



**INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH  
TECHNOLOGY**

**Search Engine : A Survey**

**Ankur Singh Bist**

Govind Ballabh Pant University of Agriculture And Technology, Pant, India

[ankur1990bist@gmail.com](mailto:ankur1990bist@gmail.com)

**Abstract**

There is lot of information available on the internet so to fetch the relevant information is very typical task . In this paper our purpose is to make a survey on various issues related to search engines like semantic web search technology.

**Keywords:** Crawler, Ontology ..

**Introduction**

Searching is an important issue to retrieve the required data . Search engines are serving the same task . There are various challenges covered by search engines in recent years . There are lot of challenges still to be solved . Answer the intelligent questions most efficiently is one of the major issue that is to be handled by search engines semantic web searching approaches are evolving that are responsible for making an appropriate results for intelligent input query . There are also various issues related to vertical search engines that needed to be handled by applying efficient approaches .The timeline of search engines is provided below:--

Year	Engine	Current status
1993	<a href="#">W3Catalog</a>	Inactive
	<a href="#">Aliweb</a>	Inactive
	<a href="#">WebCrawler</a>	Active, Aggregator
1994	<a href="#">Go.com</a>	Active, Yahoo Search
	<a href="#">Lycos</a>	Active
	<a href="#">AltaVista</a>	Active, Yahoo Search
1995	<a href="#">Daum</a>	Active
	<a href="#">Magellan</a>	Inactive
	<a href="#">Excite</a>	Active
1996	<a href="#">SAPO</a>	Active
	<a href="#">Yahoo!</a>	Active, Launched as a directory
	<a href="#">Dogpile</a>	Active, Aggregator
1997	<a href="#">Inktomi</a>	Acquired by Yahoo!
	<a href="#">HotBot</a>	Active (lycos.com)
	<a href="#">Ask Jeeves</a>	Active (rebranded ask.com)
1998	<a href="#">Northern Light</a>	Inactive
	<a href="#">Yandex</a>	Active
1999	<a href="#">Google</a>	Active
	<a href="#">MSN Search</a>	Active as Bing

	<a href="#">AlltheWeb</a>	Inactive (URL redirected to Yahoo!)
	<a href="#">GenieKnows</a>	Active, rebranded Yellowee.com
1999	<a href="#">Naver</a>	Active
	<a href="#">Teoma</a>	Active
	<a href="#">Vivisimo</a>	Inactive
2000	<a href="#">Baidu</a>	Active
	<a href="#">Exalead</a>	Inactive
2002	<a href="#">Inktomi</a>	Acquired by Yahoo!
2003	<a href="#">Info.com</a>	Active
	<a href="#">Yahoo! Search</a>	Active, Launched own web search (see Yahoo! Directory, 1995)
2004	<a href="#">A9.com</a>	Inactive
	<a href="#">Sogou</a>	Active
	<a href="#">AOL Search</a>	Active
2005	<a href="#">Ask.com</a>	Active
	<a href="#">GoodSearch</a>	Active
	<a href="#">SearchMe</a>	Inactive
2006	<a href="#">wikiseek</a>	Inactive
	<a href="#">Quaero</a>	Active
	<a href="#">Ask.com</a>	Active
2007	<a href="#">Live Search</a>	Active as Bing, Launched as rebranded MSN Search
	<a href="#">ChaCha</a>	Active
	<a href="#">Guruji.com</a>	Active
2008	<a href="#">wikiseek</a>	Inactive
	<a href="#">Sproose</a>	Inactive
	<a href="#">Wikia Search</a>	Inactive
2009	<a href="#">Blackle.com</a>	Active
	<a href="#">Powerset</a>	Inactive (redirects to Bing)
	<a href="#">Picollator</a>	Inactive
2010	<a href="#">Viewzi</a>	Inactive
	<a href="#">Boogami</a>	Inactive
	<a href="#">LeapFish</a>	Inactive
2011	<a href="#">Forestle</a>	Inactive (redirects to Ecosia)
	<a href="#">VADLO</a>	Active

	<a href="#">DuckDuckGo</a>	Active, Aggregator
	<a href="#">Bing</a>	Active, Launched as rebranded Live Search
2009	<a href="#">Yebol</a>	Active
	<a href="#">Mugurdy</a>	Inactive due to a lack of funding
	<a href="#">Goby</a>	Active
	<a href="#">Blekkio</a>	Active
	<a href="#">Cuil</a>	Inactive
2010	<a href="#">Yandex</a>	Active, Launched global (English) search
	<a href="#">Yummlly</a>	Active
2011	<a href="#">Interred</a>	Active
	<a href="#">Yandex</a>	Active, Launched Turkey search
2012	<a href="#">Volunia</a>	Active

Timeline[ 1 ]

**Various Issues Of Search Engines**

Example of different search engines are as follows[4] :-----

- Conventional (library catalog). Search by keyword, title, author, etc[4].
- Text-based (Lexis-Nexis, Google, Yahoo!). Search by keywords. Limited search using queries in natural language.
- Multimedia (QBIC, WebSeek , SaFe) Search by visual appearance (shapes, colors,...).
- Question answering systems (Ask, NSIR, Answerbus) Search in (restricted) natural language[4]
- Clustering systems (Vivísimo, Clusty)
- Research systems (Lemur, Nutch)[4]

A search engine operates using web crawling , indexing and searching . Web search engines work by storing information about many web pages, which they obtain from the HTML itself. These pages are retrieved by a Web crawler that is an automated Web browser which follows every link on the site. The contents of each page are then analyzed to determine how it must be indexed . Data about web pages are stored in an index database for use in later queries that are raised from user for processing .

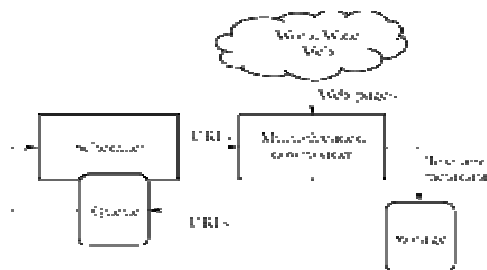


Figure . High-level architecture of a standard Web crawler [1]

Ye Wang , Zhihua Geng explained the issue of vertical search engine as they give flat results without efficient results using dolphin system architecture. Various elements of dolphin system architecture are[3] :-

1. Graphics user interface
2. Data center
3. Query processor
4. Data storage
5. Focused crawler
6. Internet

Semantic search engines are fighting with the issue of mapping right results to intelligent query as well as with the problem of finding correct and efficient results in various distributed results .

Elements of semantic web framework[2]:-

1. User interface and application
2. Trust
3. Proof
4. Unifying logic
5. Query
6. Resource description framework
7. XML
8. Web ontology language : OWL

Intelligent search engines are made to serve the purpose of web mining various approaches like information preference , formal concept analysis and various simulation technology are used by different authors to make various advancement in this domain . Search engine optimization techniques used by google:[6]----

- a.) Create unique, accurate page titles
- b.) Make use of the "description" meta tag

**Improving Site Structure[6]**

- a.)Improve the structure of your URLs
- b.)Make your site easier to navigate

**Optimizing Content[6]**

- a.)Offer quality content and services
- b.)Write better anchor text
- c.)Optimize your use of images
- d.)Use heading tags appropriately

**Dealing with Crawlers[6]**

- a.)Make effective use of robots.txt
- b.)Be aware of rel="nofollow" for links

**SEO for Mobile Phones[6]**

- a.)Notify Google of mobile sites
- b.)Guide mobile users accurately

**Promotions and Analysis[6]**

- a.)Promote your website in the right ways
- b.)Make use of free webmaster tools

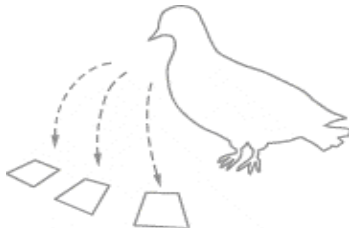


Figure . Pigeon rank[5]

### PigeonRank Technology[5]

1. Firstly a user submitted a query to Google.
2. The query is then routed to what is known as data coop.
3. When a relevant result is located by one of the pigeons in the cluster, it strikes a rubber-coated steel bar, this gives the page a Pigeon Rank value of number one[5].
4. For each peck, the Pigeon Rank value is increased.
5. The pages that get the most pecks are prioritised and are shown at the top of the user's results page[5].
6. The remaining results are displayed in order of this pecking system.

The pigeon rank methods makes it typical to amend results, aside from the Location & Frequency tricks, some try and boost rankings by including images on their pages, Google's Pigeon Rank technology is not fooled by such techniques[5].

### Conclusion

In this paper firstly we discuss about the timeline of search engines . We make a summarised look on various issues and approaches adopted in this domain . Search engine optimization and pigeon Rank technology are also discussed .

### References

- [1] [www.wikipedia.com](http://www.wikipedia.com)
- [2] G.Madhu and Dr. A. Govardhan , Dr. T. V. Rajinikanth ,”Intelligent Semantic web search engine :A brief Survey”.
- [3] Ye Wang , Zhihua Geng , Sheng Huang ,”Academic Web Search Engine- generating a survey automatically”.
- [4] Prof. Dragomir R. Radev ,”Search Engine Technology”.
- [5] Sylvia King, ”Commercial systems”.
- [6] Search Engine Optimization Starter Guide Google.