Customising Location of Knowledge

Ann Apps MIMAS, The University of Manchester, UK, Tel: +44 161 2756039, ann.apps@manchester.ac.uk

Ross MacIntyre MIMAS, The University of Manchester, UK. ross.macintyre@manchester.ac.uk

The process of scholarly research involves the discovery of material of interest, followed by its location and subsequent perusal. Zetoc is a current awareness citation service available to UK researchers for the initial discovery step. This paper highlights some recent discovery enhancements to Zetoc, and describes the subsequent location via OpenURL technology of articles discovered in Zetoc. The later steps within the scholarly research workflow, of location, request and delivery of research literature, are provided by further services, which may be chosen by either institution librarians or individuals. The dissemination of bibliographic citation details as interoperable components enables their use by downstream applications unknown to the source service, thus providing further opportunities for customisation of the location of knowledge.

Keywords:

OpenURL, ContextObject, COinS, bibliographic citation, RSS.

1. Introduction

With the advent of electronic publishing of research literature, and the subsequent introduction of linking from bibliographic citation discovery services to the full text of described articles, the process of scholarly research is changing. The underlying scholarly research paradigm of accessing literature is: 'discover' apparently relevant knowledge; 'locate' where the information is available; 'request' a copy; and take 'delivery' of it for reading [1]. The previous model of discovering relevant literature within one's institution's library is moving to a more personalised approach. Some researchers may have their own way of working, using various applications which they prefer, rather than using a route to appropriate literature designated by their library. For the many researchers who do still use resources set up by their institution, the library will wish to customise their routes through various services to provide a consistent environment to their members. Whether the process is chosen by an individual or by a library it will involve a selection of various services available, rather than a single monolithic literature service.

A previous paper presented at a Dublin Core conference, in 2003, described the location of bibliographic resources discovered via a citation service [2]. This paper reports more recent developments, but for completeness sets the scene by briefly recapping the service infrastructure in sections 2 and 3.

2. Discovering Articles Using Zetoc

Zetoc [3] is a current awareness service available to researchers in the UK to 'discover' articles of interest. It provides access to the British Library's Electronic Table of Contents of journal articles and conference papers, holding details of around 20,000 current journals and approximately 16,000 conference proceedings per year, covering all

subject domains, data being available from 1993 and updated daily. The service is available to researchers, learners and teachers in UK higher and further education institutions. Searches using the Zetoc Web interface yield bibliographic citation details of the discovered articles. Zetoc includes an email alerting service that provides tables of contents of new issues of journals. Additionally alerts can be requested that supply details of articles, including conference papers, which match certain criteria, either keywords in their titles or by particular authors. This popular alerting service enhances the discovery of relevant knowledge by providing timely information directly to a researcher. A recent addition to the Zetoc portfolio is RSS feeds for journals [4], which syndicate the arrival of recent articles into Zetoc. Both email and RSS alerts contain a URL for each article enabling direct access to its full record within the Zetoc Web interface, and hence location of the article.

3. Locating Articles Via Zetoc

Once a researcher has discovered an article of interest, they will wish to read it, and therefore to 'locate', and request delivery of, its full text. Zetoc provides a document delivery service from the British Library, and also assistance with requesting articles from an institution's library via traditional Inter-Library routes. But researchers prefer to locate an electronic copy of the article to be delivered directly to their desktop. Rather than Zetoc itself providing location of electronic articles, it utilises a service appropriate to the user where possible, a facility driven by OpenURL technology [5].

Within its traditional domain of the scholarly information community, the OpenURL Framework, ANSI/NISO standard Z39.88-2004 [6], provides a means of transporting standard details of referenced information objects along with, optionally, details about where the reference was made and by whom. This composite package describing the reference and its context is called a ContextObject. The OpenURL Framework, its development, and justification are explained more fully in [7]. Implementation Guidelines [8] are also available. OpenURL provision by Zetoc is described in more detail in [9].

For users whose institution has an OpenURL resolver, Zetoc creates an OpenURL source link containing bibliographic details of the particular article. This is the first link below 'Further Options' in the Zetoc full record pages shown in Figures 1 and 4. When a user accesses this link they are taken to their local OpenURL resolution service, which will provide them with a menu of options for locating the full text of the article along with other helpful, related services. An example menu is shown in Figure 2. Thus 'location' of the article is provided by a researcher's institution's OpenURL resolver service. Typically the library will have populated the 'knowledge base' of the OpenURL resolver with information about the institution's subscriptions to various electronic journals and other resources, ensuring that its members locate an article only where they have appropriate access. The resolver's knowledge base also contains details of how to link to many publishers' electronic journal sites, typically populated by the resolver vendor. Information may also be included in the knowledge base about linking to open access journals or the institution's repository. When a user selects the link from their institution's OpenURL resolver to the full text of the article, 'request' and 'delivery' of the article is provided by the indicated publisher's or open access service.

For users who do not have access to an OpenURL resolver, Zetoc provides access to other services that may enable a user to locate an article such as the publicly available search engine Scirus [10].



Figure 1. A Zetoc Full Record page for a Discovered Article



Figure 2. The University of Manchester's SFX 'FindIt' Menu for Article Location

4. Institution Customisation of Zetoc Article Location

Zetoc maintains an internal registry of OpenURL resolvers, populated with details supplied by its constituent libraries. The presumed affiliation of a user is deduced during authentication, either from their machine address if they are on campus, or from their login. The initial implementation of Zetoc OpenURL article location showed a common, neutrally worded and unbranded link, which said "More information about this article". However, many librarians wished to customise this link, generally with an image 'button' and their own choice of wording. They prefer to have consistent links from all the resources their members use, hopefully simplifying the users' discovery experience so that they make best use of the library resources. There is also a wish to show their institution's branding to reinforce to their readers that the seemingly free location and delivery of articles is in fact available because the library has purchased subscriptions to the resources. Thus Zetoc has provided a facility to register an image and customised text for the OpenURL link on a full record page, extending the schema of its internal resolver registry. This option has proved popular, most institutions now having customised the OpenURL source link for their users. Figure 1 shows the customised link seen at The University of Manchester.

5. Personal Customisation of Zetoc Article Location

For the majority of researchers the route to article location from Zetoc via their institution's OpenURL resolver is satisfactory. But for some users this prescription may be unacceptable or inappropriate. A user, possibly at an institution that does not provide an OpenURL resolution service, may have access to an OpenURL resolver of which Zetoc is unaware. But they will be shown the Zetoc 'default' options for those with no OpenURL resolver. A researcher who has more than one OpenURL resolver available to them, maybe because they are affiliated to more than one institution, or because they have one available through a professional subscription, may wish to choose which to use. But they will be linked to the resolver indicated by their login. Or an academic may

be visiting and using a machine at another organisation, but prefer to use their home institution's resolver. However they will be shown the resolver of the visited institution because of the automatic Zetoc login from an academic machine address. Or someone may simply prefer to make their own choice about which service to use for article location.

5.1 Personalisation Support with COinS

To help address these issues, and to provide opportunities for future additional services, Zetoc embeds an article's details 'hidden' as a 'latent' OpenURL [11] in the full record screen of its Web interface. The bibliographic citation details are included within a ContextObject within an HTML 'span' element, known as a COinS [12]. The COinS is invisible to human readers of the Web page, most of whom would regard it as too cryptic. But any COinS-aware application is able to generate an active, actionable OpenURL simply by prefixing a resolver address.

An extension to the Firefox Web browser, such as Openly's OpenURL Referrer [13], allows a user to personalise their OpenURL linking via their chosen OpenURL resolver. A list of resolver profiles, each including either a button image or link text, can be registered in the OpenURL Referrer, then a particular one selected for current use, a choice that can be changed easily to suit the occasion and environment. After this selection, a COinS-aware application will show a link to the resolver of choice. Activating this link will send the ContextObject, containing the citation details, to the chosen resolution service. Figure 3 shows the process of personalisation by selection of The University of Manchester's 'FindIt@JRUL' SFX resolver.



Figure 3. Selecting a Resolver with OpenURL Referrer

Within Zetoc, if OpenURL Referrer has been profiled with a user's choice of OpenURL resolver, using Firefox they will see the relevant button or link at the side of the OpenURL link. Clicking on this button, rather than their institution's (if one is shown) will take them to their chosen resolution service. Figure 4 shows a Zetoc full record page, as would be seen at an institution with no resolver, with a 'FindIt@JRUL' button, which will send the hidden ContextObject to the University of Manchester's resolver.



Figure 4. A Clickable COinS Button and the Hidden ContextObject

At present this personalisation is applicable to only a small number of users who understand OpenURL linking and to those who have access to a personal Firefox Web browser. But one may anticipate future applications and installations that extend this functionality to naïve users. COinS provides a means of disseminating bibliographic citation details of information resources for use by unknown applications, which possibly pass them on further. In fact it isn't possible for Zetoc to track when its COinS are used. Significantly COinS captures bibliographic citations in a standard way, ContextObject being part of Z39.88-2004 and also recommended by the Dublin Core Metadata Initiative's (DCMI) recommendation for capturing bibliographic citation information within a Dublin Core metadata description [14]. Thus COinS is providing interoperable components, making its constituent information available for serendipitous use by any application that wishes to collect it.

Increasing numbers of applications are providing COinS. For example they are available within Copac, the UK research libraries' online union catalogue [15], Google Scholar [16], CiteULike [17], a free service that helps academics to share, store and organise the academic papers they are reading, and Citebase [18], a search and analysis tools for free, online research literature. Further developments are appearing such as 'WAG the Dog' [19], which generates COinS using screen scraping, refined by holdings or ISBN lookups, in various applications, to provide an OpenURL-enabled 'more like this' service based on journal categories.

6. Resolver Registries

Registries of OpenURL Resolvers are appearing, including the UK OpenURL Router [20], and the OCLC OpenURL Resolver Registry [21] with mainly US coverage. Institutions register their OpenURL resolver, along with relevant machine address ranges and customisation details such as an image. The router passes on OpenURLs to the resolver that is appropriate for a user. Thus a source service can send all OpenURLs to a single router instead of determining a user's affiliation, having looked up the appropriate image for the source link. An institution can register their resolver, and make any subsequent updates, with one service only, ensuring consistent customisation of its OpenURL links within all source services. Services that still need to maintain internal registries, maybe because of a mismatch with their users' domain, a need to capture additional details, or a desire for their own default option, can part-populate them from the public registries.

At present coverage within the registries is patchy, making exclusive use of them inadvisable. However a library can provide their 'router resolver address' to source OpenURL services, ensuring a single point for later updates, as in the example shown in Figure 3. Resolver registries provide for institution rather than personal customisation, although the OCLC registry includes the option of an Openly OpenURL Referrer setting. Determining the appropriate resolver for a user is an area where there are still issues to resolve, including global identification of institutions and organisations, and of OpenURL resolvers.

7. Locating Articles via Alerts

7.1 Email Alerts

Zetoc email alerts contain bibliographic citation details for journal articles for human reading within the text of the email. Email alerts may also include details of conference papers not published in journals if requested by search criteria. The alerts contain the URL of the Zetoc full record page for the article. Generally a click on this URL will go directly to that page, depending on a user's email software making the URL actionable and on the user's authentication with Zetoc. From here it is possible to locate the article via a resolution service, or to avail of COinS functionality as described above.

7.2 RSS Alerts

Zetoc recently introduced RSS syndication of new articles within journals, available to authorised users via the list of journals that Zetoc contains. Zetoc feeds conform to RSS 1.0, RDF Site Summary [22]. This allows extensions, including elements from

additional namespaces, to include citation information in a more detailed way than is provided by the basic RSS elements (title, link, and description). However, because many RSS readers display the values of only the basic elements, significant information is repeated within the description field. Even though the more specifically tagged details are not currently utilised by many RSS readers, they are available for use by future more intelligent RSS services.

For each journal the RSS feed contains its details, including title, ISSN and publisher, as well as Zetoc terms and conditions, described using simple Dublin Core properties. The Zetoc logo is included for those RSS readers that are able to display it. The RSS channel items are details of articles in the journal, which have recently been loaded into Zetoc. Article details (title, authors and publication year) are again described using simple Dublin Core properties. For each article the link element contains the URL of the Zetoc full record page for the article, providing the possibility of locating the article via Zetoc functionality, similarly to the email alert. Additionally the basic bibliographic citation details of the article are included within a 'bibliographicCitation' element from the qualified Dublin Core ('dcterms') namespace, as a text string, consistent with the DCMI recommendation for capturing citations. These details include the journal name, volume and issue information and page numbers. A sample RSS channel item for a single article is shown in Figure 5 (the 'dc' and 'dcterms' namespaces having been defined in the leading 'rdf:RDF' element of the channel).

```
<item
rdf:about="http://zetoc.mimas.ac.uk/wzgw?terms=RN184284855&field=zid">
    <title>Glutamate receptors at atomic resolution</title>
    <description>
    Mayer, M. L.; NATURE -LONDON- VOL 440; NUMBER 7083 pp. 456-462
    </description>
    link>http://zetoc.mimas.ac.uk/wzgw?terms=RN184284855&field=zid
    dc:title>Glutamate receptors at atomic resolution
    dcterms:bibliographicCitation>
    NATURE -LONDON- VOL 440; NUMBER 7083 pp. 456-462
    </dcterms:bibliographicCitation>
    <dc:date>2006-04-01</dc;date>
    <dc:identifier>http://zetoc.mimas.ac.uk/wzgw?terms=RN184284855&field=zid</dc:identifier>
    <dc:creator>Mayer, M. L.</dc:creator>
    </item>
```

Figure 5. A Zetoc RSS Channel Item

A further enhancement has been the inclusion of a ContextObject containing the citation details of the article, a replication of that supplied as a COinS within the Zetoc Web interface. This allows consuming applications to process the citation details of an article supplied through an RSS feed. The ContextObject is supplied within a 'bibliographicCitation' element, using the Z39.88-2004 Key/Encoded-Value format ContextObject as a syntax encoding scheme. As with the COinS implementation, the introduction of bibliographic citation information into an RSS feed disseminates location details into the internet for use by downstream applications of which Zetoc is unaware. An example of a 'bibliographicCitation' property containing a ContextObject for inclusion in the RSS channel item of Figure 5 is shown in Figure 6 (the 'kevco' namespace having been defined in the leading 'rdf:RDF' element of the channel as 'info:ofi/fmt:kev:mtx:').

Figure 6. An Example ContextObject within an RSS Channel Item (Note that line breaks within the ContextObject are for readability only.)

8. Location Dissemination in Other Contexts

8.1 The Z39.88-2004 XML Format

The original OpenURL carried its metadata description of a referent as a set of 'key and value' pairs on the URL. This was taken forward as the Key/Encoded-Value (KEV) format ContextObject in the Z39.88-2004 standard. The uses of ContextObjects described above all employ the KEV format. The standard also includes an XML serialisation, opening up the possibility of using ContextObjects for 'server to server' applications.

Zetoc uses some of the elements from the XML ContextObject journal metadata format to describe journal article bibliographic fields in the response record from its Web Services SOAP interface [23].

A proposed RSS module allows the embedding of an XML ContextObject in an RSS feed [24]. This would enable the passing of contextual as well as bibliographic information to a downstream application, which could then provide appropriate services to a consumer of the feed. Because Zetoc has little contextual information about the user, the KEV ContextObject was chosen in preference over the more complex XML ContextObject for inclusion in its RSS feeds. But for applications that do have more details of the context of the reference, an XML ContextObject within an RSS feed is another example of casting interoperable components into the internet to be caught downstream by another application that can make use of them.

8.2 Locating Non-Bibliographic Resources

The original OpenURL defined a scholarly information set of metadata covering only books, journals and conference proceedings. The OpenURL Framework standard widens the scope to a potentially diverse range of domains by defining metadata formats in an extensible registry.

An alternative metadata format, currently with status 'experimental' in the OpenURL Registry, is for Simple Dublin Core [25]. This format opens up the possibility of indicating the 'subject' of a described entity, although the more specific bibliographic details cannot be included in this metadata format. The Simple Dublin Core metadata format could be used to describe non-text scholarly information items, such as images maps and sound, because it introduces the keys 'type' (or genre) and 'format'. Figure 7 shows an example of an image described using the Simple Dublin Core KEV metadata format. This would provide a more detailed description of an image than the previous 'best practice' of using the book metadata format with genre 'unknown' and no possibility of capturing format.

```
&rft_val_fmt=info:ofi/fmt:kev:mtx:dc
&rft.title=Cherry Blossom&rft.creator=Ann Apps
&rft.type=Image&rft.format=image/jpeg
```

Figure 7. The Referent Section of a ContextObject Describing a Photograph Using Simple Dublin Core (Line breaks are for readability.)

The JISC Information Environment Service Registry (IESR) [26] OpenURL Link-To Resolver yields an XML description of a collection, service or agent entity corresponding to a particular identifier. In fact this resolver implements resolution of IESR URIs (Uniform Resource Identifiers). It provides an additional option to request the entity type and result format using the Z39.88-2004 Simple Dublin Core metadata format to describe the ContextObject service requested. This OpenURL Link-To Resolver could be extended in the future to provide, amongst other options, a simple subject search of IESR restricted by entity type. For example the ContextObject shown in Figure 8 would return a set of XML Collection entities that have subject 'music'.

```
http://iesr.ac.uk/service/ourllinkto?
url_ver=Z39.88-2004&url_ctx_fmt=info:ofi/fmt:kev:mtx:ctx
&rft_val_fmt=info:ofi/fmt:kev:mtx:dc
&rft.subject=music
&svc_val_fmt=info:ofi/fmt:kev:mtx:dc&svc.format=text/xml&svc.type=Collection
```

Figure 8. A Possible ContextObject Using the Simple Dublin Core Format (Line breaks are for readability.)

9. Conclusion

The Zetoc citation discovery service has been used as an example to demonstrate the finding of bibliographic citation information and how this is subsequently used to determine the location of an article. There are many other discovery services which researchers may use as a starting point to extend their knowledge, including following references from the bibliographies at the end of electronic journal articles. The OpenURL ContextObject provides a standard interoperable means to disseminate bibliographic citation and hence location information. A ContextObject that describes an article by metadata is neutral with respect to where the appropriate copy of an article may be found by a particular researcher.

The initial discovery service passes the citation information to another service to determine the location of an appropriate copy for the reader's context. This may be the OpenURL resolver of the researcher's institution. Or it may be a service chosen by the user. Generally details are passed on to a further service, such as an electronic journals application, to provide delivery of an article to the user. Thus the introduction of the standard, interoperable description of resources provided by the Z39.88-2004 ContextObject has enabled a 'service oriented' paradigm of research, or knowledge location, rather than previous proprietary solutions. Further it allows customisation, either by an institution on behalf of their members, or by an individual, of the particular services chosen during the process of discovery, location and request of scholarly knowledge.

Such opportunities of personalisation would allow a researcher to build a literature search workflow according to their preferred research knowledge discovery process. This was recently described by Chudnov [27] as a change from resolver rules and service pre-coordination by libraries, to enabling the dynamic post-coordination of services by a user. In 'Principles of Web 2.0' Miller [28] suggests that applications should work for the user "rather than forcing us to conform to the paths laid out for us". Further exciting, potential developments can be seen with the dissemination of bibliographic citation details about scholarly research literature, encoded interoperably within ContetxtObjects, via COinS and within RSS feeds. These are available to downstream applications within a service workflow choreographed by a researcher. Effectively discovery services are throwing bibliographic citation details as interoperable components to be caught and used by serendipitous services unknown to the source service.

These recent developments seem to indicate an evolution taking place in the process of scholarly research. Further indications may be seen in the way researchers are sharing knowledge through social bookmarking tools such as Connotea [29]. For dissemination and sharing of bibliographic citation information to be most effective, it has to be via standard, interoperable data description, which the Z39.88-2004 ContextObject

provides. Hopefully the ultimate benefit will be improvement of the means by which researchers find and access material in the furtherance of their work, thus assisting in the advancement of knowledge.

10. Acknowledgements

The authors wish to acknowledge the contribution of colleagues in the Zetoc team at MIMAS to the developments described in this paper: Ashley Sanders, Jane Stevenson, Andrew Weeks and Suzanne Williams. The Zetoc service, including its development, is supported by the Joint Information Systems Committee (JISC) of the UK Higher and Further Education Funding Councils. The Zetoc data is provided by the British Library.

11. References:

- 1. MODELS Moving to Distributed Environments for Library Services. http://www.ukoln.ac.uk/dlis/models/
- 2. A. Apps and R. MacIntyre. Using the OpenURL Framework to Locate Bibliographic Resources. In: *Proceedings of the 2003 Dublin Core Conference (DC2003), Seattle, Washington, USAm 28 September 2 October 2003.* ISBN 0-9745303-0-1, 2003, pp. 143-152.
- 3. Zetoc Electronic Table of Contents. http://zetoc.mimas.ac.uk/
- 4. Zetoc RSS. http://zetoc.mimas.ac.uk/rss.html
- 5. H. Van de Sompel and O. Beit-Arie. Open Linking in the Scholarly Information Environment Using the OpenURL Framework. *D-Lib Magazine*, 2001, vol. 7, n. 3, march2001-vandesompel.
- 6. ANSI/NISO Z39.88-2004, The OpenURL Framework for Context-Sensitive Services. http://www.niso.org/standards/standard_detail.cfm?std_id=783
- 7. A. Apps and R. MacIntyre. Why OpenURL? *D-Lib Magazine*, 2006, vol. 12, n. 5, may2006-apps.
- 8. Z39.88-2004: The Key/Encoded-Value Format Implementation Guidelines. http://www.openurl.info/registry/docs/implementation_guidelines/
- 9. A. Apps and R. MacIntyre. Emerging Uses for the OpenURL Framework. In: *Proceedings of the Ninth ICCC International Conference on Electronic Publishing* (*ELPUB2005*), *Leuven*, *Belgium*, 8-10 June 2005. Leuven: Peeters Publishing, 2005, pp. 283-289.
- 10. Scirus. http://www.scirus.com
- 11. D. Chudnov, R. Cameron, J. Frumkin, R. Singer and R. Yee. Opening up OpenURLs with Autodiscovery. *Ariadne*, 2005, vol. 43, /issue43/chudnov/
- 12. E. Hellman. OpenURL COinS: A Convention to Embed Bibliographic Metadata in HTML. http://ocoins.info/
- 13. OpenURL Referrer from Openly Informatics. http://www.openly.com/openurlref/
- 14. A. Apps. Guidelines for Encoding Bibliographic Citation Information in Dublin Core Metadata. http://www.dublincore.org/documents/dc-citation-guidelines/
- 15. The Copac research library online union catalogue service. http://copac.ac.uk/
- 16. Google Scholar. http://scholar.google.com/
- 17. CiteULike. http://www.citeulike.org/
- 18. Citebase. http://www.citebase.org/
- 19. WAG the Dog. http://www.library.gatech.edu/research_help/WAG_faq.html
- 20. The OpenURL Router. http://openurl.ac.uk/doc/
- 21. OCLC OpenURL Resolver Registry. http://www.oclc.org/productworks/urlresolver.htm.
- 22. RDF Site Summary (RSS) 1.0. http://web.resource.org/rss/1.0/
- 23. A. Apps. zetoc SOAP: a Web Services Interface for a Digital Library Resource. *Lecture Notes in Computer Science*, 2004, vol. 3232, pp. 198-208.
- 24. T. Hammond, T. Hannay and B. Lund. The Role of RSS in Science Publishing. D-

- Lib Magazine, 2004, vol. 10, n. 12, december 2004-hammond.
- 25. OpenURL Key/Encoded-Value Metadata Format for Dublin Core. http://www.openurl.info/registry/docs/mtx/info:ofi/fmt:kev:mtx:dc
- 26. A. Apps. Disseminating Service Registry Records. In: *Proceedings of the Tenth International Conference on Electronic Publishing (ELPUB2006), Bansko, Bulgaria, 14-16 June 2006.* Sofia: FOI-COMMERCE, 2006, pp. 37-47.
- 27. D. Chudnov and J. Frumkin. Library Dialtone: Bootstrapping with Autodiscovery and Service Registries. In: *Access* 2005, *Edmonton, Canada,* 17-19 October 2005. 2005.
- 28. P. Miller. Web 2.0: Building the New Library. *Ariadne*, 2005, vol. 45, /issue45/miller/
- 29. B. Lund, T. Hammond, M. Flack and T. Hannay. Social Bookmarking Tools (II). *D-Lib Magazine*, 2005, vol. 11, n. 4, april2005-lund