

**AN EXPOSURE TOWARDS MONITORING OF SEARCH ENGINES****P.Pramoda Rao<sup>1</sup>, K.Anusha<sup>2</sup>**<sup>1</sup>M.Tech Student, Dept of CSE, RRS College of Engineering & Technology, Muthangi (V), Patancheru (M), Hyderabad, T.S, India<sup>2</sup>Assistant Professor, Dept of CSE, RRS College of Engineering & Technology, Muthangi (V), Patancheru (M), Hyderabad, T.S, India**ABSTRACT:**

In the present days network has encountered the problems of information explosion information overload strange information and so on. By means of rapid development concerning Internet, rising information constantly comes into network and hence quantity of websites and web pages also immediately increase. It has become more and more complicated to use the traditional search engines to recover the related information. Intellectual search engines have to assemble users' actual needs and can make available individuals with helpful references, and assist the users discover the appropriate information. Because of difficult selection of keywords, or web pages enclose too much non-constructive information; numerous web pages will come into view when user type keyword to search fetch users much trouble. A huge amount of outmoded information as well as trash messages subsist when someone contain an information search. For the large amounts of web pages, the users have to spend much time on these results to seek for the useful information. Many of image recovery occurrences are used to exemplify the intricacy, suitability and assortment which make available people by means of references for studying rational search engines.

***KEYWORDS: Websites, Intellectual search engines, Web pages, Search engines, Keywords.*****1. INTRODUCTION:**

There is a requirement to everyday get better the competence and the performance of a website in common. Intense and growing

investments in the design of website, is still exposed, and conversely identification of desired information in a website is not simple and designing effectual websites is not an insignificant task [4]. When users are

searching information on the search engines, they do not hope too much feedback results, but to see the results whether live up to their requirements. Many references only introduce the given keywords as well as target number concerning search engines; do not analyze the useful information present in target pages [8]. The searching approaches directly provided to the users by the search engines are mostly based on the Boolean logic match of key words, so the search results far exceed the users' abilities to absorb and take advantage of in number [1]. People often search information in the enormous ocean of information with the help of search engines. It has become more and more complicated to use the traditional search engines to recover the related information. Because the sources of the information of the network are difficult, and there exist the phenomenon of the users' locality, informality, different levels of knowledge, different levels of quality, different professional fields, widely different social environments, different reliefs, the different purposes of surfing the internet, so all of these lead to the network information explosion [11]. It is necessary to decide effectual characteristics and amass the eigenvectors concerning images in case of

employ of searching online based on difficulty of the image information; by means of content-based recovery process to restore retrieve based on keywords, to recover the recovery accuracy and intelligence level concerning searchers [3]. Due to impediment of image information, it can resourcefully keep away from outmoded information and get better search accuracy by construction of feature vector, moving information base, taking benefit of HTML information sharply. In the meantime, the quality, value and reliability of the information have been problems that users worry about [14].

## 2. METHODOLOGY:

Modern network has encountered the problems of information explosion information overload strange information and so on. All of these require that the technology of the search engine must be improved and perfect with the development of the Internet. For the large amounts of web pages, the users have to spend much time on these results to seek for the useful information [9]. In today, with explosive growth of network information, it is unfeasible for people to spend so much time and energy in browsing the every web page searched. Many references do not analyze

the functional information contained in target pages and bring in specified key words as well as target number concerning search engines [7]. With the rapid development and universality of the Internet, more and more information constantly comes into the network and the quantity of websites and web pages also immediately increase. Besides, large portion of the pages cannot communicate the very information what users expect or they are only some unrelated information [2]. As a result, it has been one of the development trends of the search engine to provide users with a good excellence and proper number results, which promotes the development of the technology of information filtering, such as methods of filtering information have been put forward [16]. The common search engines have similar principle of work, they all have disadvantages. Because of difficult selection of keywords, or web pages enclose too much non-constructive information; numerous web pages will come into view when user type keyword to search fetch users much trouble. Intellectual search engines shown in fig1 have to assemble users' actual needs and can make available individuals with helpful references, and assist the users discover the appropriate information [12].

Accordingly of complication of image information, it can efficiently keep away from outmoded information and get better search accuracy by construction of feature vector, moving information base, taking benefit of HTML information sharply. Network information multiply geometrically as time goes on which put in to intricacy of network information; the search process based on keywords exist immense variables; intricacy as well as unpredictability of network information have an effect on the search accuracy; greatly practical information is submerged into ocean of data [5].

### **3. REINFORCEMENT OF INTELLIGENCE CONCERNING SEARCH ENGINE:**

To recover exactitude of information search results and consequence of search, we have to get better accuracy of keywords, improve the logic match ability of keywords and put in common as well as personalized search [15]. Numerous references only set up specified keywords and target number concerning search engines and do not measure the useful information present within target pages. Many of image recovery occurrences are used to exemplify the

intricacy, suitability and assortment which make available people by means of references for studying rational search engines [10]. On version of non-structure as well as difficulty of the image information, it is essential to decide effectual characteristics and amass the eigenvectors concerning images in case of employ of searching online; by means of content-based recovery process to restore retrieve based on keywords, to recover the recovery accuracy and intelligence level concerning searchers. In accordance with dissimilar pictures, take on dissimilar ways of data cleaning, for instance medial alligation, analysis of correlation. Construct a representation by means of genetic algorithm as well as neural network to include data cleaning; or move up system of clustering web documents, as well as pattern detection and examination used in mining of web [6]. Dropping time relevance extent and dataset proportions and building appropriate principles will put in to data maintenance. Get better the search accuracy and faultless knowledge base with aid of information response. Design precise and rapid classification concerning information as well as matching algorithm [13]. Set up SVM as well as neural network into search engine to pick up the intellect of

search engine. Include severe data cleaning and sort out unpractical information. A huge amount of outmoded information as well as trash messages subsist when someone contain an information search. Devoid of maintenance of un practical information, it undoubtedly will manipulate the classification, exploration, drawing out of data.

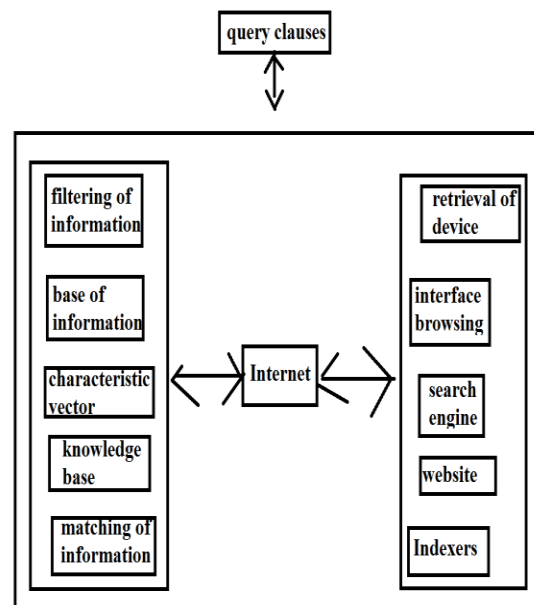


Fig1: An overview of intellectual search engine

#### 4. CONCLUSION:

By means of unpredictable growth of network information, it is unfeasible for people to spend so much time and energy in browsing the every web page searched. To recover exactitude of information search results and consequence of search, we have to get better accuracy of keywords, improve

the logic match ability of keywords and put in common as well as personalized search. Large portion of the pages cannot communicate the very information what users expect or they are only some unrelated information. Many references only introduce the given keywords as well as target number concerning search engines; do not analyze the useful information present in target pages. The searching approaches directly provided to the users by the search engines are mostly based on the Boolean logic match of key words, so the search results far exceed the users' abilities to absorb and take advantage of in number. Devoid of maintenance of un practical information, it undoubtedly will manipulate the classification, exploration, drawing out of data. it has been one of the development trends of the search engine to provide users with a good excellence and proper number results, which promotes the development of the technology of information filtering, such as methods of filtering information have been put forward. When users are searching information on the search engines, they do not hope too much feedback results, but to see the results whether live up to their requirements.

## REFERENCES:

- [1] Qin Hua, Su Yidan, Li Taoshen, A Data Cleaning Method Based on Genetic Algorithm and Neural Network. Computer Engineering and Applications, Vol 3, pp. 45-46, January 2004 (In Chinese).
- [2] Xu Jiakun, Studying by Comparison the Four Searching Engines in Common Use in the Research of Network Information Measurement. New Technology of Library and Information Service, Vol 11, pp. 46-48, November 2004 (In Chinese).
- [3] Fang Zhijian, Zhang Ruilin, Tong Xiaosu, Recently research and future development of search engine. Computer Engineering and Design, Vol 28, pp. 4038-4041, August 2007 (In Chinese).
- [4] Problems and Solutions of Web Search Engines  
Wang Liangshen, Hou Jie, Xie Zaiyu, Wang Xiaochen, Qu Caiyue, Li Hui, 2011
- [5] Seikyung Jung, Jonathan L. Herlocker and Janet Webster, Click data as implicit relevance feedback in web search. Information Processing & Management, Elsevier, Vol 43, pp. 791-807, March 2007.
- [6] Wang Weiling, Liu Peiyu, Liu Kefei, A Feature Selection Algorithm for Web Documents Clustering. Computer Applications and Software, Vol 24, pp. 154-156, January 2007 (In Chinese).
- [7] Michael Chau, Hsinchun Chen, A machine learning approach to webpage filtering using content and structure analysis. Decision Support Systems, Elsevier, Vol 44, pp. 482-494, February 2008
- [8] Zhu Zhiguo, Design of Architecture of Web Usage Pattern Mining System. Information Studies: Theory & Application, Vol 33, pp. 97-101, April 2010 (In Chinese).
- [9] Tan Hansong, Li Hong, Web Mining on Information Filtering. Computer Engineering and Applications, Vol 30, pp. 186-187, October 2003 (In Chinese).
- [10] Du Yajun, Qiu Xiaoping, Xu Yang, Inquiry Intellectual Capacity into Chinese Search Engine. Application Research of Computers, Vol 4, pp. 29-31, 35, April 2004 (In Chinese).

[11] Liao Kaiji, Yi Cong, The Study of Web Business Information ExtractionBased on Regular Expressions. Journal of Intelligence, Vol 29, pp 159-162, May 2010 (In Chinese).

[12] Zhang Fan, Lin Jian, Research on Filtering Mechanism in IntelligentSearch Engine. Library and Information, Vol 4, pp. 52-56, April2007(In Chinese).

[13] Zhu Zhiguo, Deng Guishi, Analysis and research on Web usage mining.Application Research of Computers, Vol 25, pp. 29-32, 36, January2008 (In Chinese).

[14] Chen Jihong, Qing Xiao, A Comparative Study of Four Search Engines.Information Science, Vol 21, pp. 1084-1087, October 2003 (InChinese).

[15] Lai Yonghao, Xie Zanfu, Research on Anti-jamming Bad Web FilterAlgorithm. Computer Engineering, Vol 33, pp. 98-99, November 2007 (In Chinese).

[16] Liu Shuchao, Li Yongchen, Wu Hongping, Research and Discussion ofWeb Data Mining. Manufacturing Automation, Vol 32, pp. 163-166, September 2010 (In Chinese).