Academic research and its relevance to enterprise search success

Martin White, Intranet Focus

martin.white@intranetfocus.com

At a time when maintaining a technology-employee balance is becoming increasingly difficult it is important to be able to take advantage of the substantial amount of academic research that is being undertaken on topics that include information retrieval, digital assistants, enterprise social networks, information systems adoption and collaboration effectiveness. I should say here that there is virtually no research on intranets. Academic research is certainly not being carried out only within universities; many excellent case studies have been published over the last few years. I have a collection of over 200research papers across the topics I've mentioned above, but I do have the benefit of access to the extensive digital resources of the University of Sheffield.

There are three challenges in using this research in practice. The first is finding the research papers. Google Scholar is my first choice because it offers date range search and also lists open source versions of published papers where they are available. Microsoft Academic and BASE (Bielefeld Academic Search Engine) also have their merits. The ACM Digital Library provides access to ACM journals and conference proceedings and IEEE Xplore does the same for the IEEE. Elsevier offers the Scopus database service for its own journals and books and there is also Web of Science from Clarivate. However, that leads on to the access problem. The ACM and IEEE databases are only available to members and Elsevier journals are behind a subscription pay wall. That is where Google Scholar is so useful in providing open source versions. If there appears not to be an open source equivalent putting the title of the research paper into Google as a search query sometimes highlights options that are not listed in Google Scholar.

Reading between the lines

The third challenge is in working out how to read an academic paper. My advice is to read the introduction and then jump to the bibliography at the end. The main purpose in doing this is to see what the date is of the most recently cited paper. If this is more than around three years old the chances are that the paper is the published version of a PhD thesis. Then move backwards to the section that gives an assessment of the extent to which the research scales, and the factors that might have influenced the outcomes of the research. This is usually a very honest assessment. One of the hot topics in academic research at present is "replicability". Too many papers publish results that other teams cannot emulate. Still working backwards you will come to the conclusions and a discussion of the research results. It can be helpful to read these final sections before reading all the

experimental methodology and outcomes. Without having a sense of the outcomes it is easy to become lost. The final step in the evaluation of the paper is to check on the authors to see how much previous work they have undertaken in the subject area of the paper.

In my experience people who see little value in academic research often do so on the basis of anecdotal evidence and not from personal experience. It is not easy to access this information (though it is usually possible to buy an individual paper) but the benefits can be quite considerable. Academic research may not provide definitive answers to very difficult issues but it can provide a vendor-independent framework for discussion and inspiration. In this article I am going to highlight some exceptionally valuable academic research that has been published over the last few months that has significant implications for enterprise search success.

Enterprise search satisfaction

Over the last two decades especially Microsoft Research has undertaken a significant amount of research into how people search on the web. One visible sign of this research effort is 'Interactions with Search Systems' by Ryen White, leader of the Cortana Research team at Microsoft. The bibliography lists over one thousand six hundred research papers but none of these addresses the way in which employees interact with enterprise search applications. The reason for this is that there have been no papers published on the topic except for a few examples that take a very narrow view of the topic. A number of search consultants have good anecdotal stories to tell about enterprise search behaviours but they are not based on anything approaching a rigorous base of sampling and analysis. As a result there are potential dangers in extrapolating results from web search and applying them to enterprise research.

Now at last Dr Paul Cleverley and Professor Simon Burnett (Robert Gordon University) have published (in the Journal of Information Science) what is without doubt a landmark research paper on the factors that influence user satisfaction with enterprise search applications. The Journal of Information Science is a subscription research journal published by Sage but there is an open access Author Accepted version on OpenAir@RGU, which is the open access repository of Robert Gordon University, Aberdeen.

There are three reasons for me applying the "landmark" label to this paper. The first is the scale, with more than one thousand users in a large multinational company providing feedback over a period of two years. Nothing on this scale has ever been undertaken. Over the last sixteen years the company has implemented three different search applications.

The second reason is that Paul Cleverley is a geophysicist who moved into information science roles in the oil and gas industry and then last year was awarded a PhD for his work on the use of filters and facets in enterprise search. So here is a discipline expert with a very solid understanding of research methodologies applying all his experience and expertise to understanding enterprise search behaviours. Moreover, since six out of the

ten largest companies in the world are in this sector, there is at least a reasonable expectation that the outcomes will be similar in other large multi-national companies.

The third reason is that this paper proves that it is possible for research with an academic rigour to be undertaken within an organisation. Academics come up with all sorts of reasons for not attempting research within organisations – now this paper and its methodology shows that it can be done, and how it can be done. Hopefully others will now follow this lead.

The methodology is what is usually referred to as a longitudinal mixed methods approach. First feedback was obtained from the search user-interface to gauge satisfaction with the search outcomes. Second interviews were carried out with members of the thirteen internal and contract staff supporting the search application. The two data sets were then triangulated to highlight areas of agreement (all but two), dissonance (none) and silence (two). The study was longitudinal, with the same group of users being monitored over a period of two years. The interviews were coded so that a clear differentiation could be created between satisfaction and dissatisfaction. The areas of dissatisfaction were Technology Factors, Information Factors and Literacy Factors. The analysis of the outcomes of the research is very thorough, and the paper closes with a definitive bibliography of almost one hundred and fifty research papers, reports and books.

Trying to summarise the outcomes of a twenty-four-page research paper is very difficult. Nevertheless, there are some outcomes from the research project that stand out as being especially worthy of close attention by the enterprise search community. In commenting on them in this article I run the danger of taking them out of context of the research and analysis so I would encourage you to read the full paper.

In the paper the factors identified that predominately influenced satisfaction were technology, information quality, information literacy and task utility. The technology factors include search tool reliability, search ranking and query syntax handling. In total these factors were the largest single group (38%) and that could be used as a justification for investing further in search technology. However together information factors (36%) and literacy factors (26%) accounted for 62% of the reasons for dissatisfaction and to me that indicates that technology investment on its own is not going to make a significant difference to search satisfaction.

Moving on to search-level metrics, the search application was used by around seventy thousand staff each month and generated over four hundred and fifty thousand search queries. The average query length was 1.89 words and the top thirty most frequent searches fell from 14% of all search queries at the start of the project to just 8% at the end of the project two years later when of course users had gained substantially more experience with the application. This confirms anecdotal evidence that the tail of low frequency queries is very long in the enterprise environment. In my view this has significant implications for "cognitive search" because there will be such low levels of use data from the majority of the queries to be able to predict optimal results. The percentage of results with "no results" decreased from 0.4% to 0.3% over the same period.

Throughout the section in which verbatim comments are included it is very clear that resolving these comments requires knowledge of the technology, the content and the use case. This requires a skilled search support team, but from the outcomes of the research this team also has an important role to play in overcoming the problems of a lack of search literacy. The literacy factors include the selection of a query term, results scanning and familiarity with the search application. As over 25% of all dissatisfaction events were attributed to poor search literacy there has to be a major question mark over the view that when technology and information content are optimised search outcomes will take care of themselves.

My final comment is so well stated by the authors that I will reproduce it from the paper.

"The importance of configuration in enterprise search was evident, where an unintentional change biasing documents over web pages let to sub-optimal results. With an average query length of approximately two words, made by users to dynamic growing corpus sizes, it is unlikely that many information needs will be met without constant configuration, promotion of authoritative (trusted) corporate information and monitoring of performance."

With all research projects there is always the danger that the outcomes are not scalable and extensible to other organisations. Reading this paper, so many of the comments aligned with the experiences I have had with probably one hundred or more enterprise search-related projects that I have every confidence that the outcomes will translate (with due care and attention) to any organisation that depends on being able to offer complete and effective access to global information repositories.

No matter how large your organisation, if you have responsibility for search management you should be taking this remarkable paper, marking it up paragraph by paragraph, and then using it to benchmark your approach to achieving the levels of search satisfaction that your employees expect.

The importance of Search Results Page (SERP) formatting

Over the years I have been involved in many usability tests where employees are given a search task to perform, such as, "Find the technical support manager for air compressors in Argentina." While an apparently simple task, the diversity of approaches employees take becomes visible very quickly. Multiple start points are immediately apparent, reflecting the experience and expertise of the employee performing the search.

An enormous amount of research has gone into information seeking over the last few decades. A survey of this research published in 2007 ran to over four hundred pages, and the pace has accelerated since then.

When assessing enterprise search performance, the focus is always on counting clicks, worrying about "precision at k," mean reciprocal rank, and other formulae that assume

users work their way sequentially through the ranked list of results. These clicks do not reveal an element of the search process: the stopping strategy for the search.

Relatively little research has been carried out into what might cause a user to stop a search. In enterprise search this could be something as simple as the date shown in the result snippet. One user may decide anything older than 2016 is not going to be relevant, while another user may stop at 2017.

Click traffic will not make this stopping strategy apparent, especially in cases where a session is halted and then resumed with a different query some time later.

The scent of a SERP

Peter Pirolli and Stuart Card developed the information foraging model for information seeking while working at Xerox PARC in the early 1990s. Ed Chi, a fellow Xerox PARC employee, further developed the model in the late '90s. The concept of an "information scent" refers to the way (for example) pigs can find truffles even though they are well hidden.

So what's the connection between truffle hunting and Search Engine Results Pages (SERPs)? The answer is a search user's view of results pages is informed by a wide range of proximal clues, which together create an information scent in the mind of the searcher. For example, a glance at ten PowerPoint files listed on the first page of results could bring a search to an abrupt halt before it has even started.

David Maxwell, a PhD student in computer science at the University of Glasgow and Leif Azzopardi, associate professor at University of Strathclyde, presented a paper at the 40th European Conference on Information Retrieval in March, which prompted this column. You can download the paper (along with many other interesting papers) from Maxwell's personal website. In their paper, Maxwell and Azzopardi hypothesize, model and then validate the impact the information scent of a SERP has on stopping strategies and therefore, search performance. In summary (and there is a substantial amount of data and analysis in the paper), they believe the role the quality of SERP presentation has had on search effectiveness and satisfaction has been significantly underestimated.

The paper goes on to discuss the search ability of users. Again, in the "click count" world, all users are assumed to have equal search proficiency and an equal command of the languages being used on the SERP. The paper shows search proficiency influences opinions about the usefulness of the page based on information clues from SERP, and the authors set out some potential categories of user proficiency. Another paper by Leif Azzopardi, this time with Paul Thomas and Nick Craswell (both eminent members of Microsoft Research) takes up the SERP topic. It is entitled Measuring the Utility of Search Engine Result Pages: An Information Foraging Based Measure and can be downloaded from the Microsoft Research site. A SERP is typically constructed from:

- A header, where the query box and query statistics are displayed.
- The core, where the main set of algorithmic results is shown along with advertisements and other answers e.g. navigational entity cards, image, video and news elements.
- Often, a right rail, where entity cards, advertisements, related searches, etc. are shown.
- A footer with navigational cues to the next or previous page.

Now this of course is a typical web site SERP but the principles apply to enterprise search applications as well. The essence of their argument is that traditional approaches to search metrics treat each result in isolation. Their research suggests that as a user works their way down a set of results they learn from the process and there is a cumulative effect that may lead them to take a different perspective on the rank order. This affects stopping strategies and the quality of information snippets in results lists may also be having a significant impact on the user's assessment of the results. The end result is that all the work that has been carried out in "precision at n" may not be a good indicator of search performance.

Implications for Enterprise Search

As with any research, the outcomes presented in this paper should not be generalized without carefully considering the methodology and analysis. The authors rightly set out where further research is required to understand more clearly the impact of information scent on stopping point determination. This research will undoubtedly lead to a more reliable assessment of information seeking behaviours in an organisation.

Even so, I believe all enterprise search managers can take away some lessons from the current research:

- Relying only on search click traffic analysis is rather like assessing a holiday beach from a monochrome print.
- Usability studies provide essential information about how the user is performing, not just how the system is performing.
- SERP presentation values are likely to have a significant impact on achieving high levels of search satisfaction. Further research (at an organisation level) will be needed to assess the improvement in performance.
- If this proves to be the case, then using cognitive search applications to present a small number of highly personalized results could be counter-productive.
- Key performance indicators, such as "precision at k" calculations, may potentially need to be completely reconsidered.

There are many other examples where academics are working on real-world problems. Earlier this year I wrote three posts about research into enterprise social networks, a topic on which there is a lot of passion but (in my view) very little insight. For the last few years I have been writing a Perspectives column for Business Information Review in which I scan through all the Sage journals to find academic research that could be of interest to

information and knowledge managers and then write a summary that conveys the potential impact of the research on organisations. Each column takes about a week to research and write but it (like eLucidate) is a task I enjoy, and from which I have gained many insights into novel approaches to the organisational management of information and knowledge.