

Exposing your catalogue content to a discovery layer

GARNER, HJ

Available from Sheffield Hallam University Research Archive (SHURA) at:

<http://shura.shu.ac.uk/22582/>

This document is the author deposited version. You are advised to consult the publisher's version if you wish to cite from it.

Published version

GARNER, HJ (2013). Exposing your catalogue content to a discovery layer. Catalogue and Index, 170, 22-28.

Copyright and re-use policy

See <http://shura.shu.ac.uk/information.html>

In this article, I will describe how we integrated our catalogue data within the Summon discovery layer. I will look at how it is indexed, search and displayed. I will give an overview of the searching and modifying options available and examine how well these features work with library catalogue data. I will also look at usage analytics and usability testing as an approach to evaluating a discovery layer as a catalogue search tool.

At Hallam, we are keen to ensure that our students can access our information resources using the latest discovery options available. In 2010 we evaluated the discovery system market and after careful consideration decided to choose the Serials Solution Summon discovery platform. We specifically liked the idea of the "Single search box" which had been identified in student surveys as something students wanted. It would be able to offer a unified search for journal articles, library catalogue data and other resources. Our library catalogue data has approximately 420,000 bibliographic titles.

Marc mapping

Like most Summon installations, we decided to include Library Catalogue data in the discovery layer. In order to do this, we first needed to complete a MARC21 mapping exercise to map the MARC21 fields we used with those available in Summon. Summon does not support the full range of MARC21 tags, so we therefore needed to find suitable alternative fields to map some data to. For example, in Summon, there was no equivalent 245|h [gmd]. We have used the gmd option extensively for format information and so regarded it as important enough to be mapped to a 500 note field. This meant that it would be keyword searchable and also display.

Content mapping

We used the fixed field material types in Millennium to map to the 41 content types which are available to Summon customers. You cannot set your own content facet values, but it is now possible to rename content facets locally. Mapping was quite straightforward: an exception was the need to have a streaming video facet. Our collection includes a significant amount of online video content and so needed to push the user group to support the introduction of a "streaming video" content type. We also requested that our full-text availability facet be expanded to include all online content such as online video. This also ensured that the resource icon would have the online "sunburst" symbol and made online video easy to identify.

The content type facet you map to will determine the icon displayed with the resource. Visually, icons are not always that easy to identify but results do also have a content label under the resource description. All customers use the default content icons.

Panorama: The secret drone war



12/2012, 1 online resource (30 min)
[VOD] This Video on Demand resource
are available. This recording is to
Ethics, Unmanned aerial vehicles

Video / VoD: Available Online

Daily updates

We needed to find a mechanism for sending daily catalogue updates to Summon for processing. We wanted to send a MARC21 update file to Summon 15 minutes after staff services closed each day. We purchased the Millennium Scheduler software to facilitate this process. This allowed us to set up a daily task which would search for new records, updated records and records with a delete code and then ftp the data to the Summon site for processing. We hoped that this would be a fully automated process. However, due to not being able to express a relative date in the Saved search in Millennium and the need to supply a unique file name to Summon, this operation requires routine human intervention. We discovered that it can take about 72 hours for catalogue data to be updated in our Summon instance. It is quite frustrating not to have real time updates. There can also be specific problems with this. For example, the publisher withdrew the digital rights to a popular eBook. It was flagged for deletion in the ILS, but showed in Summon for a number of days. There is also no notification service from Summon to inform libraries that their data has been received and processed. In order to minimise any data slippage, customers are recommended to send a fresh load of catalogue data every quarter.

Matching and merging of records in Summon

The Summon service contains content from a variety of providers, including over 7000 publishers and over 100,000 journal and periodical titles. It uses a match and merge system to deliver a single Summon record for articles. The Summon service will also apply a match and merge process to print and electronic book content and print and electronic journal titles. In order to merge two records the following criteria must be met:

1. A unique identification number must match, e.g. ISBN, OCLC, LCCN.
2. The publication year must match and the record must have only one publication year.
3. The title must be at least a partial match

Searching in Summon

A screenshot of a library search interface. It features the text 'Library Search' in a large, bold, black font. Below this is a search input field containing the placeholder text 'find books, articles, video and more'. To the right of the input field is a 'Search' button. Below the input field is a link for 'Advanced Search'.

There are two approaches to searching in Summon, known locally as “Library Search”. Most users start with the “Single box” from the Library gateway page, enter in a search term, view results and then opt to refine by a set of facets on the left hand side e.g. Items available online, Peer-reviewed, Content type, Publication date, Library location, Date, Author, Subject terms, Language, Genre, Region, Time period. Facets are not hyperlinks, they are boxes that require ticking to select. Most searches will return a *high number* of results, so use of the facets would seem necessary. This creates a two-step approach for the user. Each time the user performs a search or adds a facet which includes catalogue data, Summon will do a live look up of circulation data from the ILS, which can sometimes slow the search down.

Using facets to refine a results list is central to the Summon experience. Users need to do quite a lot of scrolling to view them all, e.g. our subject facets are quite low down the list. The content type facet is a key refinement. Locally, our default search will exclude both Book reviews and Newspaper articles. The initial result set will deliver the 6 highest result content types. All other content types will require the user to select the “More” button to pull up the full list for the search. Some library catalogue content such as video or audio material will often appear in the “More” section. We have set the content type list to alphabetical order rather than by highest results order, to create a consistent list order. The Usability Committee’s recommendations for Summon improvement (2011) available from the Summon Community wiki, comment that facets look “more like navigation tools than limiters” and suggests some improvement.

“Advanced search” offers more specific types of searching. You can enter search terms such as subjects, words from the title, author/creator, ISBN/ISSN, journal title or date range. You can also limit to results available online, peer reviewed or items from the catalogue. It can generate a more precise result set with less need for use of facets. Using the Advanced search for author/creators will deliver much better results. Summon supports phrase searching and the use of Boolean operators.

Summon and Library of Congress subject terms

Summon’s approach to Library of Congress Subject headings has some limitations over the standard catalogue. The terms from the Library catalogue data are keyword searchable in Summon. From an initial search it will then take all the 6xx subject strings, for example:

Architecture -- United States -- History -- 1900-2000

Then separates each element into the following facets:

Subject – Region – Subject –time period

This will create a range of facets which can be used to refine a search to produce a shorter result set. However, the relationship between the topic and its subdivisions is lost. Controlled vocabulary from other

resources such as journal articles will be blended into these facets and can create synonyms in the list.

When you look at a specific resource in the result set you will see that it offers links to Subject terms. However, it does not retain the relationship between the topic and its subdivision. The term "Tourism -- Environmental aspects" from the example below is now two headings, if I select "Environmental aspects" it will launch a new search for that term.

Sustainable tourism: a global perspective



by Harris, Rob; Griffin, Tony; Williams, Peter
2002, ISBN 9780750689465, p. 311

... This book offers an introduction to the concepts of **sustainable tourism**. Tourism, Ecotourism, Environmental aspects

- Book: AVAILABLE , 1 week, 910.133372 HA (LEVEL 6), Adsetts Main Co. ► Show 6 more...
- Book: Full Text Online
- Book: Full Text Online

Searching in Summon will not expose users to any Library of Congress Authority records or links. Sheffield Hallam's standard catalogue offers Library of Congress Authority control for names and subject terms. However, users who input a non-preferred term into Summon will not be directed to the actual Library of Congress term. This can be a problem in certain searches. At Hallam, users have been searching for "film scripts" and it would be helpful for them to be instructed to use "Motion picture plays".

Google-like searching

Some Google-like features are available in Summon. An autocomplete feature can be switched on which will make suggestions as it matches against the searching history of the entire Summon customer base. We have decided against using this as it could suggest titles that we do not have in the collection. This could cause issues with "known item" searching rather than topic searching.



A Google-like feature which we have activated is the related search suggestion list. From the user's search, Summon will offer suggestions based on real-time Summon usage data across its entire customer base and might guide users to related concepts and expanded queries. See example:

Searches related to **twitter**

twitter microphone for the masses
twitter marketing
twitter power tweets as electronic word of mouth
twitter journalism

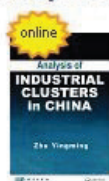
twitter power
twitter politics
twitter and journalism
twitter arab spring

It does not always deliver suggestions, but it can generate some interesting search terms.

Book, eBook and full-text indexing

We export our MARC21 eBook records from the catalogue into Summon. However, we have also switched on our holdings for major eBook collections in the Summon knowledge base. This means that the titles are then exposed to full-text or table of content indexing where possible from the Summon service. This does mean that we often have a duplicate eBook link, like this one below. The first link goes direct to the resource, the second one goes to the link resolver. This doesn't seem to be a problem for users.

Analysis of industrial clusters in China



by Yingming, Zhu
2009, ISBN 1420089196, vii, 234
[ELECTRONIC BOOK] / Available on CRCnetBASE (formerly known as ENGnetBASE) Technology:
General Issues, Industrial clusters
eBook: [Full Text Online](#)
eBook: [Full Text Online](#)

The Summon service has indexed the full-text of books from the HathiTrust Digital Library, a group of libraries who are participating in the Google Books scanning project and other mass digitisation initiatives. This means that any of our library catalogue content in Summon, which is within the HathiTrust collection, will be full-text searched. At present, we have just over 46,000 titles which are being full-text searched in this way. This is definitely a good thing for our users as it will provide a huge number of additional search terms. When results are returned specifically due to full-text searching rather than the bibliographic data, the Preview display will state "**why is your search result here?** Your query matched the full-text indexing" This can help explain any strange results from full-text searching.

Titles exposed to full-text indexing will appear as a merged record for both the eBook and Print book title.

Testing complex and embedded systems



by Pries, Kim H and Quigley, Jon M
2011, ISBN 1439821402, p. 312
"An Aurbach book" -- t.p. Using combinatorial approaches, this book aims testing organizations to perform meaningful testing. The text... Computati
computer systems, Testing
Book: AVAILABLE , 1 week, 004.21 PR (LEVEL 2), Adsetts Main Collection
eBook: [Full Text Online](#)

Discipline scoped searching

Summon contains a huge quantity of indexed content, at Hallam we have over 55 million online results. The size of the collections can create issues for searching. This could be addressed by activating the Discipline search option. There are 59 disciplines which come from a combination of discipline classification sources: Columbia University's Library of Congress Classification hierarchy, Ulrich's and the Serials Solution knowledgebase. It would enable you to create a scoped search box widget and might be used in a subject page. We have not opted for this at Hallam as we felt the disciplines do not match our subject profiles and not all content was indexed. In theory, if you applied a search for "stress fracture" using the Medicine discipline you would then get very different results to using the Engineering discipline.

Presenting catalogue records in Summon



The preview icon will allow users to see further information about the resource. It can either be set to show bibliographic data from the Library Catalogue record, or a Summon "Details page" can be activated and configured. This can include a LibraryThing rating, Table of contents Summary and/or reviews from Syndetics (if the library has a subscription). Custom messages such as lending policies can also be displayed.

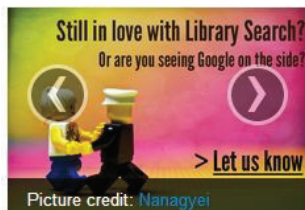
We have opted to have the preview content from the Library Catalogue as we felt that it was important for users to see bibliographic content from the local catalogue record.

Summon uses direct linking to resources when the content type is eBook or eJournal and there is just one 856 field in the record. Whilst this is very helpful to most users, it means that when there is a single 856 field it is possible to navigate to the Library Catalogue record.

Usage analytics and usability testing

It is really interesting to analyse the usage data that is available from the Summon Support Center. It will allow you to see how users are searching for catalogue content. You can search by date range and look at the top 500 searches. This will reveal whether users have used single or multiple word terms, phrase searching, wildcards, book title, author and word from title or a Dewey number search. Unsurprisingly, I noticed that the most popular search is "Research methods".

We are also using a Google analytics tool developed by Matthew Reidsrow from Grand Valley State University in order to further examine user's search behaviour. It provides information such as the type of link user's select (Document, Image, Preview, Summary), clicks based on position in the results list (84.9 % select the highest positioned result) and the content type they select (56% journals, 21% books, 13% eBooks). (Data from February 2013)



Picture credit: Nanagyei

Usability testing is also a really important way to find out the strengths and weaknesses of a Discovery layer such as Summon. Matthew Reidsrow (2011) describes how usability testing might be conducted and why it is important to help improve the search results screen for users. At Hallam, we are conducting similar usability tests on a regular basis by asking users to perform a range of searching tasks.

What next?

At Hallam, we made a strategic decision to promote Summon “Library Search” as the main catalogue interface. It is even used on the stand alone Catalogue terminals within the library. Whilst, the standard catalogue is available and is used, it is not considered helpful to promote two search tools. I feel that the integration of a library's resources into a single Discovery layer is the desired goal. It also has many benefits such as full-text searching of book content. However, I would like to see our current discovery layer evolve further to ensure that a library's catalogue data and local collections are easily discoverable within the vast result sets on offer in the discovery layer. Serials Solution announced Summon 2.0 in March 2013, it will be a major refresh to the Summon search interface and will offer significant improvements.

Reference list

Usability Committee's recommendations for Summon improvement. (2011)[online]. Last modified 25 Feb 2013. Last accessed 6 March 2013 at:
http://community.summon.serialsolutions.com/index.php/Usability_Studies_-_Reports_and_Presentations

REIDSROW, Matthew (2011). Why we do usability testing.[online].Posted 25 October . Last accessed 6 March 2013 at:
<http://matthew.reidsrow.com/articles/12>

Further information

Sheffield Hallam University library gateway
<http://library.shu.ac.uk/>

Columbia University Libraries Hierarchical Interface to Library of Congress Classification
<https://www1.columbia.edu/sec/cu/libraries/bts/hilcc/categories.html>

[Hathi Trust Digital Library](http://www.hathitrust.org/)
<http://www.hathitrust.org/>