



Information Retrieval  
Specialist Group

# Search Solutions 2012



29<sup>th</sup> November 2012  
BCS London Office, UK

Dyaa Albakour, Udo Kruschwitz, Andrew  
MacFarlane, Tony Russell-Rose, John Tait and  
Nilesh Thatte (Editors),  
BCS Information Retrieval Specialist Group

# Programme

	Page
9:30 - 10:00 Registration and coffee	
<i>Session 1: Web Search Challenges</i>	
10:00 Introduction	4
10:10 Milad Shokouhi, Applied Researcher, <b>Microsoft Research</b> , <i>Time-Sensitive Query Trends in Bing Logs</i>	5
10:40 Alex Jaimes, Senior Research Scientist, <b>Yahoo! Barcelona</b> , <i>What Can Search Tell Us? A Human-Centered Perspective</i>	6
<b>11.10 – 11.30 Coffee Break</b>	
<i>Session 2: Putting the user first</i>	
11:30 Martin White, Managing Director, <b>Intranet Focus Limited</b> , <i>It's not about technology, it's about people!</i>	7
12:00 Emma Bayne, Head of Systems Development, <b>National Archives</b> , <i>Enhancing the Search Experience through Interface Design</i>	8
12:30 Kristian Norling, CTO, <b>Findwise</b> , <i>What practitioners think about Enterprise Search.</i>	9
<b>13:00 - 14:00 Lunch</b>	
<i>Session 3: Beyond Keyword matching: searching with taxonomies, thesauri, ontologies</i>	
14:00 Stella Dextre Clarke, Chair, <b>ISKO UK</b> , <i>Thesauri, interoperability and the role of ISO 25964</i>	10
14:30 Leo Sauermann, CEO and Founder, <b>Gnowsis EU</b> , <i>Making Search Personal with NEPOMUK and Refinder</i>	11
15:00 Charlie Hull, CEO, <b>Flax</b> , <i>Search,plus - building taxonomy, autotaxonomy and media monitoring tools with open source search software</i>	12

**15:30 - 16:00 Coffee Break**

*Session 4: The future of search*

- |       |  |    |
|-------|--|----|
| 16:00 | Nicholas Kemp, Principal Scientist, <b>DSTL</b> , <i>Technology Landscaping in DSTL</i>                  | 13 |
| 16:30 | Richard Morgan, Client Services and Support Manager, <b>Funnelback</b> , <i>Relevance for the masses</i> | 14 |

- |       |                             |    |
|-------|-----------------------------|----|
| 17:00 | Fishbowl Session            | 15 |
| to    |                             |    |
| 17.30 | <i>Hot Topics in Search</i> |    |

Want to discuss your own hot topic? Join in, it's a fishbowl session!

18:00 to 19:00 **IRSG AGM**

17:30 to 20.00 Drinks Reception

## **Introduction**

*Dyaa Albakour, Udo Kruschwitz, Andrew MacFarlane, Tony  
Russell-Rose, John Tait and Nilesh Thatte  
BCS IRSG*

Search Solutions is a special one-day event dedicated to the latest innovations in web & enterprise search. In contrast to other major industry events, Search Solutions aims to be highly interactive, with attendance strictly limited.

The programme includes presentations, panels and keynote talks by influential industry leaders on novel and emerging applications in search and information retrieval.

Milad Shokouhi,  
Applied Researcher,  
**Microsoft Research**,  
*Time-Sensitive Query Trends in Bing Logs*

Freshness matters and all key elements of a search system ranging from auto-completion and query suggestion to result ranking and vertical selection should consider the time-sensitivity of queries. In this talk, I describe some of the key challenges in identifying time-sensitive queries and building freshness-aware ranking models, and summarize some of the state-of-the-art models developed for addressing such issues.

Alex Jaimes,  
Senior Research Scientist,  
**Yahoo! Barcelona,**  
*What Can Search Tell Us? A Human-Centered  
Perspective*

Search plays a major role in many human activities, helping people make decisions: from shopping to voting, and from health to entertainment, among others. In spite of the importance of search, however, the process itself is not well understood and it is often viewed as a single, contextless action resulting in a single click. The implications of this are many, but include the use of wrong computational metrics to judge relevance, inadequate search interfaces, and less than ideal display of results. In this talk, I will discuss search from a Human-Centered view and examine the process itself as well as talk about the many important insights we can get from search in terms of computational and social factors, describing the implications of findings on the design and deployment of real-world search functionalities.

Martin White  
Managing Director  
**Intranet Focus Limited**

*It's not about technology, it's about people!*

When enterprise search does not work as expected it is rarely the technology that is at fault but instead failure to invest in a skilled search support team. Finding people with the right skills and experience seems to be a challenge for search vendors, search implementers and organisations trying to get the best from search applications. Based on work carried out by the author for the European Commission in 2011, together with some recent market surveys, this paper will look at the people-related challenges and opportunities for the search community.

Emma Bayne  
Head of Systems Development  
**National Archives**

*Enhancing the Search Experience through Interface  
Design*

This presentation will demonstrate Discovery, The National Archives' new catalogue. Discovery is about using search technology with intuitive design to open up The National Archives' ever-growing and wide-ranging collection, spanning 1,000 years of history. In this presentation, Bayne and Phillips share their experience of developing a new subject taxonomy covering category selection and rules development for over 20 million objects, and explain how they have introduced user tagging alongside new search logic to make the collection more findable. They will discuss how user interface improvements complement search technology enhancements to make searching the collection easier for everyone, from first-time users to experienced researchers. Delegates will gain practical knowledge of:

- Using taxonomies to improve search performance.
- Marrying user interface improvements with enhancements to search technology.



Kristian Norling  
CTO  
**Findwise,**  
*What practitioners think about Enterprise Search*

Earlier this year a survey was conducted by Findwise, asking practitioners within the enterprise search field, what they think and do with search. The results did corroborate findings from many earlier studies. However what was interesting was to understand what the group of practitioners did that were (very) satisfied with their current search solution and what the dissatisfied group does (not do). Kristian will present both best practice, based on what leading practitioners do, and pinpoint what the critical areas for improvement are for those that would like to improve their search solutions.

Stella Dextre Clarke  
Project Leader, **ISO NP 25964**  
Chair, **ISKO UK**,  
*Thesauri, interoperability and the role of ISO 25964*

In the twenty-first century, retrieval thesauri are designed for networking and seamless integration with other semantic tools. ISO 25964 has arrived to guide their construction and use, leading to withdrawal in 2011 of the previous standards ISO 2788 and ISO 5964 (for monolingual and multilingual thesauri respectively). The new standard has two parts as follows:

*ISO 25964: Thesauri and interoperability with other vocabularies*

- *Part 1: Thesauri for information retrieval* (published 2011)
- *Part 2: Interoperability with other vocabularies* (for publication Jan 2013)
- 

The “other vocabularies” considered include classification schemes, taxonomies, subject headings schemes, name authority lists, synonym rings, terminologies and ontologies. To meet the “interoperability” requirement, new features of the standard include:

- A data model and XML format for data exchange;
- Guidelines for mapping between one vocabulary and another;
- Recommendations on complementary use of vocabularies.

Stella Dextre Clarke specializes in the design and implementation of knowledge structures for information retrieval applications, including thesauri and taxonomies. Her consultancy clients have ranged from public bodies such as the Cabinet Office, to private sector publishers such as Elsevier Science. Since 2000 she has led the Working Groups responsible for revision of first the British and then the International Standards for thesauri. In recognition of the impact of her work she was the 2006 winner of the Tony Kent Strix Award for outstanding achievement in information retrieval. Currently she chairs the UK Chapter of the International Society for Knowledge Organization.

Leo Sauermann  
CEO and Founder  
**Gnowsis EU**

*Making Search Personal with NEPOMUK and  
Refinder*

Finding your own information is a challenge, when it is spread across multiple applications and formats. This talk is about two personal search solutions intended for personal information management. Both are results of the NEPOMUK EU research project. First is NEPOMUK-KDE, the desktop search engine of KDE 4.0. A semantic storage system is used to index the data of a Linux desktop and make it searchable and annotateable. Data is represented in the integrated RDF data model and specialized NEPOMUK Ontologies, which is different from conventional key/value data formats. NEPOMUK-KDE has become a widespread Semantic Search engine on which developers created applications, examples will be presented. Second is the Refinder Cloud Search engine. It provides a personal search engine across multiple cloud applications. Based upon the same standards as NEPOMUK-KDE, Refinder Cloud Search allows end-users to index their cloud data. Developers can use it to build their own solutions. RDF and ontologies enabled us to index across multiple data sources and formats. The developed formats have strong parallels with Schema.org, which opens a potential for merging standards in the future.

Charlie Hull,  
CEO, **Flax**,  
*Search,plus - building taxonomy, autotclassification  
and media monitoring tools with open source  
search software*

"Search" projects we encounter are often more complex than just providing a search box on a website or intranet - increasingly there's a need to help users classify content, add metadata and monitor for interesting or relevant input. Luckily, open source search engines and related software can be used to build tools and systems to answer this need, often on very tight timescales and at low cost. I'll demonstrate how we built a powerful media monitoring system capable of applying tens of thousands of stored queries every second, and a system for automatically classifying documents into a taxonomy with an easy-to-use GUI, both entirely built on open source software. Search isn't just about "search" any more.

Nicholas Kemp  
Principal Scientist, **DSTL**,  
*Technology Landscaping in DSTL*

Science and technology developments can have significant implications for UK defence and security. They have the potential to enhance or disrupt our capability and freedom to act. The Dstl Futures and Innovation Domain sponsors an S&T Futures programme which identifies emerging S&T through horizon scanning and engages with a broad range of collaborators to enhance understanding of the developing S&T landscape and its implications for defence and security. The programme then brings forward outputs which stimulate innovation and inform planning at various levels from individual research plans right up to senior level strategy.

Richard Morgan,  
Client Services and Support Manager,  
**Funnelback**,  
*Relevance for the masses*

There is a lot more to search than the relevance ranking algorithm including design, personalisation and the use of search technologies as a broader toolkit. But relevance still sits at the core and will only grow in importance in a world of big data and statistical analysis. This talk is about how we can use tuning, content optimisation and analytics to open up the black box of relevance to searchers and empower them.

## **Fishbowl session: "Hot topics in Search"**

Want to discuss your own hot topic?

It's a fishbowl session!