5.1 Introduction

The searching subsystem is one of the major subsystems of an information retrieval system. In this subsystem, users' queries are received and interpreted, appropriate search statements are formulated, and the actual search (i.e. matching queries with the document profile or database) is conducted with a view to retrieving the required information. All these tasks can be performed manually or can be automated. The development of cheaper direct access mass storage devices, magnetic disks and drums, the associated software, and advancements in electronic communication systems brought about the possibility of more dynamic searching via online methods. The concept of online searching has occupied a large and significant area in the study and recharge of modern information retrieval. However, a user often faces difficulties in approaching an online search system, especially in formulating an appropriate search statement. The cost of searching a database, whether in-house or external, can be reduced significantly if an appropriate strategy for searching is followed. The search strategy helps the user select the optimum path for searching a file or a database. This involves a number of measures that are to be taken before and during a search. This chapter discusses the basic concept of the search strategy and describes the actual searching process in the context of information retrieval.
One major task in the searching process relates to the coordination of items in order to formulate the actual search statement. The result of the search depends largely on how adequately the search terms are combined. Boolean search techniques have been used widely since the beginning of mechanized information retrieval. Significant differences between the major crawler-based search engines are summarized in this part of the research. Information on this page has been drawn from the help pages of each search engine; along with data gained from articles, reviews, books, independent research, tips from others and additional content received directly from the various search engines.

The search strategy encompasses several steps and levels of work in information retrieval. There are many issues that need to be considered while formulating an appropriate search statement. These are:

i) the concepts or facets to be searched and their order

ii) the term(s) that appropriately represent(s) the search concept

iii) the feature(s) of the retrieval system concerned, and

iv) the measures to be taken in revising a search statement.

5.2 Searching Strategies of Selected Search Engines

Searching for information from the Internet today has been made easier by the widely available search engines. It is of considerable importance for the designer to develop quality search engines and for the users to select the most appropriate ones for their use. On the basis of the survey conducted by some of the website analysis organizations, like, “OneStat.com”, Nielson/Net Rating, Search Engines Watch, the
five most widely used Search Engines have been selected in order to have a comparative study of their searching strategy and performance in retrieving Web information resources. These five Search Engines are: Google, Yahoo, Live, Ask, and AOL.

The five search services in this study offer some combination of standard IR search facilities. These include the abilities: to form Boolean queries; to specify that a term should (not) or must (not) appear in a Web page; to allow the user to use wildcards and truncation in search statements as well as for search algorithms to use automatic; to search for phrases rather than individual words; to specify the importance of case sensitivity; and to do proximity searching. More advanced capabilities available in some engines include the abilities: to allow the user to write a query in the form of a complete sentence or short paragraph, which the search engine then parses and exploits; to suggest to the user additional words or phrases to include to refine an initial query or to allow the user to specify certain already retrieved documents to serve as examples of the type he or she would like to see; and to show the user groupings of retrieved documents that reflect how various concepts occur among them. The precise algorithms that search engines use for retrieval are not publicized, but one can infer their approximate workings by reading the Help, Hint or FAQ pages that accompany them as well as by being familiar with the field of IR. Further, some search engines allow the user to specify, for example, that the pages he or she is interested in should contain a plug-in, script, Java applet or embedded real audio.

The matching algorithms that search engines employ also often embrace certain principles that apply to Web-based searching. In short, search engines employ a variety of relatively advanced IR techniques, and there are some peculiar aspects of
search engines that make searching the Web different than more conventional information retrieval. In the following section a comparative study has been carried out using different search features of the above five search engines.

5.2.1 Search Interface

(a) **Google** has both simple and advanced search interface. The simple interface is a single search box with two search buttons: "Google Search" and "I'm Feeling Lucky". The ‘I’m feeling Lucky’ button in Google takes users directly to the first web page Google returned for a query. The advanced interface provides boxes for "all the words", "exact phrase", "any of the words", and "without the words", pull-down menus to limit by location on the page (anywhere, title or URL), language and domain, radio buttons to filter results using "SafeSearch". With the emergence of number of competitive search engines, Google has developed new advanced interfaces in different time. Google's approach, which it calls "universal search," is a modest thing, a first step in what it says will be a long effort to break down barriers between different types of information a user may be seeking, such as Web links, images and news. Google combines these different types of content into one list. To make this simpler, Google's Universal Search now groups these different kinds of results within its single, familiar main list of results. The ultimate goal of universal search is to break down the silos of information that exist on the web and provide the very best answer every time a user enters a query. Google's vision for universal search is to ultimately search across all its content sources, compare and rank all the information in real time, and deliver a single, integrated set of search results that offers users precisely what they are looking for. Google's homepage and a number of applications have also been updated with a new navigation bar to provide easier
access to popular Google products. Now, instead of having links above the Google.com homepage search box, users will see a navigation bar on the top left side of the page with various Google search properties and popular products including Gmail, Google Calendar, Google Docs & Spreadsheets, and Picasa Web Albums. The images coming at the top of the page are part of Google's new Universal Search, a system designed to automatically blend listings from the web along with those from specialized search resources from news, video, local, book and image search databases.

**Figure - 5.1**

Google Search Page

(b) **Yahoo** has both Simple and Advanced interfaces. The simple search provides a single search box. The advanced interface includes all of the features in the simple interface, plus the ability to limit by language, domain, country, file format, number of results, location of search terms in the document (text, title, link name, URL and link to the URL), and "word filters". The advanced interfaces can be change going through the Yahoo Search preferences. SafeSearch can be set to filter recognized
explicit, adult-oriented content and images out of the search results. SafeSearch may be set to one of three modes and can be password protected. Recently Yahoo has changed the search interface with in-build Search Assist option to suggest the search keywords.

**Figure - 5.2**

Yahoo Search Page

(c) **Live Search** also has both Simple and Advanced interfaces. The simple search provides a single search box. The advanced search interface includes the ability to limit the results by language, domain, geographical location of the source, and keywords appearances. The search box, and "see also" (images, news, msn, etc) have moved down into the page. The "most popular products" showing images and ratings, coming from MSN Shopping, is new. Microsoft publically showed for the first time some aspects of Live Search 2.0, including a quantum leap in search relevance, with big improvements in mobile and local (maps) search, and a focus on four key verticals, i.e. Health, Shopping, Local, and Entertainment. Notably, the user
interface for Windows Live Search is cleaner and less cluttered than before. The service no longer offers an infinite scroll of results, but a new set of "related searches" links offer some compelling query refinement tools for many queries. In a nod to conventions, Microsoft has also renamed its "search builder" tool as a more traditional "advanced" search link. Clicking this link opens up the search builder box directly beneath the search box. There is now an Autospell correction options that will suggest an alternative word incase of misspelled word.

**Figure – 5.3**

*Live Search Page*

(d) *Ask* took greater strides with cool features such as previews of the pages it listed, lots of summary information at the top of the page and prominent suggestions for narrowing or broadening searches. The search page consists a simple search box with the option of advanced search interfaces. The advanced search box gives the ability to limit the search by location of words, domain, language, country and the
Ask's new system, called "Ask3D," is a much bolder and better advance in unifying different kinds of results and presenting them in a more effective manner. It shows, once again, that Ask places a higher priority than its competitors do on making search results easy to navigate and use. Ask puts them on one page in separate sections, which is found to be the superior approach, because each type of result is displayed more effectively; it's easier to see at a glance what you have.

Figure – 5.4
Ask Search Page

(e) AOL has added some unobtrusive AOL user specific links at the very top of the page, but that's about it. Like Google, AOL has tab-like links help to select between web search, images, video, audio, news and shopping. The design of the site is transparent and easy to read and easy to use. The new features do not interfere with
the searching. And may be this is a sensible approach. Users do like search interfaces that work, they are not there to be impressed by the Ajax and Flash skills of the search engine programmers. Like the other top search engines AOL has also provides the advanced search facility with various search limiting options along with the Page-Specific and Topic-Specific searches. The search preference allow to on/off the search history and the safe search.

**Figure - 5.5**

AOL Search Page

![AOL Search Page](image)

5.2.2 Search Features

(a) **Google** defaults to Boolean operator AND, and also supports use of the Boolean OR. It supports quotation marks for phrase searching. Google also supports + (plus) and − (minus) to require and exclude search terms. Google is case insensitive, meaning that 'gOoGle' is the same as 'GOOGLE' and 'google'. The search operators are case sensitive, meaning that 'OR' is not the same as 'or'. Without using any operators, Google will show pages with all words first, trying to find the words in
order. Google excludes common words (known as stop words) like 'I', 'the', 'a' etc. Following are the basic operators supports by Google search.

**Google’s Basic Search Operators**

+ **Required:** Forces Google to include the word. Can be used with stop words.

- **Excluded:** Excludes the word, ensuring that it will not appear of result, as in 'pig-bacon'.

**OR Operator:** Gives you a choice. You can have 'one' OR 'the other', as in 'cats OR dogs'. This operator must be in all CAPS. You can also use the symbol for OR, '|', known as the pipe.

" **Complete Phrase:** Putting words in quotes tells Google to search for the complete phrase. This can also include stop words, as in "the search engine".

~ **Similar Words:** Search for similar words, or synonyms. Searching search ~tips will returns results with 'help', 'guide', 'tutorial' etc.

* **Wildcard Search:** The '*' is called a wildcard. Searching for 'looking for *' will return results like 'looking for dogs', 'looking for cats' etc.

.. **Number Range:** Search within a range of numbers. Searching for 'computer $500..1000' will find computers between $500 and $1000 dollars.

(b) Yahoo also has its crawler and ranking system, which gives it a unique search voice that differs from Google. Yahoo is unique among the major search engines in running a paid inclusion program that means site owners can pay to have URLs listed within the same database used for finding web search matches. Having to mention paid inclusion at all is necessary but distracts from Yahoo still being an
excellent search engine with the most mature search technology next to Google. Indeed, since Yahoo's search technology incorporates that from the old Inktomi and AltaVista search engines that predate Google, Yahoo arguably is even more mature than Google. Of all the major search engines, Yahoo has by far the busiest home page.

(c) Live Search (formerly Windows Live Search) is accessible through Microsoft's Live.com and MSN web portal. The new search engine offers some innovative features, such as the ability to view additional search results on the same web page (instead of needing to click through to subsequent search result pages) and the ability to dynamically adjust the amount of information displayed for each search-result (i.e. just the title, a short summary, or a longer summary). It also allows the user to save searches and see them updated automatically on Live.com. Microsoft has made a number of changes to its Live Search offerings, including pushing some services out of beta and adding new features to others. It is working toward unifying all of its various platforms into a single service. The related searches not only make it easier for users to continue to use shorter queries, but also help Microsoft better understand searcher intent when people use them to drill down on the subject they're looking for. The new version of Live Search delivers improved results organized onto a single page so that customers can spend less time clicking through search results and other related Web sites and readily find what they are looking for instead. Recent significant updates include improvements to the core technology powering search, improvements in user experience and presentation of results, and deep investments in specialized content in key vertical search areas, including shopping, multimedia, health, entertainment, mapping and local search. Collectively, these
improvements mark a milestone in the company’s efforts toward delivering a highly differentiated search offering for both consumers and advertisers.

(d) Ask is working on new search algorithm Edison. Edison is the code name behind merging of Ask.com’s two different search technologies they own, i.e. Teoma and Direct Hit. It is a next generation algorithm that, among many other things, synthesizes modernized versions of Teoma and DirectHit technologies. Ask actually does something different than what the other major search engines do. With the rest they look at all the links from across the web to determine which are important, and then they use those values as part of the ranking process. With Ask's ExpertRank, when someone do a search, they examine the linkage patterns only within the documents that matched the words and concepts that searched for originally.

(e) AOL has probably lots -- enough to make it well worth considering for a web search. The key thing is that AOL can control the ranking algorithm to some degree. AOL has a little module or area at the top of the results with drop-down boxes. AOL also has modules that show up for a range of searches designed to provide direct answers and information. The existing modules are similar to Google Universal Search units that appear.

5.2.3 Advanced Search Options

(a) Google’s Advanced Search: Google supports the following advanced operators.

*date:* Last X Months Search for 'new' pages Google added in the last X months, such as 'google date:3'
**daterange:** Between Two Dates - You may limit your search to an exact date of a "range" of dates that a page was added by Google. Dates must be entered in julian format.

**filetype:** Filetypes - Google will return files of this type. Currently officially supported file types are pdf, ps, wk1, wk2, wk3, wk4, wk5, wki, wks, wku, lwp, mw, xls, ppt, doc, wks, wps, wdb, wri, rtf, swf, ans, txt, but other are supported as well, like xml, cpp, java etc.

**site:** Restrict To Site - Restricts the results to the given domain.

**cache:** Google Cache - Displays the page from Google's saved cache. Other words in the search will be highlighted in the returned page.

**link:** Sites That Link To – Returns pages that link to the your search.

**related:** Related Sites - Returns pages that Google thinks are related.

**info:** Website Information - returns some information about the given web page. Typically website and description.

**define:** Define Word - Searched for the the definition of a word.

**phonebook:** Google Phonebook - search in residential phone book. Google provides the phone number, address and a link to a Google map.

**stocks:** Stock Symbols - Returns stock information for the symbol provided.

**weather:** Weather Forcast -Returns the weather forcast.

**music:** Music Search - Returns track and album information for the search term.

**movie:** Movie Search - Returns all movies related to the search term given.

**allintitle:** All Words In Title - Search for documents with all the words in their title. This operator cannot be combined with others.
**intitle:** Word In Title - Search for documents with the first word after the intitle operator in their title.

**allintext:** All Words In Text - Search for documents with all the in their text. This operator cannot be combined with others.

**intext:** Word in Text - Search for documents with the first word after the intext operator in their text.

**allinurl:** All Words In URL - Search for documents with all the words in their url. This operator cannot be combined with others.

**inurl:** Word In URL - Search for documents with the first word after the inurl operator in their url.

**allinanchor:** All Words In Anchor - Search for documents with all the words in an anchor. This operator cannot be combined with others.

**inanchor:** Word In Anchor - Search for documents with the first word after the operator in an anchor.

(b) Yahoo Advanced Search helps to find sites that match very specific criteria. To use advanced search options, Yahoo provides the advanced link to the right of the search box. The more restrictions we set on the search, the fewer results are returned. Advanced search box provides the following options,

**all of these words:** Includes all of the words typed in the search box. This is similar to inserting "AND" between words or the symbol "+" before a word.

**at least one of these words:** Searches for results that match either one or more of the words. This is similar to inserting "OR" between the words.

**the exact phrase:** Searches for the words in exactly the order we enter them. This is similar to putting quotes (" ") around a set of words.
none of these words: Excludes words from the search. This is similar to inserting "NOT" between the words or the symbol "-" before a word.

Updated: To limit results to sites those have been updated within the past 3, 6, or 12 months.

Site/Domain: Restrict search to sites with a specific domain suffix, such as .com, .edu, or .gov. To specify more than one domain, or one not listed in the drop-down menu, click the upper radio button and type the domains into the text field, separated by commas.

Creative Commons Search Beta: This special Yahoo! Search finds pages that have content with a Creative Commons license. It helps to find content on the Web—text, books, educational material, etc.—that is free to share or build upon.

File Format Restriction: We can choose to restrict search by a specific file type, other than the default setting (all file-types). Currently supported formats are: .htm, .html, .pdf, .xls, .ppt, .doc, .xml, .rdf, .rss, .txt.

SafeSearch Filter: We can choose to have explicit, adult-oriented content filtered out of the search results. We can keep this filter in place for all the web searches by selecting the option in "Preferences".

Country: Yahoo normally searches through web sites from all over the world. We can also choose to see sites just from a particular country.

Language: To limit the search to certain languages, we can select them here.

Search Subscriptions Beta: By partnering with publishers Yahoo! has made popular subscription content searchable. We can search subscription content from a range of publications, including Consumer Reports, Forrester

**Number of Results:** Yahoo Search normally shows 20 web results per page. We can choose to see a different number, between 10 and 100.

**Search Meta Words:** Here is a list of special keywords that give results unique to that special keyword instruction. We can enter these Meta words directly into the Yahoo Search box. These keywords help to find web pages that are similar to, or that link to a particular web site. These are:

- **site:** This allows one to find all documents within a particular domain and all its subdomains.
- **hostname:** This allows one to find all documents from a particular host only.
- **link:** This allows one to find documents that link to a particular URL.
- **url:** This allows one to find a specific document in our index.
- **inurl:** This allows one to find a specific keyword as part of indexed URLs.
- **intitle:** This allows one to find a specific keyword as part of the indexed titles.

(c) Ask's advanced search form gives the power to specify exactly the results the searcher wants. Following are the advanced options provides by Ask Advanced Search,

**Find results with:** Allows to narrow search with term(s) we specify or exact terms we specify. We can specify term(s) to include or exclude from the search.

**Location of words:** Allows to specify where the search terms appear: Anywhere on the page, In page title, or In URL.
**Domain:** Allows to search for term(s) only within a particular domain.

**Language:** Allows to search for term(s) only within Web sites that appear in a particular language. We can select one language at a time.

**Country:** Allows to search for term(s) only within Web sites from a particular country.

**Page modified:** Allows searching for term(s) only within Web sites that have been updated within a certain amount of time.

Any advanced Web search operators can be used in any Web search boxes and trigger an advanced Web search. Below is a list of search operators supports by Ask.com:

" ": The terms surrounded by quotation marks. This will search for exact phrase entered.

+: Plus sign in front of the term(s) specified. This will include specified term(s) in addition to the original term(s) entered.

-: Minus sign in front of the term(s) specified. This will exclude specified term(s) from the original term(s) entered.

**OR:** OR (uppercase) between two terms. This will include either of specified multiple terms in addition to the original term(s) entered.

**Intitle:** intitle: followed by the specified term(s). This will search for documents whose page title has the specified term(s).

**Inurl:** inurl: followed by the specified term(s). This will search for documents whose URL has the specified term(s).

**Site:** The term(s) followed by site: and domain. This will search for documents found only within the specified domain.
Last: The term(s) followed by last: and time variable. This will search for documents that are found within the specified period.

Supported time period:
Within one week = last:week
Within two weeks = last:2weeks
Within one month = last:month
Within six months = last:6months
Within one year = last:year
Within two years = last:2years

Afterdate: The term(s) followed by afterdate: and date in yyyy-mm-dd format. This will search for documents that are found after the specified date.

Beforedate: The term(s) followed by beforedate: and date in yyyy-mm-dd format. This will search for documents that are found before the specified date.

Betweendate: The term(s) followed by betweendate:, starting date, a comma, and ending date. Dates must be in yyyy-mm-dd format. This will search for documents that are found between the specified starting date and ending date.

Inlink: inlink: followed by the specified term(s). This will search for documents that contain the specified term(s) in anchor text.

(d) AOL did until recently excel in the field of advanced search. The present AOL Search does make use of operators like OR, AND, and NEAR. There is also a well-
designed form based page for advanced search. We can use the Advanced Search feature to get more precise results or control over the results.

- Include or exclude specific words in your search results.
- Specify language of sites that you would like to see in your search results.
- Specify file format, such as PDF, Word, Excel or Shockwave Flash.
- Sites with specific numeric range.
- Search for keywords that you're searching for anywhere on the page, such as in the title, text, URL or links.
- Include or exclude specific Websites or domains in your search results.
- Filter search results based on usage rights.
- Filter search results using SafeSearch.

*Domain Specific Search*: We can specify search within a specific domain name only. We can also specify a domain to exclude from the search. Instead of selecting "only" in the drop down menu, select "Don't." This domain specific search feature is also available without going to the Advanced Search page. Just enter the search term followed by site:domain (no space between site: and domain) on AOL Search home page or search results pages.

*Usage rights*: We can use the usage rights feature to find content that can be share or modify, such as for music, images and movies. If we leave the filter as "not filtered by license," then we'll get search results regardless of the rights. The usage rights filter is based on Creative Commons license. We can filter by usage rights for results that are:

Free to use or share; Free to use or share, even commercially; Free to use, share or modify; Free to use, share or modify, even commercially.
**SafeSearch:** The optional SafeSearch feature automatically filters out adult sites with explicit content. SafeSearch status is available right below the AOL Search box, or in the Preferences.

**Preferences:** We can update AOL Search preferences to turn on or turn off the Search History and Safe Search.

**Search History:** The Search History feature keeps track of the searches for the last 30 days, making it easy to get back to previous search results. We can also see the search results that clicked on earlier.

**AOL Toolbar:** AOL Search toolbar allows users to quickly search the Web, block unwanted pop-ups and have easy access to their AOL or AIM® mail.

5.2.4 Limit Options

All the five major search engines have provided different links to limit the searches.

**Google** has provided the following option to limit the search:

language, domain, date, file format, occurrences, usage rights, and safe search.

**Yahoo** has given the option to limit the search by language, domain, country, file format, number of results, location of search terms in the document (text, title, link name, URL and link to the URL), and "word filters".

**Live** has given the option to limit the search by language, domain, geographical location of the source, and keywords appearances.
Ask has given the option to limit the search by location of words, domain, language, country and the page modification date.

AOL has also provided the various search limiting options along with the Page-Specific and Topic-Specific searches.

5.2.5 Results Display

(a) Google's results include document title, first few words of text, url, size and a link to a previously cached version of the page. It also analyzes the frequency of search terms and their proximity to each other in the document. When there is more than one result from a site, Google groups the results by site, displaying only the first two results from that site followed by a link to find more results from that site. Clicking the "Similar Pages" link retrieves pages that are "related" to the current result. To make the search simpler, Google's universal search now groups different kinds of results within its single, familiar main list of results. At the top of the page, under the Google logo, are the two most relevant search categories -- in this case, Web and News -- if someone wants to separate types of results. A Google universal search for a prominent person might bring up images at the top but also a video, which can be played without leaving the search page. At the bottom of the page is a list of related searches.
Google’s supplemental index

According to Google’s FAQ page, supplemental results are part of Google’s auxiliary index (main results are drawn from the main index) and pages, which appear on the supplemental listing, have “fewer restrictions” than those that appear on the main results page. They further say that the inclusion of sites on the main or supplemental index is purely automated and does not affect page rank at all. Supplemental results are only included when the main index fails. In truth however, pages that appear on the main index will almost always show up first in a search. Supplemental search results will only show up if there are very few or no results at all in the main index. Plenty of older web sites also tend to populate the supplemental results page.
(b) **Yahoo** result displays title, brief summary and url for each document found, number of results, time taken for finding the results. Yahoo returns comprehensive results that match the words into four main sections: Inside Yahoo, Directory Category Matches, Sponsor Matches and Web Matches. Yahoo does not provide refining results in it result page. Using the “Page Display & Layout Settings” we can set the open result in a new window, number of result per page, instant search, subscriptions, and multimedia search.

**Figure – 5.7**
Yahoo Search Results Page

(c) **Live Search** delivers traditional search complemented by Images, News and Local query features. Additionally, the search engine provides users with the possibility of tapping resources associated with services that are still in beta phase, including QnA, Video, Academic and Feeds search. Live Search results include document title, first few words of text, url of the document, and a link to a previously cached version of the page. At the top of the result page display the
number of the results founded. When there is more than one result from a site, it groups the results by site, displaying only the first two results from that site followed by a link to find more results from that site. Clicking the "Similar Pages" link retrieves pages that are "related" to the current result. At the bottom of the page there is a link to submit searchers view on Live Search results. Live Search delivers the results they are looking for simply with a unique design and advanced tools for helping people quickly find, view, organize and preview search results. The new version of Live Search delivers improved results organized onto a single page so that customers can spend less time clicking through search results and other related Web sites and readily find what they are looking for instead.

Figure – 5.8
Live Search Result Page

(d) Ask result presents the title of the documents, few words from the documents, and url of the resulted document. Instead of sticking with a single list of search results that mixes various types of material, Ask presents result in a page divided
into three panels. The middle column is the largest one and could be called the Results Panel. It's where we can see the results. The left panel could be called the Query Panel. The thinner left panel contains the search box and links to other sections of Ask, but mainly it displays suggestions for refining search, or for making related searches. The left panel shows suggestions for how to narrow or broaden the search. A somewhat wider right panel contains search results that go beyond Web links, such as images, news headlines, encyclopedia articles, videos, weather information, local time and, where relevant, music clips one can play without leaving the page. This could be called the Blending Panel, because this is where Ask blends results from other search databases it has. In this case, the column starts by showing image search results, and one can hover to easily make any of the images that appear there larger. If one hover over the image thumbnails, they enlarge. If one hover over the video thumbnails, they play. Ask also gives larger previews of the pages it lists, visible by just hovering over a binocular icon. And we can get to the advanced search panel without leaving the page. It listed, lots of summary information at the top of the page and prominent suggestions for narrowing or broadening searches. Ask's new system, called "Ask3D," is a much bolder and better advance in unifying different kinds of results and presenting them in a more effective manner. It shows, once again, that Ask places a higher priority than its competitors do on making search results easy to navigate and use. Ask display the sponsored results at the top and bottom of the result pages.
(e) AOL Search Results:
A new survey from McAfee finds that AOL has the "safest" search results in terms of not listing sites that might be somehow be risky or unsafe for searchers. Like other search engines, Ask results page also provides the title of the documents, url of the documents, few words of the documents, and a link of similar pages. Web results obtained using AOL Search are Websites listed in order of relevance, with listings of the highest relevance appearing first. These listings are administered, sorted and maintained by Google, AOL's search partner. AOL delivers comprehensive web results from Google along with One Click access to relevant videos, pictures, local maps, news, stock quotes and more. One Click is a great time saving feature that helps to discover targeted information on what is looking for -- with fewer clicks and fewer hassles. One Click appears above the general Web search results and highlights relevant content from AOL and its partners. Sponsored links that appear
in AOL Search results or on AOL channels are listings that have been purchased by companies to have their businesses or Web sites appear for specific search terms related to their services. Sponsored links, also sometimes referred to as pay-per-click ads, are administered, sorted and maintained by Google and AOL. Like the Ask, the AOL also display the sponsored results at the top and bottom of the result pages.

![AOL Search Result Page](image)

**Figure – 5.10**

AOL Search Result Page

5.2.6 Risk Factor

A new survey from McAfee finds that AOL has the "safest" search results in terms of not listing sites that might be somehow be risky or unsafe for searchers. In terms of unsafe results, most of the major search engines are pretty clean. McAfee ran 2,500 queries, looked to five pages deep, then assessed the "safety" of the sites that came up. The overall risk breakdown is, Yahoo: 5.1%; MSN: 4.6%; Ask: 4.2%; Google: 4.2%; AOL: 3.6%. 
The vast majority of results are perfectly safe, a pretty good job given the automated nature of these tools. The figures above include both "red" dangerous sites (such as giving out adware, sending you spam if you give an email address) and "yellow" sites that may deserve caution (pop-up ads, try to change your browser settings). Here's the breakdown of percentage of red sites in the results, i.e. Yahoo: 3.1%; Google: 2.7%; Ask: 2.6%; MSN: 2.5%; AOL: 2.3%. The report also looked at the safety of results from a paid and non-paid perspective. It's a huge difference, and one that's especially worrisome given that the paid results are supposed to be human reviewed at some point. In terms of paid results safety by search engine, it breaks down percentage is like, MSN: 10.7%; AOL: 8.1%; Yahoo: 8.0%; Google: 7.3%; Ask: 6.5%.

On average, each page of results generally has about the same percentage of unsafe listings (4.3 to 4.5 percent). Ranking position is also covered. In general, a top ranked site isn't any more or less safer than a site in position 10. Sponsored results show more difference, especially for ads that have the number one position across the top or on the right hand side of the page. These are generally much safer.

5.2.7 Other Features

In addition to running the world's most popular search services, Google and its rival search engines, i.e. Yahoo, Live, Ask, AOL, also runs a large number of vertical search services. These services also make the above search engine more interesting and luring. The services offered by these search engines are described below.
Google


**Google Book Search:** On October 6, 2004, Google launched Google Book Search to search the full text of books.

**Google Scholar:** On November 18, 2004, Google launched Google Scholar, an academic search program.

**Google Blog Search:** On September 14, 2005, Google announced Google Blog Search.

**Google Base:** On November 15, 2005, Google announced the launch of Google Base, a database of uploaded information describing online or offline content, products, or services.

**Google Video:** On January 6, 2006, Google announced Google Video.

**Single-Word Translations:** Its translation service now offers single-word translations. One can do the word translations at http://translate.google.com/translate_dict?hl=en

**Google Code Search:** To find out code snippets and other geeky programming stuff. The content of the search has been expanded to include individual code files and snippets from Web pages.

**Google Voice Search 'Map It':** Google's free directory assistance/voice search offering has added a "map it" feature. Users can say "map it" after receiving the desired listing and get a text message with a link to a map (which opens a WAP browser).
**Google Experimental:** Google add a new experimental version of its popular search service called Google Experimental, available on Google Labs. This new test site provides users an opportunity to try out some of the latest search experiments and innovations and provide Google with feedback.

**Google’s Web History:** Web History lets you look back in time, revisit the already browsed sites, and search over the full text of pages already seen. Google uses the Toolbar to associate the visited pages with the Google Account, and builds up a database of search and browsing history.

**Google File Search:** Google ends up with an unprotected directory that lists all of its files and subdirectories, when directly accessed from a browser. If someone links to the directory or submits it to Google, it becomes available to anyone who performs a search.

**Google Gmail:** While Gmail dubiously maintains its “Beta” tag, Google’s email service put the company on the map in being much more than a search engine and going after the big boys. The company has also launched Gtalk in an effort to join the instant messaging space, though the product lags behind competitors in adoption. Maintaining webmail and instant messaging services requires massive amounts of manpower and hardware, making it for the most part a four-way race between the big boys: Yahoo, Microsoft, and AOL.

**Google Webmaster Central:** It includes a site status wizard which can be used to check if a site is being spidered by Google, Webmaster tools, Info on how to submit your site to Google, A new Google Blog for Webmasters with information on index updates, crawlers, and other technical issues, then there is the Google discussion group for webmasters, and, finally, a webmaster help center, which gather many of the most relevant help pages.
(b) Yahoo

**Automatic Search Query Suggestions:** It works like when one do a search on Yahoo, it will automatically be given suggestions based on what have been typed. So, not only does this limit the number of words one need to type into a query, it may also present similar queries to shorten the searching time and get to what someone is looking for quicker. The important thing of search suggestions is that it is not default at http://search.yahoo.com. One can turn them off by clicking on "more" and then selecting "Disable Search Suggestions" in the Yahoo search page.

**Image Search:** Yahoo Images gives matches that come from Yahoo's crawling of the web, which means it will have a different collection than what Google gathers. In addition, Yahoo Images also provides matches that come out of its Flickr photo-sharing site. Image search also allows to narrow results by size, including those matching wallpaper formats, plus to select either color or black & white.

**News Search:** Yahoo has one of the oldest news search sites. A highlight is the long-standing Full Coverage area, where human editors are involved in creating custom pages for various news topics. Indeed, the human component key difference between Yahoo News and Google News.

**Yahoo Local & Maps:** Yahoo Local has been constantly expanded over the years to provide tons of information about local businesses and events. Yahoo Maps emphasis on maps, though local information is also offered. Yahoo was the first major search engine to offer it.

**Yahoo Answers:** Yahoo Answers is a huge, thriving community of volunteers who answer questions put to them. One can browse through a vast collection of questions already answered or put out own question for others to answer.
**Yahoo Shopping:** Yahoo Shopping help to find matching products from across the web.

**Yahoo Video:** Video content from the web, uploaded to Yahoo and from other sources.

**Yahoo Audio:** Find music, podcasts and more from across the web. Yahoo Podcasts is a dedicated area for podcasts. Yahoo Music provides music clips, videos, songs for purchase, information and more -- including lyric search that was recently added.

**Yahoo Directory:** The web's oldest guide to web sites that are organized into categories by humans, rather than through the work of automation.

**Yahoo Travel:** Find hotel, flight and other information relating to travel.

**Yahoo Mobile:** Either access Yahoo's blended mobile search results, Yahoo oneSearch, using phones internet browser or download the dedicated Yahoo Go mobile search tool.

**Yahoo Site Explorer:** This is Yahoo's dedicated area just for site owners looking to learn more about how they are listed. The Yahoo Site Explorer lets to analyze the standing of one's own site. This is a popular tool among search engine marketers as it gives a much more accurate picture of your number of back-links than Google. Google's search engine is totally unreliable in this respect. If you sign in and add a small text file to your site’s server, you may also use this service to find out if Yahoo! is able to find your site. You may also give it the URL of your RSS feed, helping Yahoo’s spiders to index your new pages.

**Yahoo Mail:** Yahoo Mail recently upgraded to offer an Outlook-like interface that includes the ability to chat with Yahoo Messenger buddies from within user account. Yahoo’s IM product is among the top commercially run products.
**Yahoo Webmaster Help:** A link to the Yahoo Search Blog, Search Help, Search Content Guidelines, form for reporting spam, Support Feedback form, separate suggestion form, and finally, there is a submit page, which gives information on how to submit a site for free, the pay per click text ad programme and Yahoo’s Search Submit Express, where one can pay Yahoo to do something they should do for free anyway.

**Yahoo Companion:** Allows users to search Yahoo, access Yahoo Mail, checks on stocks and more via a toolbar within the browser.

(c) Live Search

**Image Search:** Live has its own image database from crawling the web that gives it a different collection than what Google has. Aside from its own database, Live has one of the nicest image search interfaces. Hover over an image, and it gets slightly larger plus provides info such as the file name, dimensions and URL. A slider also lets to see many images at the same time shown in very small thumbnails moving up toward a few images in larger thumbnails. Live Search Image has increased the image index and changed its user interface. With this increase, consumers are now able to search across millions of images and receive precise results with smart scroll, thumbnail image-sizing customization, filmstrip view and the ability to see the full image without leaving the results page.

**News Search:** Live search users have a pass on using Live news search over Google. That’s because it doesn’t seem to hit as many sources as Google but rather sticks with large, mainstream media. In addition, it annoyingly insists on showing only UK news sources.
**Local & Maps:** Formerly called Virtual Earth, Windows Live Local provides unique 45 degree "Bird's Eye Views" of select areas that allows to see buildings and locations not from typical looking straight down aerial view but instead often glorious, detailed side views. Windows Live has long offered easy to use capabilities like this. Live Search Maps include a new user interface, better ways to map local business listings, and quicker, more accurate directions with better printing features.

**Academic Search:** Using by the Academic button one can search against scholarly material that Microsoft collects through partnerships with publishers. Academic searches allow consumers to search through academic journals and published papers.

**Feed Search:** Again, by selecting Feeds in the More Button option, one can do search through web feed content that Microsoft has collected. Live Search Feeds searches for RSS feeds based on the user's search query.

**Book Search:** Book search allows book search against in copyright books gathered in cooperation with publishers, as well as out-of-copyright books. Live Search Books is advancing the way people search online by digitizing and indexing information from printed material from around the world.

**Video Search:** Video Search is powered by AOL-owned Truveo. Also offer MSN Video or the beta version of MSN Video's future look, known as MSN Soapbox. Live Search Video allows users to search for and browse online video content, and watch high-quality videos from the Web. With a new smart motion preview feature, consumers can now preview video within the search results by rolling over the thumbnail of the video.

**Live QnA Answer Search:** Live QnA is Microsoft's challenge to the powerful Yahoo Answers service, where people ask questions and get answers back from others in the community there. QnA lets users to search from a database of questions
and answers, and at the same time allow users to post their questions online on to the site where the public can answer them.

**Live Search Mobile:** We can access Live Search Mobile in a variety of ways.

**Live Mail:** Microsoft owns the original big boy in webmail – Hotmail – which recently received a major facelift and became Windows Live Hotmail.

**Live Search Health** helps consumers refine health searches quickly using related medical concepts to get relevant health information from a broad spectrum of perspectives, and helps them navigate complex topics with inline article results from experts like the Mayo Clinic — all on a single page.

**Live Search Macros** is a powerful tool that allows consumers to create their own custom search engines or use customizations created by others who share their interests to make their searches more relevant.

**Live Search Products** helps consumers easily find and discover products, reviews, guides, prices and other relevant information — all within the Live Search interface, giving them confidence in their buying decisions.

**Live Search Scopes** allow people to see search results from Web pages, local maps and directions, images, videos, and more without the need to re-enter the same query each time or to leave the page.

**Live Search Instant Answers** provides answers to questions at the top of the algorithmic results, so consumers spend less time searching for the information they need. Several new and updated Instant Answer features include Live Search traffic, weather, maps, entertainment, news, product, stocks and health.

**Live Search Entertainment** helps consumers stay informed of the latest entertainment news with inline celebrity answers accompanied by images, video, facts and buzz. In addition, Live Search’s new xRank feature allows consumers to
find out how their favorite celebrities rank in regard to total Web searches via a daily updated ranking.

(d) Ask

*AskCity*: It is the local search tool of Ask.com, which has included drawing tools to mark maps with text. It integrates the best local information on the Web with a distinctive, "all-in-one" user interface, which also allows search for Businesses, Events, Movies, Directions, Reviews, the ability to buy tickets and make reservations, and a lot more. Ask remains the only major search engine that gives consumers the walking directions too.

*Ask for Kids*: Formerly known as Ask Jeeves for Kids, is a fast, easy and kid-friendly way for kids to search online. Designed to be a fun destination site focused on learning and "edu-tainment," Ask for Kids uses natural-language technology that allows kids to ask questions and perform web searches.

*Blogs & Feeds*: The Blogosphere combines the Ask.com search technology and Bloglines subscription data. It indexes the most popular blogs that people across the web subscribe to daily to ensure higher-quality results and minimize blog spam.

*Conversions*: The Ask Conversion Tool converts world currency, temperature, weight, length, area, and even cooking/volume conversions. There is no need to search the web for multiple conversion tools when one can do it quickly on Ask.com.

*Dictionary Search*: Dictionary search provides the definition or pronunciation of a word, origin of a word or a new word altogether.

*Movie Search*: Quick access to the latest information for upcoming and recent film releases.
**News Search:** To quickly find all the breaking news – local and international. Like Google and Yahoo and others, no list is provided in News Search, but it is far more than simply tapping into wire services.

**Image Search:** Just as Ask crawls the web for pages, it also crawls to build its own unique image search engine to find celebrity images, pictures of vacation destination. It has a unique image search "voice," as opposed to Google.

**Shopping Search:** Find the designer dress, unusual gift, occasion gifts and the location of your favorite stores. Search for the latest electronics, clothes and products based on reviews, features, price, and store location.

**Smart Answers:** Ask's Smart Answers provide key information right at the top of the page. Smart Answers can be found on the web results page of many searches, like for a band or famous person, and contain a brief bit of information and links to more information. Smart Answers include information you might get by opening up an encyclopedia, dictionary or another authoritative and generally accepted resource. The sites providing this type of information to Ask are predominantly big informational databases like the World Fact Book, Wikipedia, The American Heritage Dictionary, Roget's Thesaurus, etc.

**White Pages Search:** Quickly find phone numbers and addresses for people, businesses, government offices, doctors, and schools in the U.S.

**Zoom Related Search:** To narrow or broaden search with possible alternative search terms which appear on the right hand side of the Ask results page.

**Narrow Your Search:** helps to drill down into topics that are specifically related to the search

**Expand Your Search:** Allow to explore topics that are conceptually related to the search
**Mobile Search:** Ask.com has launched a mobile search product at m.ask.com or mobile.ask.com. Ask Mobile is optimized for mobile devices and uses Skweezer throughout to improve readability and download speed on the mobile device. Ask Mobile offers links to key search categories directly from the home page, making it easier to navigate to relevant results.

**Health Smart Answers:** Ask introduced new Health Smart Answers to help searchers find the trusted health information they are looking for faster.

**Ask Blog & Feed Search:** Yet another unique search voice from Ask is to search against blog and feed content. By default it searches against posts, ranked by relevance. Users can also sort purely by most recent or by popularity, which includes the number of readers to a particular blog via Ask-owned Bloglines. There's a drop-down box that lets to subscribe to the blog the post is from via several popular services.

(e) AOL

**Related searches:** As soon as one start entering search query in the search box, AOL will immediately try to guess what the user is looking for, presenting related web searches in a small window below the search form. This feature is somewhat related to previous experiments where the search engines try to fill out the search form for the searchers. This is less interfering and less confusing.

**Snapshot modules:** The regular search engine result pages are clean, uncluttered and not the least original. For some searches, however, interesting things start to happen. A search for the pop group the Corrs brings up pictures, a presentation and video clips in the right hand corner. This is what AOL calls a snapshot module. The
system will apparently also bring up relevant news content for "current event" queries.

**History button:** There is another difference from Google's regular results though. AOL has added a history button for searchers logged in as AOL users. Hovering over the search button (no click is needed) brings out a list of recent searches to repeat or edit any searches.

**AOL Video Portal:** AOL has upgraded and relaunched a range of features and sites on its network. The latest on the list to get a revamp is the Video Portal. The redesigned site offers a new home page, an upgraded video search experience via its video engine Truveo and an embedded playback experience. There's also a new Windows Vista gadget that offers AOL video content on the desktop, has customizable "skins" and a number of dynamic features including a "real-time ticker" that indicates when new videos are added to the Truveo index. Video search results include multimedia streams and files (Real Media, Windows Media, QuickTime, MP3 and Flash) from the Web sorted by relevance. Video search results are powered by AOL Video and Truveo.

**Local and Shopping search:** AOL has also launched a beta Local Search, leveraging its ownership of MapQuest and AOL's CityGuide. The new local search service offers geo-targeted advertising and reviews of local businesses, bars, restaurants, and more. And AOL Shopping has dumped longtime partner Shopzilla in favor of new shopping partner Pricegrabber.com. Shopping Search results are powered by PriceGrabber, and are drawn from across the Web.

**AOL Paid Search Service:** AOL has introduced a paid search service recently backed by Google's technology to help advertisers' better target AOL users. AOL Search Marketplace is part of a December 2005 agreement with Google in which the
Web search leader purchased a $1 billion stake in AOL. AOL Search Marketplace uses a white-label version of Google AdWords, providing advertisers the ability to buy search advertising that solely targets the AOL Search audience on the AOL client applications and AOL.com. AOL Search Marketplace, a customized version of AdWords, now lets advertisers reach AOL Search customers more effectively.

**Image Search:** AOL's image search is powered by Google, but as with web search, searcher might find slight differences because the exact same image database might not be hit and there might be ranking differences. Image search results are sorted by relevance, with the most relevant images appearing first. These listings are administered, sorted, and maintained by Google. Google considers a number of factors when determining whether an image is relevant to your search request. Because these methods are not entirely foolproof, it's possible some inappropriate pictures may be included among the images you see. AOL® image search uses a mature content filter called SafeSearch by default.

**News Search:** AOL recently revamped its AOL News site, which allows to narrow the search to Business, Sports and Entertainment news. But that's mainly because of the fewer sources being checked. News search results include news articles from an extensive database of news and multimedia content from news providers such as Associated Press®, Reuters®, CNN®, People Magazine and Fortune®. The articles that are presented within the Top News category come from all the providers and sources offered on AOL News search. These sources are updated regularly. The other categories—Entertainment, Sports and Business—contain articles that are relevant in those categories.

**AOL Audio Search:** The Audio search offers not just podcasts, it scans the web for free music as well as sound clips.
**AOL Mail:** Millions of people remain attached to their AOL email addresses. AOL has made it easy to do, allowing users to keep their address even if they cancel their AOL accounts. This makes AOL.com’s webmail product one of the more popular, albeit less advanced email services. Meanwhile, AOL Instant Messenger (AIM) is a major player in IM, particularly in the US.

### 5.3 Common Features and Functionalities of the Five Search Engines

Almost all search engines described above provides basic text search facilities like Boolean search, proximity search, phrase search, truncation, field-specific (often called meta tag) search and limiting search (by language, domain, country, and so on). The search operators vary from one search engine to another. However, the search engine interfaces have improved significantly in ease of use compared with the interfaces of many traditional text retrieval systems. It may be noted that major improvements in the web search environment has been in the design interfaces that allow users to conduct complex search queries without having to remember or use complex search syntax or operators. Users can formulate comparatively complex search expressions simply by choosing the appropriate options and entering the query terms in the advanced search of different search engines. Many advanced search screens also provide enough help information for user to complete their search on the screen itself. In addition to those retrieval features that are common to the traditional text retrieval systems, many search engines provide some unique search facilities that have been designed specifically for web information retrieval. The following table shows the search features and functionalities available in the above five search engines.
## Table 5.1
### Features and Functionalities of the Five Search Engines

<table>
<thead>
<tr>
<th>Search Mechanism</th>
<th>Google</th>
<th>Yahoo</th>
<th>Live</th>
<th>Ask</th>
<th>AOL</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Indexing Method</strong></td>
<td>Latent Semantic Indexing</td>
<td>keyword specific indexing</td>
<td>keyword specific indexing</td>
<td>keyword specific indexing</td>
<td>keyword specific indexing</td>
</tr>
<tr>
<td><strong>Relevance Ranking</strong></td>
<td>Based on a special algorithm for page popularity measured in links to it from other pages</td>
<td>By analyzing the web page text, title and description accuracy as well as its source, associated links, and other unique document characteristics</td>
<td>Use Neural networks which is based on a large number of factors to make decisions</td>
<td>Use algorithm ExpertRank and link patterns on the web in the ranking of results.</td>
<td>Based on Subject-Specific Popularity, links to a page by related pages.</td>
</tr>
<tr>
<td><strong>Crawler and Indexer Features</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Robot Exclusion Standard Support</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Duplicate Page Detection</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>File Format Indexed</strong></td>
<td>html, htm, text, shtml, PDF, embedded Shockwave flash objects, doc, rtf, LaTex/Tex, WordPerfect, Xml, and MPEG Layer 3</td>
<td>html, htm, text, shtml, PDF, embedded Shockwave flash objects, doc, rtf, LaTex/Tex, WordPerfect, Xml, and MPEG Layer 3</td>
<td>html, htm, text, shtml, PDF, embedded Shockwave flash objects, doc, rtf, LaTex/Tex, WordPerfect, Xml, and MPEG Layer 3</td>
<td>html, htm, text, shtml, PDF, embedded Shockwave flash objects, doc, rtf, LaTex/Tex, WordPerfect, Xml, and MPEG Layer 3</td>
<td>html, htm, text, shtml, PDF, embedded Shockwave flash objects, doc, rtf, LaTex/Tex, WordPerfect, Xml, and MPEG Layer 3</td>
</tr>
<tr>
<td><strong>Index Protected Server</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Searching Features</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Boolean Search</strong></td>
<td>Partial. AND assumed between words. Support OR. No ( ) or nesting.</td>
<td>Accepts AND, OR, NOT or AND NOT, and ().</td>
<td>Accepts AND, OR, AND NOT, and ().</td>
<td>Partial. AND assumed between words. Support OR. No ( ) or nesting.</td>
<td>Accepts AND, OR, NOT, and ()</td>
</tr>
<tr>
<td><strong>Phrase Matching</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Fuzzy Search</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Feature</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>Word Forms</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regular Expression</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Numeric Data Search</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Case Sensitivity</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Natural Language Query</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Field Filtering</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Wildcard</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Truncation</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Word Stemming</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>+Requires/-Excludes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Title Search</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Other Features</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language Filtering</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Limit by age of documents</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Domain</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Customizable Result Formatting</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Limit by file types</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Stop Words Search</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Family Filters</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Page translation</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Search Within</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Related Searches</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Clustering</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Near</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>URL Search</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Link Search</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

### 5.4 Conclusion

Recent search engine market share figures show Google with 57.4%, Yahoo with 22.9%, MSN with 8.8% and Ask.com with approximately 3.6% of the search market. These are average figures for data collected by Hitwise, NNR and comScore. Google is the indisputable leader of the search engine market with a share of more than half according to the combined statistics from
Nielsen//NetRatings, comScore and Hitwise. The challenge facing most search engines is getting Internet searchers to break their Google habit. Google has become so inextricably connected with the concept and practice of search, Yahoo has to work a lot harder to get on people's radar screens. Yahoo is starting by trying to change the search habits of the hundreds of millions of people who visit its broad stable of websites. Yahoo is the latest search player to overhaul the way it presents search results in an effort to chip away at Google's lead. Ask.com, which draws 20 million to 25 million users a month, relaunched a dramatically different looking interface, and Microsoft Corporation showed off its new method for displaying search results. To deliver a better experience, technologists pore over behavioral patterns and complex algorithms to find new ways to help people more quickly and easily find the information they are seeking. More and more search engines experiment with integrated interfaces that facilitate the access to more resources without opening new pages. After all, the purpose of a search engine is to find answers. Some of the answers are simple facts; others are entire paragraphs from a web page, chapters from a book, songs or inferences. The answers are everywhere, surfers just need a tool that understands the question and knows where to find the answers. While the way the search engines present these answers, the real issue search engines need to address is that queries are just a bunch of words that rarely create meaningful entities.