

Netrank White Paper

The Future of Search, May 2007



Table of Contents

Abstract	3
A Brief History of Search	4
Direction	8
Handling the Change	10
Conclusion	13
What Netrank can do to help	14
References	17

Abstract

With the increase in online spending and with search being the largest and the fastest growing source of referrals, no company can afford to ignore the changes in the search dynamic.

The increasing emphasis on personalisation, geo-targeting and 'web 2.0' within search engines (and particularly Google) is changing the face of online PR and marketing and as accurate position reporting becomes harder and web analytics rise to the forefront of many business development decisions, an understanding of the current position and a good idea of the direction search is taking is essential.

This paper will look at where search is now and how it has arrived at its current position. It will follow these trends on to extrapolate a likely direction for future developments, as well as discussing how to embrace the advent of Web 2.0 and turn it to your marketing advantage.

A Brief History of Search

Traditionally search engines have been part of a simple two way process. A user tells the engine what to search for and the engine finds the most relevant results it can for the keywords provided. Algorithms beavered away, determining what page was most relevant for the keywords, which pages were spamming, how authoritative a page was and a myriad of other factors. Essentially the search engine was a massive database, which the user could retrieve data from based on the quality of their search.

While the Internet was still largely in the domain of non-commercial enterprises, such as technical and university staff this was fine. Search engines were largely used as expert tools and anyone with any skill in using them could find what they wanted fairly quickly. The Internet explosion has changed all of this. In 1998 there were a few dozen results for a search, where as now there might be hundreds of thousands, since almost everyone is using the Internet as a resource and search engines have moved on from being expert tools into the realm of the regular user.

The fact that this growth will continue exponentially is unavoidable. As the Internet itself grows, so too will the amount of information and disinformation available on any given subject increase. Unlike traditional forms of archiving, almost everything is saved (with no mind to quality) and as the web grows, the task of searching or of being ranked well for a given search becomes ever harder.

The initial response to this includes improved advanced tools and speciality search engines for the user and improved algorithms for the engines. The largest change was the shift of emphasis from on-page to off-page factors. Any information can be manipulated to give a false impression and thus, where previously META data had served to categorise pages, now the meta tags were being spammed. Relevant content (quite rightly) remained 'king', without content a page was unlikely to get indexed for a search term, but this was heavily open to abuse, cloaking became a common method of affecting search ranking, either with hidden text, or underhand redirects.

This all changed with the advent of link-analysis algorithms. Rather than considering each individual web page as an isolated island, link analysis algorithms treat the web as an interconnected "spiders web", with the hyperlinks the strands connecting each page to each other and forming a relationship. The concept of link analysis opened up an entirely new way of determining the relevance and importance of pages, one where the linking habits of other pages could affect the ranking of your own site.

The most famous link analysis algorithm is PageRank. Invented by Google's founders Larry Page and Sergey Brin, it was one of the most important factors in the relevance of Google's results compared to other search engines at the time. This was a major contributing factor to their meteoric rise in market share to become the dominant search engine in most of the world today.

PageRank, like many early link analysis algorithms paid no attention to the contents of the linking pages. Without context, links became a commodity and link farms and exchanges grew. Buying links or spamming sites that allowed

posting became a common way to abuse this system.

Link contextualisation, that is, not just the number of inbound links, but where they came from, the authority and theme of those pages and the content which surrounded the links were introduced to algorithmically measure the popularity of pages. Links from, and information placed on trusted or 'authoritative' sites such as BBC, DMOZ or The Economist were weighted more highly.

Unfortunately, these new techniques merely provided a short term solution. The financial rewards associated with performing well in the SERPs are high and relevancy began to slip once again.

Link contextualisation immediately became prone to abuse. "Google Bombing" is an excellent example of how this was being manipulated where websites were made to rank highly for unrelated terms. Usage tracking opened up a new form of click fraud. Some Super Affiliates managed to become 'authoritative' and quality resources like DMOZ and Wikipedia were regularly manipulated, to the extent that DMOZ

closed its doors to submissions until very recently and Wikipedia have taken the step of adding the “rel=nofollow” attribute to all their external linksⁱⁱ.

Tackling off-site content resulted in spammers setting up several sites, with content designed to draw in users and spiders to core sites. This is increasingly becoming a more common practice. The problems the engines have are exactly the problems they had with on-site content, merely distributed over many sites. An entire sub-culture of 'splogs' (Spam Blogs) appeared where robots automatically produce blogs and comment on other blogs, for the sole purpose of affecting the index.

Search has now become more personalised. The earliest form of search personalisation is perhaps geo-targeting, where the information about the user's location can be used to deliver more relevant results. Now search engines have access to many other pieces of information about their users, and are beginning to use this to provide more focused results. User behaviour is measured more closely. Some search engines monitor which search results users

click on, which is likely to be used to determine future search results. Users are also encouraged to use search engine toolbars, which report back behaviour patterns to the engines and give implicit data on page popularity.

More recently sub-engines, such as Trexyⁱⁱⁱ and younanimous^{iv} utilise the intersection of individual search patterns to predict best fit results and Google search has become personalised across the board for all logged in users. The advent of Web 2.0^v saw a huge boom in user interaction and associated search repercussions.

To date, the contest between the black hat spammers and search engines has been a continual arms race. For every method introduced by the search engines, the spammers have worked out a way around it. This looks set to continue into the future although following past trends will become increasingly difficult and above the capability of all, but the most technically competent and determined hackers.

With the increasing trends towards personalisation, the signs are that black hat SEO will move increasingly towards malicious controlling of users' computers

and website accounts to enable their tactics to work in a world where the users themselves are an increasingly important part of the algorithm. Google have already begun working towards helping users fight badware^{vi}, presumably with the impact to their relevancy in mind.

Where search goes from here is by no means set in stone. It seems clear that we are in a transitional phase in the development of web usage. While it appears that search is going to play a massive part in the future of the Internet, it is likely that we will see huge paradigm shifts before things settle, as the huge popularity of the social search movement inflicts a heavy toll upon the relevancy of current search methods.

Direction

Tim Berners-Lee's vision was of a participatory web. His idea was that a browser could also edit and the whole web would be collaborative. Web 2.0, however inappropriate a term it may be is a reality. Social networks such as MySpace, Digg, YouTube, Flickr, del.icio.us are hugely popular and their sheer weight of numbers renders them as powerful forces in the online media environment. The problem which these social networks pose to engines is that there is a marked difference between what users profess to wanting and what they actually use. Social bookmarking is big business. Votes, comments, tags, blogs, shared user behaviour and submitted choices allow users to collectively mould search to their needs. Yahoo! have been strongly advertising their increased emphasis on social search. Their Yahoo! Answers are being heavily marketed as a way of bringing the raw materials of search together to produce search 'art' through the application of "collective wisdom". Yahoo! have purchased both del.icio.us and Flickr recently and their published goal is "to change the game of search" and "tap the untapped authority" of users^{vii}, '.

Whatever the overt position offered by the engines is, user behaviour, whether measured explicitly or implicitly is the future of search relevancy. An algorithm cannot ever recognise spam as well as a human, but the principles of social search are inherently flawed in two ways. Popularity does not necessarily indicate accuracy of information and users are notoriously bad at recognising their own behaviour patterns.^{viii} This is likely to mean that while social search is going to continue to feature heavily in assisting algorithms, it will be reduced to a smaller proportion of those calculations.

Social bookmarking has many excellent uses. In particular, it allows algorithms to pick up on dialect specific expressions. For example, a British page about [lifts] may not ever turn up without an element of human intervention for an American user looking for the search term [elevators]. People who use similar bookmarks may well be looking for similar items. We already see this in e-commerce sites, such as Amazon's hugely successful recommendations engine: "*Products tagged "bees" are also tagged:"*" (followed by a list of associated tags)^{ix} and increasingly within SERPs. For example

with Google OneBox Results^x or Ask's Smart Answers^{xi} specific answers are supplied to the question most often associated with a search.

As search engines move more into other media channels, such as digitisation of books, papers and magazines, cable TV, streaming radio, computer game chatter, instant messaging and software chat, activities in offline media may come to influence search results. For example, a heavily advertised product in magazines may well be the result a searcher is looking for online rather than the result which might have otherwise been top of the SERPs.

Additionally, increased AI (Artificial Intelligence) and advances in neural net technologies coupled with the increased personal nature of search, may well allow cross pollination of trend analysis methods. This will allow the searches of groups of users who communicate or who follow similar patterns of usage both online and within other monitored media channels to have SERPS that are adaptive to their usage and that of their peers, thus tailoring what they see to more closely match what the majority of users and more specifically what their associates or "similar" classifications of users prefer, resulting in a truly organic and self adaptive search paradigm.

Handling the Change

As the number of users electing to utilise social search grows and searching across all engines shifts towards a less algorithmic result set, companies will begin to rank for keywords in ways that will be much less affected by linking programmes or page content. Essentially rankings will be available based on how 'cool', funny or controversial your page is perceived to be. In general this will mean that those web pages which play an early part in providing information, knowledge and entertainment to their users will have a huge advantage in the social SERPs.

It should be stressed that inferred popularity and algorithmic indexing are not about to be replaced by an entirely social voted ranking system. Spam recognition aside, algorithms and user behaviour monitoring should eventually win out, as users look for relevance and accuracy of results over perceived popularity. However the major search players are already incorporating democratic social elements into their results.

For companies eager to rank highly, search is becoming an ever more holistic process. Keywords and back-links are

becoming the 'noise' in ranking algorithms, so allowing enough tagging users through that noise to your message, in order to gain rankings within social search platforms will require an ever increasing amount of more traditional PR within an online environment.

Offline publicity will continue to play a huge role in generating online interest but as users spend more of their time online, standard advertising will be more available within, for example, the gaming industry. The 'World of Warcraft' game alone has over 8.5 million subscribers^{xii} at present and Google acquired Adscape, who deliver dynamic advertisements for video games, for \$23 million^{xiii}, generating another channel for branding and advertising.

Pay per click advertising is not enough on its own, but will become an even more valuable resource in search. Organic results are responsible for over 60% of traffic^{xiv}, but around 62% of users only follow links from the first page of results^{xv}. If you cannot guarantee a front page position for your page in any individual's personal or geo-targeted SERPs then two thirds of users will never even see your

position. Juxtaposed to this, eye-tracking studies clearly show^{xvi} that there is a huge visibility advantage to having a top 5 natural search result over paid listings and search users are up to six times more likely to click on the first few organic results as they are to choose any of the paid listings^{xvii}.

The case for a combination of both good organic search positions and paid listings is strong. Most of the larger e-commerce companies use PPC effectively^{xviii} and brand trust can be increased by seeing the same results in either section of the SERPs to the extent that a blend of paid search and organic search will produce an ROI greater than the combined value of each done independently.

Whether it is measured explicitly or implicitly, people talking about your pages is a requirement for rankings. Creating a company blog is important, but appearing in the blogs and feeds of others is going to increase traffic twofold, both through direct referrals and through associated ranking increase. Link-bait and viral marketing are already big business^{xix} and this is a trend which can only increase. As the amount of content on the web is increasing, so the battle to

be the most interesting, cool, or quirky - that is, the battle to get the most links and the most traffic, becomes ever more difficult. More and more a traditional offline marketing approach is being adapted for online use, although this is not at the expense of technological expertise.

Search engines are becoming ever more adept at handling pages, but they are seeking to remain relevant. Returning a whole page of near identical content is not providing a good user experience, so canonicalisation is at the forefront of recent algorithmic changes.

Inappropriate redirects, poor use of robot exclusions, pages reached from multiple urls and similar technical issues are rife and ignoring these can cost a page dozens, or even hundreds, of SERP positions.

Recognising which search terms are converting is an area of search which is widely mismanaged. Ranking for trophy phrases may be good to show the board of directors, but the importance of web analytics in recognising the long tail conversions cannot be overstated. An example of this might be a search for [car] against a search for [approved used estate car Croydon]. While there may be hundreds of times as many searches for the generic trophy term, the long tail term is more likely to result in conversions because the searcher is more likely to have found what they are looking for.

Conclusion

Above all else it is important that the advent of Web 2.0 is something not to be feared as a threat to existing business models, but to be embraced as an opportunity to create entirely new revenue streams, and as a way to engage more closely with a wider number of potential and existing customers.

Making sure your site is optimised for engines, technically suited to search and accessible to engines is the first (and required!) step towards ranking. However as Web 2.0 and the personalisation of search become important factors in the search industry it is imperative for companies that want to have a significant visible presence in the search engines and high ranking SERPs across a large range of keywords, especially the quality traffic that is delivered by the long-tail, to consider that search is not just about your SEO agency or perhaps your Marketing and IT departments.

A much more holistic approach across your company will be needed to include your PR and Corporate Affairs teams, and even your Customer Service teams. And more importantly engaging with your customers and turning them into online advocates of what you do.

What Netrank can do to help

While the advent of personalised search will require a number of new approaches, it is important to remember that these are combined with, rather than replacing, the more traditional methods by which the ranking of sites is improved.

Ensuring that search engines spiders can navigate your site is and will remain the fundamental first step that sites require to be listed in search engines. If a search engine cannot find your pages in the first place, then they will not be listed.

- Netrank can produce in-depth SEO Audits for sites which includes (amongst many other areas) determining whether the site suffers spidering problems with details of how this can be fixed.
- For complex sites which suffer from spidering problems, our Search Visibility Engine is a powerful and cost efficient way of making content visible to ensure that search engines can find your key pages.

On-page factors, such as the content of the pages and the markup used to represent it, will also continue to factor

into the equation. The importance of writing good content is, if anything, increasing – content that is “sticky”, that people will want to share, “linkbait” (content which is deliberately written in such a way that people will link to it) and other such viral content may be one of the key drivers to success in a world where what users think about your content becomes part of the algorithm.

- Netrank's online PR consultancy service - Search PR helps companies understand how traditional PR skills can be adapted to embrace the web 2.0 phenomenon.
- Our dedicated copywriting department provides content that is optimised for search from a technical perspective and which is compelling and engaging for your readers to encourage tagging and social bookmarking.
- Our in-depth SEO audits also include optimising the on-page aspects of a site to achieve maximum effect.

- If developing a search friendly and well-optimised site in-house is difficult, Netrank can develop professional sites for you, from microsites to full-blown brand sites, all designed from the ground up to be effective for search. We can also offer a revenue sharing model to offset the initial cost of site production.

All the existing forms of back-link building remain valuable – having authoritative and relevant sites linking to your site will still form an important part of the algorithm for determining whether your site is authoritative or relevant to a given topic. Conventional “non-social” backlinks may even form a part of personalisation algorithms. For example, if a user is identified as liking a particular site, other sites that this site links to may be considered as potentially more relevant for that user.

- Netrank believe that our traditional link building services will continue to be an important part of the overall optimisation strategy. Our trained back-link building experts can locate suitable linking opportunities and cultivate a

broad range of effective links on your behalf.

- Our Search PR service helps you to develop your own back-links through implementing PR campaigns online.
- Our Linkscape DNA product provides clients with a map of their online landscape, detailing the most authoritative sites for a particular keyword which can be targeted for linking opportunities.

The opportunities that the new social and personalised search paradigm presents are only just beginning to be understood. Netrank are one of the first search engine optimisation companies which are reaching out to explore this new search realm, giving us a firm lead over our competitors in this area.

- Our sophisticated iReputation brand management tool enables you to monitor your brand performance across Usenet, user groups, forums, blogs and online news sites and identify any potential PR issues. iReputation gives you access to what

consumers are saying right now about your brand, products and services making it easy to respond to the appropriate medium if desired.

- Our Search PR service includes advice about engaging with social mediums such as blogs while avoiding the common pitfalls that often catch out newcomers to this new medium.
- As a wholly owned subsidiary of radio group UKRD, Netrank are in a unique position to offer competitive rates and unparalleled access to the ears of the general public via this cost-effective and ubiquitous medium, providing even more ways to get your message heard.

For more details please contact the Netrank team:

Netrank

Broadwalk House South
Southernhay West
Exeter
EX1 1TS

+44 1392 411299

enquiries@netrank.co.uk

References

- i Nielsen - via PCWorld.com [<http://www.pcworld.com/article/id,124688-page,1/article.html>]
- ii 20 January 2007 [<http://lists.wikimedia.org/pipermail/wikien-l/2007-January/061137.html>]
- iii Trexy.com [<http://trexy.com/>]
- iv youanamous [<http://www.youanamous.com/about/>]
- v "Web 2.0 is the business revolution in the computer industry caused by the move to the internet as platform, and an attempt to understand the rules for success on that new platform. Chief among those rules is this: Build applications that harness network effects to get better the more people use them." Tim O'Reilly [http://radar.oreilly.com/archives/2006/12/web_20_compact.html]
- vi Stopbadware.org [<http://www.stopbadware.org/home/help>]
- vii Jeff Weiner [<http://googlesystem.blogspot.com/2006/06/yahoo-to-become-social-search-engine.html>]
- viii Jakob Nielsen [<http://www.useit.com/>]
- ix A search for the tag [bees] [http://www.amazon.com/tag/bees/ref=tag_tdp_jmp/]
- x As described by Google [<http://googlesystem.blogspot.com/2006/07/google-onebox-results.html>]
- xi As described by Ask [http://sp.ask.com/docs/about/site_features.shtml#smartanswer]
- xii As of 07/03/2007 [<http://www.blizzard.com/press/070307.shtml>]
- xiii February 2007 [<http://www.redherring.com/Article.aspx?a=21323>]
- xiv iProspect [<http://www.iprospect.com/premiumPDFs/iProspectSurveyComplete.pdf>]
- xv Online Media Daily [http://64.233.167.104/search?q=cache:CxtCG1jDZWoj:publications.mediapost.com/index.cfm%3Ffuseaction%3DArticles.showArticle%26art_aid]
- xvi Did-it, Enquiro & Eyetools [<http://www.prweb.com/releases/2005/3/prweb213516.htm>]
- xvii OneUpWeb [http://www.bb2e.net/google_topten.pdf]
- xviii Amazon UK received 13% of its search traffic from paid listings in the four weeks to 21st October 2006 and eBay UK received 37% of its search traffic from paid listings. [http://weblogs.hitwise.com/heather-hopkins/2006/11/search_engine_landcape_in_the.html]
- xix Examples include Sony Bravia [<http://www.bravia-advert.com/>] and Guinness [<http://landingpage.guinness.com/Gateway-Hands.htm>]