</sup>.

Pattern commented, “This isn’t about breaching borrower/patron privacy — the data we’ve released is thoroughly aggregated and anonymised. This is about sharing potentially useful data to a much wider community and attaching as few strings as possible.”

### **The vision for the next five to ten years**

“The impetus for change will come from students themselves as the behaviours and approaches apparent now become more deeply embedded in subsequent cohorts of entrants and the most positive of them – the experimentation, networking and collaboration, for example – are encouraged and reinforced through a school system seeking, in a reformed curriculum, to place greater emphasis on such dispositions.”<sup>70</sup>

The key for the medium term is to provide Google-like interfaces with Google-like results; there is a race to provide this through systems such as Summon, Primo, WorldCat Local, which will all use preharvested data rather than federated searching. WorldCat Local already provides social-networking tools.

### *Open source*

This chapter has so far not touched on resource discovery via library management systems and next generation catalogues in any detail, as it is anticipated that open-source Library Management Systems such as Evergreen<sup>71</sup> and KOHA<sup>72</sup> and next generation catalogue/library resource portals such as VuFind<sup>73</sup> will come to dominate the market over the next five to ten years. The implementation of open-source systems is gathering pace in the US; and in Europe, with the establishment of companies such as PTFS Europe<sup>74</sup> and the addition of previously missing modules such as Serials and Inter Library Loans soon to be released, these systems will soon

become a reality. Some existing library management vendors, such as TALIS, already have an eye on the future with the Juice Project. Its aim is to make it easy to enhance library catalogues without duplication of effort, or as Richard Wallis puts it “slap some mash on your OPAC”<sup>75,76</sup>.

### *Mashups*

“The Mashup, where data from one or more sites is brought together to add value to the data on another site, is only four years old”<sup>77</sup>.

The truth is, we will not know how the next Google will look until it comes along – it will certainly include mashups with a whole lot more personalisation and Web 2.0 functionality, such as blogs, personalized accounts, alerts, plug-ins and APIs, some of which may fall by the wayside as new features are developed<sup>78,79</sup>.

We may even be in a situation where we can do it ourselves or use free stuff, e.g. Tony Hirst’s work with Yahoo Pipes<sup>80</sup>, where he uses Scott Wilson’s jOPML (Journals into OPML)<sup>81</sup>, an application built using the data exported from the ticTOCS project<sup>82</sup> to:

- pull in a list of journals in a particular subject area based on user-provided keywords into the Yahoo Pipes environment
- pull in the most recent table of contents from those journals into that environment
- then filter those recent articles to only display articles on a particular subtopic

Ideally, harvested searches plugged into open-source library management systems/institutional repositories, etc., will become the preferred choice allowing libraries to make sense of the anarchy by choosing separate, interoperable modules from a variety of existing suppliers, open-source communities and developers.

### *Visualization*

“The main goal of *data visualization* is its ability to *visualize* data, communicating information clearly and effectively. It doesn’t mean that data visualization needs to look boring to be functional or extremely sophisticated to look beautiful.”<sup>83</sup>

Carol Tenopir asks is the “search box solution really the best for facilitating research?”<sup>84</sup> The work of the Human-Computer Interaction Lab<sup>85</sup> and Edward Tufte<sup>86</sup> amongst others has led to data visualization gaining acceptance in recent years. Recently, Brendan Dawes (perform any search on his website to watch your results move across the pages as a number of insects<sup>87</sup>) created DoodleBuzz.com, a way to read the news through an experimental interface that allows you to create topographic maps of current news stories<sup>88</sup>. He was also commissioned by the BBC to redesign their Memoryshare website<sup>89</sup>; when this was shown at the recent Mashed Library 2009 event<sup>90</sup> there was a palpable ‘sharp intake of breath’ from the librarians in the room at the replacement of a perfectly functional search box with a data visualisation page. This is an evolving area and will certainly be a part of the resource-discovery process in years to come. Certainly today, in its simplest form, it can transform a two-dimensional repository into a resource of obvious appeal to potential users, as the recent JISC funded Kultur project shows<sup>91</sup>, or create a virtual bookshelf to reintroduce serendipity into library catalogue<sup>92</sup>.

### *What can’t we afford not to do?*

This chapter started by considering how A&I databases were being slashed in order to balance the budgets. However, there is also a cost in going to any new system, whether it is the new

generation of harvested searches or the “free beer, free kittens” approach that some see as being represented by the open source systems<sup>93</sup>.

Although the cost is there, can we afford to stay as we are? Two very different sessions at the 2009 UKSG Conference asked, “*Why is Google so easy and the library so hard?*”<sup>94</sup> and “*Why do we want to teach our users to be librarians?*”<sup>95</sup> They are important questions, and in themselves are justifications for investing in new systems even in a period of financial hardship. With increasing attention being paid to National Student Surveys, resource discovery and e-resources provision can often be an open goal regarding negative student satisfaction.

“We’re facing challenging budget years ahead. It’s essential that we raise the profile of the library and demonstrate real value”<sup>96</sup>

New systems should allow us to avoid those duplications of A-Z lists on the web pages, the library catalogue, the link resolver, while bringing in isolated collections in other parts of the university, such as the repository, etc. What our users want is a Google search with Google-like results<sup>97</sup>. We must not make the mistake of assuming that Google and library resource-discovery systems are mutually exclusive, neither should we assume that just because Google is the first point of call it is the only one - one size does not fit all<sup>98</sup>.

“Google indexes thousands of pages every day that are outside the realms of ‘traditional’ literature and academic discourse...”

...There is information out there that doesn’t exist anywhere else. Google is a superb tool for locating it.”<sup>99</sup>

Our job over the next five to ten years is to provide a way to access these valuable resources in an intuitive, easy to use one-stop shop, and not to be afraid of running a continual beta test where new services and functions can be added as and when necessary. To do this we need flexible, interoperable resource-discovery systems based on open source software. In addition, we must keep evaluating users’ needs and reach out by adapting our systems to fit their requirements, rather than expecting them to come to us; indeed our very future depends on it.

## References

1. Nicholas, D, Huntington, P, Jamali, H R, Rowlands, I and Fieldhouse, M, Student digital information-seeking behaviour in context, *Journal of Documentation*, 2009, 65(1), 106-132. DOI: 10.1108/00220410910926149
2. Research Information Network, *Ensuring a bright future for research libraries: A guide for vice-chancellors and senior institutional managers*, 2008, London, RIN.  
[http://www.rin.ac.uk/files/Ensuring\\_bright\\_future\\_guide\\_Nov08@.pdf](http://www.rin.ac.uk/files/Ensuring_bright_future_guide_Nov08@.pdf) (accessed 28 July 2009)
3. Corbyn, Z, Expect few new titles in library as sterling's fall pounds acquisitions, *Times Higher Education*, 21 May 2009.  
<http://www.timeshighereducation.co.uk/story.asp?storycode=406612> (accessed 28 July 2009)
4. Great Britain Parliament House of Commons Science and Technology Committee. *Scientific publications. free for all?. tenth report of session 2003-04*. House of Commons papers 2003-04 399-II.  
<http://www.publications.parliament.uk/pa/cm200304/cmselect/cmsctech/399/39902.htm> (accessed 28 July 2009)
5. CIBER, *Evaluating the usage and impact of e-journals in the UK. Journal spending, use and research outcomes: A UK institutional analysis*, CIBER working paper 2, 2008, University College London, CIBER.
6. JISC Academic Database Assessment Tool (ADAT):  
<http://www.jisc-adat.com/adat/home.pl> (accessed 28 July 2009)
7. Research Excellence Framework:  
<http://www.hefce.ac.uk/Research/ref/> (accessed 28 July 2009)
8. Summary of Latest Enhancements in Scopus:  
[http://info.scopus.com/june\\_07/](http://info.scopus.com/june_07/) (accessed 28 July 2009)
9. Thomson Reuters Links Article Match Retrieval:  
<http://isiwebofknowledge.com/directlinks/amrfaq/> (accessed 28 July 2009)
10. Thomson Reuters to participate in the Summon Service:  
[http://www.serialssolutions.com/assets/attach\\_news/PR-Summon-WoS.pdf](http://www.serialssolutions.com/assets/attach_news/PR-Summon-WoS.pdf) (accessed 28 July 2009)
11. Cameron, R D, *A universal citation database as a catalyst for reform in scholarly communication*.  
<http://elib.cs.sfu.ca/project/papers/citebase/citebase.html> (accessed 28 July 2009)

12. Canós Cerdá, J H, Nieto, E M and Campos, M L, What's Wrong with Citation Counts? *D-Lib Magazine*, 2009, 15(3/4). DOI:10.1045/march2009-canos
13. JISC Digitisation Programme:  
<http://www.jisc.ac.uk/digitisation> (accessed 28 July 2009)
14. JISC Digitisation Programme: Digi posters: At-a-glance guide to all the digitisation projects:  
<http://www.jisc.ac.uk/media/documents/programmes/digitisation/digiposters.pdf> (accessed 28 July 2009)
15. Enriching Digital Resources, JISC:  
<http://www.jisc.ac.uk/whatwedo/programmes/digitisation/enrichingdigi.aspx> (accessed 28 July 2009)
16. UKRDS, *The UK Research Data Service Feasibility Study*. Report and recommendations to HEFCE, 19<sup>th</sup> December 2008.  
<http://www.ukrds.ac.uk/HEFCE%20UKRDS%20Final%20Report%20V%201.1.doc>  
(accessed 28 July 2009)
17. Data Audit Framework:  
<http://www.data-audit.eu/> (accessed 28 July 2009)
18. Research Information Network, *Researchers and discovery services: behaviour, perceptions and needs*. A study commissioned by the Research Information Network, 2006, London, RIN.  
<http://www.rin.ac.uk/researchers-discovery-services> (accessed 28 July 2009)
19. Tonkin, E, Persistent identifiers: considering the options, *Ariadne*, 2008, (56).  
<http://www.ariadne.ac.uk/issue56/tonkin/> (accessed 28 July 2009)
20. JISC Standards Catalogue:  
[http://standards-catalogue.ukoln.ac.uk/index/Standards\\_Entries](http://standards-catalogue.ukoln.ac.uk/index/Standards_Entries) (accessed 28 July 2009)
21. DOI System Overview: Introduction:  
[http://www.doi.org/overview/sys\\_overview\\_021601.html](http://www.doi.org/overview/sys_overview_021601.html) (accessed 28 July 2009)
22. CrossRef:  
<http://www.crossref.org/> (accessed 28 July 2009)
23. Research Assessment Exercise:  
<http://www.rae.ac.uk/> (accessed 28 July 2009)
24. Beit-Arie, O, Blake, M, Caplan, P, Flecker, D, Ingoldsby, T, Lannom, L W, Mischo, W H, Pentz, E, Rogers, S and Van de Sompel, H, Linking to the appropriate copy, *D-Lib Magazine*, 2001, 7(9). DOI:10.1045/september2001-caplan
25. Culling, James, *Link resolvers and the serials supply chain: final project report for UKSG*, Oxford, Scholarly Information Strategies, 2007:  
[http://www.uksg.org/sites/uksg.org/files/uksg\\_link\\_resolvers\\_final\\_report.pdf](http://www.uksg.org/sites/uksg.org/files/uksg_link_resolvers_final_report.pdf) (accessed 28 July 2009)

26. Van de Sompel, H, Hochstenbach, P and Beit-Arie, O, OpenURL Syntax Description, Draft Version 0.1, 2000.  
<http://www.openurl.info/registry/docs/pdf/openurl-01.pdf> (accessed 28 July 2009)
27. ANSI/NISO Z39.88-2004 - The OpenURL Framework for Context-Sensitive Services, National Information Standards Organization (NISO), 2004.  
[http://www.techstreet.com/cgi-bin/detail?product\\_id=1262106](http://www.techstreet.com/cgi-bin/detail?product_id=1262106) (accessed 28 July 2009)
28. UKSG:  
<http://www.uksg.org/> (accessed 28 July 2009)
29. Apps, A and MacIntyre, R, Why OpenURL? *D-Lib Magazine*, 2006, 12(5). DOI:10.1045/may2006-apps
30. FRBRization in the Open Library:  
<http://openlibrary.org/about/frbrization> (accessed 28 July 2009)
31. Berners-Lee, T, *Cool URIs don't change*, 1998  
<http://www.w3.org/Provider/Style/URI> (accessed 28 July 2009)
32. Meyer, C A, Reference accuracy: best practices for making the links, *Journal of electronic publishing*, 2008, 11(2). DOI: 10.3998/3336451.0011.206
33. KBART: Knowledge Bases And Related Tools working group:  
<http://www.uksg.org/kbart> (accessed 28 July 2009)
34. KBART Monitoring Group Report February 2009:  
[http://www.uksg.org/sites/uksg.org/files/kbart\\_monitors\\_report\\_feb09.pdf](http://www.uksg.org/sites/uksg.org/files/kbart_monitors_report_feb09.pdf) (accessed 28 July 2009)
35. WebFeat: a brief history:  
<http://www.webfeat.org/about.htm> (accessed 28 July 2009)
36. Web services librarian:  
<http://webserviceslibrarian.blogspot.com/> (accessed 28 July 2009)
37. Tenopir, C, Online databases: Can Johnny search? *Library Journal*, 2007, 132(2), 30.  
<http://www.libraryjournal.com/article/CA6407784.html?industryid=47130> (accessed 28 July 2009)
38. King, D, Many libraries have gone to federated searching to win users back from Google. Is it working? *Journal of Electronic Resources Librarianship*, 2008, 20(4), 213-227. DOI: 10.1080/19411260802554520

39. Walker, D, Building custom Metasearch interfaces and services using the MetaLib X-server, *Internet Reference Services Quarterly*, 2007, 12(3/4), 325-339. DOI: 10.1300/J136v12n03\_06
40. Lampert, L D and Dabbour, K S, Librarian perspectives on teaching Metasearch and federated search technologies. *Internet Reference Services Quarterly*, 2007, 12(3/4), 253-278. DOI: 10.1300/J136v12n03\_02
41. Helfer, D and Wakimoto, J, *Federated search: the good, the bad and the ugly*, Special Libraries Association Annual Conference, 4-6 June 2007, Denver, Colorado  
<http://www.slideshare.net/dorishelfer/federated-search-the-good-the-bad-and-the-ugly>  
(accessed 28 July 2009)
42. Tenopir, C. ref. 37
43. Gibson, I, Goddard, L and Gordon, S, One box to search them all: implementing federated search at an academic library, *Library Hi Tech*, 2009, 27(1), 118-133. DOI 10.1108/07378830910942973
44. Summon:  
<http://www.serialssolutions.com/summon/> (accessed 28 July 2009)
45. Primo:  
<http://www.exlibrisgroup.com/category/PrimoOverview> (accessed 28 July 2009)
46. Worldcat Local:  
<http://www.oclc.org/worldcatlocal/> (accessed 28 July 2009)
47. Lederman, S, Beyond federated search?  
<http://federatedsearchblog.com/2009/03/19/beyond-federated-search/#more-241> (accessed 28 July 2009)
48. Rochkind, J, Beyond federated search?: comments  
<http://federatedsearchblog.com/2009/03/19/beyond-federated-search/#comments>  
(accessed 28 July 2009)
49. LexisNexis Signs on to the Summon Service:  
[http://www.serialssolutions.com/assets/attach\\_news/PR-Summon-LexisNexis.pdf](http://www.serialssolutions.com/assets/attach_news/PR-Summon-LexisNexis.pdf)  
(accessed 28 July 2009)
50. Nygren, E, Haya, G and Widmark, W, *Students experience of Metalib and Google Scholar*, 2006, Stockholm, universitetsbiblioteket.  
<http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-1264> (accessed 28 July 2009)
51. Lederman, S, Federated search roadmap: Part II - Identifying benefits.  
<http://federatedsearchblog.com/2008/06/30/federated-search-roadmap-part-ii-identifying-benefits/> (accessed 28 July 2009)

52. Lauridsen, H and Stone, G, The 21st-century library: a whole new ball game? *Serials*, 2009, 22(2), 141-145. DOI: 10.1629/22141
53. Nicholas, D, If we do not understand our users, we will certainly fail. *E-Resources Management Handbook*, Eds. Stone, G, Anderson, R and Feinstein, J, 2008, Oxford, UKSG. DOI: 10.1629/9552448\_0\_3.13.1
54. Walker, J, Key issue: Mining information-seeking behaviour data to enhance library services, *Serials*, 2009, 22(2), 182-184. DOI: 10.1629/22182
55. CIBER, *Information Behaviour of the Researcher of the Future ("Google Generation" Project)*, 2007, University College London, CIBER.  
<http://www.ucl.ac.uk/infostudies/research/ciber/downloads/ggexecutive.pdf> (accessed 28 July 2009)
56. Pattern, D, *OPAC 2.0 and beyond*, 32nd UKSG Annual Conference and Exhibition, 30 March - 1 April 2009, Riviera International Conference Centre, Torquay.  
<http://eprints.hud.ac.uk/4143/> (accessed 28 July 2009)
57. xISBN:  
[www.oclc.org/research/projects/xisbn](http://www.oclc.org/research/projects/xisbn) (accessed 28 July 2009)
58. LibraryThing APIs:  
[www.librarything.com/api](http://www.librarything.com/api) (accessed 28 July 2009)
59. University of Huddersfield, Applied economics letters, table of contents & recent articles:  
<http://library.hud.ac.uk/catlink/bib/592251/cls/> (accessed 28 July 2009)
60. bX recommender service: overview:  
<http://www.exlibrisgroup.com/category/bXOverview> (accessed 28 July 2009)
61. University of Huddersfield: British journal of midwifery, people who looked at this journal, also looked at these journals:  
<http://library.hud.ac.uk/catlink/bib/241308/cls/> (accessed 28 August 2009)
62. King, D. ref. 38
63. Howell, F, What price customer loyalty? Supermarket shoppers will get discounts under glitzy new schemes. But Big Brother may come to rule the till, *Independent*, Saturday, 21 October 1995.  
<http://www.independent.co.uk/money/what-price-customer-loyalty-supermarket-shoppers-will-get-discounts-under-glitzy-new-schemes-but-big-brother-may-come-to-rule-the-till-1578705.html> (accessed 28 July 2009)
64. Hirst, T, Why Private Browsing Isn't...  
<http://ouseful.wordpress.com/2009/07/15/why-private-browsing-isnt/> (accessed 28 July 2009)

65. Walker, J. ref. 54
66. JISC MOSAIC:  
<http://www.sero.co.uk/jisc-mosaic.html> (accessed 28 July 2009)
67. Towards Implementation of Library 2.0 and the e-Framework (TILE):  
<http://www.jisc.ac.uk/whatwedo/programmes/resourcediscovery/tile.aspx> (accessed 28 July 2009)
68. Pattern, D, Free book usage data from the University of Huddersfield:  
<http://www.daveyp.com/blog/archives/528> (accessed 28 July 2009)
69. Open Data Commons:  
<http://www.opendatacommons.org/> (accessed 28 July 2009)
70. JISC, *Higher Education in a Web 2.0 World*, Report of an independent Committee of Inquiry into the impact on higher education of students' widespread use of Web 2.0 technologies, 2009, London, JISC.  
<http://www.jisc.ac.uk/publications/documents/heweb2.aspx> (accessed 28 July 2009)
71. Evergreen:  
<http://www.open-ils.org/> (accessed 28 July 2009)
72. KOHA:  
<http://koha.org/> (accessed 28 July 2009)
73. VuFind:  
<http://www.vufind.org/> (accessed 28 July 2009)
74. PTFS Europe:  
<http://www.ptfs-europe.com/> (accessed 28 July 2009)
75. Wallis, R, Juice up your OPAC, *Panlibus*, 2009 (12), 12-13.
76. Wallis, R, *Juice Project*, Mash Oop North, University of Huddersfield, 7 July 2009.  
<http://mashlib09.wordpress.com/> (accessed 28 July 2009)
77. The Library 2.0 gang:  
<http://librarygang.talis.com/> (accessed 28 July 2009)
78. Nicholas, D. ref. 1
79. King, D. ref. 38
80. Mashlib Pipes Tutorial: 2D Journal Search:  
<http://ouseful.wordpress.com/2009/07/09/mashlib-pipes-tutorial-2d-journal-search/>  
(accessed 28 July 2009)
81. Scott Wilson, jOPML:

- <http://jopml.org/> (accessed 28 July 2009)
82. TicTocs: journal tables of contents service:  
<http://www.tictocs.ac.uk/> (accessed 28 July 2009)
83. Smashing editorial, Data visualization and infographics, *Smashing magazine*, 14<sup>th</sup> January 2008.  
<http://www.smashingmagazine.com/2008/01/14/monday-inspiration-data-visualization-and-infographics/> (accessed 28 July 2009)
84. Tenopir, C. ref. 37
85. Human-Computer Interaction Lab, University of Maryland Institute:  
<http://www.cs.umd.edu/hcil/> (accessed 28 July 2009)
86. The Work of Edward Tufte and Graphics Press:  
<http://www.edwardtufte.com/tufte/> (accessed 28 July 2009)
87. brendandawes.com:  
<http://www.brendandawes.com/> (accessed 28 July 2009)
88. DoodleBuzz: Typographic News Explorer:  
<http://www.doodlebuzz.com/> (accessed 28 July 2009)
89. BBC Memoryshare:  
<http://www.bbc.co.uk/dna/memoryshare/> (accessed 28 July 2009)
90. Dawes, B, *Somewhere I have never travelled*, Mash Oop North, University of Huddersfield, 7 July 2009. [http://www.archive.org/details/MashOopNorth\\_BrendanDawes/](http://www.archive.org/details/MashOopNorth_BrendanDawes/) (accessed 28 July 2009)
91. UAL Research Online:  
<http://kultur.arts.ac.uk> (accessed 28 July 2009)
92. University of Huddersfield: 1960s fashion print : a sourcebook:  
<http://library.hud.ac.uk/catlink/bib/614544/cls/> (accessed 28 July 2009)
93. Pattern, D, *Did someone say "Free Beer"?! Open Source: "Free Speech, Free Beer and Free Kittens!"*, University of Huddersfield 26 June 2009.  
<http://eprints.hud.ac.uk/4850/> (accessed 28 July 2009)
94. Duddy, C, *A student perspective on accessing academic information in the Google era*, 32nd UKSG Annual Conference and Exhibition, 30 March - 1 April 2009, Riviera International Conference Centre, Torquay.
95. Pattern, D, *OPAC 2.0 and beyond*. 32nd UKSG Annual Conference and Exhibition, 30 March - 1 April 2009, Riviera International Conference Centre, Torquay.  
<http://eprints.hud.ac.uk/4143/> (accessed 28 July 2009)

96. Shipp, J, University Librarian at the University of Sydney, comment at the Sydney Online conference January 2009
97. Lauridsen, H. ref. 52
98. King, D. ref. 38
99. Duddy, C, A personal perspective on accessing academic information in the Google era, or 'How I learned to stop worrying and love Google', *Serials*, 2009, 22(2), 132-135. DOI: 10.1629/22131