

# Information without frontiers – barriers and solutions

*Based on summaries of three of the lightning talks from the UKSG Forum, London, November 2014*

Following the success of the first UKSG Forum, it was decided to hold a second – this time in a larger, central venue: the Grand Connaught Rooms in London. Under the umbrella title, also the title of this collection of three summaries of lightning talks from the event, were grouped a number of useful topics: Expanding Library Boundaries, The Human Touch, The Future of Library Discovery, Library Systems, Academic Content Beyond the Academy, UKSG Research and Innovation, Publishing Innovation, Open Access Monographs, KB+, Altmetrics, User Experience and Data Enrichment. In this article we bring you summaries of the talks by Anna Grigson, Catherine McManamon and Sam Herbert.

## Managing systems change at LSE Library

On 1 August 2014, LSE Library went live with ExLibris' Alma as our new library management system (LMS), replacing our previous Voyager system. Although the technical implementation of the LMS was a major project, it was only part of a wider 'systems change project', and in many ways it was the most straightforward part of the process. Far more challenging – but equally important to the ultimate success of the project – was managing the other aspects of change associated with the project, including managing business process change and managing our people through change.

### Managing business process change

The business case for the new LMS had set out the high-level benefits we were seeking from the system, one of which was to develop and improve our services by improving the efficiency and effectiveness of our business processes. In addition, the detailed requirements analysis created as part of the tender process had identified the specific areas where the functionality of our legacy system was no longer sufficient to meet our needs (such as e-resource management) and where the improved capabilities of a new system were necessary to improve our processes.

Having selected Alma as the best fit for our requirements, we were confident that its implementation would deliver some beneficial changes. But whilst the new system could give us the capacity to improve our processes, it would not be sufficient to create that change by itself. The very flexibility of Alma meant that design of many of the workflows would be down to us, which gave us the opportunity to improve our process design, but also carried the risk that we could bring inefficient processes with us from our old system.

We knew at a general level that some of our existing processes were probably not as efficient as they could be, perhaps because they had originated as complex workarounds shaped by the particular functionality of Voyager, or because they were based on assumptions that had not been re-examined for some time. But we also knew that identifying and changing these processes would be difficult – not because of a particular resistance to change from staff, but simply because our processes had become so familiar that it was difficult to spot exactly where and how they were inefficient.



ANNA GRIGSON

Head of Collection  
Services Group  
LSE Library

63 So we needed a way to step back from the day-to-day level and rediscover the purpose underlying our processes, to help us drop the unnecessary elements and redesign new processes to be more efficient. To do this, we decided to use the LEAN 'rapid improvement' methodology, which has been widely used for business process re-engineering<sup>1</sup>.

'we decided to use the LEAN "rapid improvement" methodology'

We started by reviewing the overall scope of our service, and selected a limited number of processes which we felt would benefit most from review, concentrating initially on book acquisitions (both print and electronic), circulation, inter-library loans and authentication. We then held a series of 'process review' workshops, each of which focused on specific processes, and worked through a four-step analysis:

- Step 1 reviewed the existing 'as is' process, documenting it as a detailed flowchart
- Step 2 stepped back and looked at the value and quality that the process should deliver to our customers. 'Customer value' asked how the processes created value for a user, for example the 'e-book acquisition' process created value by adding links which made the e-book accessible to the user. 'Customer quality' asked what distinguished a good service from a poor one, for example 'reliability of links' for an e-book
- Step 3 took these insights back to our existing processes, and asked which steps in the existing process added value, and which steps were considered 'waste' or inefficiency which did not *directly* contribute to customer value or quality
- Finally, Step 4 revised the 'as is' process map, and produced a 'could be' process that minimized the number of 'waste' steps and focused more on value and quality.

These 'could be' workflows outlined by these flowcharts were by no means fully detailed and documented processes – a full redesign of processes would have taken too long, and would have required much more extensive knowledge of Alma than we had at that stage. But what we did create was a framework of how we were aiming to set up workflows in the new system, which would help us to concentrate on what we should be working towards and make sure that we left any 'legacy processes' behind with our legacy system.

### Managing people through change

The business process change project defined the changes we were aiming for, but delivering the changes depended entirely on our people. We knew this was going to be a long and at times difficult project, so it was essential to pay attention from the start to how we would support our staff through the unsettling process of change, and engage them as proactive participants in the change process.

Coping with any sustained change is difficult, so we wanted to support our staff by helping them build their resilience to cope with change. Before the start of the system implementation phase, we ran two sets of training – one for managers on how to manage their staff through change, and one for non-managers on how to manage themselves through change. Both sessions explored our emotional reactions to change, and aimed to give staff some tools and techniques for dealing with the experience of change.

We were also aware that the business change process could be potentially unsettling and challenging to staff motivation. In many cases, our staff had been working with very stable routines for a very long time, and their measure of delivering 'value and quality' was therefore based on performing their particular routine well – hitting their deadlines, or meeting standards of accuracy. If an external process review consultant came in and told them that large parts of their routine had been classed as 'waste' and inefficiency, they would feel understandably devalued, demotivated and disengaged.

One way to mitigate this risk was to engage staff in making the changes themselves. So we sought to give staff as much control as possible over the business change process by doing the majority of the work in-house, and involving as many staff as possible. Our sole use of external support was to buy in some initial LEAN training to introduce staff to the key concepts, and train some of our managers to act as 'business change champions'. These

'we sought to give staff as much control as possible'

64 change champions then ran the process review workshops in-house, and we invited most of the staff involved in each process to the workshops, which both recognized the value of their knowledge and gave them a sense of ownership and control of the process. Finally, we ensured that the outcomes of the reviews were shared with all staff, rather than being presented only as reports to senior management.

There were some downsides to this in-house approach, for example a consultant may well have been able to spot more potential improvements to our processes, and achieve even greater efficiencies. However, it has also delivered significant advantages which will last far beyond the lifetime of the project. In addition to staff engagement and motivation, we have built our capacity to support an ongoing review of our business processes, emphasizing for our staff that this was not simply a one-off exercise which ended with the go-live of the new system, but forms the foundation of a culture of continuous improvement.

'a culture of continuous improvement'

### Ongoing change

Six months after go-live, we have now completed the formal implementation of our new system, but we are still working through the wider change process. Managing people through a major change takes time and needs ongoing support. Meanwhile, our business change team has morphed into a 'continuous improvement steering group', which co-ordinates the implementation of the processes we have already redesigned, as well as initiating further reviews to ensure we continue to develop our use of Alma to exploit its full potential. We have delivered a lot of change but, more importantly, we have developed a capacity for change which will continue to support us into the future.

Competing interests: The author has declared no competing interests.

#### Reference

1. Lean manufacturing. Wikipedia: [http://en.wikipedia.org/wiki/Lean\\_manufacturing](http://en.wikipedia.org/wiki/Lean_manufacturing)

Anna Grigson  
Head of Collection Services Group  
LSE Library, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, UK  
E-mail: A.C.Grigson@lse.ac.uk

ORCID: <http://orcid.org/0000-0003-3791-6226>

## Learning from observation: end user reaction to the Summon discovery service

### Overview

User experience was a priority for the Digital Library Services (DLS) team at Manchester Metropolitan University (MMU) when the Summon discovery service was launched in 2013. The team chose to brand Summon as 'Library Search'.

In terms of its capabilities, limitations and nuances, Library Search presented marked differences from the catalogue, subject-specific databases and MMU's outgoing federated search system. The DLS team were conscious that while some users would experience these changes as improvements, for others who had established workflows for research and



CATHERINE  
McMANAMON

(Formerly Assistant Librarian at MMU)  
University of Liverpool Library

65 information discovery, the changes could instead be seen as challenges. As the changes were significant, DLS undertook a study early in the process in order to develop a more informed understanding of user experiences, and respond in a timely way.

With that in mind, the team researched the theories and principles of usability testing. Ease of use, intuitiveness and efficiency were all things we wanted to offer our users through Library Search. As the focus of this study was on both user attitudes and actual use, DLS wanted more robust data and feedback than a standalone survey could provide; thus the group settled on a threefold strategy of survey, direct observation and focus-group discussions. The study design was informed significantly by a similar investigation undertaken by Martin Philip at the University of Huddersfield<sup>1</sup>, which was carried out shortly after Huddersfield had implemented Summon; the DLS team learnt a great deal from this approach and in many respects was able to avoid reinventing the wheel with the MMU study.

'a threefold strategy of survey, direct observation and focus-group discussions'

The three elements provided different types of data and enabled the team to study a broader sample of users. A survey was the simplest and quickest way to capture key information about information-seeking behaviours and experiences of Library Search across campuses. The observed search tasks allowed DLS to monitor users actually interacting with the system and provided useful data to help determine how usable it was in practice rather than in theory. They also allowed the team to observe which features users were prioritizing, and to identify 'unknown unknowns' – instances when users may have thought they were using Library Search to its full potential, but were actually missing features that might have increased the effectiveness of their research. These 'unknowns' would not have been captured if analysis had been confined to surveys only.

Time and resources were a factor in the approach DLS adopted and, in this context, staying low-tech enabled us to understand the particular experiences of users at MMU whilst still providing the level of insight the team sought.

Participants were asked to use Library Search for 15 to 20 minutes as though they were looking for information for an assignment. They were provided with a list of sample assignment questions across different disciplines, or could use one of their own assignment topics if they preferred. Other studies have asked participants to complete very specific tasks in an allotted time, but DLS felt this would not be reflective of the way many students usually search. For this reason, we did not pre-determine or insist what participants should search for and facilitators did not interact with participants beyond asking them what broad topic they were researching at the start. This minimal interaction allowed the team to determine if any issues encountered were due to topical peculiarities, but ensured that the participants were not influenced in their behaviour by facilitating staff. Facilitators observed one or two students simultaneously, and recorded their observations on a checklist.

Finally, the study participants took part in a focus group which provided in-depth and contextual information that could not be gathered from the usability observation checklist.

## Findings

In all, 406 people were surveyed. Of those who had used Library Search, 85% strongly agreed or agreed that it was easy to use. The same number strongly agreed or agreed that Library Search returned information relevant to their search. Of respondents, 18% said that they always used the refinement features when using Library Search, with 45% stating that they used refinement features most of the time. Full Text Online was the most popular filter, with 57% of respondents (208) indicating that they used this refinement in their searches. Many more questions were asked but overall the data collected from this survey reflected positive user engagement with Library Search (see Figures 1, 2 and 3).

There were 32 participants in the observed search task stage of the study. The observation checklist sought data on behaviour, search technique, use of refinements, accessing

resources and use of additional features. In observation of search techniques, some quirks were revealed. While staff observed that most users (23) searched with multiple keywords and some users refined their search with use of filters, many ignored filters even though they would have improved their search. For example, a user was seen to type in the preferred date range of publications into the main search box and did not use either the date slider or date boxes, while another entered 'e-book' after their keywords, but did not limit by this under content type.

In terms of use of the refinement features, unsurprisingly, in the survey the Full Text Online filter was most frequently used. However, in the observation only ten users were

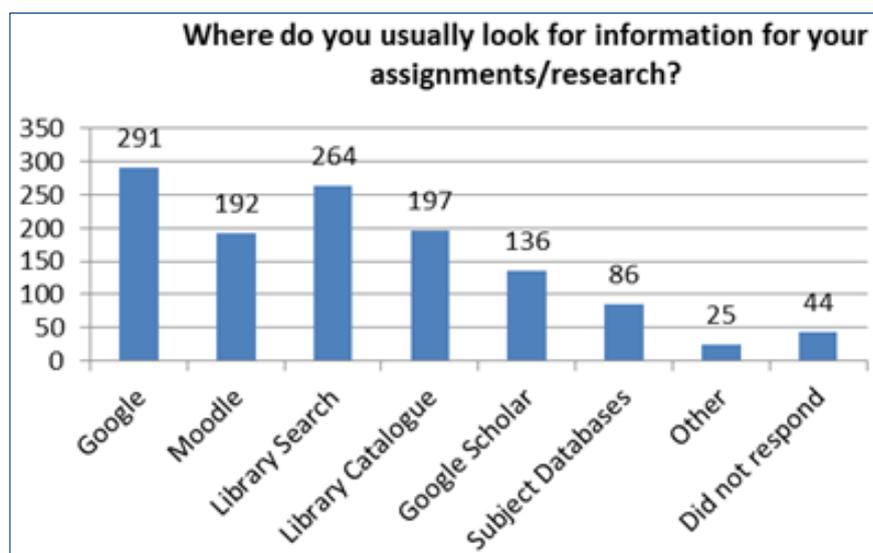


Figure 1. Respondents' engagement with different discovery tools and search engines

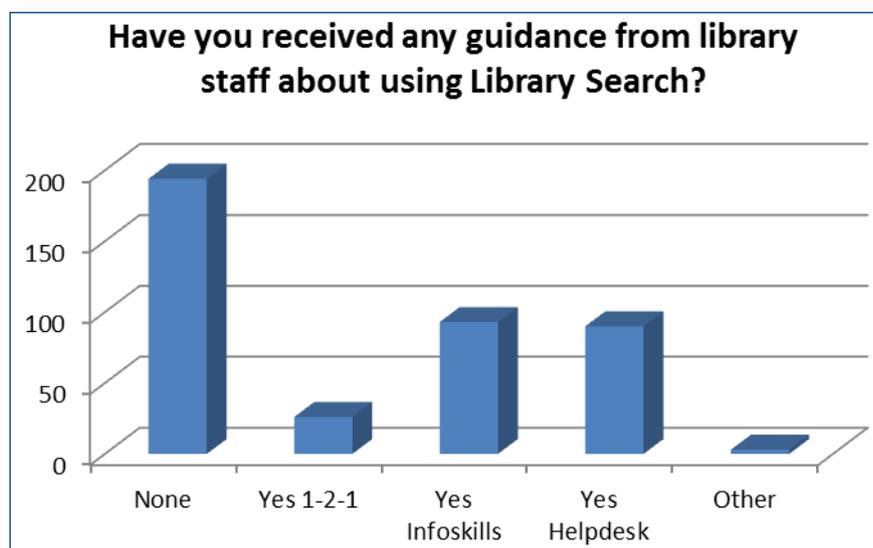


Figure 2. Type of help respondents received with Library Search

recorded explicitly limiting their search to include full text online resources, though staff noted that it was clear that many more were trying to access e-journals. This was one of the most interesting findings of the usability study – it highlighted the difference between what our users say and what they do.

The focus groups were conducted immediately after participants had undertaken their search tasks, so use of the system was fresh in their minds. The focus group questions tried to gather more detailed

'the difference between what our users say and what they do'

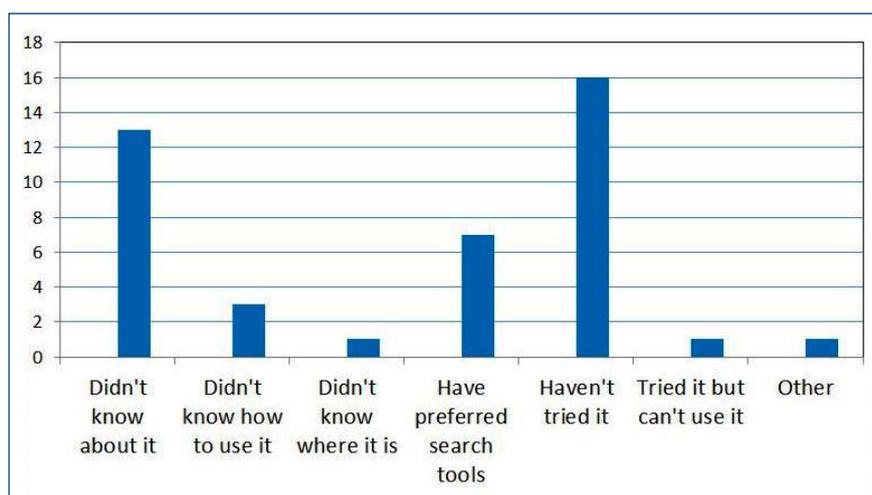


Figure 3. Reasoning of respondents who had never used Library Search

feedback relating to their impressions and experiences of the Summon layout and level of intuitiveness, the relevance of results, how it compared to other search tools, and users' overall satisfaction with Summon.

Feedback relating to layout and intuitiveness was generally positive, with repeated comments about a clean and straightforward interface and an improvement on traditional library-sponsored resources. There were many favourable comparisons with Google:

'Pretty easy, yeah. It's just like a Google search so you already know how to do it.'

'It's good to have the filtering tool on the left ... because Google doesn't.'

'It's quite similar in a way to the Google interface and it's quite comfortable in that respect.'

Relevance is an area of ongoing development for all discovery tool vendors. However, the survey and focus groups indicated generally good levels of satisfaction in relation to relevance, but this often required the user to be very systematic in their search and their use of the refinement panel.

In the main, inconsistencies or errors that users reported in the focus groups were issues already known to the DLS team. Specifically, these issues related to linking to newspapers and Citation Online records. Because of issues with the quality of metadata and issues with direct linking, DLS made the decision early on to exclude newspapers by default. DLS did not notice significant problems with newspaper content during user observation, but some user confusion about this did emerge in the focus groups: some had not noticed the exclusion, while others had and were annoyed by it. For those who did comment on newspapers, it was clear they did not know the best way to search for these resources.

Citation Online results caused the most frustration for participants. In some instances, the number of such results returned could have been lessened by a more strategic use of filters. However, the focus group discussion made very clear that some users did not understand the terminology or the point of such records and they saw them as items they should have been able to access.

## Actions

As a system that incorporates ease of use with authority, Library Search is filling a gap for MMU students with its integrated search. But there were still actions for DLS to take:

- a Library Search Awareness Week was implemented to promote the resource, with emphasis on the issues and features the study had flagged up. This has been consolidated in the long term through LibGuides<sup>2</sup> and a suite of Library Search podcasts. All subject guides feature library search widgets and additional information, as well as video tutorials

- MMU Library Services purchased ProQuest International Newsstand, which works more seamlessly with Summon. This acquisition was informed by this study and has allowed the team to remove Library Search's default exclusion of newspapers
- InfoSkills, MMU's branding for information literacy training, has put an emphasis on search-optimizing features such as the filter panel and filtering out citation online results.

The amount of useful data and information the team received from the project had made us very amenable to developing other usability studies. It was relatively low-tech, with the biggest investment being staff time. It has made clear the value of user observation. While we can and should be informed by what our users tell us, it is in the observation of real-world user behaviour that we can get to the unknown unknowns and have a truly accurate idea of how users are engaging with and using our systems and making sure they get the most out of them.

Informed consent was obtained from all participants involved in this study.

Competing interests: The author has declared no competing interests.

#### References

1. Philip, M, *Do students want a One-stop-shop to help them navigate their way around the maze of library resources? A Usability Study looking at the beta version of Summon, the new library search engine at the University of Huddersfield*: <http://eprints.hud.ac.uk/9824/> (accessed 26 January 2015).
2. MMU LibGuides: <http://libguides.mmu.ac.uk/content.php?pid=600186> (accessed 26 January 2015).

Catherine McManamon  
E-mail: [c.mcmanamon@liv.ac.uk](mailto:c.mcmanamon@liv.ac.uk)

## Semantic content enrichment: magic or myth?

### Introduction

User expectations in the academic world, as elsewhere, are often set by consumer websites such as Amazon where users can quickly and easily find the information they need, narrow it with faceted search, and get highly specific and relevant notifications. Semantic content enrichment offers a way of meeting these expectations for scholarly publishing.

This summary explains the levels of supporting information that can be added automatically to a piece of plain text using the example of a typical abstract from a biomedical journal article available in PubMed.

The article we have chosen concerns the Ebola virus, 'Ebola Virus Can Be Effectively Neutralized by Antibody Produced in Natural Human Infection'<sup>1</sup>, but any article could be processed in a similar way; this is sample content to demonstrate what is possible. We have processed the abstract of this article using content enrichment software tools.

There is now a good selection of proprietary and open-source software that can be used to support enrichment, including, for example, products from MarkLogic, TEMIS and Smartlogic. TEMIS Luxid was used to enrich the abstract for the UKSG presentation. Manual semantic content enrichment has been around for a long time, typically through the addition of keywords to articles using editorial skill and judgment. But it is relatively recently that publishers have



SAM HERBERT  
Client Services  
Director  
67 Bricks Ltd

'Semantic content enrichment offers a way of meeting ... expectations'

## Content analysis

The first content enrichment process demonstrated finds key words and phrases in the abstract, which are words or phrases that we programmatically identify as potentially interesting. This is achieved using a combination of statistical and grammatical analysis of the words and sentences in the abstract. Statistical analysis involves comparing the frequencies of words in the abstract against pre-calculated frequencies of words from a large reference corpus. Wikipedia serves as a useful general purpose corpus; more advanced applications might use a corpus tailored to the content (for example, the terms 'virus' or 'antibody' in the abstract we are looking at would appear significant when compared to a general corpus like Wikipedia, but would be less significant when compared against a corpus consisting of virology science articles). Grammatical analysis involves examining the sentence structure, assigning part-of-speech roles for the words in each sentence, and filtering the words and phrases found in the previous stage to remove those which do not appear to be nouns or noun phrases. The combination of these two steps identifies potentially useful keywords which may be added to the abstract as metadata. The advantage of this approach is that it involves minimal human effort because it is fully automated.

## Working in the background

Since such terms may not correspond cleanly to the terms a user may feel intuitively are significant, it may be better not to risk confusion by displaying them directly to a user viewing the article but instead to use them, for example, as input to an algorithm that calculates the 'relatedness' of pairs of articles and thus enables recommendations of similar articles to the user. If terms are to be shown, it is recommended to filter them in some way, e.g. to display only the highest weighted (most 'interesting') keywords.

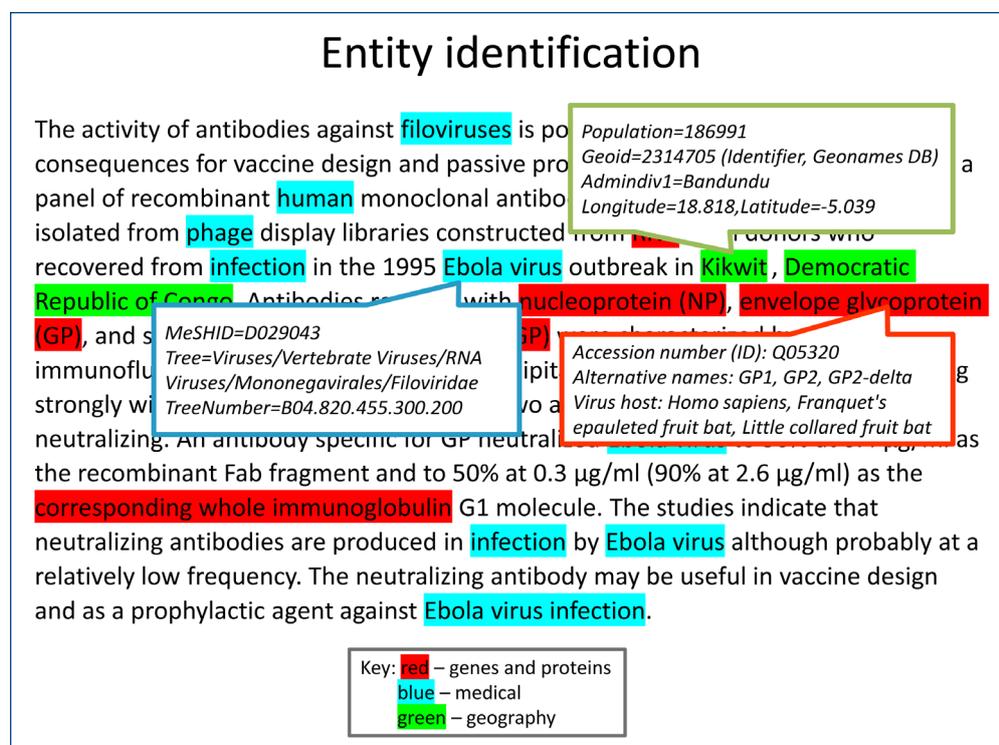


Figure 1. Entity identification against a taxonomy

## Making the most of specialist taxonomies

A more powerful way of recognizing important words is to run the article against a taxonomy, illustrated in Figure 1. This identifies entities, or validated terms from a known taxonomy (or other format of knowledge model). This approach potentially supports many features including semantic search and inline linking. In our example, we ran the abstract against three different taxonomies:

- running the content against a taxonomy of genes and proteins picked out the term 'envelope glycoprotein (GP)' amongst others and offered some additional information about this protein, such as alternative names that could be used in a semantic search, and also the host species for the viruses with this glycoprotein
- the well-known medical taxonomy, MeSH, identifies the Ebola virus itself. Here, because MeSH is hierarchical, the taxonomic term identification supports taxonomic browsing, faceted search, entity pages where all the content available about a specific subject may be gathered, and an opportunity for editorial content analysis
- running the abstract against a geographical taxonomy identified the location of the town Kikwit in Democratic Republic of Congo, with its population, administrative area, and its co-ordinates, which allowed plotting on a map.

### Identifying relationships with meaning

The next stage is to build on these concepts to try to identify relationships within the content, to understand and represent the facts being presented. This is significantly more difficult to achieve, and uses a combination of entity identification, grammatical analysis and identification of relationship terms. In our example, the software tools identified that Entity A: Envelope glycoprotein was connected to Entity B: Ebola virus by a reaction type Inhibition: Neutralise, i.e. it correctly analysed the action that envelope glycoprotein neutralises the Ebola virus. This is an example of linked data or a 'triple', the terms commonly used in information processing to refer to such relationships. This type of enrichment can deliver some really interesting and powerful capabilities for the user, for example an advanced search where the user could find all articles where there is a relationship between these two entities. When this kind of powerful enrichment is introduced into a real product, it is vitally important that it is accompanied by a carefully considered and extensive series of validation and QA tests, because false assertions would be very damaging. These tests must continue to monitor the ongoing enrichment, and should themselves be added to and improved throughout the product's lifetime.

### Training the machine

Thus far we have looked at identifying and connecting individual phrases within the text. Automated classification represents a rather different approach that considers the abstract as a whole. The abstract in our example can be classified as being about virology. To do this automatically requires a significant effort to set up, but is then very powerful when processing large numbers of articles on an ongoing basis. Starting with a core taxonomy, a subject matter expert manually assigns a training set of articles to each position or node in the taxonomy. Using statistical techniques, this allows the system to build a model of the content, by recognizing patterns of terms in the articles assigned to each node. Once the system has processed sufficient training articles, it will then be able to categorize new articles against the taxonomy without human intervention. This technique supports the ability for a publisher to create new content slices such as virtual journal issues on special topics. At least 100 training articles are needed per node, though this may vary depending on the nature of the content, and there should be minimal crossover between the content sets for each node (i.e. the assignment of articles to nodes should be as unambiguous as possible). Therefore, a smaller taxonomy with clear distinctions between the nodes will typically produce better results than a large taxonomy, and creating sample training sets will be less of a burden.

'the ability for a publisher to create new content slices'

### Semantic fingerprints

The sum of all the metadata in an article created by all these enrichment techniques may be referred to as its semantic fingerprint as it will be unique for each article. It can be used for improved relatedness compared to the simple keyword approach. Similarly, we can generate a semantic

'a range of personalized features'

71 fingerprint for the reader based on the articles they choose to view. By matching the two together, it is possible to offer a range of personalized features including tailored search and smart notifications, to present the most relevant material to the time-pressed researcher. Of course, researchers' interests change over time and so the system must constantly review and update the user's semantic fingerprint and, for example, give more weight to newer articles than older ones.

## Conclusion

In summary, the rich array of semantic content enrichment methods now available offers scholarly publishers the opportunity to meet users' expectations in the academic world. Semantic content enrichment is neither magic nor myth but is the intelligent application of modern content processing tools which, when done well, has the ability to greatly improve the user experience. We propose that it has become a necessary core competency for scholarly publishers to meet increasing user expectations.

'the ability to greatly improve the user experience'

Competing interests: The author has declared no competing interests.

### Reference

1. Maruyama, T, Rodriguez, L L, Jahrling, P B, Sanchez, A, Khan, A S, Nichol, S T, Peters, C J, Parren, P W H I and Burton, D R, Ebola Virus Can Be Effectively Neutralized by Antibody Produced in Natural Human Infection, *Journal of Virology*, 1999, 73(7), 6024-30 (PMID: 10364354).

Sam Herbert  
Client Services Director  
67 Bricks Ltd, Jam Factory, 27 Park End Street, Oxford OX1 1HU, UK  
Tel: +44 (0)7734 138274 | E-mail: [contact@67bricks.com](mailto:contact@67bricks.com)  
ORCID: <http://orcid.org/0000-0002-6590-1091>

**Article copyright: © 2015 Anna Grigson, Catherine McManamon and Sam Herbert. This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use and distribution provided the original author and source are credited.**



To cite this article:

Grigson, A, McManamon, C and Herbert, S, Information without frontiers – barriers and solutions, *Insights*, 2015, 28(1), 62-72; DOI: <http://dx.doi.org/10.1629/uksg.176>

Published by UKSG and Ubiquity Press on 5 March 2015

## Save the date!

Don't miss the next UKSG Forum

18 November 2015



**Watch the UKSG website for details: <http://www.uksg.org/event>**