
Features in the Web Search Interface: How Effective Are They ?

Deepak P

Sandeep Parameswaran

Abstract

With web search getting to be more and more popular, and predictions that they would be even more popular in the coming years, given the exorbitant growth of the web in recent years, search engines, in their quest to be branded the best, have regularly been providing additional features. The users of web search interfaces are typically diverse and have wide ranging interests. This is in contrast to other application interfaces which cater to a specialized group of people. In this study, we examine the influence the interface manifestation of such features has, in their usability and effectiveness.

Keywords : Web Search, Web Search Interface, Search Engines.

0. Introduction

It has been acknowledged [1] that web application interfaces are fundamentally different from common GUI based applications and hence have to be designed taking a lot of additional factors in mind. Further, web searches are directed towards fulfillment of widely varying requirements [2] by users of widely varying backgrounds that using assumptions on the awareness and the knowledge level of the typical user to optimize the interface may often turn out to be harmful. With the competition among search engines for popularity escalating by the minute, more and more search engines provide value-added features [3] that aid the web user. It has been opined that feature addition to web search (interface) is hard [4], given that added features are often seen by the hasty user as more of a nuisance than an aid. In this paper, we propose to evaluate how effective the add-on features in web search engines are (in usability), by means of an objective evaluation and a survey on a pool of users who use the Google [5] search engine, which is, by far the most popular search engine.

Section 2 reviews the more relevant works in web search interfaces. Section 3 is a study on the features that are present in contemporary search engines, focusing almost entirely on their interface manifestation and usability. It further enumerates the way the features can be incorporated in the search engine interface, how they could be placed to aid the user in the search, and a list of the more common features present in contemporary search engines. Section 4 presents a survey conducted to evaluate the usability and effectiveness of the features in the web search interface. Section 5 lists the conclusions derived from the study and generalizations from the opinions gathered in the survey. The list of references forms Section 6.

1. Web Search Interfaces

Web search interfaces have undergone a drastic change through the years. The initial text-box only interface, has almost completely returned back after being replaced unsuccessfully, temporarily and partly by a graphical interface [4]. In a web application like the search engine, the user typically does not devote his entire attention at a single application for long periods of time. Rather, he switches back and forth between applications and web sites. Thus, response time and quality of results are very critical issues for a search engine, as they decide as to whether the user would return to it. Although not as important as the above two factors, the interface design is also a major issue for the search engine. A search engine that provides a load of features, must ensure that they do not clutter up the screen real estate, making the search engine difficult to use. One among the major guiding principles is that the

single text box, the place where the user has to enter the search query should be center of attention in a web search interface. Web application interface designers need to understand, accommodate and support the extensive freedom that a user enjoys while using a web application [1]. Sometimes, we can force users through set paths and prevent people from linking to certain pages, but sites that do so feel harsh and dominating.

Web search engine users are casual users, who due to some reason or intention, approach the application to get a URL to a web page that they want to visit. Thus, the web search engine is a means to the end. The importance of the search engine interface has been acknowledged even in very early days [6]. A study conducted way back in 1995 [6] evaluating the interface of the then search engines, branded OpenText as the best citing its “powerful search interface”. Further, they opined that WebCrawler had the easiest-to-use interface for novices. Another study [7] concludes that the best search engine would be the one that can produce “accurate results from easy-to-use interfaces”. It has been postulated [8] that “interfaces and documentation” are the two most important factors that affect users efforts in learning and using a search engine.

The usability of computer systems are often classified based on five components [9]: learnability, efficiency, memorability, errors (error rate) and satisfaction. Of these, the most important factors that concern search engine interfaces are learnability and memorability. Web search interfaces have to be instantly learnable as the users are mostly casual users. Interfaces that are easiest to learn usually turn out to be the most memorable as well. The same study [10] concludes that among the specialized interfaces for web search engines, the “connector menu interface”, which consists of a series of text boxes with the option of connecting the words by boolean operators is a very usable and preferred one.

The main usability questions to ask when reviewing a search vendor or evaluating the user experience of your own search solution are, “Can you customize the results” and “Can you simplify the search interface”. These two things matter for ease of use and value to the user. Search results are as important as the interface [11]. These arguments, postulated keeping site search in mind, also apply equally well to web search.

A review classifies search interfaces into 4 classes [12]. A simple search button is regarded as a “passive” interface. A “standard” interface has an edit box, a search button and a link to a more specialized kind of search. Surfacing interfaces, more prevalent in site search, includes a combo box, which contains the different filters or categories for search. A qualifying interface, which also includes the combo box contains non-category filters like time or location filters for narrowing down the search. Many contemporary search engines use the language filter to direct searches towards documents in a specified language.

2. Features in Search Engines

As already opined, the ever-escalating competition among search engines to attract user attention and to draw more audience, has led to search engines providing much more [13] than a single text edit box in their interfaces. Most of these features are aimed at refining the search to provide the user with better relevant results. The web search is typically described as a process which includes stages such as query formulation, deployment, a systematic review of the search results followed by possible query reformulation and re-deployment (perhaps, multiple times) [14]. The usual features that search engines provide are those that aid searching and speed-up convergence to the relevant results. But, of late, search engines have ventured into providing features, which are unrelated to the searching process (e.g., the calculator, web page translator and stock quote features of Google [15]).

2.1 How do search engines provide additional features on the interface ?

Incorporating new features to an application interface is often a tough and challenging task. We might have to undo certain careful decisions taken earlier regarding the structure of the interface and perhaps, might even have to redo the entire interface. As already mentioned, add-on features to the search engine can either be totally unrelated to the search process or can aid the search process. Whatever be the type, there are commonly two options to incorporate it in the interface

- ✍ Include it as a link, explicitly in the interface
- ✍ Introduce a new operator for the purpose and include documentation for the same in the 'Help' page

The first option is often unwelcome as it uses up screen real estate. Such incorporation of features often aid in cluttering up the page rather than aiding the user [4]. Yet certain important features such as the introduction of the service in a vernacular language are often presented this way [16]. The second option is perhaps the most widely used option. The most widely used illustration of this option is the filetype filter, whereby a user who requires results of type pdf only can append a "filetype:pdf" string to the search query. But users are often unaware of such features, as the presence of such a feature is not evident on the main search page. Flash alerts on the main page just after the introduction of a feature would however, aid in spreading the news to the regular users.

2.2 Where (in the search process) do search engines provide additional features ?

This section is concerned only with features that aid the searching process. Features may aid the searching process during its various phases

- ✍ Query Formulation/Query Creation: The surfacing and qualifying interfaces for search engines [12] as discussed in Section 2 aid in imposing filters before the search query is actually presented to the system
- ✍ Search Result Review/Selection of a Search Result: The search results are usually taken to be the set of URLs returned by the system. But seldom do search engines provide such plain results. Almost all the search engines of today provide short descriptions of each search result along with the search result and some even cluster search results [17].
- ✍ Refinement of the Search: The help in this phase is usually provided by means of queries like, "do you want to see the omitted results", "similar pages" etc [5].
- ✍ Reformulation of the query: The help here is provided by means of features such as the spell checker of Google [5] (the "did you mean" feature)
- ✍ Examining Selected Item: Help can also be provided when the user is examining a result. Google provides "search query highlighting" in the cached copies of results
- ✍ Other situations: Google automatically provides an option to view the (old) cached copy of a page in cases when the result page is no longer alive. Such a condition is in fact, a failure on the part of the search engine to provide recent results.

The additional features can be provided on

- ✍ The main search interface OR
- ✍ Results pages

Evidently, only features that aid query creation or formulation can appear in the main search interface. They are often duplicated in the results pages as well, given that the phase of reformulation of the query has to be given at least as much support as the initial query creation.

2.3 Features in the interfaces of contemporary search engines

Here we would like to enumerate a list of the more common features in contemporary search engines and examine how they are represented in the interface. We restrict the investigation to features that aid the search process directly. This is, by now means, an exhaustive list, but ventures to provide a taste of the common functionalities of the current search engines.

2.3.1 Features that aid query formulation

- ✍ Surfacing interfaces: A combo box is provided to narrow down the search to a particular category. Yahoo [18] provides a combo box by the side of the search box with options such as searching in yellow pages, images, products etc. Google [5] also provides a category specific narrow down option to categories such as images, groups and news, but by means of specific links to separate sites. Other filters are provided in the advanced search interface as well.
- ✍ Qualifying interfaces: These aid narrow down by time and location filters. A lot of contemporary search engines have these filters in their advanced query interfaces. e.g., site search, inward link search etc.
- ✍ Other filters: Filters can be provided by means of specialized operators as well. There are often operators to restrict the search to results of a particular type (filetype:), to documents that have the query in the title (intitle:) or URL (inurl:) etc. These are advantageous from the interface designers point of view in that they do not consume real estate in the screen (refer Section 3.1)
- ✍ Google has a feature [15] (“I’m feeling lucky”) which directs the user to a single result page which presents itself as a separate button in the main interface.

2.3.2 Features that aid (or complement) search result review/selection

- ✍ Descriptions: Almost all search engines provide descriptions for each search result, which may either be derived from the pages themselves or from pages with inward links to the page in question.
- ✍ Context Clustering: Vivisimo [17] provides result clustering on the fly, so that the user is provided with groups of search results. The visual appeal of such clustered results can often be inviting. Algorithms for such context clustering are available in literature [19].
- ✍ Resources: Teoma [20] provides a collection of relevant links from “experts and enthusiasts” along with the search results. It is provided in a separate column, and hence does not clutter up the results page too much.
- ✍ News headlines: Of late, Google [5] provides any relevant news headlines along with the search results. Although this does not aid the search process, it does complement it. The news headlines take up less space compared to the results themselves, given that they do not have descriptions.

2.3.3 Features that aid search refinement

- ✍ Similar pages: Google [5] provides a similar pages link with almost every result page, to refine the search to pages that are similar to it (represented by an explicit link with each search result)

-
- ✍ See Omitted Results: Google automatically removes similar pages from a site to avoid redundancy in search results. But it provides an option to view the omitted pages in order to refine the search towards the direction of a specific site (implicitly).
 - ✍ Context Clustering: This also acts as an implicit search refinement towards different clusters of results.
 - ✍ Refine: The refine feature of Teoma is best discussed under the next section.

2.3.4 Features that aid query reformulation

- ✍ Refine: Teoma [20] provides a listing of possible query refinements or reformulations which could be used to narrow down the search towards a specific direction.
- ✍ “Did you mean”: The did you mean feature of Google lists possible reformulations of the query to alleviate spelling mistakes and typo errors.

2.3.5 Miscellaneous Features

- ✍ Cached copy: As the current search methodology searches on indexed collections rather than the current state of the web, the results can be stale. Such imperfections are usually masked by providing a cached copy.
- ✍ Search Query Highlighting: A further advantage of providing the cached copy is that the search engine can provide features such as search query highlighting in the cached copy of the results (as they are held entirely by the search engine)
- ✍ Web Page translation: The user who requires results in a particular language can be satisfied by means of translating the results originally in different languages, to the language in question. Yet, this is often regarded as something totally unrelated to the search process.

As can be seen from the above, most of the additional features introduce one or more links in the interface. Thus the usability of such features, given that screen real estate is hot property in search engine interfaces is a matter of interest. The following sections describe a survey on the usability of additional features in a search engine.

3. A Survey on the Usability of Additional Features in Web Search Engines

3.1 Choice of the search engine

We, in our attempt to restrict ourselves to questions such as “how do you find that feature”, chose to restrict ourselves to a particular popular search engine. The obvious choice was Google, which is by far the most popular search engine, at least for the less tech-savvy population in developing countries like India. Google [5] is reported to be searching 4k million web pages. Further, it provides vernacular interfaces which makes it popular in a multilingual population like India.

3.2 The audience of the survey

As the evaluation of the usability and effectiveness of addition of features in the web search interface of Google was the subject of the survey, we chose people who use the Google search engine frequently. We took care to include as much variety in the audience. People with ages ranging from 15-50 were chosen for the purpose. With an almost equal balance between men and women, the population was chosen

from different walks of life, ranging from non-tech-savvy people like shopkeepers who are just casual computer users to academicians in the discipline of computer science. Each member of the population was held at an equal weighting.

3.3 The features to be evaluated

The following six features of the Google search engine were chosen for evaluation on the survey. The trade-off was between including features of as much variety, at the same time, keeping the questionnaire to a reasonable size.

- ✍ Cached Links: A per-result link which leads to the cached copy of the result
- ✍ File Types: A pre-search filter which enables searching for only specific types of files. This feature, which can be deployed using the filetype: operator does not manifest itself as a link
- ✍ "I am feeling lucky": A prominent button in the main search interface, which automatically leads the user to the most prominent result of the search query, without having to go through the search engine results pages
- ✍ Similar Pages: A per-result link which refines the search to the pages similar to the specific result
- ✍ Site Search: A pre-search filter like the implemented by the operator site: which directs the search to pages on a specific site
- ✍ Boolean operators: Boolean operators can be used between search query words either by explicit mention in the search query or by using the fields in the advanced search option.

The options given against each feature measures how effectively the user has deployed the feature. It also tries to measure why it has not been deployed, if at all the user has not used the feature. A remarks column per feature allows the user to enter data likely to aid the survey.

3.4 The Questionnaire

The table representing the questionnaire is given as below

| Feature | Alternatives | Tick | Remarks |
|--------------|---|------|---------|
| CACHED LINKS | Frequent User Infrequent User Never use it, infact never felt the need Used to use it Who needs to see an old version What is caching ✍ | | |
| FILE TYPES | Frequent User Infrequent User Used to use it I need only web pages Never knew about the feature Never knew about the feature, used to try alternative methods Cant understand the meaning | | |

| | | | |
|--------------------|---|--|--|
| I AM FEELING LUCKY | Frequent User Infrequent User Used to use it Never knew about the feature Cant understand the meaning | | |
| SIMILAR PAGES | Frequent User Infrequent User Used to use it Never knew about the feature Cant understand its meaning | | |
| SITE SEARCH | Frequent User Infrequent User Used to use it Never knew about the feature Never knew about the feature, used to try alternative methods Cant understand the meaning | | |
| BOOLEAN OPS | Frequent User Infrequent User Used to use it Never knew about the feature Cant understand the meaning | | |

3.5 Results of the survey

The table representing the results of the survey with each the percentage of people who chose each option recorded against each option is given as below.

| Feature | Alternatives | Responses % |
|--------------------|--|-------------|
| CACHED LINKS | Frequent User | 10 |
| | Infrequent User | 13 |
| | Never use it, infact never felt the need | 28 |
| | Used to use it | 00 |
| | Who needs to see an old version | 20 |
| | What is caching L | 29 |
| FILE TYPES | Frequent User | 26 |
| | Infrequent User | 04 |
| | Used to use it | 00 |
| | I need only web pages | 27 |
| | Never knew about the feature | 11 |
| | Never knew about the feature, used to try alternative methods | 27 |
| | Cant understand the meaning | 05 |
| I AM FEELING LUCKY | Frequent User | 07 |
| | Infrequent User | 09 |
| | Used to use it | 00 |
| | Never knew about the feature | 05 |
| | Cant understand the meaning | 79 |

| | | |
|---------------|---|----|
| SIMILAR PAGES | Frequent User | 40 |
| | Infrequent User | 25 |
| | Used to use it | 10 |
| | Never knew about the feature | 15 |
| | Cant understand its meaning | 10 |
| SITE SEARCH | Frequent User | 18 |
| | Infrequent User | 10 |
| | Used to use it | 02 |
| | Never knew about the feature | 40 |
| | Never knew about the feature, used to try alternative methods | 21 |
| | Cant understand the meaning | 09 |
| BOOLEAN OPS | Frequent User | 74 |
| | Infrequent User | 12 |
| | Used to use it | 00 |
| | Never knew about the feature | 05 |
| | Cant understand the meaning | 09 |

3.6 Brief Per-Feature observations

3.6.1 Cached links:

The feature manifests itself as a link with anchor text “cached” with each search result. 30% of the respondents did not understand the feature. This might just be a special case, as understandability of technical jargons cannot be expected in an Indian population. It can be seen that only 10% of the population actually use the feature frequently. This is a feature whose presence is enforced by the imperfections of the current search methodology, whereby the search engine cannot guarantee that every link provided by the search engine would be alive. Thus this feature is not effective and usable, given that many people do not understand it, and further, people who understand what the link stands for, do not seem to comprehend the necessity of such a link perhaps due to unawareness of the current web search methodology.

3.6.2 File Types :

As discussed in section 3.2, absence of an explicit link to represent the feature does hinder usability. Users are unaware of the feature unless they see the explicit representation. Many opined in the remarks column that they in fact had to narrow down the search to files of a specific type more than once, but that they resorted to appending the type of the file in the search query (e.g., “anatomy pdf” as a query). This shows that absence of an explicit filter in the main interface does hinder usability a lot. This feature, however has an explicit manifestation in the advanced search interface.

3.6.3 “I am feeling lucky” :

Although this is an explicit button in the main search interface, about 80% of the users could not understand the meaning. Many opined that the name of the feature was too cryptic. Many frequent users thought that the name ought to be changed to something like “best result”. But users who once read about the feature have resorted to using it frequently.

3.6.4 Similar Pages :

A large volume of the users in fact do use the similar pages feature. This is a link that appears with each result with the same font and prominence as “cached”. A simple expressive phrase like “similar pages” does the trick here.

3.6.5 Site Search :

This is implemented by means of an operator. Akin to the filetype feature, people who came to know of the feature through the survey immediately acknowledged the usefulness of the feature. Further, people confessed having used a lot of other alternatives (just as in Section 4.6.2) to achieve the same result. This feature has an explicit manifestation in the advanced search interface.

3.6.6 Boolean Operators :

This feature, is a much celebrated feature in search engines. It is extensively used, that it has spread without the need for an explicit manifestation in the main page (it does have a representation in the advanced search interface). Boolean operators are widely regarded as a “cannot-do-without” feature.

As a general observation, very few people ticked the “used to use it” option. It can be concluded that the features were in fact very relevant and useful that a person who once used it did use it whenever demand arose. It is a further indication that the learn-ability of the features is good. Further, many opined that a link to the features page [15] from the main search page would be of extreme help in understanding the features. Although Boolean operators, file types and site search have explicit manifestations in the advanced search interface, the others are not as popular as the boolean operators. Boolean operators occur as the first entries in the advanced search interface. This points to the need to ensure that a lot of features should not clutter up the advanced search interface as well. Each one of the search interface is to be designed so as to include only the absolutely necessary features. Including more would hinder the usability of all the features. Further, we can also conclude that understandable descriptions are more inviting than eye-catching ones like “I am feeling lucky”.

4. Conclusion

The generalizable conclusions, mostly based on the survey and partly based on the evaluation of the interfaces, show that web search users prefer the easiest interface to the most powerful interface. Learnability and memorability form two very important criteria for web search interfaces, primarily due to the fact that most of the web search users are casual users who come to the search engine interface with widely varying intentions [2]. Further, from the remarks of the people who attended the survey, it can be generally opined that most users who come to the search engine, prefer to actually type in the search query as soon as possible (and refine the results later on, if needed) to sitting down and using the pre-query formulation features to generate a query more probable of getting them accurate results. Further generalizations are presented as below.

Initial query formulation features aid more in cluttering up the search interface. Reducing the main search interface to include only very useful and essential features (or filters) would be a good choice. Advanced search interface is typically not used extensively. Giving a useful feature an explicit and expressive manifestation in the advanced search interface in addition to an operator based access, would perhaps be the best choice. People seldom like going back to the advanced search interface each time, and prefer to use the operator once they have understood the feature. Current representations of useful

features are in-fact very much learnable. This is testified by the observation that people who have used the feature continue to use it. Understandable short descriptions are preferable to eye-catchy ones to represent features. The survey remarks showed that people generally preferred the features that aid refinement of search and reformulation of the query. If the validity of such a hypothesis can be proved, more and more research has to go into finding how initial query formulation features like time and location filter can be adapted to useful features to aid subsequent phases based on information gathered from the search process.

5. References

1. Jakob Nielsen, "The difference between Web design and GUI design", May 1997, <http://www.useit.com/alertbox/9705a.html>
2. Rose & Levinson, "Understanding user goals in web search", Proceedings of the 13th World Wide Web Conference, NY, pp.13-20
3. "Search Engine Features Chart", April 2004, <http://www.searchengineshowdown.com/features/>
4. Krishna Bharat & Bay-Wei Chang, "Web Search Engines: Algorithms and User Interfaces", CHI 2003 Tutorial, <http://www.chi2003.org/docs/t20.pdf>
5. Google Search Engine, <http://www.google.com/>
6. Courtis et. al., "Cool Tools for searching the web: A Performance Evaluation", Online, 19(6), November 1995, pp. 14-32
7. Scoville, Richard, "Find it on the Net!", PC World, January 1996
8. Chu and Rosenthal, "Search engines for the world wide web: A comparative study and evaluation methodology", ASIS 1996 Annual Conference Proceedings, October 1996
9. Nielsen, "Usability Engineering", Cambridge MA, 1993, Academic Press
10. Peterson, Michael, "Designing World Wide Web search engine interfaces that Aid users in implementing boolean modified search queries", Term Paper for a course on Human-Computer Interaction, Indiana University at Bloomington, <http://www.monroe.lib.in.us/~bpmchi/scholarship/peterson/peterson.html>
11. Frank, "Site search can be flattened by usability", http://experiencedynamics.blogs.com/site_search_usability/2004/01/the_power_of_si.html
12. "Search interfaces", Article on bobulate.com, http://www.bobulate.com/popups_full/search_p1.html
13. Notess, "Search Engine Features Chart", April 2004, <http://www.searchengineshowdown.com/features/>
14. Quesenbery, Whitney, "Designing usable search", Article on WQusability.com, 2002
15. Google Web Search Features, <http://www.google.com/help/features.html>
16. Google India, <http://www.google.co.in>
17. Vivisimo Search Engine, <http://www.vivisimo.com/>
18. Yahoo! Search Engine, <http://www.yahoo.com/>
19. Deepak P, Sandeep Parameswaran, "Context Disambiguation in Web Search Results", Journal of the Indian Institute of Science, Sept-Dec 2003, pp. 93-102
20. Teoma Search Engine, <http://www.teoma.com>

About Authors

Deepak P is working in Department of Computer Science and Engineering, Indian Institute of Technology Madras, Chennai.

E-mail : deepakswallet@yahoo.com

Sandeep Parameswaran is working in IBM Global Services (India) Pvt. Ltd., Embassy Golf Links Campus, Bangalore.

E-mail : sandeep_potty@yahoo.com