



# International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: [www.ijirccce.com](http://www.ijirccce.com)

Vol. 5, Issue 5, May 2017

## Integrating Personalization and Image based Search System to Enhance Search Results

Sneha A. Taksande<sup>1</sup>, Prof. A. V. Deorankar<sup>2</sup>

M. Tech Student, Dept. of Computer Science and Engineering, Government College of Engineering, Amravati, Maharashtra, India. <sup>1</sup>

Associate Professor, Dept. of Information Technology, Government College of Engineering Amravati, Maharashtra, India. <sup>2</sup>

**ABSTRACT:** In today's situation, information is the need of society as it has significant importance for contemporary success in every field thus a search engine come to play a crucial role in information retrieval. Sometimes the search engines may not give the accurate search results when queries are short and ambiguous. So the proposed system gives image based more relevant web pages. And users query is assisted by the system internally, which formulate the query and further submitted to the google. And the more relevant results are obtained according to the query. The accuracy of web search results can be improve by considering the textual contents of the images in correlation with the personalization. For image based search system the images are obtained from google using google API. These images are categorized using JSON of googles API. For personalized search the data is fetched from google, using googles API according to users' profession. This method retrieve the more relevant information for professional search. The proposed search system is work in association with google as it is taking the data from google and performing re-ranking on this data. Thus the integration of both the mechanism implies to enhance the search results.

**KEYWORDS:** Search engine, Web search results, Image based search, personalized search, Relevant Information.

### I. INTRODUCTION

The World Wide Web contains a vast amount of interlinked web documents. Retrieving information from such a huge collection is easy using various search engine, but retrieving relevant information is still a challenging task. Also traditional web search engines do not use the images in the HTML pages to find relevant documents for a given query. "A picture is worth a thousand words." Despite this old saying, modern Web search engines ignore the pictures in HTML pages. It retrieve the documents merely by comparing the query keywords with the text in the documents. Of course this text includes the words in image captions and markup tags.

In this work, we propose a novel document retrieval approach that uses the content of the pictures in the Web pages to boost the accuracy of pure text-based search engines. The proposed system will analyze the specified query and suggest user more relevant query on the basis of his profile. For personalized search, when user specify search query, system will fetch the data from google database with the help of googles API then system will formulate the given query by adding the users profession to the specified query by using the facility of this API. This query is then submits to the google internally and data relevant to this query is fetched and displayed to the users. This will gives the more relevant information to the user according to user's profile. Also system ranks the data according to search results and preferences. Preference wise search facility is also provided. In that, when user gives any query system will search the data according to the preferences. The data is fetched from google database using API, the snippets and tiles from the searched links is given to NLP which splits the snippets into nouns, verb, pronouns, etc. The pronouns extracted will be the keywords. These keywords will be considered as preferences which will stored in system database to find out the search results according to these preferences. The web links fetched from google database will be filter out on the basis of historical search. When user specifies query in image search system, the categories of these images is provided by JSON reader. When user selects any image then the results for the respective image according to its

# International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: [www.ijirccce.com](http://www.ijirccce.com)

Vol. 5, Issue 5, May 2017

categories is shown to the user. The links will be then re-ranked with the help of ratings given by the other users and Re-ranking of links will be done on the basis of user's profile, user's preferences and search history.

Proposed system has focused on image based search and re-ranking strategy for personalized search. The contribution of this method is retrieving the most relevant information for the users query. The results which are more relevant to the user but low in rank are up-ranked by proposed method by maintaining user browsing history and preferences. Also the images in HTML pages are used to improve the web document search by considering their categories using JSON libraries of google API and showing search results for these images. The system will work in association with google using googles API as it is taking the google data.

## II. RELATED WORK

In paper [1] authors gives personalization as a vital tool to meet the user's information need, as it gives the results which are more relevant to the user. The advantage of proposed system is that it reduces the language gap between the user and the search engine. Further, the results which are more relevant to the users query but low in rank are up-ranked by proposed method by maintaining user browsing history. So, this made another advantage of relevancy increase. In paper [3] it gives the method for Re-ranking of documents using images. It present query document features for image-based re-ranking i.e. text features, visual metadata features, query visualness features. Paper [2] addresses a way to overcome the semantic gap and the intent gap simultaneously, it integrate multiple visual modalities and click-through data with learning image similarity and typicality, and present a novel image search re-ranking approach, named spectral clustering re-ranking with click-based similarity and typicality (SCCST). This paper [4] proposed a new methodology for rank improvement using search engine query logs. The most important part of this architecture is the use of Panda algorithm to find the relevancy of URLs based on the relevancy of content corresponding to them. This paper [5] contains a survey on privacy protection in personalized web search that model user preferences as user profiles. Also a personalized web search framework is discussed in this paper, which adaptively generalizes profiles while considering user specified privacy requirements.

## III. PROPOSED METHOD

As the proposed search system is based on personalization and image based, it consists of two methods. The system architecture of the proposed system for personalized search is depicted in following figure.

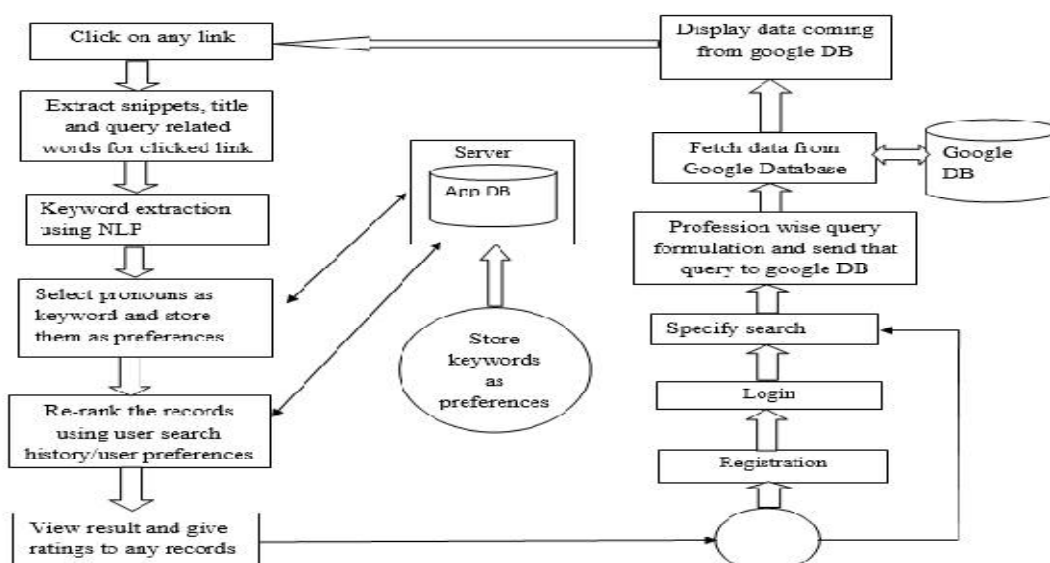


Fig. 1. Architecture of personalized search system

# International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: [www.ijircce.com](http://www.ijircce.com)

Vol. 5, Issue 5, May 2017

1. Specify query in search box.
2. If user is logged in user then users profile wise query reformation is done. It means it will add the users' profession to that query term using facility of googles API.
3. Send this formulated query to google database and get profession wise relevant data form google.
4. For preference wise search when user clicks on any links the keywords are extracted from the snippets or titles of the web pages. This will provide as a input to the NLP i.e. natural language processing which splits these snippets and divides the words as nouns, verbs, pronouns, etc. the pronouns are considered as keywords and they stored in the System database as preferences.
5. When users performs history preferences wise search then the results will re-ranked according to the preferences of the users.
6. And finally user will get the rank of his searched information link which is small or equal to the rank of googles link for the respective query.

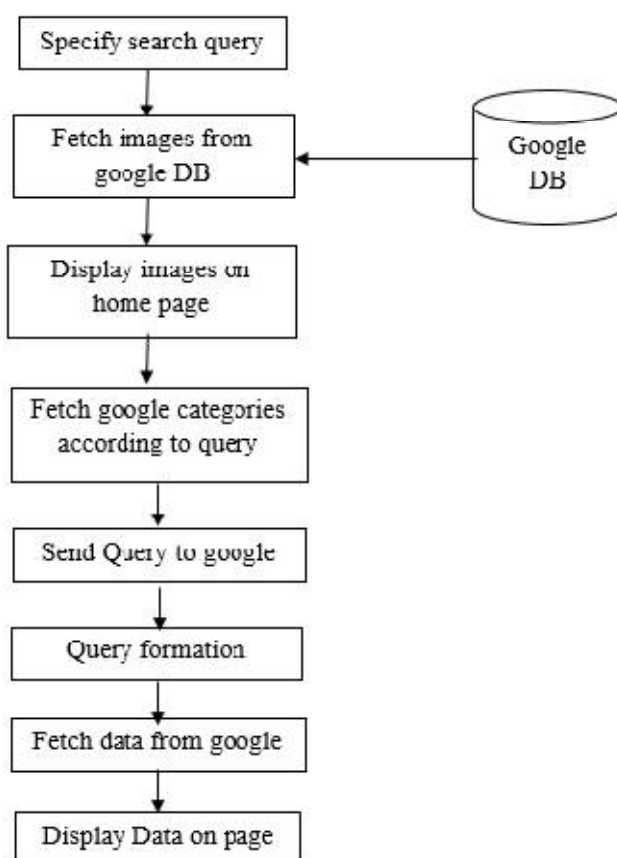


Fig. 2. Workflow of Image based Search System

The working for the above workflow is as given below:

1. When user specifies query, system will display images fetched from google database using googles API for image extraction.
2. These images are categorized using JSON of googles API. System will reformulate the query by adding these categories to the query keyword and submit it to the google.
3. Google then fetch the information according to these categories.

# International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: [www.ijircce.com](http://www.ijircce.com)

Vol. 5, Issue 5, May 2017

4. When user select any image then the information related to that image category is displayed by the system.
5. This information will be more relevant to the users query, because for some queries user might have to search by checking more web pages which takes more time for searching.
6. But as this approach shows the images on the home pages user can get the idea and by clicking that image user can get his intended search results more easily.

## IV. SIMULATED RESULTS

### • Image Based Search Results

When user specifies query for a simple search, system will display images of matching web pages as shown. For the user query of apple the images are displayed by the system as shown in figure below. User will select any one image. System will find out the details of selected image. With the help of image details, system will filter unwanted data links and display the relevant links only. This can be done using categories provided by google. When user select one of the retrieved images the information relevant to that image category is retrieved to the user. This approach suits for the users who are not having good hand at searching or who are new to the search system. They might expect the most relevant information in their first search but it may not happened every time. So it will be helpful for them if system displays the relevant images for the search query.

They may get idea of their intended search and get the more relevant search results.

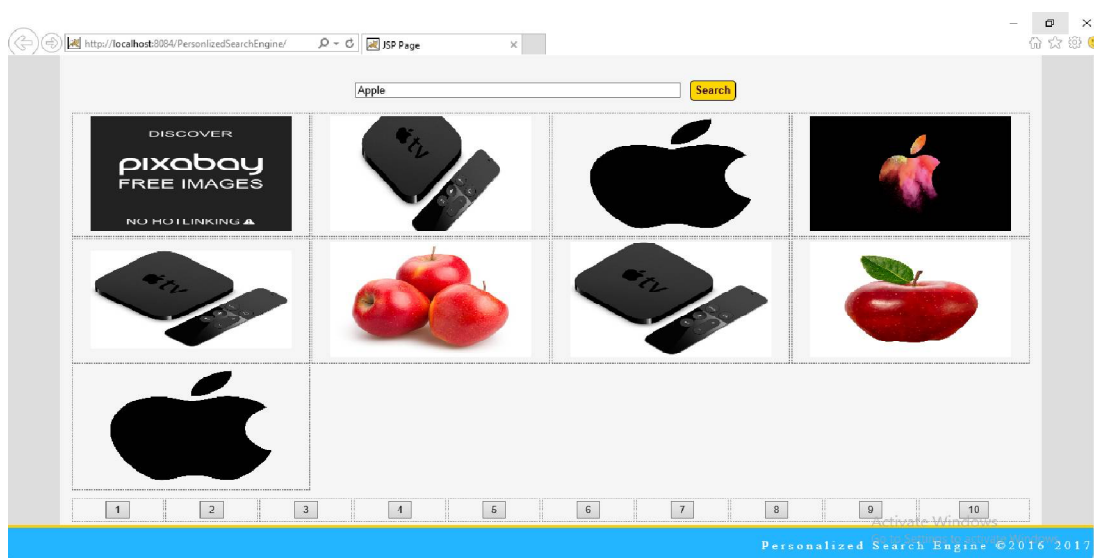


Fig. 3. Image search for specified query Apple

For example, when user selects 6th image from the above screenshot, the information related to that image will be shown as below i.e. the information related to apple fruit. Whatever the information is available on google for that image category, will be fetched by the system from google and display to the user.

# International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: [www.ijirccce.com](http://www.ijirccce.com)

Vol. 5, Issue 5, May 2017

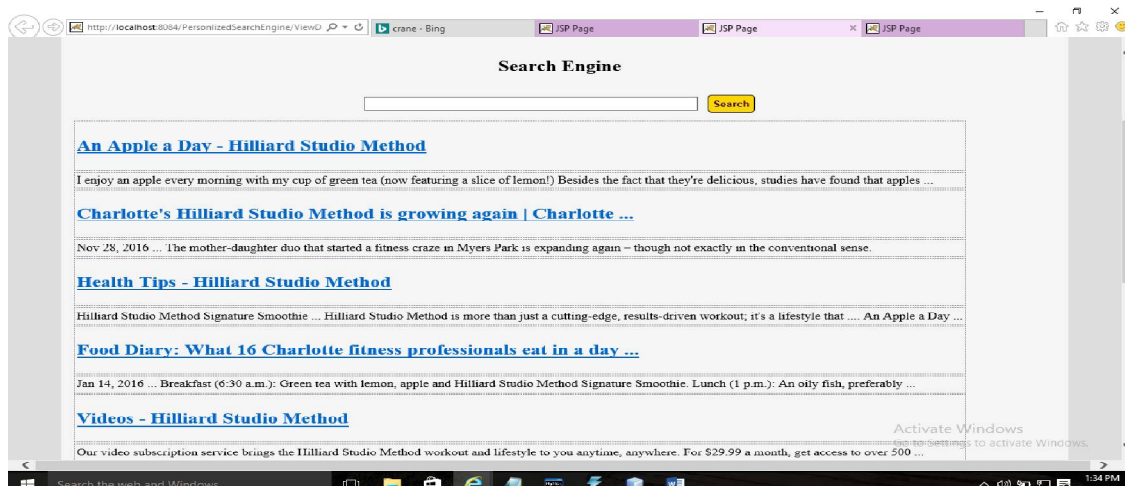


Fig. 4. Web links extracted related to the image of apple fruit.

When user selects 4th image from the above screenshot, the information related to that image category of apple TV will be shown as below i.e. the information related to TV's of Apple Company as shown in figure below.

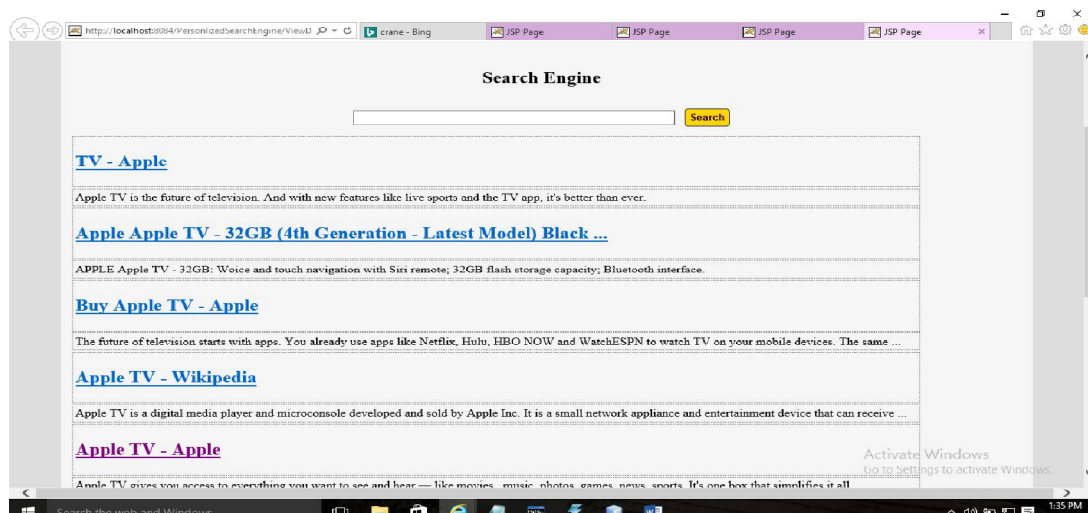


Fig. 5. Web links extracted related to the image of TV's of Apple Company

## • Personalized search System

Personalized search system is based on user's profile. The proposed System provides the user registration facility through which system will collect the user's data and will store that data in system database. After logging in to the users account, when user specifies the search query then the search results will be displayed according to users profile i.e. the results can be retrieved according his/her profession, branch. User will have the different options for intended search, i. e. profession wise search, preferences wise search, history wise search and if user don't want any of these, then he/she can do a simple search.

# International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: [www.ijirccce.com](http://www.ijirccce.com)

Vol. 5, Issue 5, May 2017

Suppose user wants the search result for the search query “crane”. Now crane can be the ambiguous query as crane is the name of bird, crane is one if the constructing equipment and also there is a book of author whose name has crane word in his name, but when user specify this query using the profession wise search then the system will retrieve the more relevant information according to his profession or branch. So if the user is from civil engineering background then user will get the information about the constructing equipment crane as shown in figure below, or else the result will as simple as the google. This can be illustrated as shown in figure below. In search history the analysis can be done based on the ranking of the links retrieved according to search query. As shown in below figure the user is from civil background, so user will get more relevant results according to his profile for this query term.

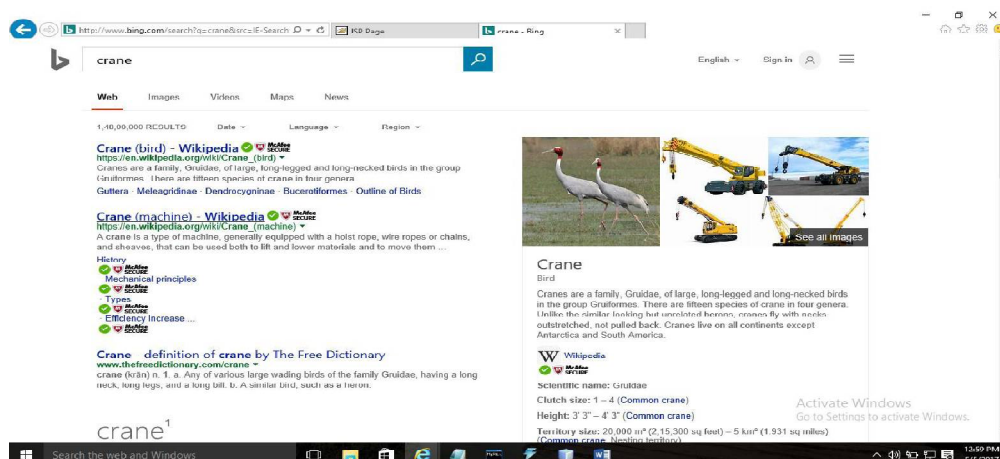


Fig. 6. Search results obtained using Bing

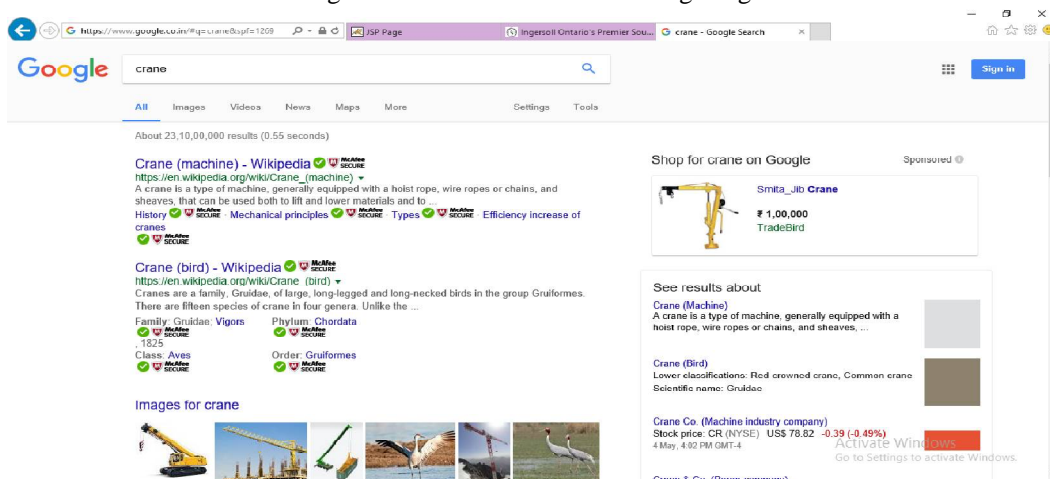


Fig. 7. Search results obtained using Google

The above generalized search shows all the information related to query keyword “crane”. These shows the search results for crane equipment, crane bird, etc. Thus user have to search more number of web pages to get his intended information which may require more time and labor for the search. But this can be minimize using proposed method.

# International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: [www.ijirccce.com](http://www.ijirccce.com)

Vol. 5, Issue 5, May 2017

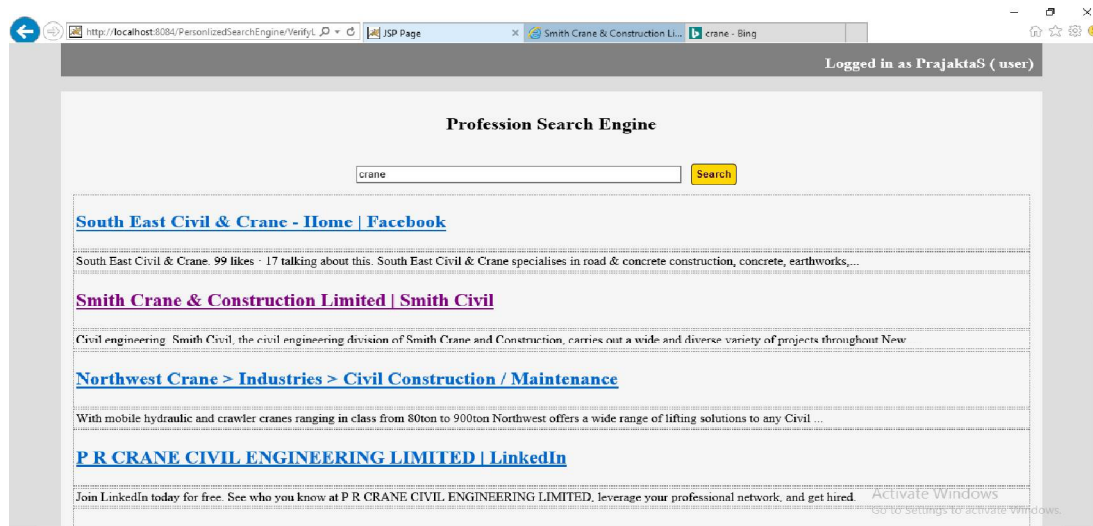


Fig. 8. Search results for crane obtained using proposed search system

If user want to search for more relevant information according to his profession then he/she can get it using the proposed approach. As shown in above figure if user performs profession wise search then it will get more relevant results for his/her intended search.

## V. CONCLUSION

Different from traditional web search system relying on the simple keyword matching technique, a user's profile based search system is proposed. For the new users who don't have good hand at searching, it become difficult for them to reach to their need. There is wastage of time and labor, but the proposed method reduces the time and labor for them. So this architecture will result into the most relevant information at the topmost position of the result list of search engine. It also increase the users' interests in searching. The integration of two methods i.e. personalization and image based search has been illustrated, that leads to improve the search result.

This research work aims to improve the accuracy of web search results using contents of images in HTML pages, to design a personalized search system based on user interest to improve the result delivery and to implement image based more relevant web pages search system.

## REFERENCES

1. ShilpaSethi and Ashutosh Dixit, "Design of Personalised Search System Based On User Interest and Query Structuring", 2nd International Conference on Computing for Sustainable Global Development, IEEE, 2015.
2. Xiaopeng Yang, Tao Mei, Yongdong Zhang, Jie Liu, and Shinichi Satoh, "Web Image Search Re-Ranking With Click-Based Similarity and Typicality", IEEE transactions on image processing, vol. 25, no. 10, October 2016.
3. Andrew W. Fitzgibbon, Sergio Rodriguez-Vaamonde and Lorenzo Torresani, "What Can Pictures Tell Us About Web Pages? Improving Document Search Using Images", IEEE, transactions on pattern analysis and machine intelligence, vol. 37, no. 6, June 2015.
4. ShipraKataria and PoojaSapra, "A Novel Approach for Rank Optimization using Search Engine Transaction Logs", International Conference on Computing for Sustainable Global Development, IEEE, 2016.
5. Ramitha A T and Dr. Jayasudha J S, "Personalization And Privacy In Profile-Based Web Search", International Conference on Research Advances in Integrated Navigation Systems (RAINS), IEEE, April 06-07, 2016.
6. ZhongJi, Yanwei Pang, and Xuelong Li, "Relevance Preserving Projection and Ranking for Web Image Search Reranking", IEEE Transactions on Image Processing, 2015.
7. N. Buvanewari and S.Bose, "Comprehensive Personalized Recommendation Technologies", Fifth International Conference on Recent Trends in Information Technology, IEEE, 2016.
8. Dong Zhou, Xuan Wu, Wenyu Zhao, Séamus Lawless, and Jianxun Liu, "Query Expansion with Enriched User Profiles for Personalized Search Utilizing Folksonomy Data", IEEE Transactions On Knowledge And Data Engineering, 2017.