ICOMP Search Working Group White Paper: How Search Works

Introduction:

The first port of call for consumers seeking information online is a search engine, and for most consumers this means Google. Indeed Google enjoys a monopoly having more than a 90% market share across Europe[i] and handling in excess of 3 billion search queries a day.[ii] Search engines are largely funded by the companies who advertise on their results pages and pay a fee for each user click they receive.

As the Internet grows, offering an ever greater range of opportunities, products and services, our dependence on it also increases. So too does our reliance on search engines to find and access the plethora of information available online. As the effective gateway to the Internet, search engines play a critical role in steering traffic and navigating users around the hundreds of millions of live websites, providing users with hassle free access to the information they seek at the push of a button.

As search engines enable consumers to access information online, search engine ranking is immensely important to online businesses: the impact of a high ranking in the organic results on the search engine results page (SERP) cannot be overstated. Research shows that almost 90% of clicks take place on results on the first page.[iii]

But how do search engines function? What are the mechanisms which allow them to almost instantly deliver useful and relevant links to any given user query?

This paper seeks to shed light on the inner workings of a search engine, explaining how SERPs are produced and links are ranked therein. It will also highlight the issues that arise when a search engine's business interests conflict with those of the Internet user.

Ranking Search Results

As noted above, up to 90% of consumers only click on results which appear on the first SERP. Websites know that their click-through rate (CTR) will be dramatically higher if they appear on the first SERP and some companies invest a huge amount of time and money to ensure a high search result ranking. It is also inevitable that the higher a company ranks on the first page the greater the CTR, given that the top two search results typically attract nearly half of all user clicks.
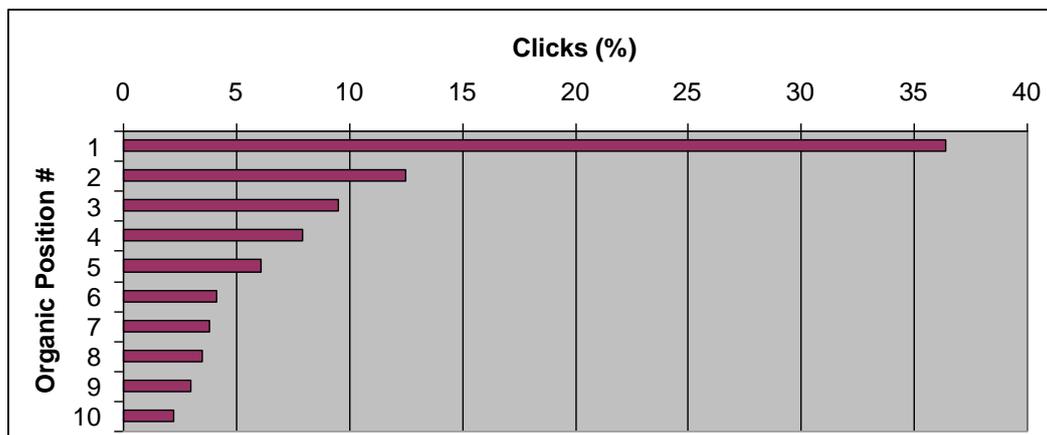
**User Click-through Rates on the First Google SERP[iv]**

To further illustrate this point, a study was commissioned by the online mapping company, hot-map.com and carried out by the Institute of Communication and Media Research at the German Sport University in Cologne. The Study measured exactly how a user's eye moves across the SERP, analysing what they look at for the longest period of time and where they click. Coloured heat-maps were generated using an advanced technology known as "Eye-Tracker" revealing and recording the views and clicks of participants. The results of the study illustrate how important it is for websites to occupy the top search results.[v] As shown below, the coloured areas of the screen highlight the "hot-spots" of where the users' gaze was concentrated, the white and red mouse icons where clicks occurred. The areas most viewed are highlighted in red, with the colour fading to yellow and then green as the concentration diminishes. It is worth noting that the search engine has placed its own product in the area of the screen that is most likely to attract users' attention.
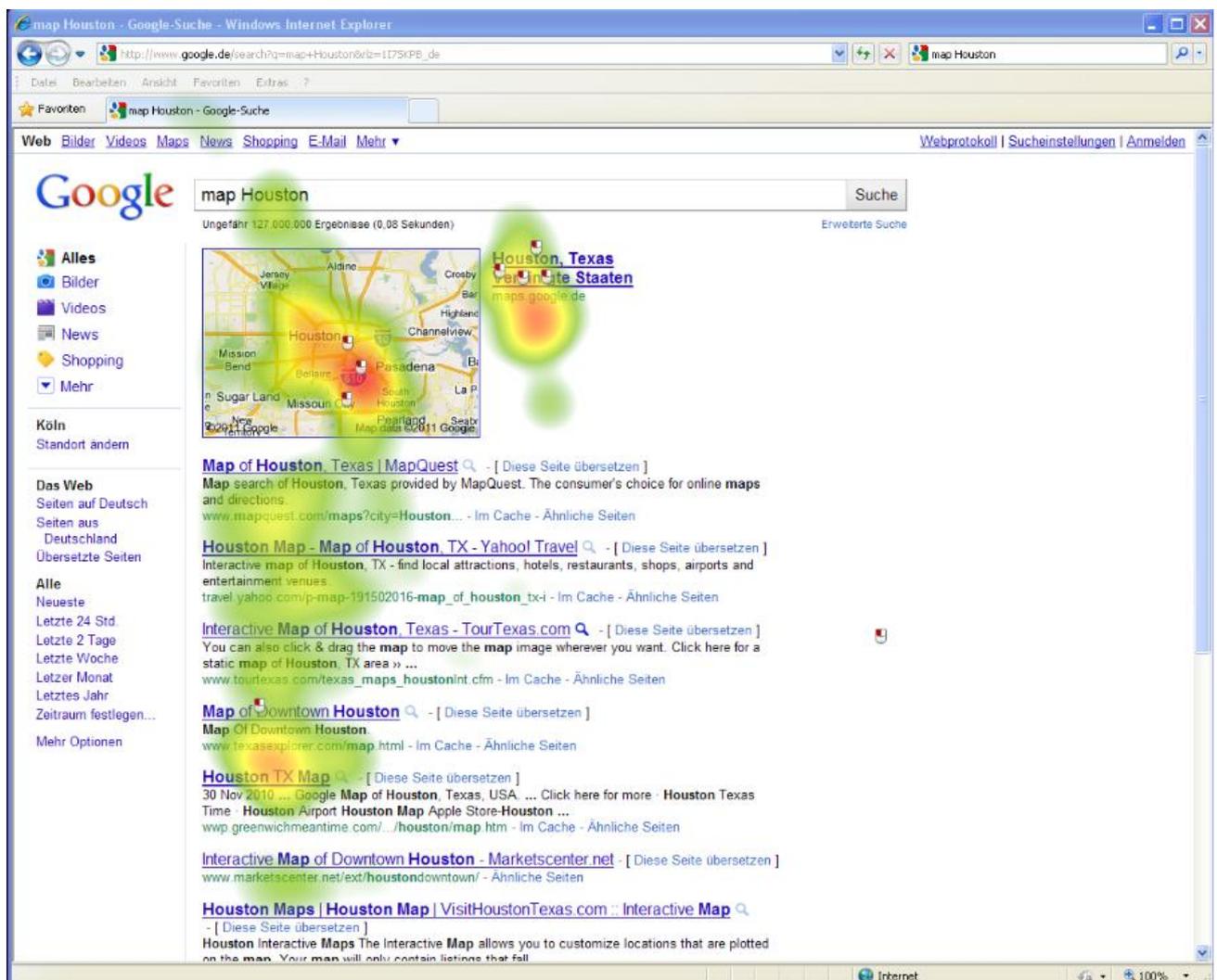
Eye-Tracking study results, July 2012

For those websites not ranked on the first SERP, the effects can be devastating, especially since the organic click through rate on the second SERP is typically a mere 1.5%. Since most websites are reliant on Google Search to direct traffic to their content, this can mean they effectively disappear from the Internet. Websites sometimes find themselves relegated in search rankings, often termed

"penalties". As a result, such websites can experience a sharp drop in traffic. Such issues will be discussed in detail later in this paper.
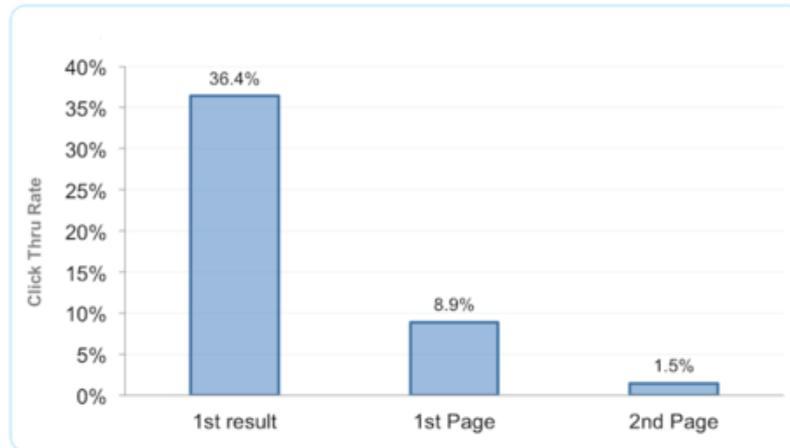

Organic Click-through Rate, #1, 1st Page, 2nd Page[vi]

SERP ranking is critical to the success of an online business. For this reason it is useful to understand how search engines determine their SERP rankings as well as to understand the mechanisms behind the transformation of a user query into a search result.

First a search engine uses index matching to identify all the available websites whose indexed pages match the search term entered by the user. In order to index websites, search engines employ a web crawler which browses the Internet for key words, phrases and other relevant signals from websites. This is normally carried out regularly to update the data and to ensure the search results remain current. Google, for example, crawls 20 billion pages a day.[vii]

Once all "relevant" website results have been identified, they are ranked in order of relevance, as perceived by the search engine. The final process applied to organic results is to remove from the SERP any website with a penalty, be it legitimate or otherwise.

Inner Workings of a Search Engine
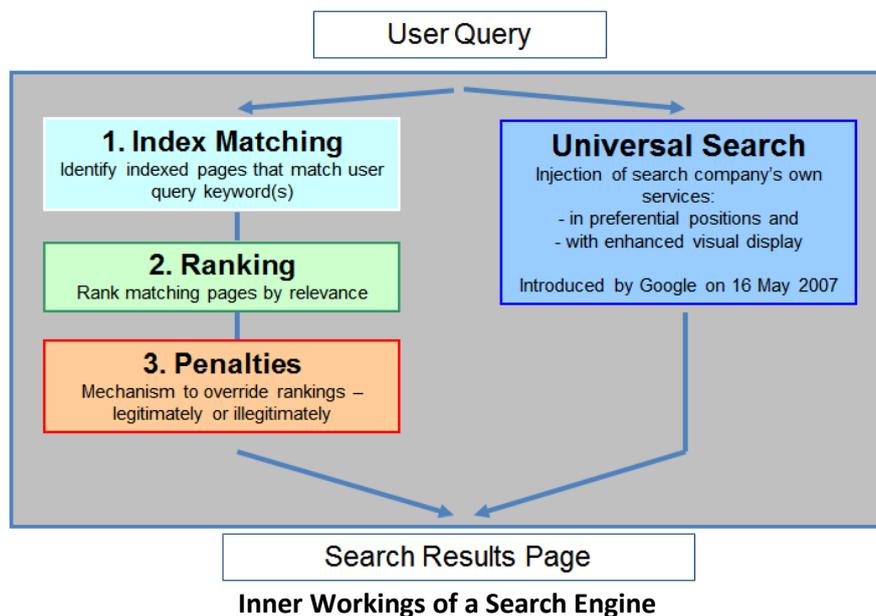
Before a user is served with the results of given query some search engines insert injection of their own products and services, often in a preferential position and with an enhanced visual display, in the result page. For Google, this comes in the form of *"Universal Search"*, which from 16 May 2007 has ensured Google's own products and services rank at or near the top of the Google SERP.

Case Study: Streetmap and Universal Search

Founded in 1997 and an early pioneer of online mapping, Streetmap quickly became a leading and well-recognised web brand, providing UK consumers with accurate, familiar and user-friendly maps in response to searches for place names, streets, grid references and postcodes. At its peak over 90,000 people per hour visited the site. In 2004, Streetmap was recognised by Media Momentum as one of the top 50 most dynamic companies at its annual awards. In 2006 a Superbrands survey voted Streetmap the 4th most trusted UK Internet brand, and by 2007 it was generating nearly £2.25 million in annual revenues.

The launch of Google Maps in 2005 had no effect on Streetmap's consumer website and even with the introduction, in early 2007, of a "Maps New" icon above the search box, Streetmap continued to thrive and grow even with Google as a competitor.

However, Google's launch of Universal Search in May 2007 changed the game. Using a separate algorithm, Universal Search automatically places Google's own products, such as Maps, YouTube, Images and News, at the top of the search results page and above any natural results. This means that customers, instead of being presented with a list of links to a number of competing mapping services, would see a Google Map at the top of most relevant searches.

This drastic change to the presentation of users' search results pages inevitably resulted in an increase in clicks through to Google's own products and a decrease to those of competitors. In the first month after the launch of Universal Search, Streetmap's advertising revenue fell by 40 per cent.

As a result of Google's unfair manipulation of search and the placement of Google Maps at the top of the results page, to the detriment of competing services, Streetmap and other similar services have been lost to consumers searching for postcodes and other geographic search terms. In a market where visibility is everything, consumers struggle to discover Streetmap in search results and as such visitor numbers have seen a sharp decline.

Universal Search artificially inserts Google's own vertical services into its organic horizontal search results using a different algorithm and ranking mechanism to that which is applied to other websites which compete with Google's services. The distortive effects are two-fold: firstly, as consumers are largely unaware that a separate algorithm has been used to place these results into the SERP, they view the implanted results as they would organic results. As we have already seen, almost 50% of users click on the first two results and so Universal Search deliberately leads consumers to believe the most relevant result for their query comes from a horizontal Google service rather than the organic results. Secondly, by taking up prime real estate on the SERP, Universal Search pushes other organic search results down the first page and onto subsequent pages, where they are extremely unlikely to be noticed by users. Since the introduction of Universal Search, Google have in fact introduced further measures to ensure users' search results contain as much Google owned information and properties as possible. Examples of this include the insertion of *Google+* results and also the *Google Knowledge Graph*, both of which result in less available space for organic search results containing Google's competitors.

As outlined above, some search engines impose "penalties" on websites and this has a negative effect on a website's final ranking. While many penalties are applied for good reasons, such as to relegate spam websites into obscurity, in a large number of cases penalties are often applied to legitimate websites, for very obscure and even illegitimate reasons. For websites subject to penalties, in particular the unexplained Google penalty, the consequences are severe and can result in huge drops in traffic and corresponding advertising revenues. Websites are rarely informed of the fact that they have been subject to a penalty and in many cases it can be virtually impossible to confirm the nature of the penalty and why it has been applied. Furthermore, the mechanisms put in place to confirm and challenge such penalties are far from transparent as search engines, in particular Google, provide very limited information in the public domain and in response to emails.

Furthermore, updates to Google's algorithm such as *Panda* have made it clear that those websites that are penalised and moved down the search ranking include websites that are viewed to have little or no unique content. The purpose of many vertical search companies, such as price comparisons websites, is to provide users with the best price for a specific product or service, for example. Therefore, by their very nature vertical search companies do not produce "unique" content. This is as true of Google products, which are inserted at the top of search results, as it is of the competitors who may have been the subject of a penalty and moved down the SERP.

As with the insertion of a search engine's own services, (e.g. Google's *Universal Search*), Internet users are more often than not unaware of the existence of penalties and unbeknown to a user the best answer to their search term may have been removed from the SERP.

Search Results – Reality Gap:

Despite the lack of transparency surrounding penalties and search preferencing, Internet users trust search engines to provide them with the most relevant, impartial results to any query. This is particularly true of Google, the search market leader where a market share in excess of 90% across Europe has been built on users' historical expectation that results are technology and mathematically driven.

Although the search box has remained almost unchanged since it was launched, Google's search results page has evolved a great deal over time. Increasingly, Google's SERP has benefited the company's own expanding commercial services, directing users to its own vertical services and enabling even-more space for online advertising. The upshot of this has been that there is far less real estate on Google's SERP in terms of relevant, objective and, most importantly, organic search results: something which a user would expect to see. The below screen shots depict the results of a search for "casinos" from 2001 and 2011 and illustrate this lack real estate on the Google Search homepage (note, the colour blocking has been added to distinguish the different areas).

Google Search Results for 'Casinos' 2001

Google Search Results for 'Casinos' 2011

It is clear from the above two screen shots that the SERP in 2001 differed tremendously to modern day SERP. In 2001, two thirds of the results page was devoted to organic search results, with the remainder of the page being covered by clearly labelled advertising, marked as *"Sponsored Links"*. Furthermore, there was a clear margin of separation between the organic search results and the paid-for-advertising.

Moving forward ten years to 2011, the page delivered as a result of the same search term differs significantly different. A user is no longer presented with organic results coupled with clearly labelled advertising to which they had become accustomed. Sponsored links, now labelled subtly as *"Ads"*, cover more than half (53%) of the results page. Since advertising revenues are the primary source of income for most search engines (in Google's case such revenues represent 96% of total revenues), there has been a significant increase in space dedicated to advertising.[viii] It therefore comes as no surprise that many search engines, in particular Google, now use nearly all "white space" available on their search pages for online advertising.

In the image "Google Search Results for 'Casinos' 2011", it can be seen that Universal Search inserts *Google News* results for "Casinos" (generically labelled as *News*) in a prime position on the SERP, below sponsored links and above the organic search results. Those results highlighted in the image in blue are results which are "blended into" the search results and appear where a user would expect to see organic search results. The remaining space to the left and top of the SERP (highlighted in yellow) feature the search engine's own services, such that 29% of the total page is taken up by Google's own services.

It is also worth remarking, that the Google-sponsored knowledge website Wikipedia tends to fetch page one SERP top organic rankings for most any query, further taking away space on SERPs from all other websites trying to compete.

The combined effect of advertising and preferencing, not to mention the removal of websites through penalties, means that organic search results are pushed further down the page, featuring only on the latter results pages or at all. A mere 11% of the first results page is devoted to organic search results, i.e., the search results which a user would expect to see.

With online advertising and search engines' offerings dominating much of the results page, a discrepancy arises between what a user expects to see and the reality of the SERP. Although users consider (wrongly) that results are normally organic, objective and not driven by commercial interest, the reality is quite the opposite: results are largely sponsored or promotional links and the search engine is heavily biased towards its own products – a result of the search engine's purely commercially driven interests.

For search engines, as with any business, the desire to increase profits plays a fundamental role in determining strategy. While fulfilling such desire is not *per se* wrong and for many online companies makes commercial sense, it can encourage highly illegal conduct, where a search engine is dominant, such as Google.

As we have seen, advertising revenue is a key source of income for search engines and indeed advertisers now spend more on advertising online than on TV[ix] and print[x]. An increase in the amount of space available on a SERP to advertisers also raises the likelihood of a user clicking on the advert and also the likelihood that the search engine will be paid by the advertiser. Placing adverts at the top of the search results page, where user eyeballs are most likely to land, further increases this possibility.

In expanding the number of vertical offerings and ensuring that these rank at the top of the results pages, search engines can benefit in a number of ways. They increase the number of pages available to users, thereby expanding the potential page space they can sell to advertisers. This in turn has the effect of boosting advertising sales and profitability potential. In addition, the fact that users relinquish even more personal information means that the accuracy of user profiling is improved allowing more targeted adverts to be served to users. This is important because search engines' business models are based on advertising revenues and the search engine able to collect the most user data, and therefore provide the most targeted and likely to be clicked on adverts, is likely to attract the most investment from advertisers.

Being in a position to provide more targeted advertising enables search engines to raise advertising costs, as companies recognise that a user more likely to click on an advert that directly relates to their interests is also more likely carry out a transaction on the advertisers website. For Google, which is hugely dominant in the horizontal search market, the benefits are even greater. Businesses, recognising Google's monopoly not only in market share but also user data, will more likely invest their advertising budgets on Google's platform. The increased competition for relevant search terms results in higher sale prices further increasing profits.

Conclusions

As outlined above, both the way search results are ranked and the gap between what a user perceives to be the rationale for SERPS and the actual reality of SERPs, have a number of consequences for both consumers and businesses: indeed the two are closely linked.

Consumers are being misled into believing that the results they see are purely based on relevance. The reality is that organic results are significantly distorted by the insertion of the search engine's own products (through, for example, Universal Search) and also by obscure penalties which operate in the background and impact consumers in ways they cannot imagine. Given that most consumers fail to click through to the second SERP page and that the vast majority of the first page is taken up by the search engine's own products and advertising, the range of products and services offered to consumers is significantly reduced.

Presenting consumers with minimal organic results on the SERP also has a detrimental effect on those businesses whose websites would otherwise have ranked at or near the top of the search results. Many businesses, in order to ensure that they are visible to users, are likely to resort to advertising as a means to gain user visibility. Given that Google controls over 90% of the search market, the channels for advertising spend are obvious. Google is profiting from the very businesses it has harmed through Universal Search and penalties by populating search engine real estate with online advertising, paid for by the very companies which have lacked visibility as a result of Google's conduct.

These factors all impact on the wider Internet economy and increase barriers to entry. Investors are already far less willing to support entrepreneurs with business proposals in areas where search engine giants are active as the risks are high and the chances of prolonged success are low and uncertain. In turn, this effects innovation, not only *vis à vis* existing companies but also among new businesses whose products could potentially bring huge benefits to consumers, in terms of new technology and choice.

Perhaps the greatest harm caused to online businesses, and with them consumer's right to choice, has been through the search manipulation joint strategy consisting of the introduction of Universal Search combined with penalties, often aimed at competing services. By giving preferential treatment to Google's own vertical offerings, (such as Maps, Images, YouTube, News, Books, Shopping, Travel, Finance etc.), and steering users to them, other companies in these sectors have experienced a severe drop in, if not a total disappearance of, their traffic. The preferential treatment, by a company already hugely dominant, combined with penalties which are most often unexplainable, should be viewed as joint and mutually reinforcing parts of a single overall anti-competitive strategy. Given Google's market shares in horizontal search and online advertising, it has a tremendous competitive advantage and conducts itself in a way which constitutes a breach of EU competition laws. Indeed, the European Union (as well as other regulators across the globe) is undertaking detailed investigations into numerous complaints that Google has abused its market power in search to achieve dominance in adjacent vertical markets.

[i] StatCounter Global Stats: *Top 5 search engines in Europe from Nov 2011 to Oct 2012*, November 2012
[ii] Search Engine Watch: *Spotlight Keynote With Matt Cutts*, August 2012
[iii] Optify: *Study: How the New Face of SERPs Has Altered the CTR Curve*, 2011
[iv] Search Engine Watch: *Top Google Result Gets 36.4% of Clicks [Study]*, April 2011
[v] Hot Maps: *Attention and selection behavior on "Universal Search" result pages*, July 2012
[vi] Search Engine Watch: *Top Google Result Gets 36.4% of Clicks [Study]*, April 2011
[vii] Search Engine Watch: *Spotlight Keynote With Matt Cutts*, August 2012
[viii] Venture Beat: *96 percent of Google's revenue is advertising, who buys it?*, January 2012
[ix] Digital Strategy Consulting: *UK advertisers 'now spending more on web than TV'- IAB*, October 2011
[x] Mashable: *Online Ad Spending to Surpass Print for First Time in 2012*, January 2012