The Changing Face of Search

When we talk about search, we talk about Google. Google has transformed our lives and even our everyday language (if we need to look up information on the latest BCS events, we 'Google' it). Yet only a decade ago few people had come across the term. The question is: where are we heading? What will search look like in 5, 10 or 50 years time?

One way of trying to predict what the future holds is to solicit the views of the leading figures and companies in the search business. The BCS Information Retrieval Specialist Group (IRSG) has been organising such events for a number of years, and Search Solutions 2010, held in October at BCS London, is the latest such example. What follows is a summary of that discussion punctuated by our own observations and reflections.

One of the dominant themes at Search Solutions 2010 was the growing importance of "freshness" as a concept in search. In the past, most measures of search effectiveness have been primarily concerned with relevance, but a document that is considered relevant for a given query at this precise moment may not be so tomorrow or even 30 minutes hence. But freshness is not simply about identifying the most recent documents, as these are not necessarily the most authoritative results. In addition, not all queries require the same level of freshness.

Another issue that is becoming increasingly prominent is the use of context in search. Few people are aware that major search engines already make use of a variety of contextual information when returning search results, such as the user’s previous searches as well as the physical location from which the query is submitted. This is all closely related to another area where significant progress can be expected over the next few years: personalisation. Personalisation does not mean that users will be required to explicitly declare their interests (this is exactly what most users do not want to do!); instead, the search engine tries to infer users’ interests from implicit cues, e.g. time spent viewing a document, the fact that a document that has been selected in preference to another ranked higher in the results list, and so on. Personalised search results can be tailored to individual searchers and also to groups of similar users ("groupisation").

One of the most promising technologies that could radically change search as we know it is linguistic analysis and natural language processing (NLP). It has been argued for many years that NLP will eventually find its way from the research lab into the mainstream, and we are now seeing signs of it scaling to address web-scale search in the form of information extraction and question-answering services. The growth of social media and user-generated content (in the form of blogs, wikis, product reviews, etc.) is a further driver for the use of NLP.

The landscape of the search industry will change as well. At the moment, web search is dominated by three major players: Google, Microsoft (Bing) and Yahoo. Few would doubt that these three will continue to dominate in some form over the coming years, but we shouldn’t underestimate the disruptive effect of start-ups such as Blekko. In web search, the competition is no more than a click away.

By contrast, the world of enterprise search (i.e. the application of search technology to information within an organization) is much more volatile. As with web search, the landscape is currently dominated by three major players (Microsoft (FAST), Autonomy and Endeca), but in this case there are a great many smaller players, each trying to carve out their own identity in a complex market.
doubt by 2015 much consolidation will have taken place, with mergers, acquisitions and new entrants adding to a changing picture.

A major catalyst for this change is a growing awareness of open source alternatives, in the form of search platforms such as Solr. This platform in particular has reached something of a tipping point in maturity and stability, and has also the necessary support ecosystem to be considered a serious alternative to commercial offerings.

A further driver for change is a growing focus on the user experience, and the expectation that search engines should do more than simply deliver ten blue links. For sure, relevance is important, and most of the search vendors already do a decent job of utilising the cues available to present highly relevant results. But there is a growing realisation that the real value in search (particularly in its enterprise setting) is in embedding it in a wider discovery context, so that in addition to supporting basic lookup tasks (such as known-item search and fact retrieval), the system provides support for much more complex, exploratory search tasks, such as comparison, aggregation, analysis, synthesis, evaluation, and so on. Clearly, for these sorts of activity a much richer kind of interaction or dialogue between system and end user is required.

A much ignored issue in search is that of accessibility and the needs of disabled users. There has been a little work for Blind people, but very little in other directions, such as users with cognitive deficits (such as dyslexia) or physical problems (inability to use limbs etc).

The Disability Rights Commission conducted a major research project into interaction with various types of websites, including search. They found that the most effect type of user were Blind people who had severe difficulties from a group of 50 users including the blind, partially sighted, dyslexic people, people with hearing impairments and the physically disabled. More recently some work at City University has shown that dyslexia clear effect on directed and undirected search for information. It is estimated that around 10% of the popular have a form of dyslexia. Work on assistive technologies is useful to improve the experience of searching. For blind and partially sighted users, more intelligent ways of picking content from websites is desirable, otherwise screen readers can be overwhelmed with information, and cannot report research results to users very efficiently. For dyslexic people there is a need to understand the searching behaviour of such users, and build personalised interfaces which react to the type of dyslexia and learn from their interaction with the user interface. Severely physically disabled people would benefit from better audio recognition technology, which would allow them to enter queries using their voice, or using face gestures for those who cannot speak.

A further problem is the digital divide - providing access to information in third world countries, often with oppressive governments. The use of mobile devices can be a help, and there is growing evidence that in many poor countries, mobile phones are becoming a very useful way of exchanging information. The challenge for search designers on mobile devices is dealing with the screen size, the network (what can be transferred to the user) and the language to be used. For some parts of the world these are significant barriers.

Finally, when most people talk about search, they typically envisage a web page with a search box and a results list. But search is increasingly becoming a ubiquitous part of our daily lives, helping us make sense of the world around us. Search is the means by which we are able to cope with our overflowing email inboxes, to generate insights from masses of corporate data, and to discover new restaurants in an unfamiliar city armed only with a Smartphone and an Internet connection. Search will be everywhere, but invisible, contextualised, and personalised.