

## **New architectures for presenting search results based on Web search engines users experience**

**[E.J. Martínez](#), [J.A. Pastor](#), [J.V. Rodríguez](#), [Rosana López](#) and [J.V. Rodríguez Jr.](#)**

**Department of Information and Documentation. University of Murcia. Campus of Espinardo, Murcia, Spain.**

**Introduction.** *The Internet is a dynamic environment which is continuously being updated. Search engines have been, currently are and in all probability will continue to be the most popular systems in this information cosmos.*

**Method.** *In this work, special attention has been paid to the series of changes made to search engines up to this point, which are currently in common usage; we also consider the objectives set for an immediate future in which, undoubtedly, searches will be increasingly attuned to user needs.*

**Analysis and Results.** *Since they originated, these information recovery systems have seen developments in their search algorithms and interfaces for presentation of results whilst their users' information seeking behaviour has changed substantially, making it possible to distinguish up to three generations (the latest of these is still being developed). Changes in users' habits and routines when interacting with these systems show a road map which is increasingly concerned with improving the searching experience.*

**Conclusions.** *Therefore, we are dealing not only with technological development, but also with an evolving process with regard to the use of information, in turn leading to a new Internet search paradigm.*

### **Introduction**

The platform for the majority of applications and Internet services, the Web, without a doubt, forms part of daily life for a huge number of people,

*searching for information on the Web, is for most, people, a daily activity. Search and communication are by far the most popular uses of the computer. Not surprisingly, many people in companies and universities are trying to improve search by coming up with easier and faster ways to find the right information (Croft et al. 2010).*

Internet search engines (Google, Yahoo! and Bing being among the most notable) and social networks (Facebook or Myspace) are the most popular applications by a significant margin, followed by e-mail accounts (Google mail, Yahoo! mail or Hotmail).

This dominant role played by search engines is nothing new and has been the case since they were first used more than fifteen years ago; during this period they have always been one of the most recognised and widely-used applications amongst Web users. Their importance is such that two of the current main information technology companies, Yahoo! and Google, have grown and consolidated around them, emerging from the development and popularity of their search systems: Yahoo! Search (initially a directory and, since 2004, mainly a search engine) and the Google engine itself. The high level of user loyalty to these search systems is also notable, as is the difficulty in changing from one engine to another or using several engines at once. Although most (more than 60%) claim to manage several ([Crowell 2006](#)), it is also true that when a search does not yield the required results, a much higher percentage of users (87%) prefer to change the search query rather than use another engine ([iProspect 2006](#)); there are also extreme cases of those users who use only Google because they do not know any other.

Since their creation, search engines have grouped predominantly around the current main technology companies, the two previously cited and Microsoft (owner of the Bing engine, the third important by a long margin from the aforementioned). The number on offer is reduced by the disappearance of some, the absorption of others and the almost total abandonment of the rest by the community of Web users. This grouping is the result of historically established competition amongst the owner companies.

The focus has also changed with regard to evaluating how these systems function. Initially, search response time and the size of the search engine's index were the most significant aspects when evaluating ([Gwizdka and Chignell 1999](#)), ([Oppenheim 2000](#)); later, the efficiency of information retrieval, by evaluating relevance, gained special attention ([Leighton and Svristava 1997](#)), ([Martinez and Rodriguez 2003](#)). These evaluations tend to focus on the search engine's level of precision (hit) when it responds; an extremely complicated task, 'certainly, with a few query words and short searching times associated, this task is more difficult' ([Baeza-Yates et al. 2005](#)). Recently, increasing importance has been given to improving user experience when interacting with the system, 'with increasing popularity of search engines, implicit feedback (i.e., the actions users take when interacting with the search engine) can be used to improve the rankings' ([Agichtein et al. 2006](#)) both in terms of improving information retrieval efficiency as well as improving user satisfaction with regard to usability and usefulness ([Fox et al. 2005](#)).

### **The first generation: classic search engines**

There were originally three types of information recovery systems available on the Web: directories (or indexes), search engines and meta-search engines. Although for a time each of these systems had their own Web space, directories have since fallen into disuse (especially since February 2004 when Yahoo! Search began functioning as a search engine using Altavista technology) and meta-search engines have never come to represent a valid alternative. Therefore, some fifteen years after they first appeared '*on the Web, search engines clearly dominate. Other approaches such as Web directories, social bookmarking or question answering services only play an underpart*' ([Lewandowski 2008](#)). The following illustrations show statistics on the current most visited sites while the search engines' predominant role compared to other applications can clearly be seen ([Hitwise 2010](#)).

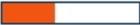
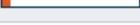
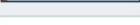
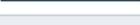
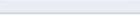
| Rank | Website           | Visits  |
|------|-------------------|---|
| 1.   | Facebook          | 9.74%    |
| 2.   | Google            | 7.30%    |
| 3.   | Yahoo!            | 3.75%    |
| 4.   | Yahoo! Mail       | 3.54%    |
| 5.   | YouTube           | 2.72%    |
| 6.   | msn               | 1.70%    |
| 7.   | Windows Live Mail | 1.53%    |
| 8.   | Yahoo! Search     | 1.29%    |
| 9.   | Bing              | 1.20%    |
| 10.  | MySpace           | 1.09%    |
| 11.  | Gmail             | 0.89%    |
| 12.  | AOL               | 0.79%    |
| 13.  | eBay              | 0.77%    |
| 14.  | Aol Mail          | 0.60%    |
| 15.  | Wikipedia         | 0.42%    |
| 16.  | My Yahoo!         | 0.38%    |
| 17.  | Yahoo! News       | 0.37%   |
| 18.  | Amazon.com        | 0.36%  |
| 19.  | Pogo              | 0.31%  |
| 20.  | Google Maps       | 0.28%  |

Figure 1: Top 20 Websites data updated to 11th September 2010.

Source: [Hitwise United States](#)

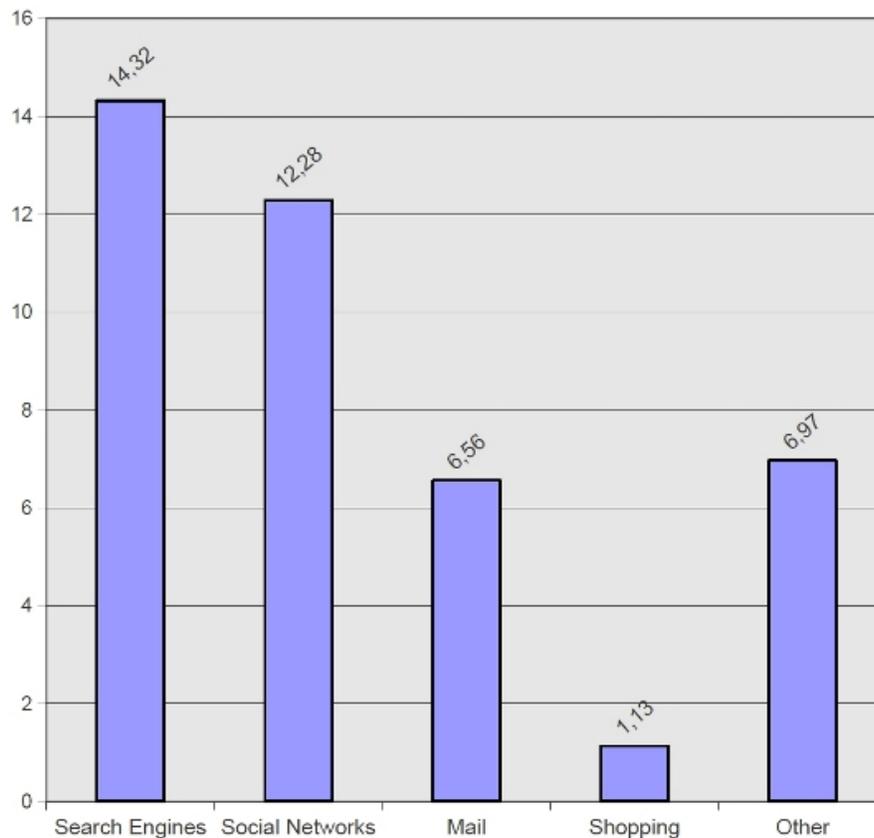


Figure 2: Grouping of the 20 most visited sites in the US by category.  
Source: [Hitwise United States](#)

The reduction in the supply of information sources implies a high concentration of searches around the trio formed by Google, Yahoo! Search and Bing (more than 90% of searches made anywhere in the world).

| Rank | Search Engine     | Searches |
|------|-------------------|----------|
| 1.   | www.google.com    | 71.27%   |
| 2.   | search.yahoo.com  | 14.26%   |
| 3.   | www.bing.com      | 10.14%   |
| 4.   | www.ask.com       | 2.35%    |
| 5.   | www.aolsearch.com | 1.22%    |

Figure 3: Top Search Engines data updated to 11th september 2010.  
Source: [Hitwise United States](#)

Over this time, several authors have analysed the objectives of users of these systems, which

*can be classified into at least two categories: navigational and information. A query is considered navigational when a user has a particular Web page in mind and is primarily interested in visiting the page. Informational queries, on the other hand, refer to the queries where the user does not have a particular page in mind or intends to visit multiple pages to learn about a topic (Lee et al. 2005: 461).*

Other authors refer to transactional searches, that is, those which aim to locate sites in order to perform a procedure (buy a plane ticket, for example) what is more, Baeza-Yates (2005) affirms that in later studies, some authors have even divided information and transaction queries into ten

sub-classes. Apart from the objective, other authors, notably Jansen and Spink (2006), have analysed how users search the Web, what type of operators we tend to use, how many terms form our search queries and how many pages of results we usually consult.

### Now: the second generation

When users have in mind the nature of the information item required to meet their needs (an advertising video, an article on a scientific innovation, a city map, etc.), the general search engines are no longer the best alternative, because of the wide heterogeneity of the information objects being managed; they also require different types of tools which have found niche space little by little. Thus,

*in the last few years, companies like Google or Microsoft created special search engines for scientific contents (such as Google Scholar and Windows Live Academic), where users can easily find the documents that they want without changing their searching behaviour' (Lewandowski 2008: 2).*

As the Hitwise statistics above show, or those supplied by Alexa.com (2010), each day sees an increase in the use of other extremely popular information retrieval systems on the Web, such as Youtube (videos) for example, Yahoo! News (Yahoo! News search) or baidu.com (music search). The growing daily importance of these search versions in Hindi, Chinese or Japanese should also be taken into account (Yahoo! in Japan or Google in India) or even the increase in search engines developed in China (e.g., baidu.com) this process has come about due to a consistent growth in access by citizens of those countries emerging to form part of the information society. As Figure 1 shows, this group of Websites is among the twenty currently most visited (Alexa 2010) and all of them apply the same search technology although the structure and nature of the data are different.

There has been a gradual introduction of second-generation search tools on the Web. These new information retrieval possibilities go hand in hand with certain changes in user habits and customs. When the Web was first developed, it brought a new problem: finding useful information therein, a task which was slow and complicated because the user had to wander through a tangle of links searching for interesting information, notwithstanding users' lack of skill at that time, meaning that locating information was often an immense or useless task. The lack of a data model to support the Web became the main obstacle (interestingly, locating information today is easy thanks to the abundance of information available, but the underlying problem, that of users wasting time discerning whether a document is useful or not, has still not been resolved). The appearance of search engines purported to improve this situation, with partial results. When first developed, retrieval and browse tasks were independent pulling actions and users' information requirements were resolved interactively, instead of automatically and permanently, with software agents pushing information towards the user.

At first, if a user wanted to see an image of the Eiffel Tower in Paris, they entered the required key words into the search engine to retrieve Web pages with information on this tourist site and with a bit of luck, there would be an image of it on one of the pages. Nowadays, this has changed significantly:

*that process may soon get a radical overhaul. Numerous companies — including the major search providers — are working to improve the search experience. For someone such as me who spends much of his life on the Internet, that's a tantalizing prospect (Wolverton 2009).*

Now the user requires much more interactivity from the search engine and is not content to read pages and extract their content; the user wants to access the required content directly. Search engine designers detect this demand and implement it, meaning that it is, in fact, the users themselves who are developing the Internet.

This synergy encourages the emergence of new systemic data retrieval possibilities; it also enables new user searching habits which are gradually reflected in a series of changes introduced both in the search engines' internal algorithms and their user interfaces. By using the search engine technology, the user can indicate, before carrying out the search, the nature of the required information object (photo, video, news, etc.). The major search engines have the ability to change the tab and redirect the search depending on the type of object required for retrieval.



Figure 4: Search on Yahoo Search! With the normal Web page search tab activated

Thus, the search for CAMINO DE SANTIAGO, shown in the previous image, returns a group of Web pages with information related to the question topic. However, the search entered in the following image will retrieve images instead of references to complete documents (note that the colour of the search tab changes from blue to black and the text on the Web form button).



Figure 5: Search in Yahoo Search! With the image search tab activated

Thus, if the user wishes to see an image of the front of the de la Gloria Portico on the Cathedral at Santiago de Compostela (or any other Spanish Romanesque architectural gem) it is no longer necessary to locate documents including photos of these places, as it is possible to access the image directly and even watch a video or read the latest news on the subject in services such as Google News. In this way, the first generation of search engines evolved from the *regular search* towards the *vertical search*, through systems that applied the same search technology to different data sources and whose queries were integrated into the same user interface, although the result was still given separately (Sullivan 2007).

The changes do not stop there: the generalist search engines have incorporated documents from other information sources into the results (generally the property of the same company as the search engine or through agreement). Thus, Google incorporates Youtube videos in its results (as well as those in Google Video) or the image portal [Freebase.com](http://Freebase.com) supplies Bing. This convergence of information sources is leading us toward a second generation of searches: the *universal global search*, simplifying the information retrieval process and covering an increasingly larger number of information sources. In presenting this new technology, Marisa Mayer (Vice-president of Google for searches and user experience) stated '*we're so excited about taking all these different silos of information and making them all into one*' (Sullivan 2007). This, as well as improving the user experience through interactivity and the collective disposal of several information sources, enriches the quality of the search engine results" (Lewandowski 2008).

In Yahoo! Search (the engine which at least visually appears to be leading the way in this field) the experiments with users were carried out in the summer of 2009 and, since then, several changes have been made to the search and presentation of results intended to improve user experience of search engines (Paczkowski 2009; Yahoo 2009a; Yahoo 2009b). We propose, without a doubt, that currently a second generation of engines is being configured, one in which the designers opt to converge data sources in the results, combining references to Web pages with reproduction of videos or image visualisation; meanwhile, user experience is seeing substantial improvements through provision of better search assistants, recommendations for similar searches and direct access to other resource sources that are widely circulated amongst the community of Web users (Wikipedia, for example). Just as some time ago, search engines developed advertising techniques in context, '*advertisements are selected for display based on the context of the pages*' (Croft

2010: 222) and reserved different areas on the results window for these advertisements, they now dedicate other areas to present (when possible) information items of a different nature and origin next to the results documents. In this way, if a user searches for information on the Irish music group U2, their favourite group, the search engine returns a similar screen to that shown in the following illustration:

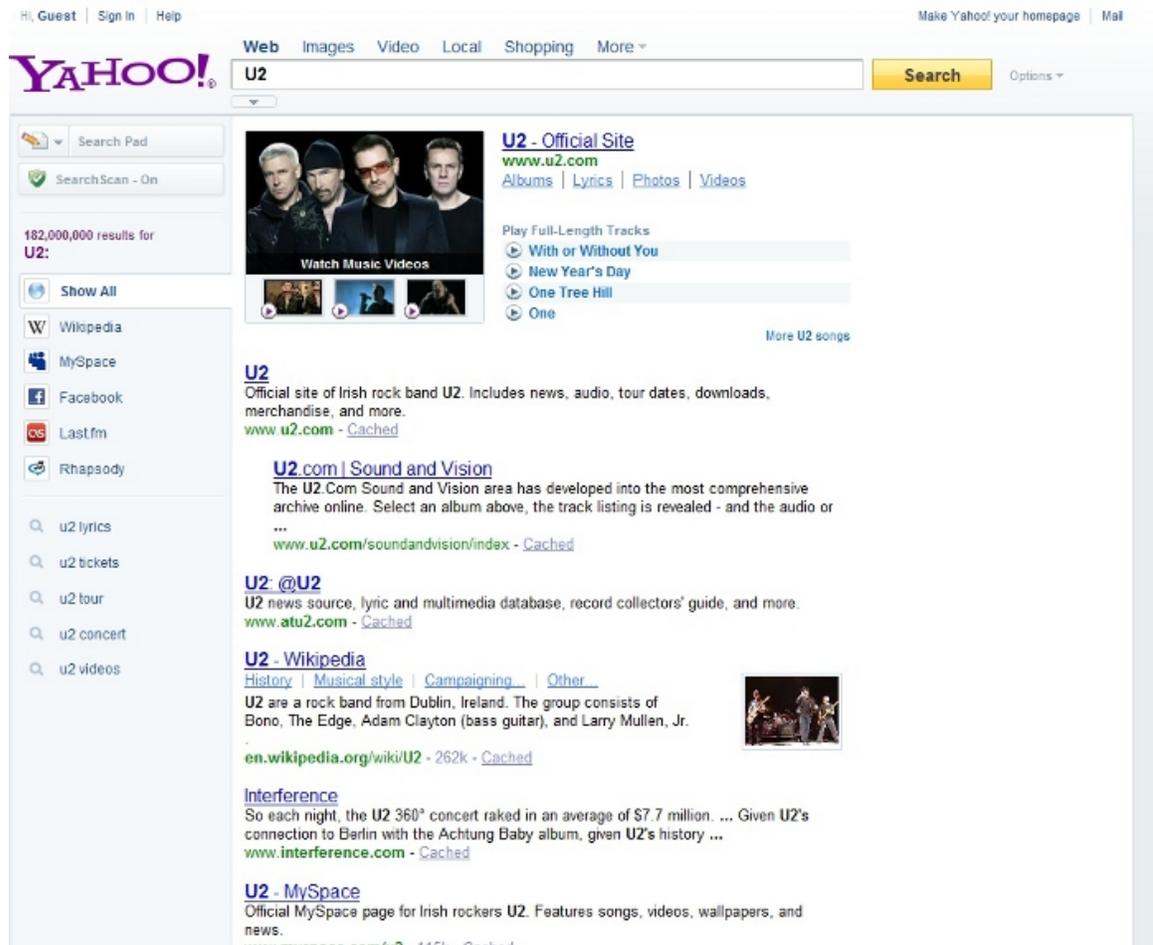


Figure 6: Presentation format in Yahoo! Search

This results window could appear, at first sight, to be a thematic portal dedicated to this band. Not only does it display references to Web pages with information on the band ([www.u2.com](http://www.u2.com) or [www.atu2.com](http://www.atu2.com), for example), but it also allows users to play videos or songs (*With or without you*, *New Year's Day*, etc.). Furthermore, in the same results area there are tabs which are adapted to the search theme, such as *lyrics* and *albums*, which appear when the results (and therefore the system interface) are personalised depending on the information needs of the user (focused on the music in this example).

Going into greater depth, another important innovation stands out: the column to the left of the results window is occupied by a selection of sites corresponding to portals with a wide range of information (of any nature and format) relating to the user question, so that the search engines '*help people explore the results that matter most to them through sites they know and love*' ([Yahoo 2009a](#)).

In our example of a query about U2, the system selects an encyclopaedia: Wikipedia, two online music sites: [last.fm](http://last.fm) and [Rhapsody](http://Rhapsody), and two social networking sites, one specialising in music: [MySpace](http://MySpace) and another with a more general reach: [Facebook](http://Facebook). This personalising, intended to improve the user experience, is not only the result of an improvement on the system interface, but is also due to the development of a series of associated technologies which converge in this new

way to present results. Search engine designers' approach is that one can interact with all the services and engine sites, turning the portal into something created by users and developers.

We can therefore witness a not insignificant transformation on the Web: the search engine no longer simply provides the user with a list of unconnected references or independent images or videos; it now attempts to design a dynamic, current and personalised user-information requirement portal, *pre-offering* cognitive structures on the user's subject of interest. What drives these changes? The answer seems obvious: the need to meet new user demands, requiring more dynamic, interactive and elaborate search results, far from the traditional list of references and Web pages, whilst simultaneously attempting to attract new users, more accustomed to image or video searches. This high level of interaction means search engines of a much more advanced level in terms of conceptualisation and data retrieve capacity and thus, results.

Although this is a common trend, not all engines have reached the same level of development and the way these advances are implemented differs substantially (Google, for example, tends not to present photographs and videos at the start of the page but mixes them amongst the search results). At the moment, Yahoo! Search, perhaps in an attempt to provide new ideas on how best to manage searches and gain new followers seems to have gained an advantage over its main competitors, Google and Bing, at least in result visualisation and usefulness, usability and interface; the new page is designed to make it easier to find and explore the most important elements. There are many authors, like Sullivan (2009), who believe that this advance is not so important as many of these innovations were already present in other search engines. To be sure of this a more profound analysis is required, both of how the results are represented on screen as well as the associated technology.

### **Review of the most important innovations**

Focusing analysis on the three main search engines (Google, Yahoo! Search, Bing), we can detect certain differences with regard to the redesign of search results presentation interfaces and also the use of some search technology, although these differences are not so substantial as to distance them from current design trends: convergence of information sources (highlighted search engine results, news, lyrics, etc.) and media types (mainly text, images and video) to improve user experience.

We therefore carried out a comparative analysis of the user interface in the systems (of the results window) to later highlight the innovations that we considered to be indicative of a true second generation of search engines. It is important to point out that some of these innovations still do not have a global reach and are only operative in versions of search engines in the United States. It must also be pointed out that the collection of documents in Bing in the United States is very different to the rest of the world (although they use the Bing interface, searches are still done using the older Live Search, also by Microsoft). It is foreseeable that, from being accepted by the user community, all of these innovations will soon be operational in the rest of the world.

Some of the innovations provided by Yahoo! Search are shown in Figure 6 above. One of these is completely visual: presenting the results page in a three-column format.

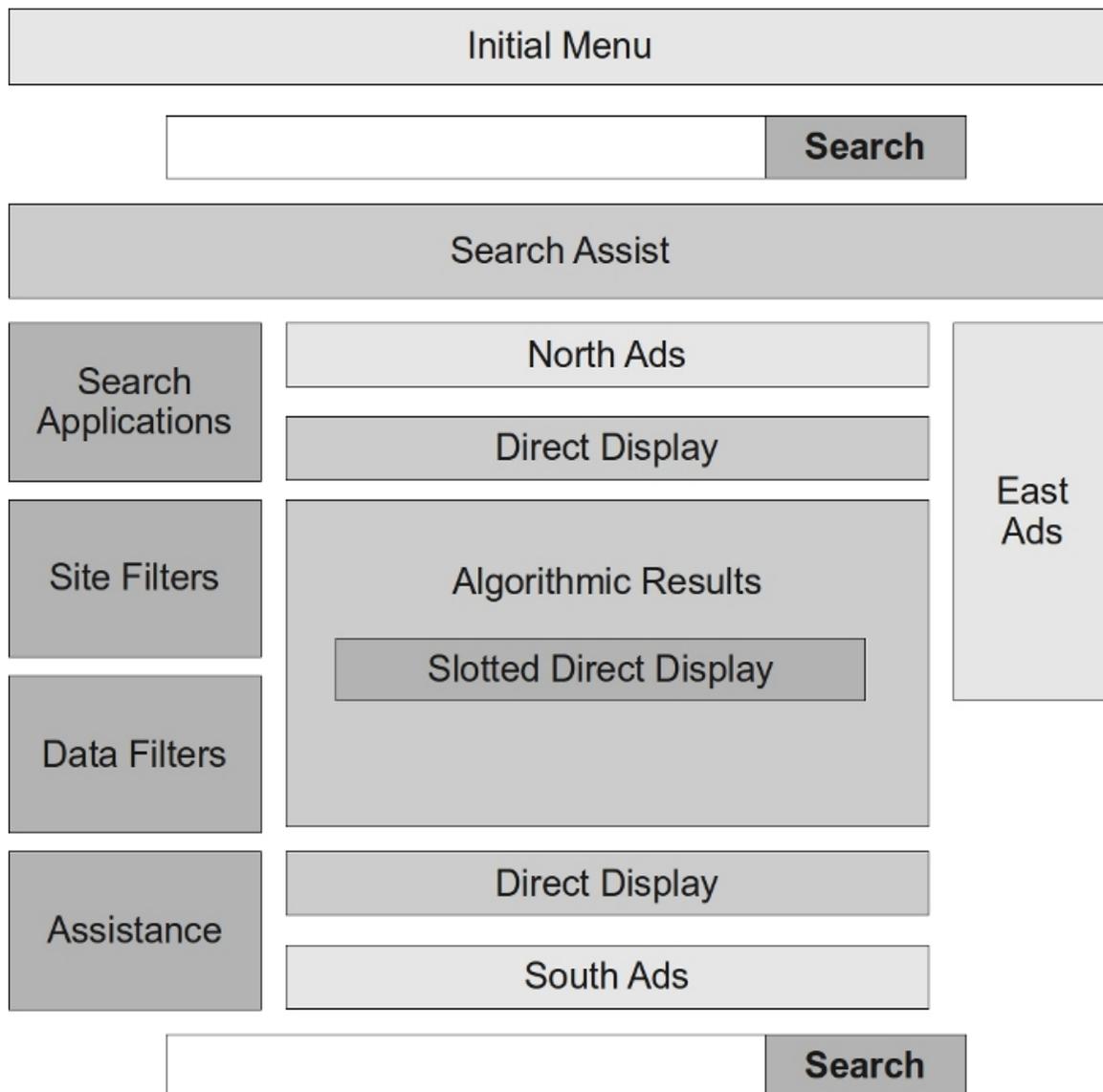


Figure 7: New-three column design of the results screen for Yahoo Search!  
 Source: (Sullivan 2009)

Sullivan comments that this method was adopted some time ago by the main search engines:

*the Ask 3D design started this trend in June 2007 (and dropped that about a year later). Google rolled out a 3-column look in early May when its “Show Option” feature is selected (see also here). Bing rolled out a 3 column look in late May (Sullivan 2009).*

It seems clear that all engines are following a similar tendency on this point, although the distribution of content in the columns is different in all engines. Focusing on the different areas in the results window that appears in the previous image, we have constructed a table to present a summary of how the three main search engines use them:

| Yahoo! Search   | Google  | Bing  |
|---|---|---|
| <b>Initial menus</b>  |   |   |
| Access to other applications (mail) or search systems (images, videos, maps, news, etc.). | Access to other applications (mail) or search systems (images, videos, maps, news, etc.). | Access to other applications (mail) or search systems (images, videos, maps, news, etc.). |
| <b>Search assist (with Top Search box)</b>  |   |   |
| Opens out below the search  | Opens out below the search  | Opens out below the search  |

|   |   |   |
|---|---|---|
| box. Shows terms that coincide with characters typed into the search alphabetically. Also shows phrases where those characters appear although they are not at the beginning ('disn' displays 'Playhouse Disney'). Allows exploration of linked terms ('Bono' with 'U2'). Also try. ... able for syntax errors. | box. Shows terms that coincide with characters typed into the search alphabetically. <b>Did you mean</b> ... able.  | box. Shows terms that coincide with characters typed into the search alphabetically (works in a similar way to the 'Did you mean' by Google).       |
| <b>Search applications</b>  |   |   |
| <b>Notes on line and Searchscan.</b> The second allows users to protect their search.   | Allows redirection of the search to other search engines from Google: Images, Videos, News, Blogs, Updates, Books and Forums  | Suggests related searches (for example, 'Videos of U2' for the query U2)  |
| <b>Site Filters</b>   |   |   |
| <b>Safemonkey:</b> application that allows access to results compiled in Yahoo! From Websites specialising in the query subject (if the query is about music it would take us to myspace or youtube for example).   | Shopping sites.   |   |
| <b>Data filters</b>   |   |   |
| Links specialist data depending on the subject query (if it is music it links us to song lyrics, purchasing concert tickets, etc.)  | Related searches. <b>Wonder wheel</b> (graphical representation of terms relationships). <b>Timelime</b> (statistics on the publication of news). <b>Date of publication</b> (latest 24 hrs, etc)                                   | Related searches.   |
| <b>Assistance</b>   |   |   |
| Directed links to search assist suggestions.  | Translated searches. Get images of the pages.   | Search history.   |
| <b>North Ads (advertising)</b>  |   |   |
| Shows ads below the search box.   | Shows ads below the search box.   | Shows ads below the search box.   |
| <b>Direct display</b>   |   |   |
| Shows a set of images, videos or mp3 links related to the subject of the search.  | Show the results ordered by the ranking algorithm. You can find set of images or videos available at the end of the page.   | Shows images, news and schedule. We can also find the first algorithmic result included in this area.   |
| <b>Algorithmic results</b>  |   |   |
| Displays text results in order of their ranking algorithm.  | Displays all media in order of their ranking algorithm without distinguishing the nature of the information resource (text, photo, image, etc.). We can find resources grouped by their source (blogs, videos, maps, images, etc.). | Displays text results in order of their ranking algorithm but classified by their information goal: common pages, news, speeches, biographies, etc. |
| <b>South Ads</b>  |   |   |
|   | This area is usually ready for related searches.  | This are is usually ready for related searches.   |
| <b>East Ads</b>   |   |   |
| Advertising   | Advertising   | Advertising   |
| <b>Down search box</b>  |   |   |

|  |           |      |
|--|-----------|------|
|  | Able Able | Able |
|--|-----------|------|

**Table 1: Analysis of the current search interface at Yahoo Search, Google and Bing.**  
Source: author's

Without a doubt, Yahoo! Search stands out significantly, at least visually, from the other search engines. This clearly contributes to improving user experience, although its advances are not only restricted to formal aspects, but also to the technology itself. The search assist option on its engine is much more user-friendly while Google makes up for this lack with its 'Did you mean ...' tool it introduced some time ago.

The SearchMonkey tool on Yahoo! Search integrates structured information in the results in a more dynamic and interactive way than Google or Bing, engines that only introduce static links to other data sources in the results, meaning one would have to complete another search to follow them, whilst Yahoo! Search has already done that and introduces the document in the results. The search assist tools are varied and different (direct access to suggestions, translated searches or search history); none stand out from the rest.

With regard to results ranking, Google has always used ranking algorithms (Pagerank at first, Trustrank currently) when ordering documents in the results screen. Following its traditional routine, this engine currently presents the results documents (whatever the type of data) ordered according to ranking, that is, it does not present videos and then news (like Bing); rather, documents appear mixed according to relevance. The final result has perhaps less impact for the user in terms of usability, but it is still highly reliable in terms of search efficiency.

In the presentation of the new search features, Yahoo! Search schematically displays the key points of its new strategy designed 'to make search more personally relevant' ([Yahoo 2009b](#)) using the perspective of a unified design to allow for speedier access to valuable search instruments, such as:

- **Intelligent search results:** allows user to explore results from key sites and narrow results using different types of SearchMonkey structured data. Over the past few months, even more enhanced results for product, local, entertainment, reference, social, and tech sites have been displayed automatically. With the new search page design, Yahoo! made it easier to see these rich results from an increasing number of sites.
- **Feature-rich experience** - provides quick access to search features that make people's online lives safer and easier, including Search Scan/SafeSearch (which helps protect you from viruses, spyware and spam in your searching) and Search Pad.
- **Search assist expansion:** powerful query assistance is still available directly below the search box and it's also incorporated into the left-hand column for quick access lower on the page. We can use this column to easily explore and discover concepts related to our query.

The *intelligent results search* is the aforementioned feature which allows us to directly access results from prestigious Websites (such as Wikipedia) that make up part of the documents compiled by the engine. This is possible thanks to the SearchMonkey technology which enhances 'Yahoo! search results with additional data and structure, such as images, key/value pairs, and additional links. Yahoo! Search users can add SearchMonkey applications to their profile on an opt-in basis' ([Yahoo 2008](#)). The following figure shows this utility:



Figure 8: SearchMonkey performance example.

Source: [http://developer.yahoo.com/searchmonkey/smguide/sm\\_overview.html](http://developer.yahoo.com/searchmonkey/smguide/sm_overview.html)

SearchMonkey is not so much a tool as a development framework for small PHP applications that interact with the search engine and freely processes designers' results. The structure of the data used is known as [DataRSS](#) and includes the following main elements:

- **adjunct:** fundamental unit of the DataRSS structure, it contains the rest of the retrieved elements, including item, type and meta.
- **item:** represents information about retrieved objects. These objects can have a specific URI or, as with RDF, could be blank nodes associated with an item. This type of unit can contain other item, type and meta elements.
- **type:** shows the type of data or resource associated with a particular item or meta element.
- **meta:** associates a literal to specific value of a particular DataRSS element. This mechanism assigns metadata by mapping with data services offered by SearchMonkey or specified by the developers themselves. This metadata can use any presentation model such as Dublin Core, FOAF (Friend of a friend), RDF (Resource Description Framework), etc.

This service is flexible enough to be able to indicate the relationship between two resources. Therefore, it is possible to show that page references to Wikipedia retrieved during a search are *definitions*, and can be given a corresponding Universal Resource Indicator with a library of images of our Web site. It is also possible to define a relationship known as *related documents* which would allow for the inclusion of a list of links to PDF files whose references are associated to the said concept/term/expression defined by the said Wikipedia page (shown in Image 7).

We are therefore dealing with a development framework and Web resource organisation designed to offer a series of services that will allow for all types of search tools to be freely designed. From a developer's point of view, the advantages are clear, as they can create search tools adapted to a specific site with total freedom regarding objects displayed, in terms of structure, content and interrelation with other types of resources.

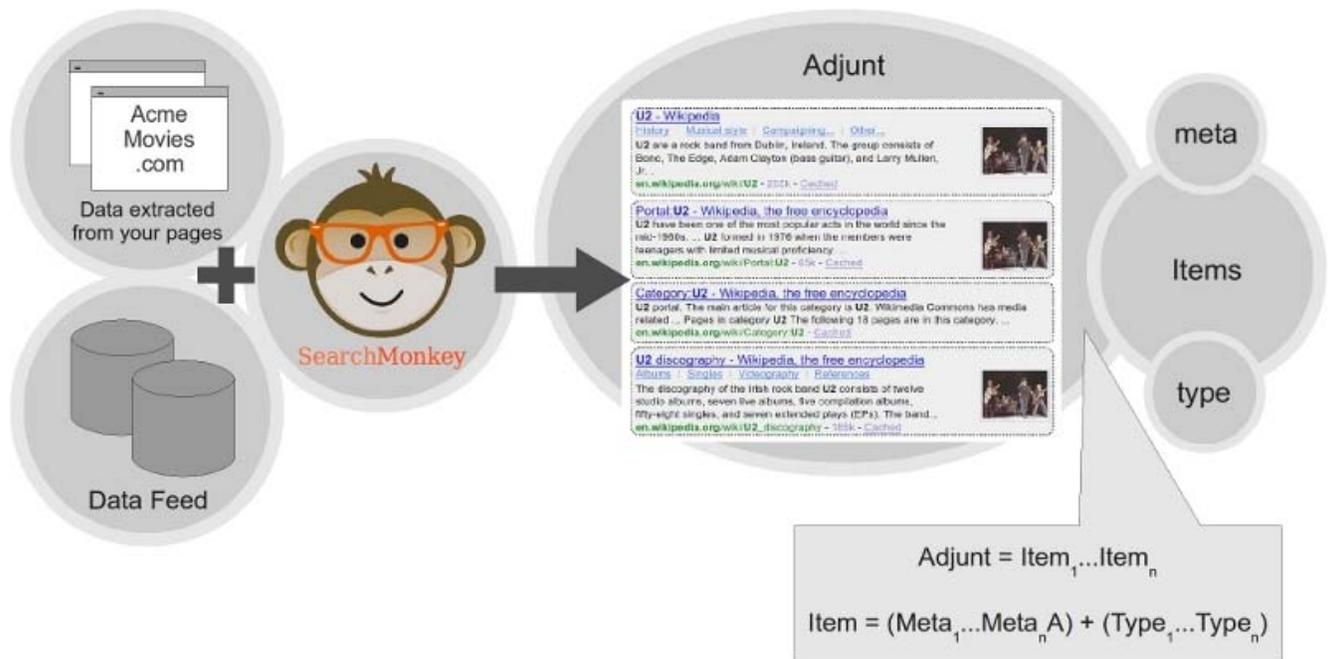


Figure 9: How SearchMonkey works, general idea.

## The future

It is important to remember that this technology has been created to provide a service to the end user. To the advantages that the integration of a search engine creates for a Web site, we must also add greater efficiency in the information retrieval process because of the application of technology related to the semantic Web. This means that results will be, without a doubt, more pertinent and exact thanks to the capacity to identify certain data that are semantic and, therefore, more representative, in order to meet users' information needs. It is also important to bear in mind that this technology enables the use of semantic networks and concept schemes which are extremely useful for expanding and enclosing queries or redirecting user searches through subject suggestions or associated term queries.

For some time now, third parties have been able to reuse the main search engines. Google Labs was the first (quite a few years ago) to offer an API to integrate its search engine within any Website and then other engines followed suit (Rodriguez 2008). This initiative has improved user experience considerably in terms of searching for information on countless sites: the integration of a familiar search engine facilitates the search and the feedback and query expansion dynamic.

This method has many obvious advantages; however, it is limited by the margin to which the results are personalised. At times, the application of CSS sheets to obtain a visual style adapted to the image of the Website that integrates the search engine may not be enough. For this reason, there are countless small libraries or scripts written in languages such as PHP or Java designed for parsing the search engine results. These are mini-applications that allow retrieved results to be modified during the search.

Thus, it is possible to select a specific group of data types to include on a results page (title, summary, type of resource) alter the order of references retrieved, and design search tools combining general search engines with queries in data sources on the Website itself. In this way, the searches will be more efficient and satisfactory, given that the visualised data can be adapted more specifically to the profile of potential Website users.

In the same way simultaneous queries can be made both on the Web and using corporate data (intranet) in one operation. Therefore, user experience of a Website will be more satisfactory because they are using a more familiar search mechanism (like a search engine) even for their

own data resources, which simplifies and reduces the number of user actions needed to retrieve information. It is even possible to adapt the information from the query to the structure of the information architecture on the Website itself, which increases the perception of homogeneity and allows the user to negotiate the site more easily.

In spite of the usefulness of this technique, its weak point lies in the fact that it parses on a URI supplied by the search engine itself, on an HTML page. In other words, results obtained after making a query are in a format lacking any type of framework or structure with semantic information. Any change by a search engine to the HTML code on a results page means modification of the parser libraries. Searcher changes to HTML code are not made arbitrarily by system designers: often, the aim is to improve accessibility and usability in order to increase search process output from a user's point of view. Search engine designers cannot bear all third-party developed solutions in mind when processing HTML code. The solution is to include semantic information that can be automatically processed and unmistakably interpreted. Consequently, search engines should offer results including this type of information.

This could be achieved by implementing Web service search engines that return results using XML or XML(S) together with a representation model adequate for query results in a search engine like [Opensearch RSS](#), [ATOM](#) or even the [Resource Description Framework](#) (RDF). Clearly, this solution means publishing the corresponding XML Schema or DTD that allows designers to understand the structure of the XML document generated by the Web service. Another option would be to use micro-formats or an RDF application (Embedded RDF (eRDF) or Resource Description Framework-in-attributes (RDFa)) within the HTML code itself.

It is foreseeable that we will soon witness the arrival of a new generation of search engines, whose main characteristic is third party exploitability thanks to the application of semantic Web technology. These tools would allow, amongst other possibilities, to build results visualisation formats, group results by content or to discriminate by origin of Website.

## Conclusions

The current stage at which search engines are developing is both a new paradigm and a new business model. The first generation of these tools focused on increasing coverage, optimising indexing process, speed of query response and the development of efficient, relevant algorithms. The second generation opts for the integration of multiple data sources and results personalisation in order to improve user experience, whilst also establishing the basis for totally free exploitation of search engine data by third parties; not only does this remove the limitations with regard to the amount of data to show and its structure, it also offers tools to facilitate the designers' work and data services with semantic information.

In our opinion, the development of the *open* exploitation of search engines will not stop here; the next logical step would be totally free, personalised engagement by the end user. In this third, new generation of search engines, each user would have personal accounts, with simple assistants, to build predefined queries with visualisation formats adapted to their tastes, selecting results from the whole Web, from a specific group of sites or limiting references to specific types of objects or content.

In this way, the possibilities for these new engines will go far beyond searching; they could eventually be made up by constantly updated semantic data sources with which to build pages with completely personalised content. Thus, users could define areas whose content varied depending on the results of a predefined query (something similar to what is now adding content through RSS links but much more personalised and with better data integration).

In this new environment, the user will have greater control (in a bidirectional sense) over the search engine; slowly but surely there is a substantial paradigm shift in Web information retrieval.

The user, does not use the search engine to *see what they might have* but knowing that the engine *is going to show the user* what it can offer to satisfy the user's information needs.

## About the authors

Francisco Javier Martinez Mendez is Lecturer of information technology in the Department of Information and Documentation of the University of Murcia, of which he is currently Head. A Graduate and PhD in Information and Documentation. Professor of the UNITWIN Chair in 'Information Management in Organizations', sponsored by UNESCO. He is also the author of the blog, [recuperación de información en la Web](#). He can be contacted at [javima@um.es](mailto:javima@um.es)

Juan-Antonio Pastor-Sanchez has a BA and PhD in Information Science from the University of Murcia. He has worked since 1994 as Web Manager in the University of Murcia and, currently, also Adjunct Lecturer in Digital Information Services in the Department of Information and Documentation. His research has always been related to thesauri and information technology. He can be contacted at [pastor@um.es](mailto:pastor@um.es)

José Vicente Rodríguez-Muñoz is Professor in the Information and Documentation Department, Faculty of Information Science, University of Murcia, Spain. Graduate in Chemistry, PhD in Computing Science. Guest Professor in the University of La Habana, Cuba. Co-ordinator of the UNITWIN Chair in 'Information Management in Organizations', sponsored by UNESCO. His research area, projects and publications, covers information management, information retrieval and evaluation of Web searching. Currently Dean of the Faculty of Information Science. He can be contacted at [jovi@um.es](mailto:jovi@um.es)

Rosana Lopez-Carreño has a BA and PhD in Information Science from the University of Murcia. She has worked since 1999 as Content Manager in the Regional Government of Murcia and, since 2009 and is Adjunct Lecturer of Information Services and Resources in the Department of Information and Documentation. Her research relates to information resources in Web portals and in the use of thesauri into the management of administrative archives. She can be contacted at [rosanalc@um.es](mailto:rosanalc@um.es)

Jose-Vicente Rodriguez-Caceres has a BE in Telecommunications from the Polytechnic University of Valencia and is now a PhD candidate in the Department of Information and Documentation of the University of Murcia. His doctoral research is in the field of the effectiveness of the Web information retrieval.

- Agichtein, E., Brill, E. & Dumais, S. (2006). [Improving Web search ranking by incorporating user behavior information](#). In *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval (SIGIR'06)*, August 6–11 2006, Seattle, Washington, USA. (pp. 19-26). New York, NY: ACM Press. Retrieved 12 February, 2011 from <http://research.microsoft.com/en-us/um/people/sdumais/sigir2006-fp345-ranking-agichtein.pdf> (Archived by WebCite® at <http://www.webcitation.org/5wRWoLMUd>)
- Alexa.com. (2010). [Alexa top 500 global sites](#). Retrieved 12 February, 2011 from <http://www.alexa.com/topsites/global> (Archived by WebCite® at <http://www.webcitation.org/5wRWuXUQO>)
- Baeza-Yates, R., Hurtado, C. Mendoza, M. & Dupret, G. (2005). [Modeling user search behavior](#). In *Proceedings of the Third Latin American Web Congress (LA-WEB'05) 2005*. (p. 242). Washington, DC: IEEE Computer Society. Retrieved 12 February, 2011 from [http://lahuen.dcc.uchile.cl/~gdupret/pubs/final\\_paper.pdf](http://lahuen.dcc.uchile.cl/~gdupret/pubs/final_paper.pdf) (Archived by WebCite® at <http://www.webcitation.org/5wRXAqZ8z>)
- Croft, W.B., Metzler, D. & Strohman, T. (2010). *Search engines: information retrieval in practice*.

Reading, MA: Addison-Wesley.

- Crowell, G. (2006, June 14). [Understanding searcher behavior](#). Searchenginewatch.com. Retrieved 12 February 2011 from <http://searchenginewatch.com/3613291> (Archived by WebCite® at <http://www.webcitation.org/5wRXfmex>)
- Fox, S., Karnawat, K., Mydland, M., Dumais, S. & White, T. (2005). [Evaluating implicit measures to improve Web search](#). *ACM Transactions on Information Systems*, **23**(2) 147-168. Retrieved 12 February, 2011 from <http://research.microsoft.com/en-us/um/people/sdumais/tois-p147-fox.pdf> . (Archived by WebCite® at <http://www.webcitation.org/5wRXsW2P7>)
- Gwizdka, J. & Chignell, M. (1999). [Towards information retrieval measures for evaluation of Web search engines](#). Toronto: University of Toronto, Mechanical & Industrial Engineering Department. Retrieved 12 February 2011 from [http://comminfo.rutgers.edu/~jacekg/publications/fulltext/webIR\\_eval1\\_99.pdf](http://comminfo.rutgers.edu/~jacekg/publications/fulltext/webIR_eval1_99.pdf) (Archived by WebCite® at <http://www.webcitation.org/5wRYXLRoc>)
- Hitwise (2010). [Top 20 Websites data](#). Retrieved 12 February 2011 from <http://www.hitwise.com/us/datacenter/main/dashboard-10133.html> (Archived by WebCite® at <http://www.webcitation.org/5wRYgl7mp>)
- iProspect. (2006). [iProspect Search Engine user behavior study](#). Boston, MA: iprospect.com. Retrieved 4 February 2010 from [http://www.iprospect.com/premiumPDFs/WhitePaper\\_2006\\_SearchEngineUserBehavior.pdf](http://www.iprospect.com/premiumPDFs/WhitePaper_2006_SearchEngineUserBehavior.pdf) (Archived by WebCite® at <http://www.webcitation.org/5wRa5HqC4>)
- Jansen, B. & Spink, A. (2006). How are we searching the World Wide Web? A comparison of nine search engine transaction logs. *Information Processing and Management*, **42**(1), 248-263.
- Lee, U., Liu, Z. & Cho, J. (2005). [Automatic identification of user goals in Web search](#). In *Proceedings of the 14th international conference on World Wide Web, Chiba, Japan*. (pp. 391-400). New York, NY: ACM Press. Retrieved 12 February 2011 from <http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=CA242BB0036CEEFC99E59D5DF643CFBB?doi=10.1.1.59.3810&rep=rep1&type=pdf> Archived by WebCite® at <http://www.webcitation.org/5wRaSYTZc>)
- Leighton, H.V. & Srivastava, J. (1997). [Precision among World Wide Web search services \(search engines\): Alta Vista, Excite, Hotbot, Infoseek, Lycos](#). Winona, MN: Winona State University. Retrieved 4 February 2010 from <http://www.winona.edu/library/staff/vl/Webind2/Webind2.htm> (Archived by WebCite® at <http://www.webcitation.org/5wRabAYvp>)
- Lewandowski, D. (2008). Search engine user behaviour: how can users be guided to quality content? *Information Services and Use*, **28**(3-4) 261-268.
- Martínez-Méndez, F. J. & Rodríguez-Muñoz, J.V. (2003). [Síntesis y crítica de las evaluaciones de la efectividad de los motores de búsqueda en la Web](#). [Summaries and critique of evaluations of the effectiveness of Web search engines.] *Information Research*, **8**(2), paper 148. Retrieved 12 February 2011 from <http://InformationR.net/ir/8-2/paper148.html> (Archived by WebCite® at <http://www.webcitation.org/5wRazA2Mc>)
- Oppenheim, C., Morris, A., McKnight, C. & Lowley, S. (2000). The evaluation of WWW search engines. *Journal of Documentation*, **56**(2), 190-211
- Paczkowski, J. (2009, August 29). [Yahoo: Testing a More Google-Like Search Experience](#). *Digital Daily* 25 August 2009. Retrieved 4 February 2010 from <http://digitaldaily.allthingsd.com/20090825/yahoo-testing-a-more-google-like-search-experience/> (Archived by WebCite® at <http://www.webcitation.org/5wRbCApG4>)
- Rodríguez J.V. Sr. and Martínez, F.J. & Rodríguez J.V. (2008). [Analysis of the similarity within the response of the Web search engines: an approach since users' trends](#). *Information Research*, **13**(4) paper 382. Retrieved 12 February 2011 from <http://InformationR.net/ir/13-4/paper382.html> (Archived by WebCite® at <http://www.webcitation.org/5wRbIdAp0>)

Sullivan, D. (2007, May 16). [Google 2.0: Google universal search](http://searchengineland.com/google-20-google-universal-search-11232). *Search engine land.com*. Retrieved 12 February, 2011 from <http://searchengineland.com/google-20-google-universal-search-11232> (Archived by WebCite® at <http://www.webcitation.org/5wRcIf34Q>)

- Sullivan, D. (2009). [Yahoo's new search clothes - but will it help? \(Probably not\)](http://searchengineland.com/yahoos-new-search-clothes-but-will-it-help-probably-not-24369). *Search engine land* 24 August 2009. Retrieved on 12 February, 2011 from <http://searchengineland.com/yahoos-new-search-clothes-but-will-it-help-probably-not-24369> (Archived by WebCite® at <http://www.webcitation.org/5wRcU2FEq>)
- Wolverton, T. (2009, July 15). [Promising new trends in Web search engines](http://www.physorg.com/news166903914.html). *Physorg.com*. Retrieved 12 February, 2011 from <http://www.physorg.com/news166903914.html> (Archived by WebCite® at <http://www.webcitation.org/5wRcdTnAW>)
- Yahoo! (2008). [Searchmonkey guide: a manual for SearchMonkey developers and publishers](http://developer.yahoo.com/searchmonkey/smguid/index.html). Yahoo! Inc. Retrieved 12 February 2011 from <http://developer.yahoo.com/searchmonkey/smguid/index.html> (Archived by WebCite® at <http://www.webcitation.org/5wRciehmF>)
- Yahoo! (2009a, August 24). [Testing a new Yahoo! search experience](http://www.ysearchblog.com/2009/08/24/testing-a-new-yahoo-search-experience/). *Yahoo! Search Blog*. Retrieved 12 February 2011 from <http://www.ysearchblog.com/2009/08/24/testing-a-new-yahoo-search-experience/> (Archived by WebCite® at <http://www.webcitation.org/5wRfltnNj>)
- Yahoo! (2009b, September 22). [Welcome to the New Yahoo! Search](http://www.ysearchblog.com/2009/09/22/welcome-to-the-new-yahoo-search/). *Yahoo! Search Blog*. Retrieved 12 February 2011 from <http://www.ysearchblog.com/2009/09/22/welcome-to-the-new-yahoo-search/> (Archived by WebCite® at <http://www.webcitation.org/5wRftL0vp>)

Martínez, F.J., Pastor, J.A., Rodríguez, J.V., López, R. and Rodríguez Jr., J.V. (2011). "New architectures for presenting search results based on Web search engines users experience" *Information Research*, 16(1) paper 461. [Available at <http://InformationR.net/ir/16-1/paper461.html>]

[Find other papers on this subject](#)

Check for citations, [using Google Scholar](#)

■ [Bookmark This Page](#)



© the authors 2011.  
Last updated: 12 February, 2011



---

[Contents](#) | [Author index](#) | [Subject index](#) | [Search](#) | [Home](#)

---