

Evaluating and comparing search features of Web metasearch engines: A checklist-based approach

Alireza Isfandyari-Moghaddam and Zohreh Bahari-Movaffagh

Department of Library and Information Studies,
Islamic Azad University, Hamedan Branch, IRAN
e-mail: ali.isfandyari@gmail.com

ABSTRACT

This paper aims at evaluating and comparing search features of Web metasearch engines (MSEs). To do this, a total of 64 MSEs were identified and examined on the basis of first observations. However after examining them, about 70% of these MSEs were discarded and 19 cases, which were free of charge, accessible and compatible with the research objectives were kept for further analysis. A researcher-made checklist, composed of 50 items under five search engines' general criteria (search operators, restrictors, result presentation, search help options and others) was used for data collection. It was found that AND, Phrase, Number of results per page, Keeping query searched, and Help were features that have been fully included in all 19 MSEs, while the features Anchor search and Size have been neglected by the MSEs examined. It should be also added that features Truncation, Keywords in title, URL, and Brief text have been highly supported by 18 cases (94.7%). The features Date, Cached, and Also try were observed only in one case (5.3%). It is suggested that the checklist used in this study be restructured for studying search features of Web search engines, digital libraries and other Internet search tools.

Keywords: Search engines; User interfaces; Information research; Information retrieval; Internet search tools

INTRODUCTION

As the Internet emerged, information storage and retrieval changed radically. As a result of this, a number of tools, channels, and carriers have been invented and designed for storing as well as retrieving different information and information sources available on the Internet. This is reflected in Courtois (1995) quoted in Isfandyari-Moghaddam (2007, p. 300), "though a latecomer in the Internet family, the Web has rapidly gained popularity and become the second most widely used application of the Internet". Similarly, Liaw and Huang (2003), and Spink et al (2006) believe that Web searching is the most popular online activity, behind e-mailing. Today the Web is such an integral part of our lives, and supplementing this Bar-Ilan, in her review of the literature on the use of Web search engines in information science research, presents an alternative to the Cartesian "I think, therefore I am" (cogito, ergo sum) viz. "I have a Web presence, therefore I am" (Bar-Ilan 2004, p. 231). Because many people, companies, and organizations take this notion seriously, in addition to more substantial reasons for publishing information on the Web, the number of Web pages is in the billions and growing constantly. However, it is not

sufficient to have a Web presence; tools that enable users to locate Web pages are needed as well. Among the variety of existing Web-based search tools, the primary tools for accessing needed information on the Web are search engines, directories and Metasearch engines (MSEs). These tools compete with each other for attracting users. Yet, evaluation studies play an important role in making such information-finding facilities familiar for searchers and other users (Isfandyari-Moghaddam and Ranjbar 2008). Regardless of value and status of other search tools such as search engines and directories, the importance of MSEs in accessing relevant information has been considered since their emergence. Therefore, a few studies (Tomaiuolo 1999; Repman and Carlson 1999; Dogpile 2005; and Isfandyari-Moghaddam and Parioikh 2006) were conducted in order to determine whether they are efficient search tools and have sufficient facilities. Furthermore, some descriptive studies (Echo 1998; Liu 1999; Repman and Carlson 1999; Sherman 2002; Bradley 2003; Zhang and Cheung 2003; Sherman 2005; Notess 2006; and Isfandyari-Moghaddam 2007) introduced some MSEs and described, compared and evaluated their search features.

Isfandyari-Moghaddam (2007) in providing a checklist for evaluating search capabilities of MSEs indicated that research in this area is needed and justifiable due to the following reasons:

1. Most of the search engines cannot completely satisfy user requirements. For example, most search engines can only index and process a very small part of the Web pages on the Web. A lot of the indexing methods are based only on meta-tags within the documents and the updating period is quite long. These problems may regretfully result in incomplete information retrieval and in the end users can only get very small proportions of the information they want (Lawrence et al. 1999; quoted in Li, Wang and Oria 2001);
2. The poor interface design in many search engines hinders the full use of their advanced functions and returns search results that are often very inaccurate and irrelevant (Pollock and Hockley 1997);
3. Simultaneous searching of several search engines to retrieve more relevant results is time-consuming.
4. These search tools are relatively new and there is little comprehensive research introducing their search capabilities. As such, there are opportunities for future and further research on MSEs, in the form of user-oriented studies.

Earlier studies by Isfandyari-Moghaddam indicated the following reasons why familiarity with MSEs is important and useful and that there is a necessity of doing such an evaluation of MSEs from time to time:

1. Web users should be aware that limiting searches to single search engines results in missing substantial pieces of information ranked highly by other search engines and directories (Isfandyari-Moghaddam and Parioikh 2006);
2. Apparently, the web is dynamic, and MSEs will change, develop, grow and maybe even improve. New features will emerge and old ones, even though preferred by many, may finally be given up (Isfandyari-Moghaddam 2007); and
3. To bridge the digital divide, it is necessary to learn more about "information sources" which in turn, requires several requisite knowledge and skills especially in searching and using search engines (Aqili and Isfandyari-Moghaddam 2008).

Isfandyari-Moghaddam also suggested that "future research pays attention to additional search capabilities of MSEs" as "some features have been neglected" or not considered in his study (Isfandyari-Moghaddam 2007, p. 302). Therefore building from earlier works

mentioned above, this study aims at comparing and evaluating various search features of MSEs more comprehensively by means of a researcher-made checklist.

OBJECTIVES AND METHOD

The objectives of this study are twofold:

- a) To evaluate and compare search features of Web MSEs by the use of a checklist;
- b) To identify and describe the search features of the most successful MSEs (i.e. the most capable MSEs which have high enjoyment of search features)

The first step of the investigation was to conduct a survey on existing MSEs and to identify MSEs suitable for further examination. After comparing and evaluating the currently available MSEs, a total of 64 MSEs were identified and examined on the basis of first observations. This was done through (d) searching the Web for extant lists of MSEs¹ and referring to each introduced MSEs directly; (b) consulting related works by Bazac (2002), Bradley (2003), Sherman (2002, 2005), Isfandyari-Moghaddam and Parirokh (2006), Isfandyari-Moghaddam (2007), Isfandyari-Moghaddam and Ranjbar (2008), and Sadeghi (2009). After carefully studying them, about 70% (45) MSEs were discarded because of the following reasons:

- a) some were non-English such as Ilmotore (Italian), ApocalX (French), Metaseek.nl (Dutch), and MetaBear (Russian);
- b) some were not free, i.e. they were fee-based such as ZDNet, LexiBot and Copernic;
- c) some were pseudo MSEs such as InfoGrid and HotBot. Bazac (2002) opined that it is commendable to use MSEs called real MSEs (which aggregate or rank search results in one page), instead of pseudo ones (which send the query to the search engines, and then present the results grouped by search engines in one long, easy to read scrollable list).
- d) some of them were not accessible including Profusion, 1Blink, DugDugi, IcySpicy, Moonmist, Highway61, Widow, Myproowler, Emailpinoy, EZ2Find, Gimenei, Infonetware, and Metor.
- e) some were filtered in Iran, as such studying them was not possible. They are Argosa.de, QbSearch, Vinden.nl, Suchspider.de, 37.com, 7Metasearch, Ithaki, MetaEureka, and Pandia.
- f) Kartoo was excluded because it presents retrieved results in the form of illustration.
- g) MetaFind was merged into MetaCrawler.
- h) MetaGopher leads user to Google.
- i) Search66, Veoda, and InternetSleuth which are apparently named as MSEs did not make use of any underlying search engines.

Subsequently, from the 64 firstly identified and investigated MSEs, only 19 cases which were free of charge, accessible and compatible with the research objectives were kept. The 19 MSEs chosen for further examination are:

1. Info (<http://info.com/>)
2. SurfWax (<http://www.surfwax.com/>)
3. Search.com (<http://www.search.com/>)
4. IxQuick (<http://www.ixquick.com/>)
5. iBoogie (<http://www.iboogie.com/>)
6. ZapMeta (<http://www.zapmeta.com/>)

¹ For example: <http://www.cryer.co.uk/resources/searchengines/meta.htm>; <http://searchenginewatch.com/2160791>; <http://www.google.com/Top/Computers/Internet/Searching/Metasearch/>

7. Jux2 (<http://jux2.com/>)
8. Izito (<http://www.izito.net/>)
9. CurryGuide (<http://curryguide.com/>)
10. Seekky (<http://seekky.com/>)
11. Clusty <http://clusty.com/>
12. Dogpile (<http://www.dogpile.com/>)
13. 1Second (<http://www.1second.com/>)
14. Mamma (<http://www.mamma.com/>)
15. MetaCrawler (<http://www.metacrawler.com/>)
16. WebCrawler (<http://www.webcrawler.com/>)
17. Findelio (<http://www.findelio.com/>)
18. Vroosh (<http://www.vroosh.com/>)
19. Excite (<http://www.excite.com/>)

To collect the needed data and thus to meet the purpose of the study, i.e. comparing and evaluating various search features of MSEs more comprehensively, a researcher-made checklist, composed of 50 items, was used. The theoretical foundation of this checklist was based on Isfandyari-Moghaddam (2007) and the researchers' observations of some Web search tools particularly MSEs. The 19 MSEs included in this study were compared and analyzed based on five (5) general criteria (namely search operators; restrictors; results presentation; search help options; and other criteria). Each general criterion has its own sub-criteria indicated below:

A. Search operators (basic search functions²)

• **Boolean operators**

1. AND
2. OR
3. NOT

• **Other operators**

4. +³
5. -⁴
6. Parentheses
7. Proximity⁵
8. Phrase⁶
9. Truncation⁷

B. Restrictors (limits)⁸

² They are also called *Advanced operators* (Advanced search... 2003). For further information, refer to Smith (2000).

³ To include words, just type a plus sign + in front of each one. For example, if you want to see pages about the latest "in thing," search for "+in thing."

⁴ To exclude words, type a minus sign - in front of them. No pages containing those words show up in your results. For example, if you want to see pages about dogs with no mention of poodles, search for "dogs - poodles."

⁵ A proximity search ("WITH" and "NEAR") ensures that adjacent terms appear in a full text, or within a certain paragraph, or in the same sentence, or a field (for example, title or abstract).

⁶ Phrase searching means that words must be together in a specified order (Smith 2000).

⁷ Truncation is used to control various formats of a regular term in a query. For example, "build*" (* is defined a special truncation operator) will search for cluster of similar terms such as "build", "building", "buildings", "builder", etc. Truncation operators may vary from system to system. The most frequently used truncation operators are the asterisk (*), dollar sign (\$), or plus sign (+) (Zhang and Cheung 2003).

10. Language
 11. Date
 12. Place (search by country)
 13. Limiting search range to underlying SEs and Directories
 14. Domain (net, edu, com, etc)
 - **Document type**⁹
 15. Images (pictures)¹⁰
 16. Video
 17. Audio
 18. News
 19. Filetype (pdf, word, excel, power point, etc)
 20. Web
 21. Keywords in title¹¹
 22. Host site search¹²
 23. Anchor search¹³
 24. Links search¹⁴
 25. Words in URL¹⁵
 26. White pages¹⁶
 27. Yellow pages¹⁷
 28. Shopping¹⁸
- C. Results presentation (display)**
29. Total hits counter (total number of hits retrieved¹⁹)
 30. Number of results per page
 31. Keeping query searched²⁰
 32. Facility of determining number of results per page
 33. Viewing results based on relevance
- **Result description**²¹
 34. URL
 35. Size

⁸ Using restrictors, user can customize searching any of an array of them. In fact, he/she can control his/her preferences.

⁹ In relation to items 15, 16, and 17, see Tjondronegoro and Spink (2007).

¹⁰ To learn more about image feature, refer to Hassan and Zhang (2001).

¹¹ Is used to find a specific keyword as part of the indexed titles; title searching is a valuable tool in the searchers arsenal for getting closer to a subject search on the Web. It can be a great way to narrow results and can often give a search more of a subject focus, available at: <http://www.searchengineshowdown.com/features/title> (accessed 12 June 2009).

¹² Is used to find all documents within a particular domain and all its subdomains; results are only from the specified site. site:nasa.gov finds pages at NASA's Web site.

¹³ Some search engines allow you to search specifically within the "anchor" or "link" text that appears on a web page. For example, consider this example: Click Here For The Mars Exploration Web Site. Notice the words "Mars Exploration Web Site" are all contained within the hyperlink? This is the anchor text or the link text (see Sullivan 2001).

¹⁴ Is utilized to find documents that link to a particular URL; pages include a link to the specified URL. link:searchengineshowdown.com finds pages with links to this site.

¹⁵ It is applied to find a specific document in search tool index.

¹⁶ See <http://en.wikipedia.org/wiki/Phonebook> (accessed 12 June 2009).

¹⁷ See http://en.wikipedia.org/wiki/Yellow_pages (accessed 12 June 2009).

¹⁸ See <http://help.yahoo.com/l/us/yahoo/search/basics/basics-23.html> (accessed 12 June 2009).

¹⁹ In the field of Information storage and retrieval, it is called *recall*.

²⁰ Whether query or term (s) searched is kept to do another search or modify it.

²¹ Result description shows only the usual elements (title, short description, URL, and file size – see for details) but also additional navigational links from the homepage of the Web site .

36. Brief text

37. Cached²²

38. Similar pages²³

D. Search help options

39. Related search²⁴

40. Refine search²⁵

41. Search history (recent search)²⁶

42. Also try²⁷

E. Other criteria

43. Help (about)

44. Spelling correction

45. Case sensitivity²⁸

46. Advanced (power) search

47. SE coverage²⁹

48. Adult (family) filter

49. Sponsored link³⁰

50. Common words search³¹

As indicated in Isfandyari-Moghaddam (2007, p. 304), in order to rate and explore MSEs search features, some test searches and strategies were applied. Features such as "Phrase", "Boolean", "SE coverage", "Date", "Language", "Adult filter", "Place", etc. which were described in the "Help" or "About" of some MSEs or were included in advanced search or search preferences were not fully examined via selected queries. Some can be easily identified such as "Help", "Advanced search", "Results presentation" and "Document

²² A link to a copy of the page saved by the search engine, in case the site is unavailable. Your search terms are highlighted on the cached page to help you find the relevant content.

²³ For more information, see Hariri (2008).

²⁴ The related searches, on the other hand, are frequently searched word combinations. "Related search" intelligently suggests alternative search phrases, allowing the user to perform additional searches with a single click, available at: http://www.sli-systems.com/related_search.php (accessed 12 June 2009).

²⁵ This feature provides an easy way to narrow down your search results. When you refine a search, you are attempting to obtain better results than you obtained on an earlier search, available at: <http://writing.colostate.edu/demos/keyword/refine.cfm> (accessed 12 June 2009).

²⁶ The history screen allows you to re-execute or edit any previous searches you have performed in your current session. To access the History screen, you must first complete at least one search. After you have completed at least one search, the history button will display on the tool bar at the top of the screen the list of search and browse requests for the current session. You can rerun any of the requests, or return to one to refine it.

²⁷ When other people have done searches similar to yours, search tool will list these queries under the search box. One of these might help you narrow your results. See <http://help.yahoo.com/l/us/yahoo/search/basics/basics-23.html> (accessed 12 June 2009).

²⁸ The ability of a search tool to distinguish between upper and lowercase letters; some search tools are not case sensitive and will simply read all letters as lowercase. Others may distinguish between the word "aids" and the disease "AIDS," or the word "baker" and the name "Baker" (Web Searching Glossary n.d.)

²⁹ Refers to the number of different single SEs to which a MSE directs its queries.

³⁰ It leads user(s) to sites that pay for placement in search results on keywords that are relevant to their business. For more information, refer to Jansen and Resnick (2006), Jansen and Spink (2007), and Jansen (2007).

³¹ A common word such as "the," "of," "on," and "a."; these words are not indexed, and when used in a query, are also ignored. However, different search engines do not use the same common or stop words, so the same query at one search site may yield different results than at another.

type", ["Yellow pages", "White pages", "Sponsored link", "Related search", "Refine search", "Search history (recent search)", "Also try", "URL", "Size", "Brief text", "Cached", "Similar pages", and so on]. But to explore some capabilities including "Parentheses", "Truncation", and "Proximity", several queries namely (web or internet) and adults, (television or mass media) and women, "educat*"³², "develop*" and "univers*", "logical with positivism", "macro near virus", were searched. For studying "Case sensitive", words "INTERNET HISTORY", "internet history", "internet History", "Internet history" and "Internet History" were searched. Also, in order to rate how MSEs offer capability "Spelling correction", a couple of terms namely "Univercity" and "Adulf Hitler"³³ were searched. It should be reminded that some subcriteria included in the checklist have been examined and described earlier in this section.

RESULTS

After evaluating MSEs in terms of their search features based on the researcher-made checklist, data found were tabularized in Tables 1 to 6, where the rows indicate the MSEs and the columns are the defined features. Each row demonstrates the level of supports on each feature for the particular MSE. Table 1 summarizes the extent of support for "search operators" features by the 19 MSEs included in this study. Based on Table 1, it is apparent that features AND, -, and Phrase were supported by all the MSEs (100%, 19), whereas in order of enjoyment, features Truncation (94.7%, 18), + (89.4%, 17), OR (84.2%, 16), NOT (63.1%, 12), Proximity (63.1%, 12), and Parentheses (36.8%, 7) were included in the search features of MSEs.

Table 1: Search Operators Supported by the MSEs

MSEs	AND	OR	NOT	+	-	Parenthesis	Proximity	Phrase	Truncation
1Second	√	√	√	-	√	-	-	√	√
Clusty	√	√	-	√	√	√	-	√	√
CurryGuide	√	√	√	√	√	-	√ ADJ	√	√
Dogpile	√	√ in advanced	√ in advanced	√	√	√	-	√	√
Seekky	√	-	-	√	√	-	√ WITH	√	√
iBoogie	√	√	-	-	√	-	√ WITH	√	√
Info	√	√ in advanced	√ in advanced	√	√	-	√ WITH	√	√
IxQuick	√	√	√	√	√	√	√ ADJ	√	√
Izito	√	√ in advanced	√	√	√	-	√ ADJ	√	√
Jux2	√	√ in advanced	√ in advanced	√	√	-	√ ADJ	√	√
Search.com	√	√ in advanced	√ in advanced	√	√	√	√ ADJ	√	√
SurfWax	√	√	-	√	√	-	-	√	√
ZapMeta	√	√ in advanced	√	√	√	√	√	√	√
Mamma	√	-	-	√	√	√	√ ADJ	√	√
MetaCrawler	√	√ in advanced	√	√	√	-	√	√	√
WebCrawler	√	√	-	√	√	-	√ ADJ	√	√
Vroosh	√	√	√	√	√	√	-	√	√
Findelio	√	-	-	√	√	-	-	√	√
Excite	√	√ in advanced	√ in advanced	√	√	-	-	√	-

³² Because in literature the most commonly cited symbol is the asterisk (*), it was chosen and tested.

³³ University and Adolf Hitler

Tables 2 and 3 summarize the extent of support for "restrictors" features by the MSEs surveyed. Accordingly, features Keywords in title (94.7%, 18), Web (89.4%, 17), and Host site search (84.2%, 16) have been considered more, while features Date and Anchor search were the least considered items with 5.3% (1) and 0% (0), respectively.

Table 2: Restrictors Supported by the MSEs

MSEs	Language	Date	Place (search by country)	Limiting search range to underlying SEs and Directories	Document type				Domain (net, edu, com, etc)	Filetype: pdf, word, excel, power point, etc	Web
					News	Video	Audio	Images			
1Second	-	-	-	-	√	√	√	-	-	-	√
Clusty	√	-	-	√	√	-	-	√	√	√	√
CurryGuide	√	-	√	-	√	√	√	√	-	-	√
Dogpile	√	-	-	-	√	√	-	√	√	-	√
Seekky	-	-	-	-	√	√	-	√	-	√	√
iBoogie	√	-	-	√	√	-	-	√	√	-	√
Info	√	-	-	-	√	√	√	√	√	√	√
IxQuick	√	-	√	-	-	√	-	√	√	-	√
Izito	-	-	√	-	-	√	√	√	√	-	√
Jux2	-	-	-	-	-	-	-	-	√	-	-
Search.com	√	√	-	√	√	√	√	√	√	√	√
SurfWax	-	-	-	-	-	-	-	-	-	-	√
ZapMeta	-	-	√	-	-	√	√	√	√	-	√
Mamma	-	-	-	√	-	√	-	-	-	-	√
MetaCrawler	√	-	-	-	√	-	-	√	√	-	√
WebCrawler	√	-	-	-	√	√	-	√	-	-	√
Vroosh	√	-	√	-	-	√	√	-	-	-	√
Findelio	-	-	-	-	-	-	-	-	-	√	-
Excite	√	-	-	-	√	√	√	√	√	-	√

Table 3: Restrictors Supported by the MSEs – cont.

MSEs	Keywords in title	Host site search	Anchor search	Links search	Words in URL	White pages	Yellow pages	Shopping
1Second	√	-	-	√	√	-	-	√
Clusty	√	√	-	√	√	-	-	√
CurryGuide	√	√	-	√	-	-	√	√
Dogpile	√	√	-	-	-	√	√	-
Seekky	√	√	-	√	-	-	-	-
iBoogie	√	-	-	√	-	-	-	-
Info	√	√	-	-	-	√	√	√
IxQuick	√	√	-	√	√	√	-	-
Izito	√	√	-	√	-	-	-	-
Jux2	√	-	-	-	-	-	-	-
Search.com	√	√	-	√	-	-	-	√
SurfWax	-	√	-	-	√	-	-	-
ZapMeta	√	√	-	√	-	-	-	-
Mamma	√	√	-	√	-	√	√	√
MetaCrawler	√	√	-	√	-	√	√	-
WebCrawler	√	√	-	√	-	√	√	-
Vroosh	√	√	-	√	-	-	-	-
Findelio	√	√	-	√	√	-	-	-
Excite	√	√	-	√	-	√	√	-

Evaluating and Comparing Search Features of Web Metasearch Engines

Table 4 abridges the extent of support for "presentation of results" features by the MSEs surveyed. According to Table 4, the most considered features were Brief text and URL with 94.7% (18) respectively, whereas features Size (0%, 0) and Cached (5.3%, 1) were the most neglected ones.

Table 4: Presentation of Search Results Features

MSEs	Total hits counter (total number of hits retrieved)	Number of results per page	Keeping query searched	Facility of determining number of results per page	Viewing results based on relevance	Result description				
						URL	Size	Brief text	Cached	Similar pages
1Second	√	20	√	√	-	√	-	√	-	√
Clusty	√	10	√	√	-	√	-	√	√	-
CurryGuide	-	8	√	√	-	√	-	√	-	√
Dogpile	-	20	√	√	√	√	-	√	-	-
Seekky	-	8	√	-	-	√	-	√	-	√
iBoogie	√	10	√	√	-	√	-	√	-	-
Info	-	20	√	√	-	√	-	√	-	-
IxQuick	√	10	√	√	-	√	-	√	-	-
Izito	√	10	√	√	-	√	-	√	-	√
Jux2	√	24	√	-	-	√	-	√	-	-
Search.com	√	10	√	-	-	√	-	√	-	√
SurfWax	√	Various	√	√	√	-	-	-	-	-
ZapMeta	√	10	√	√	-	√	-	√	-	√
Mamma	√	20	√	√	-	√	-	√	-	-
MetaCrawler	√	20	√	√	√	√	-	√	-	-
WebCrawler	√	20	√	-	-	√	-	√	-	-
Vroosh	√	8	√	√	-	√	-	√	-	-
Findelio	√	15	√	-	-	√	-	√	-	-
Excite	√	10	√	√	-	√	-	√	-	-

As illustrated in Table 5, the extent of support for "search help options" features by the MSEs surveyed indicates that two features, namely Related search and Search history, were supported by 42.1% (8) MSEs. The features Refine search and Also try were supported by 15.8% (3) and 5.3% (1) MSEs respectively.

Table 5: Search Help Options Features

MSEs	Related search	Refine search	Search history (recent search)	Also try
1Second	-	-	-	-
Clusty	-	-	-	-
CurryGuide	-	-	√	-
Dogpile	-	-	√	√
Seekky	-	-	-	-
iBoogie	-	-	-	-
Info	√	-	-	-
IxQuick	-	√	-	-
Izito	√	-	√	-
Jux2	-	-	-	-
Search.com	√	√	-	-
SurfWax	-	-	-	-
ZapMeta	√	-	√	-
Mamma	-	√	√	-
MetaCrawler	√	-	√	-
WebCrawler	√	-	√	-
Vroosh	√	-	-	-

Findelio	-	-	-	-
Excite	√	-	√	-

Finally, Table 6 illustrates the extent of support for "other criteria" features by the MSEs surveyed. Help and Common words search were the most supported features with 100% (19) and 89.4% (17) respectively. On the other hand, case sensitivity feature has been considered only by 4 MSEs (21%). In summary, the feature frequency distributions for the MSEs are presented in Table 7 and Figures 1a, b and c.

Table 6: Other criteria

MSEs	Help (about)	Spelling correction	Case sensitivity	Advanced (power) search	SE coverage	Adult (family) filter	Sponsored link	Common words search
1Second	√	-	-	√	14	-	√	√
Clusty	√	√	√	√	5	√	√	-
CurryGuide	√	-	-	√	-	√	√	-
Dogpile	√	√	√	√	4	√	√	√
Seekky	√	-	-	-	3	-	-	√
iBoogie	√	√	-	√	7	√	√	√
Info	√	√	-	√	5	√	√	√
IxQuick	√	√	-	√	11	√	√	√
Izito	√	√	-	√	6	√	-	√
Jux2	√	√	-	√	3	√	√	√
Search.com	√	√	-	√	4	√	√	√
SurfWax	√	-	√	-	3	-	-	√
ZapMeta	√	√	-	√	6	√	-	√
Mamma	√	√	√	√	9	√	√	√
MetaCrawler	√	√	-	√	5	√	√	√
WebCrawler	√	√	-	-	4	√	√	√
Vroosh	√	-	-	√	3	√	-	√
Findelio	√	-	-	-	4	-	√	√
Excite	√	√	-	√	3	√	√	√

Evaluating and Comparing Search Features of Web Metasearch Engines

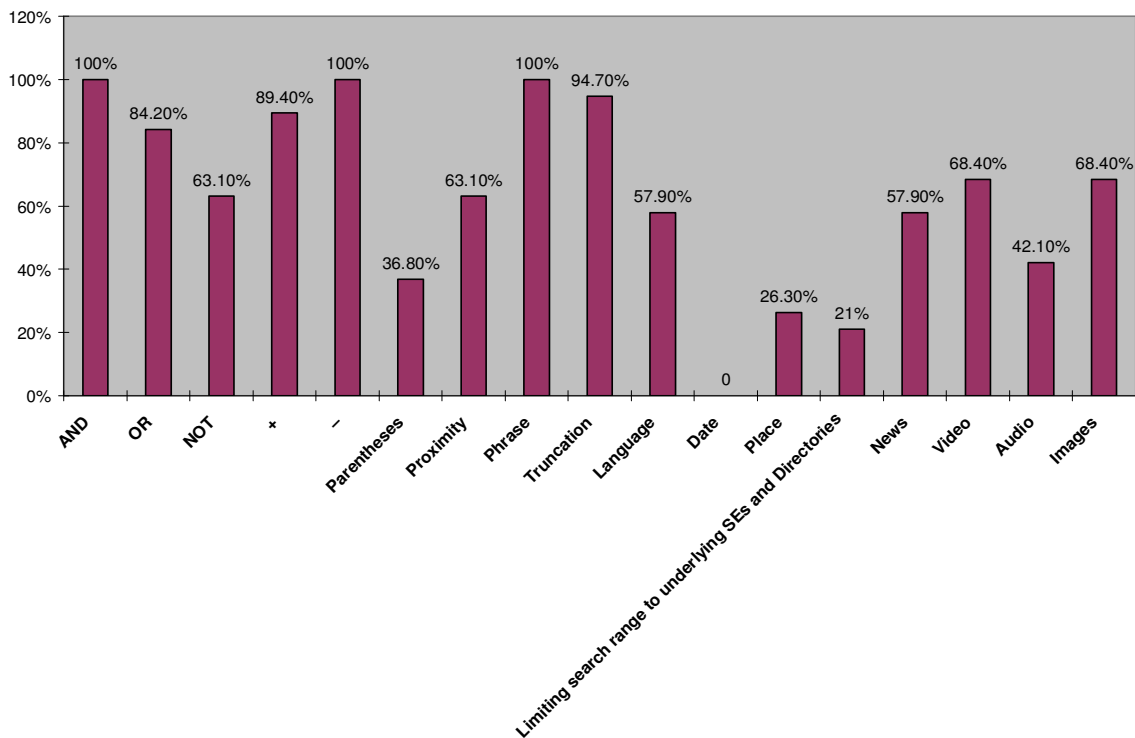


Figure 1a: Distribution of MSEs Search Features

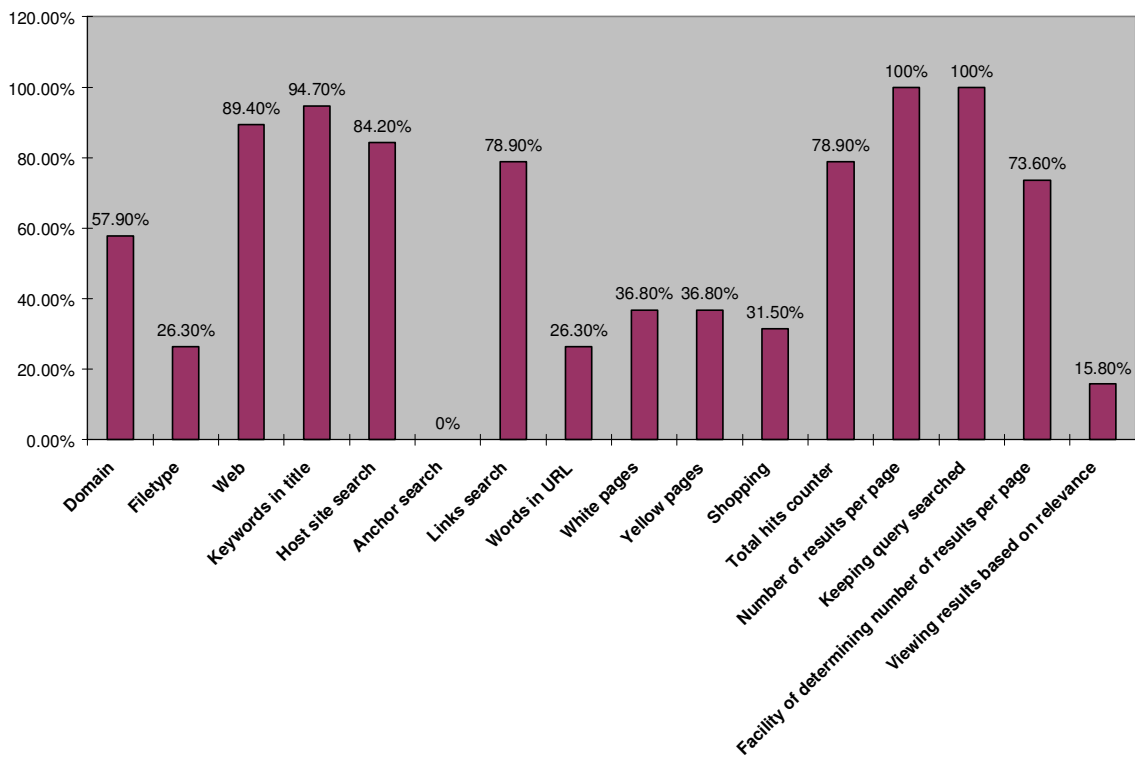


Figure 1b: Distribution of MSEs Search Features – cont.

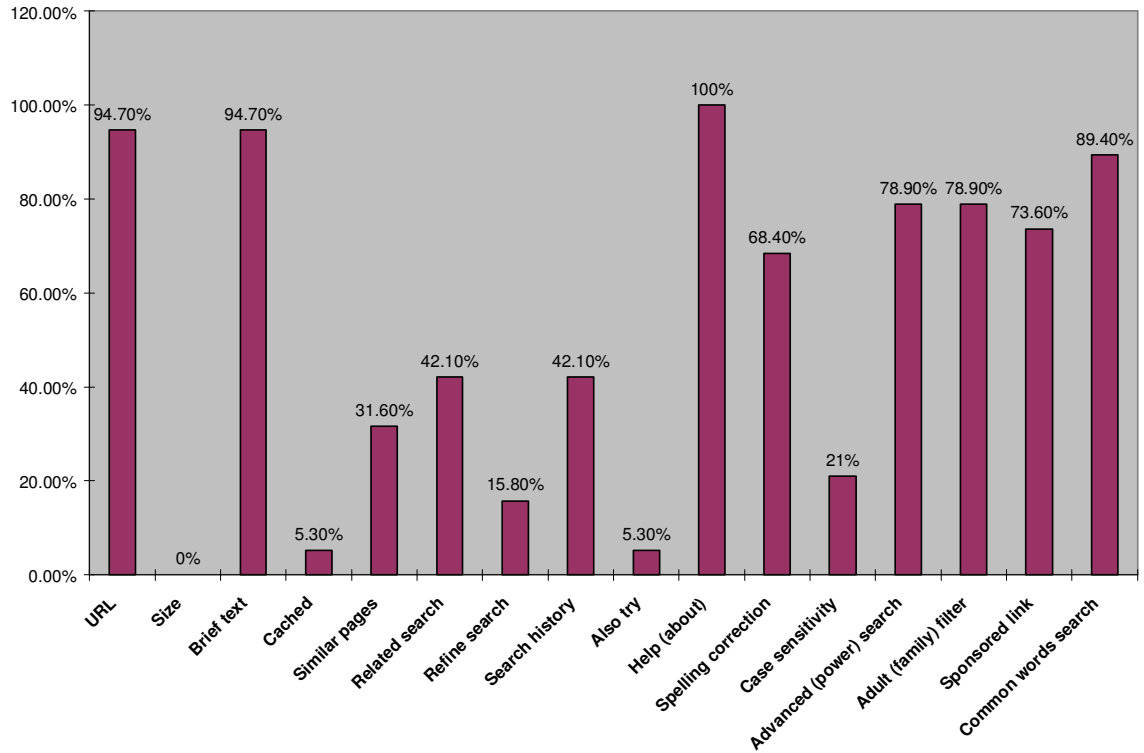


Figure 1c: Distribution of MSEs Search Features – cont.

Table 7: Summary Table of MSEs Search Features

Criteria	Subcriteria (feature)	Occurrence	Percentage
A. Search operators	AND	19	100%
	OR	16	84.2%
	NOT	12	63.1%
	+	17	89.4%
	-	19	100%
	Parentheses	7	36.8%
	Proximity	12	63.1%
	Phrase	19	100%
	Truncation	18	94.7%
B. Restrictors	Language	11	57.9%
	Date	1	5/3%
	Place	5	26.3%
	Limiting search range to underlying SEs and Directories	4	21%
	News	11	57.9%
	Video	13	68.4%
	Audio	8	42.1%
	Images	13	68.4%
	Domain	11	57.9%
	Filetype	5	26.3%
	Web	17	89.4%
	Keywords in title	18	94.7%
Host site search	16	84.2%	

Evaluating and Comparing Search Features of Web Metasearch Engines

	Anchor search	0	0%
	Links search	15	78.9%
	Words in URL	5	26.3%
	White pages	7	36.8%
	Yellow pages	7	36.8%
	Shopping	6	31.5%
C. Results presentation	Total hits counter	15	78.9%
	Number of results per page	19	100%
	Keeping query searched	19	100%
	Facility of determining number of results per page	14	73.6%
	Viewing results based on relevance	3	15.8%
	URL	18	94.7%
	Size	0	0%
	Brief text	18	94.7%
	Cached	1	5.3%
	Similar pages	6	31.6%
D. Search help options	Related search	8	42.1%
	Refine search	3	15.8%
	Search history	8	42.1%
	Also try	1	5.3%
E. Other criteria	Help (about)	19	100%
	Spelling correction	13	68.4%
	Case sensitivity	4	21%
	Advanced (power) search	15	78.9%
	Adult (family) filter	15	78.9%
	Sponsored link	14	73.6%
	Common words search	17	89.4%

DISCUSSION

Within the Search operators criteria (Table 1), AND (100%, 19), OR (84.2%, 16), NOT (63.1%, 12), + (89.4%, 17), – (100%, 19), Parentheses (36.8%, 7), Proximity (63.1%, 12), Phrase (100%, 19), and Truncation (94.7%, 18) were included in the search features of MSEs. These findings, especially regarding features AND, Phrase, Truncation, and Proximity are compatible with studies by Zhang and Cheung (2003) and Isfandyari-Moghaddam (2007). IxQuick, ZapMeta, and Search.com were the only MSEs supporting all subcriteria included under this general category, while Findelio had a poorer interface with only 5 features supported compared to other MSEs surveyed.

The Restrictors criteria supported by the MSEs were Language (57.9%, 11), Date (5.3%,1), Place (26.3%, 5), Limiting search range to underlying SEs and Directories (21%, 4), News (57.9%, 11), Video (68.4%, 13), Audio (42.1%, 8), Images (68.4%, 13), Domain (57.9%, 11), Filetype (26.3%, 5), Web (89.4%, 17), Keywords in title (94.7%, 18), Host site search (84.2%, 16), Anchor search (0%, 0), Links search (78.9%, 15), Words in URL (26.3%, 5), White pages (36.8%, 7), Yellow pages (36.8%, 7), and Shopping (31.5%, 6) (Tables 2 and 3). Among these search features, the least considered sub-criteria are Anchor search (0%, 0) and Date (1 case, 5.3%) which need to be included in poor MSEs. Notably, search feature Domain which was supported by 5 MSEs in Isfandyari-Moghaddam’s (2007) study has been currently included in search features of 11 cases. To sum up, if any MSE wants to be internationally visible, it should pay more attention to features such as Language, Date, Place, News, Video, Audio, Images, and Domain of which the MSEs investigated in this

study could have supported these features more. In this regard, it is worth saying that Info with 13 features and Clusty and CurryGuide with 12 features respectively were the most capable MSEs, while Jux2 with two features and SurfWax with three features were the poorest MSEs.

The Results presentation criteria supported by the MSEs were "Viewing results based on relevance" (15.8%, 3), URL (94.7%, 18), Size (0%, 0), Brief text (94.7%, 18), Cached (5.3%, 1), and Similar pages (31.6%, 6) (Table 4). Regardless of neglecting Size (0%, 0) and Cached (5.3%, X), it is necessary to remind that the feature "Viewing results based on relevance" which has been currently supported by only three MSEs (15.8%) should be considered as more important than ever before. This percentage is less than 50% observed in Isfandyari-Moghaddam (2007). It is also important to note that of among the 10 features considered under this category, all MSEs supported at least 5 features.

Within "Search help options" criteria (Table 5), features that were included in the MSEs under study were Related search (42.1%, 8), Refine search (15.8%, 3), Search history (42.1%, 8), and Also try (5.3%, 1). This is self-evident that features "Also try" and Refine search should not be neglected in MSEs because the more options helping users' searching are included in search tools, the more relevant and satisfactory results will be accessed. The subcriteria included under "Search help options" also needs much attention.

Finally, within "Other criteria" category (Table 6), feature Help was present in all MSEs, followed by Common words search (89.4%, 17), Advanced search (78.9%, 15), Adult filter (78.9%, 15), Sponsored link (73.6%, 14), Spelling correction (68.4%, 13), and Case sensitivity (21%, 4). All of these features except for Case sensitivity (21%, 4) have been properly supported by the MSEs under study. In this category, Dogpile and Mamma were the richest MSEs in considering all defined features, whereas Seekky had a poor performance compared to other examined MSEs.

After analyzing all features (Table 7), it was found that AND, Phrase, Number of results per page, Keeping query searched, and Help (about) were the features that have been fully included in all 19 MSEs, while features Anchor search and Size have been neglected. It should also be noted that features Truncation, Keywords in title, URL, and Brief text have been highly supported by 18 cases (94.7%). Moreover, features Date, Cached, and Also try which were observed only in one MSE (5.3%), and features Viewing results based on relevance and Refine search which were supported by three MSEs (15.8%) should be given emphasis. In other words, MSEs optimizers should include such less considered features in MSEs to facilitate users in information searching.

CONCLUSION

Rephrasing Descartes' statement as "I have a Web presence, therefore I am" (Bar-Ilan 2004, p. 231), this paper posits that for the realization of the Web presence, Chekuri et al (1997) declaration, i.e. "providing an efficient and user friendly search interface for the Web is still one of the most important challenges in making it accessible to the general public" can be currently restated. Undoubtedly, one of the main ways for increased Web presence of citizens in the third millennium is designing, optimizing and introducing Web search tools including MSEs. In this study, the search features of Web MSEs have been evaluated, compared and analyzed, and a checklist for evaluating the search features has been offered. Accordingly, it is expected that doing such studies and publishing the results can lead to better decisions both for Web optimizers, to design more user-friendly search

tools, and for users to be familiar with quality search tools and thus to access more relevant search results. Additionally, to help users select reliable and suitable MSEs, to utilize an MSE effectively and to identify the most successful (the most capable MSEs which have high enjoyment) and the weakest ones, the evaluation conducted showed that the most capable MSEs which have high enjoyment are Search.com (with 37 included features), followed by Dogpile, Info, and IxQuick (with 33 features each). The least capable MSEs are SurfWax and Findelio, with 17 and 18 features respectively. Such a finding is relatively similar to Isfandyari-Moghaddam's (2007) study in which Dogpile, IxQuick and Info were introduced among the top performing MSEs in supporting and including search features.

As for future research in the evaluation of online search tools, it is suggested that the checklist offered here be restructured for studying search features of Web search engines, digital libraries and other Internet search tools. It is also reminded that more emphasis should be given to the field of Web search tools so that problems hindering the accessibility of relevant as well as pertinent results can be alleviated. Hopefully, future researchers in this area, through his/her investigations and recommendations, can contribute to the increased usage of information and knowledge in individuals' daily life, through a step forward and toward reinforcing Web presence of all people and increasing their being!

ACKNOWLEDGMENT

The authors would like to thank Dr. Yaghoob Norouzi for his useful comments and the Islamic Azad University, Hamedan Branch for research support.

REFERENCES

- Advanced search operators. 2003. Available at: frli.law.gov.au/frlihelp/help_advanced_operators.html
- Aqili, S.V. and Isfandyari-Moghaddam, A. 2008. Bridging the digital divide: The role of librarians and information professionals in the third millennium. *The Electronic Library*, Vol. 26, no. 2: 226-237.
- Bar-Ilan, J. 2004. The use of Web search engines in information science research. *Annual Review of Information Science and Technology*, Vol. 38: 231–288.
- Bazac, D. 2002. *The meta search engines: A web searcher's best friends*. Available at: www.llrx.com/features/metasearch.htm
- Bradley, P. 2003. *Multi-search engines-a comparison*. Available at: www.philb.com/msengine.htm
- Chekuri, C., Goldwasser, M.H., Raghavan, P., and Upfal, E. 1997. Web search using automatic classification. Poster presentation at *the Sixth International World Wide Web Conference (WWW6)*. April 7-11 in Santa Clara, California, USA. Available at: <http://mathcs.slu.edu/~goldwasser/publications/WWW1997.color.pdf>
- Chu, H. and Rosenthal, M. 1996. *Search engines for the World Wide Web: A comparative study and evaluation methodology*. Available at: www.asis.org/annual-96/electronicproceedings/chu.htm

- Dogpile. 2005. *Missing pieces: A study of first page web search engine results overlap*. Available at: <http://missingpieces.dogpile.com/WhitePaper.pdf>
- Echo, G. 1998. *How they work: The metasearch engines: Covering the waterfront*. Available at: <http://www.editors-service.com/>
- Hariri, N. 2008. An investigation of the effectiveness of the “similar pages” feature of Google. *Online Information Review*, Vol. 32, no. 3: 370-378
- Hassan, I., and Zhang, J. 2001. Image search engine feature analysis. *Online Information Review*, Vol. 25, no. 2: 103-114
- Isfandyari-Moghaddam, A. 2007. Web metasearch engines: A comparative study on search capabilities using an evaluation check-list. *Online Information Review*, Vol. 31, No. 3: 300-309.
- Isfandyari-Moghaddam, A. 2009. Databases: From paper-based to Web-based. *Library Philosophy and Practice*, (March).
- Isfandyari-Moghaddam, A. and Parirokh, M. 2006. A comparative study on overlapping of search results in metasearch engines and their common underlying search engines. *Library Review*, Vol. 55, no. 5: 301-306.
- Isfandyari-Moghaddam, A., and Ranjbar, V. 2008. Difference among ranking algorithms of different Web search tools: A statistical approach. *Malaysian Journal of Library & Information Science*, Vol. 13, no. 2: 15-28
- Jansen, B.J. 2007. The comparative effectiveness of sponsored and non-sponsored results for web ecommerce queries. *ACM Transactions on the Web*, Vol. 1.
- Jansen, B.J., and Resnick, M. 2006. An examination of searcher’s perceptions of non-sponsored and sponsored links during ecommerce web searching. *Journal of the American Society for Information Science and Technology*, Vol. 57: 1949–1961.
- Jansen, B.J., and Spink, A.H. 2007. Sponsored search: Is money a motivator for providing relevant results?. *Computer*, Vol. 40, no. 8: 52 -57.
- Li, Z., Wang, Y. and Oria, V. 2001. A new architecture for web meta-search engines. *7th American Conference on Information Systems*. Available at: <http://web.njit.edu/~zxl8078/Publication/BB040.PDF>
- Liaw, S.S., and Huang, H.M. 2003. An investigation of user attitudes toward search engines as an information retrieval tool. *Computers in Human Behavior*, Vol. 19: 751–765
- Liu, J. 1999. Guide to meta-search engines. Available at: www.indiana.edu/~librcsd/search/meta.html
- Notess, G.R. 2006. Meta search engines. Available at: <http://www.searchengineshowdown.com/multi>
- Pollock, A. and Hockley, A. 1997. What’s wrong with internet searching. *D-Lib Magazine*, available at: <http://www.dlib.org/dlib/march97/bt/03pollock.html>
- Repman, J. and Carlson, R. D. 1999. Surviving the storm: Using metasearch engines effectively. *Computers in Libraries*, V. 19, no. 5: 50-55.
- Sadeghi, H. 2009. Assessing metasearch engine performance, *Online Information Review*, Vol. 33, no. 6: 1058-1065
- Sherman, C. 2005. *Metacrawlers and metasearch engines*. Available at: <http://searchenginewatch.com/links/article.php/2156241>
- Sherman, C. 2002. *The big four metasearch engines*. Available at: <http://searchenginewatch.com/searchday/article.php/2160781>
- Smith, A.G. 2000. Search features of digital libraries. *Information Research*, Vol. 5, no. 3. Available at: <http://informationr.net/ir/5-3/paper73.html>
- Spink, A., Jansen, B. J., Blakely, C., and Koshman, S. 2006. A study of results overlap and uniqueness among major web search engines. *Information Processing and Management*, Vol. 42: 1379-1391.

- Sullivan, D. 2001. Power searching for anyone. Available at: <http://searchenginewatch.com/2156031>
- Tjondronegoro, D., and Spink, A. 2008. Web search engine multimedia functionality. *Information Processing and Management*, Vol. 44: 340–357
- Tomaiuolo, N. 1999. Are metasearches better searchers? *Searcher: The magazine for database professionals*, Vol. 7, no. 1: 30-34.
- Web searching glossary. n.d. Case sensitivity. Available at: <http://www.sdst.org/shs/library/glossary.html>
- Zhang, J. and Cheung, C. 2003. Meta-search engines feature analysis. *Online Information Review*, Vol. 27, no. 6: 433-441.