Getting technical – an overview

Tom Bishop
Head of Library and Surgical Information Services
The Royal College of Surgeons of England

Introduction to Journals and E-Resources Today
October 9th, 2014
The technical landscape

- Industry initiatives
- Industry protocols, standards, codes of practice
- Internet protocols and standards
Linking mechanisms: the URL

http://www.rcseng.ac.uk/library/online-services/online-journals/neurosurgery

http://www.jstor.org/journals/10624783.html

- May be predictable e.g. derived - ISSN
- Static
- Subject to change
Linking mechanisms: the DOI

• Digital Object Identifier: A unique identifier assigned to a digital object

• A way of accessing an object without having to know its URL - the DOI identifies the object itself, not the place where it is stored.

• Persistent - as long as the object exists, so does the DOI.

• CrossRef – “the citation-linking backbone for online publications” - the official DOI link registration agency for scholarly and professional publications.
So - the DOI 10.1016/j.jvs.2014.04.034

Expressed as URL:
http://dx.doi.org/10.1016/j.jvs.2014.04.034
(or typed into a DOI resolver box, e.g. at
www.doi.org or www.crossref.org/guestquery)

Resolves to:
http://www.jvascsurg.org/article/S0741-5214%2814%2900835-0/abstract
Role of the link resolver

“Link resolvers are configured to receive OpenURL links from content provider sources. They extract metadata about the target article (or other object) from the OpenURL, then compare this information to the knowledge base, which contains data about all the content licensed by the link resolver’s owner. The knowledge base indicates whether the article is available to individuals associated with the license-holding institution; if so, where it is hosted; and how to connect the user to it. The link resolver then puts together a link to the target article.”

http://www.uksg.org/kbart/s3/stakeholders
Link resolver’s knowledge base

• Provided by vendor
• Contains ‘targets’ - titles, publishers/providers, packages, full text databases, open access repositories, etc. – and linking rules for targets
• Library activates link targets and ‘localises’ to reflect subscriptions and holdings
• Library’s own targets – print holdings, document delivery, citation-downloading tools, etc.

Knowledge base knows where content is and which versions a library’s user is entitled to access.
KBART Phase I Recommended Practice

Figure 1: User Journey via OpenURL

The OpenURL uses a link resolver (L) to transport a user from A (a citation) to Z (a copy of the cited document which is licensed by the current user), by way of an OpenURL query (Q), which appends a string (S) of metadata about the cited article to the base URL (B) of the current user’s institution (I). This is a more effective alternative than hard-coded links to other resources, such as a subscription agent gateway (G), the library’s print holdings (P), aggregated databases (D), publishers’ own websites (W), or repositories (R).

D (database)
W (publisher website)
R (repository)
B (base URL of link resolver)
I (institution)
L (link resolver’s knowledge base)
Z (target (cited) article)
Linking mechanisms: Open URL

- OpenURL is a standard designed to provide seamless, context sensitive linking, matching the user to the holdings and access rights of their library.
- OpenURL format = Base URL?Query
- Base URL = web address of the link resolver, i.e. reflecting the user’s institution
- Query (or Context Object) = the metadata that the link resolver will use to identify and link to appropriate targets

(OpenURL version 0.1)
Send DOI to CrossRef to get metadata
Check knowledge base
Generate target link, e.g.
http://bmb.oxfordjournals.org/content/100/1/23
KBART: Knowledge Bases And Related Tools working group

KBART is a joint initiative with NISO that is exploring data problems within the OpenURL supply chain.

The group was publicly launched by UKSG and NISO in January 2008.

- view the Phase I report and recommendations
- learn about KBART Phase I endorsement for content providers
- explore the information hub
- view a list of members
- sign up for updates on the group's progress
- find out about the project's background and early progress
IOTA Standing Committee
Adam Chandler (Chair)
Electronic Resources User Experience Librarian, Cornell University Library

Rafal Kasprzoki
Electronic Resources Librarian, Rice University

Susan Marcin
Head of Electronic Resources Management: Technologies & User Experience Columbia University

Oliver Pesch
Chief Product Strategist, EBSCO Information Services

Clara Rutenberg
Electronic Resources Librarian, University of Maryland

Elizabeth Winter
Electronic Resources Coordinator, Georgia Tech Library; Collection Acquisitions & Management Department

Jim Wismer
Manager, Software Engineering, Thomson Reuters

What is IOTA?

IOTA is an initiative that makes use of log files from various institutions and vendors to analyze element frequency and patterns contained within OpenURL strings.

The reports created from this analysis inform vendors about where to make improvements to their OpenURL strings so that the maximum number of OpenURL requests resolve to a correct record.
Transfer

The Transfer Code of Practice responds to the expressed needs of the scholarly journal community for consistent guidelines to help publishers ensure that journal content remains easily accessible by librarians and readers when there is a transfer between parties, and to ensure that the transfer process occurs with minimum disruption. The Code contains best practice guidelines for both the Transferring Publisher and the Receiving Publisher. Publishers are asked to endorse the Code, and to abide by its principles wherever it is commercially reasonable to do so.

The Code is governed by a group of librarians, publishers and other experts (the Transfer Working Group), and their main activities fall into three areas:

1. The Transfer Code of Practice. The current version is the UKSG Transfer Code of Practice, Version 3.0. A brief summary of the main changes between Version 2 and Version 3 can be found here. A glossary of the terms used in Version 3.0 can be found here.

A list of the publishers that currently endorse the Code can be found on the Publisher Endorsement page.
Welcome to Knowledge Base+

Knowledge Base+ (KB+) is a new service from Jisc Collections that aims to help UK libraries manage their e-resources more efficiently by providing accurate publication, subscription, licence and management information.

What can KB+ do?

KB+ aims to ensure that all parts of a library’s supply chain have access to all the information they need whenever they need it. It is remarkably easy to use – very little training is required for new users to ‘get up to speed’.

KB+ also aims to avoid duplication of effort. It is a one-stop shop for the management of e-resource information.

Specifically, KB+ provides:

- A centrally maintained and managed knowledge base in which Jisc Collections collates, verifies and updates knowledge base data, to avoid costly and wasteful duplication of effort by libraries all trying to do the same thing by themselves. More and more institutions are taking advantage of the fact that Jisc Collections will do a lot of the work involved in getting local institutional information uploaded into KB+.
- Verified, accurate and up-to-date publication information for e-journal agreements, including national and regional consortium agreements from across the UK and a growing number of non-Jisc packages. All of this is made available under an open licence and disseminated throughout the library supply chain so that the right organisations have the data they need when they need it. Currently Ex Libris, Proquest, OCLC and EBSCO all use KB+ data in their systems.

For a full list of the packages available from KB+, please visit the public exports page.
Authentication, authorisation, access

- IP address – single site, range of sites, fixed/known locations
- User name/password
  - Individual
  - Institutional
- Identity and access management
  - Secure, single sign-on to multiple resources
  - One user name/password to access resources
  - Remote access for users
  - Centralised, managed – OpenAthens MD
  - Federated, local – OpenAthens LA, Athens DA, UK Access Management Federation (Shibboleth - SAML)
OpenAthens (Eduserv)

MD: ‘Athens login’ to access. User accounts created within/uploaded to Athens administration area – user groups and permission sets used to allocate access rights.

LA (and DA): Institutional login to access. User accounts are those held in institutional user directory; communication with Athens (and other service providers in LA) matches users and permission sets.
User journey: OpenAthens MD

• User goes to resource site (service provider)
• User is prompted to login and selects Athens option
• User enters Athens user name and password at Athens Authentication point
• Credentials are sent to Athens to establish user’s rights
• Decision to allow/deny access sent from Athens to service provider
• User is allowed/denied access to resource site
• A session cookie is set to allow single-sign on access to further Athens-authenticated resources
‘Core’ user journey elements: Shibboleth

- User goes to resource site (service provider)
- User is prompted to login and selects institutional option
- WAYF (Where Are You From) prompts user to select ‘home’ institution from list/s
- ‘Home’ institution (identity provider) is contacted
- User is prompted for ‘home’ credentials at ‘home’ institution
- Credentials are verified by ‘home’ institution
- **Handle** (session identifier) is generated for the session and sent to service provider
- Attributes requested from ‘home’ institution and returned to service provider
- User is allowed/denied access to resource site
Login via your institution

You may be able to login to ScienceDirect using your institution's login credentials. Learn more
We will remember your login preference the next time you access ScienceDirect from this computer.

OpenAthens login
Royal College of Surgeons of England

Search for your institution and click the name to login.

- university of ports
- University of Portsmouth

Or choose your institution's region or group and click the name from the results below to login.

- UK Access Management Federation

View All Institutions

Please choose one of the institutions listed below. If your institution is not listed, try using the search feature above.

UK Access Management Federation
- University of Durham
- University of the Highlands and Islands
- Cardiff University
- LSE LoginPlus (FLAME Project)
- St George's, University of London
- UCL (University College London)
“Are we nearly there yet?”

• Login terminology itself can be confusing – ‘Athens’, ‘institutional login’

• WAYF terminology and process can be confusing and frustrating, particularly when used for the first time

• Different approaches on different platforms, sometimes with differing terminology, often highlight their own proprietary user name/login system

• Multiplicity of resources with different access models

• Multiple user affiliations e.g. NHS Athens ≠ university eligibilities – no overarching, cross-affiliation identity management system
Easing the journey (by varying amounts)

- Initial ‘home institution’ login (Athens DA)
- Search box on WAYF, cookie on PC to remember institution
- WAYFless URLs
- Proxy servers – e.g. EZProxy (URL re-writer), Squid (web caching)
- Virtual Private Networks (VPNs)
Discovery services

- Shift from a user searching unconnected resources and collections of resources individually – e.g. library OPACs, subject databases, e-book/e-journal collections, institutional repositories, open access archives, etc.

  TO

- A single point of access to a range of resources that allows multiple resources to be searched with a single search. ‘One stop shop’, ‘Google-like’, ‘unified search solution’ – characterised by single search box with subsequent options for search refinement.
Proprietary, e.g.

- EBSCO Discovery Service (EBSCO)
- Summon (Serials Solutions)
- Primo (Ex Libris)
- WorldCat Local (OCLC)

(and others)

Open Source, e.g. VuFinder, Blacklight
Discovery ≠ Federated search

- Federated search tools (e.g. WebFeat, MetaLib) use/d single search box to interrogate multiple databases in parallel
- Key difference is that federated searching is carried out ‘in real time’ using ‘connectors’ to each individual source, and provide source-by-source results; discovery searches utilise ‘pre-built’ harvesting and indexing, search a single unified index and deliver a single list of results
- Federated search may be seen as slower and less sophisticated… but may still have a place?
Discovery: Key elements

• ‘The discovery layer’ – search and result delivery interface is ‘unlatched’ from host systems. Faceted navigation and refined searching beyond single search box, recommendations, ‘do you mean…’, web 2.0

• Services harvest content – local library/institution resources, publisher and aggregator resources (metadata and/or full text) via agreements - and create centralised index, updated regularly. Content available ‘global’ – will go beyond institutional subscriptions if you want.
Discovery: Key elements

- Pre-indexing allows speed, de-duplication, application of a single controlled vocabulary, delivery of results ranked by relevancy based on available metadata

- Customisable – ‘look and feel’, content and focus (e.g. ‘Discovery Service for the Business School’, ‘Discovery Service for the Creative Arts’)
Web-scale Index-based Discovery

Search:

Consolidated Index

- ILS Data
- Digital Collections
- ProQuest
- EBSCOhost
- MLA Bibliography
- ABC-CLIO

Pre-built harvesting and indexing

http://www.slideshare.net/UKSG/1530-mon-lomond-breeding?next_slideshow=7
Discovery: Things to consider, debate, research or just plain argue about…

- ‘Too much information’? “577,957 results sorted by relevance”…
- Information literacy – ‘dumbing down’ or great opportunity?
- ‘Does discovery still happen in the library?’
  [link](http://www.sr.ithaka.org/sites/default/files/files/SR_Briefing_Discovery_20140924_0.pdf)
- Relationships between service provider and content providers/aggregators… Who deals with who? Vendor competition? Non- or partial participators (e.g. APA and PsycINFO?)
- What’s in the index? Full text vs metadata (and level of metadata), frequency of updates
- How does implementing discovery search affect the use of a library’s subscribed resources?
Open Discovery Initiative


See also the ODI Survey report issued in January 2013, which is referenced in Appendix A of the Recommended Practice.

BACKGROUND

The Open Discovery Initiative (ODI) aims at defining standards and/or best practices for the new generation of library discovery services that are based on indexed search. These discovery services are primarily based upon indexes derived from journals, ebooks and other electronic information of a scholarly nature. The content comes from a range of information providers and products--commercial, open access, institutional, etc. Given the growing interest and activity in the interactions between information providers and discovery services, this group is interested in establishing a more standard set of practices for the ways that content is represented in discovery services and for the interactions between the creators of these services and the information providers whose resources they represent.

The ODI working group was formed in late 2011 and held its first meeting in January 2012. As work progresses, these pages will be updated with more information about the specific plans and activities of the group.
Assessing the Impact of Library Discovery Technology on Content Usage

UKSG, with the support of Jisc, partnered with LISU to carry out the research project: Assessing the Impact of Library Discovery Technology on Content Usage.

Impact of library discovery technologies
A report for UKSG

Vladislav Speaji, Claire Chesser
University of Loughborough
Ann O'Brien
Centre for Information Management, Loughborough University
Angela Cooper
Evidence Base, Edinburgh Napier University

November 2013
About COUNTER

Last updated: September 2014

Launched in March 2002, COUNTER (Counting Online Usage of Networked Electronic Resources) is an international initiative serving librarians, publishers and intermediaries by setting standards that facilitate the recording and reporting of online usage statistics in a consistent, credible and compatible way. The first COUNTER Code of Practice, covering online journals and databases, was published in 2003. COUNTER’s coverage was extended further with the launch of the Code of Practice for online books and reference works in 2005. The body of COUNTER compliant usage statistics has steadily grown as more and more vendors have adopted the COUNTER Codes of Practice. This has contributed to the new discipline of usage bibliometrics and a great deal of work is underway to try to establish value metrics associated with usage, in which the COUNTER compliant statistics play an increasingly important role.

COUNTER does more than just set the standards for usage reports; we are co-operating with a number of organizations to develop a range of usage-related research and services. In 2006 COUNTER carried out research, sponsored by JISC (the UK Joint Information Systems Committee) on the effects of publisher platforms on usage and we are currently collaborating with the UK Serials Group on the possible development of a new Journal Usage Factor metric. Summary reports on both these projects can be found on the COUNTER website at http://www.projectcounter.org/news.html. COUNTER has also worked with NISO on SUSHI (Standardised Usage Harvesting Initiative) to develop a protocol to facilitate the automated harvesting and consolidation of usage statistics from different vendors. This protocol may be found on the NISO website at http://www.niso.org/workrooms/sushi/

COUNTER brings the following benefits to librarians, publishers and intermediaries:

Librarians are able to compare usage statistics from different vendors; derive useful metrics such as cost-per-use; make better-informed purchasing decisions; plan infrastructure more effectively.

Publishers and intermediaries are able to: provide data to customers in a format they want; compare the relative usage of different delivery channels; aggregate data for customers using multiple delivery channels; learn more about genuine usage patterns.

Future objectives
COUNTER and usage statistics

• Objective is to ensure that vendor online usage reports are credible, compatible and consistent.

• Librarians are able to compare usage statistics from different vendors; derive useful metrics such as cost-per-use; make better-informed purchasing decisions; plan infrastructure more effectively.

• Publishers and intermediaries are able to: provide data to customers in a format they want; compare the relative usage of different delivery channels; aggregate data for customers using multiple delivery channels; learn more about genuine usage patterns.
COUNTER Codes of Practice

Codes of Practice for how vendors report usage statistics to libraries

- COUNTER Code of Practice for E-Resources (release 4, published April 2012)
- COUNTER Code of Practice for Usage Factors (published April 2014)
- COUNTER Code of Practice for Articles (published March 2014)
• Standardized Usage Statistics Harvesting Initiative
• An ANSI/NISO Standard – a protocol that provides instructions to automate the collection of usage statistics reports from compliant vendors, which might otherwise be manually downloaded from a vendor website or received via email.
• Publishers put usage data into a standard format (COUNTER XML).
• Implementation is a requirement for compliance with the COUNTER Code of Practice.
JUSP

- Single gateway for libraries to access statistics from participating publishers, gateways and host intermediaries
- Contains JR1, JR1a and JR1 GOA COUNTER-compliant usage statistics
- Data from January 2009 (where available) for all participating institutions
- Free to join for all UK higher and further education institutions and research councils; charged version available for non-JISC funded bodies
- Open to all journal publishers and intermediaries able to supply COUNTER-compliant data using the SUSHI protocol
JUSP reports

• Journal-level reports
• Summary reports – e.g. SCONUL returns, summaries by publisher by year or selected date range, trends over time, highest usage
• Titles and deals – annual, across years, view and compare deals
• Profiling - compare own usage from a particular publisher with an average for all libraries in the same Jisc band region, group etc.
University of Bath OpenURL example and screenshots courtesy of and with thanks to Laurence Lockton, Systems Librarian, University of Bath.