
VERTICAL, TOPICAL, SPECIAL SEARCH ENGINES AND VORTALS TO MINE INFORMATION ON THE INTERNET

G Rathinasabapathy

Abstract

A Search Engine is a program designed to find information on the World Wide Web, inside a corporate or proprietary network or a personal computer. The very first tool used for searching on the Internet was Archie (1990) was created at McGill University and the first Web search engine Wandex (1993) was developed by Matthew Gray at MIT. Thereafter, many search engines came into use. There are a wide variety of Search Engines across the Internet and help us to find the most accurate and relevant information across the Information Superhighway. This paper attempts to mine, classify and profile the special Search Engines which are also known as Topical or Vertical Search Engines and Vortals.

Keywords: Search Engines; Vertical Search Engines; Vortal.

1. Introduction

A search engine is a program designed to find information on the World Wide Web, inside a corporate or proprietary network or a personal computer. Search engine is often used synonymously with spider and index, although these are separate components that work with the engine. The Search Engines used to find information on the Internet are known as Internet Search Engines. The very first tool used for searching on the Internet was Archie which was created in 1990 at McGill University. Then Gopher was created in 1991 by Mark McCahill at the University of Minnesota. While Archie indexed computer files, Gopher indexed plain text documents. Two other programs, Veronica (Very Easy Rodent-Oriented Net-wide Index to Computerized Archives) and Jughead (Jonzy's Universal Gopher Hierarchy Excavation And Display), searched the files stored in Gopher index systems. Then the first Web search engine Wandex was developed by Matthew Gray at MIT in 1993. Another very early search engine, Aliweb, also appeared in 1993 which still runs. The first "full text" crawler-based search engine was WebCrawler, which came out in 1994. Unlike its predecessors, it let users search for any word in any web page, which became the standard for all major search engines since. It was also the first one to be widely known by the public. Also in 1994 Lycos, created at Carnegie Mellon University came out. Thereafter, many search engines appeared which included Excite, Infoseek, Inktomi, Northern Light, and AltaVista.

2. Types of Search Engines

There are a wide variety of Search Engines across the Internet which help us to find the relevant information. The Search engines can be classified under the following major categories:

- Major Search Engines
- Metasearch Engines
- News Search Engines
- Pay Per Click Search Engines
- Multimedia Search Engines
- Kids Search Engines
- Country-Specific Search Engines
- Speciality Search Engines

Major Search Engines

The Search Engines which are either well-known or well-used are known as 'Major Search Engines'. These search engines can potentially generate so much traffic so, for webmasters, the major search engines are the most important places to be listed. As far as the searchers are concerned, well-known, commercially-backed search engines generally mean more dependable results. Therefore, these search engines are known as 'Major Search Engines'. Major Search Engines include Google (www.google.com), Yahoo (www.yahoo.com), AllTheWeb.com (www.alltheweb.com), AOL Search (<http://search.aol.com>), HotBot (www.hotbot.com), AltaVista (www.altavista.com), Gigablast (www.gigablast.com), LookSmart (www.looksmart.com).

Meta Search Engines

A 'Meta Search Engine' is a search engine that sends user requests to several other search engines and/or databases and returns the results from each one. The meta-search engines allow users to enter their search criteria only one time and access several search engines simultaneously. Since it is difficult to catalog the entire web, the idea is that by searching multiple search engines we are able to search more of the web in less time and do it with only one click. The ease of use and high probability of finding the desired page(s) make metasearch engines popular with those who are willing to weed through the lists of irrelevant 'matches'. The Meta search engines are also known as 'Metacrawlers'. Eg. Dogpile (www.dogpile.com), Vivisimo (<http://vivisimo.com>), Kartoo (www.kartoo.com), Info.com (www.info.com), MetaCrawler (www.metacrawler.com), MetaEureka (www.metaeureka.com), WebCrawler (www.webcrawler.com).

News Search Engines

The Internet offers a number of search engines exclusively for searching 'News' items. Surfers still looking for news using "normal" search engines may stop doing it and find the services below to be a much better way to search for the latest news stories from hundreds of sources on the web which includes newspapers, magazines and television stations worldwide. E.g. Google News (<http://news.google.com>), Yahoo News (<http://news.yahoo.com>), Ananova (www.ananova.com), NewsNow (www.newsnow.co.uk).

Pay Per Click Search Engines

There are some Search Engines on the Internet which offer fee based services which are known as 'Pay Per Click Search Engines (PPC)'. The services listed by them generally sell paid listings on a

cost-per-click (CPC) basis. So these search engines are also known as CPC search engines. Normally, advertisers pay for every click the search engine sends them, and those who pay the most generally get listed higher. E.g. Google AdWords (<https://adwords.google.com>), Overture (www.overture.com), Enhance (www.enhance.com), Epilot (www.epilot.com), Kanoodle (www.kanoodle.com).

Multimedia Search Engines

Those Search Engines which are helping to locate sound, image and video files as well as radio and television programmes are known as 'Multimedia Search Engines'. E.g. Ditto (www.ditto.com), Google Images (<http://images.google.com>), Kazaa Media Desktop (www.kazaa.com), MIDI Explorer (www.musicrobot.com), Picsearch (www.picsearch.com).

Kids Search Engines

A vast majority of the contents available on the Internet might find inappropriate for kids. Sites that deal with explicit sexual matters, porn sites, violence, hate speech, gambling and drug use etc., should not be made available to the kids. To cater to this need, the Internet provides a number search engines exclusively for kids. Parents may guide their children to make use of these kids search engines to find information across the web. Eg. Ask Jeeves For Kids (www.ajkids.com), KidsClick (www.kidsclick.org), Yahooligans (www.yahooligans.com),

Country-Specific Search Engines

Although a number of search engines that serve the entire world, there are also some search engines which are devoted to particular countries and regions of the world. These search engines are known as Country-specific Search Engines and these can be used to find information pertaining to a specific country or geographical location. E.g. France (www.abondance.com), European Search Engine (www.webmasterworld.com/forum18/544.htm), Japan (www.searchdesk.com).

3. Speciality Search Engine

Finding specific and more relevant information through general search engines is a difficult task. Instead of using the general search engines, it is always better to use the speciality search engines which are available on the Internet. The 'speciality search engines' are also known as 'Topical Search Engines', 'Vertical Search Engines' or 'Vortals'. These speciality search engines will help us to search through specific types of information listings in different areas. The speciality search engines can be divided into the following categories viz.

- Answer Search Engines
- Domain Search Engines
- Government Search Engines
- Invisible Search Engines
- Mailing List Search Engines
- News Group Search Engine
- Shopping Search Engines
- Travel Search Engines
- Vertical Search Engines

Answer Search Engines

If a surfer is looking for the answer to something rather than a specific web site, he/she can use the Answer Search Engines. An Answer Search Engine is a service that specializes in offering direct

answers to questions asked by the surfer rather than links to related sites. E.g. AllExperts.com (www.allexperts.com), Answers.com (www.answers.com), BrainBoost (www.brainboost.com).

Domain Name Search Engines

These search engines help us to know whether a domain name is available for booking. E.g. Domain Notes (<http://domainnotes.com>), Checkdomain.com (<http://checkdomain.com>), DomainSurfer (www.domainsurfer.com).

Government Search Engines

The Search Engines which are helping us to search for official government information is known as "Government Search Engines". E.g. India <http://goidirectory.nic.in> (although it is not a search engine, most of the information of Government of India can be browsed), Canada (www.canada.com), Europe (www.searcheurope.com), United States of America (www.firstgov.gov).

Invisible Web & Database Search Engines

There is a big problem with most of the search engines that is many people are not even aware of by the net surfers. The problem is that vast expanses of the Web are completely invisible to general purpose search engines like Google, AltaVista and HotBot. Even worse, this "Invisible Web" is in all likelihood growing significantly faster than the visible Web we are familiar with. The following search engines will help us to search the "invisible web" or "deep web." E.g. InvisibleWeb.com (www.invisibleweb.com), WebData (www.webdata.com), LexiBot (www.lexibot.com), Turbo10.com (<http://turbo10.com>), CompletePlanet (www.completeplanet.com/index.jsp)

Mailing Lists Search Engines

An electronic mailing list, a type of Internet forum, is a special usage of e-mail that allows for widespread distribution of information to many Internet users. If you are looking for a good email list to subscribe to? The following search engines will point you in the right direction. E.g. Yagoo Groups (<http://groups.yahoo.com>), Coollist (www.coollist.com), TileNet (<http://tile.net/lists>), PAML (<http://paml.net>), Nextmark (<http://lists.nextmark.com>).

News Group Search Engines

A newsgroup is a repository usually within the Usenet system, for messages posted from many users at different locations. How to find out the Newsgroup of your interest? You may use the following News Groups Search Engines. Google Groups (<http://groups.google.com>), TileNet (<http://tile.net/news>), Harley Hahn (www.harley.com/usenet), Langenberg (<http://newsgroups.langenberg.com>).

Shopping Search Engines

Internet allows us to search for products from hundreds of online vendors and facilitates online shopping. There are some exclusive search engines for this purpose which are known as 'Shopping Search Engines'. E.g. BizRate (www.bizrate.com), DealTime (www.dealtime.com), Froogle (www.froogle.com), MSN Shopping (<http://shopping.msn.com>), PriceGrabber (www.pricegrabber.com).

Travel Search Engines

If we want to perform a travel especially to abroad, it is always better to have a thorough plan. The Internet offers 'Travel Search Engines' to help us a lot in this regard. E.g. Allcheckin (www.allcheckin.com), AOL Pinpoint Travel (www.pinpointtravel.com), Dohop (www.dohop.com), IgoUgo (<http://igougo.com>).

Vertical Search Engines

Vertical search engines, part of a larger subgrouping known as "specialized search engine, is a relatively new tier in the Internet search industry consisting of search engines that focus on specific businesses. It is important to note that there are vertical search engines for doctors, veterinarians, biologists, computer professionals, lawyers, and other specialised professionals. A few selected such search engines are furnished in Table-1 furnished below.

Table-1

Search Engine	Uniform Resource Locator
<i>Agricultural Sciences</i>	
AgEcon Search	http://www.agnic.org/
AgNIC	http://www.agnic.org/
<i>Biological Sciences</i>	
BioChemWeb	http://www.biochemweb.org/general.shtml
Biocrawler	http://www.biologie.de
<i>Chemical Science</i>	
Chem Finder	http://chemfinder.cambridgesoft.com/
Chem Industry.com	http://www.chemindustry.com/
Computer Science	
Research Index	http://www.researchindex.com
Tucows	http://www.tucows.com/
<i>Engineering & Technology</i>	
Athenus	http://www.athenus.com/
Engineer.net	http://www.engineer.net/
<i>Medical Science</i>	
Med Explorer	http://www.medexplorer.com/
Medsite	http://www.medsite.com/
<i>Science-General</i>	
SciNet	http://www.scinet.cc/
Scirus	http://www.scirus.com

4. Conclusion

It can be concluded that while Google, Yahoo and other general purpose major search engines will continue to dominate the online consumer search market, research analysts say myriad specialized search engines are emerging to address the particular information needs of niche audiences and professions.

References

1. Brin, Sergey and Page, Lawrence. The Anatomy of a Large-Scale Hypertextual Web Search Engine. Available at <<http://www-db.stanford.edu/~backrub/google.html>> Accessed on <03, August 2006>
2. Green, David. The evolution of web searching. Online Information Review. Vol. 24 No. 2, 2000.
3. Sullivan, Danny. Computer Search Engines & Computer-Related Searching. Available at <www.searchenginewatch.com> <February 20, 2002> Accessed on <07, August 2006>