



## PERSONALIZED WEB SEARCH: A SURVEY

Sakshi Singh

Suresh Kumar

### Abstract:

Nowadays, the internet provides information regarding every topic in the world. And the topic varies from traditional to technical. The Internet provides information irrespective of the subject by retrieving data from the servers. Search engines collect information from the servers as required by the user. Search engines services provided that enables Internet users to search through the World Wide Web (WWW) for content. They play an important role in the retrieval of the required information. An ideal search engine should move through all WWW's web pages and the associated data based on the keyword provided to the user should be listed. But not a single search engine results in the desired output as per the user's wish. To overcome this problem personalization web search is used. Personalization search is the customization of search engine outcomes generated by a filter taking into consideration possibly appropriate data such as user history, place, social networks, preferences, etc. This article focuses on the study of various effective custom web search methods and strategies that many writers have suggested.

**Keywords:** Personalized web search, Information retrieval, Search engine, Profile based search

### 1. INTRODUCTION

The search engine presents a well-organized way to search the appropriate web-based data. However, the results of the search may not always support the user because the search engine does not acknowledge the purpose of the user behind the queries fired by the user is short and incomplete. For instance, a specific keyword "apple" a user may look for apple fruit or apple computers. Personalized web search systems are designed to deliver appropriate outcomes to the user, taking into account the interest of the user. Different customization strategies have been proposed and web search engines have been developed to minimize the search time and to be more specific. With the growth of the internet, people are becoming increasingly dependent on their multiple data requirements on web search engines. Despite the comprehensive use, the search engine has several difficulties. How to obtain the real-time information requirement of the user in a personalized search is the main problem. The user enter the query in the search engine is the key source for assessing the need for data. However, the queries are short, incomplete and ambiguous which controls the clear expression of the data demands of the user and therefore affects the quality of the search engine data. The search results are based solely on the request. When submitting the same request, the underlying search engine delivers the same search results to different users, not taking into account their distinct data requirements and preferences. Some individuals understand and have created such adaptive systems that customize data or provide users with more suitable data. Users are needed to register in these schemes, for example, private data, their interests, age, place, etc. or users are needed to provide feedback on appropriate or irrelevant outcomes, from 1 (very poor) to 5 (very nice) scale-based scores. The process takes time and users want to use new methods to minimize the problems. Personalized web search captures data for each user by taking into consideration their conditions. As the search market competition is growing quickly, a few search engines

supplied a custom seek provider. For instance, the personalized search of Google lets users define categories of the subject on the web page. Some internet users use appropriate comments to manner user requirements or asked them for preceding registration of their demographic facts to offer quality search. On account that those techniques involve users to manually become aware of their possibilities in more fields beyond search, methods ought to be created which could implicitly apprehend the person's want for statistics. The next chapter discusses the different custom web search methods that many authors have suggested. The paper is structured as follows: Section 2 covers the problems and drawbacks faced by personalized web search while gathering customized data. Section 3 consist of various personalized web search methods and strategies discussed by various authors in their respective research papers. Section 4 describes the conclusion.

### 2. PROBLEMS AND DRAWBACKS

Apart from various characteristics and advantages, personalized web search has to face several problems while gathering customized data. Garbriella Pasi [22] mentioned two basic problems faced while retrieving personalized information: 1) Context representation and 2) The process to exploit context knowledge in order to improve search outcome's quality. The other concerns are confidentiality and efficacy evaluation of personalized IR approaches. Most data recovery schemes and search engines are based on a system-centered strategy in which IRS acts as a black box, irrespective of the user context, providing the same response to the same question. The quality of the search outcomes relies heavily on the data expressed in the request of the user. The efficacy of the system therefore heavily relies on the amount and quality of data accessible with respect to the user and his interests. Two major techniques that can be used to capture the interest of the user: implicit and explicit [13, 20]. To optimize search through better knowledge and communication, the issue of active visualization of search results was considered. While the information relevant to the

needs of the user expressed in a query may have been included in the long-organized results list, it is quite difficult to locate it. The client rarely goes beyond evaluating the search results on the first two / three pages of search engine.

### 3. LITERATURE REVIEW

Mohamed Reda Bouadjenek [1] suggested a fresh document ranking function while considering the internet social context to improve the Information Retrieval (IR) model. The author's suggested method is based on social annotations connected in bookmarking schemes, as well as other social metadata from multiple platforms such as Facebook, Twitter or Youtube, such as remarks, tweets, etc. It is easy to use these social annotations to get an intuition about the web pages to which they are linked. The author suggested the following contribution to custom internet search: (1) The SoPRA feature of a Social Personalized Ranking. (2) User profile and social document weighing technique. (3) Extension of SoPRA by taking into account individual user tagging and (4) A comprehensive assessment of SoPRA. Fig. 1 shows the basic architecture of SoPRA. The approach Consider matching the user profile with the record and matching the request with the annotations of the document. By the defined architecture, SoPRA ranks records by (i) a textual context matching document score and query; (ii) a personal matching document score and query; and (iii) the user's interest score to document.

Fig.1. Basic SoPRA architecture [1]

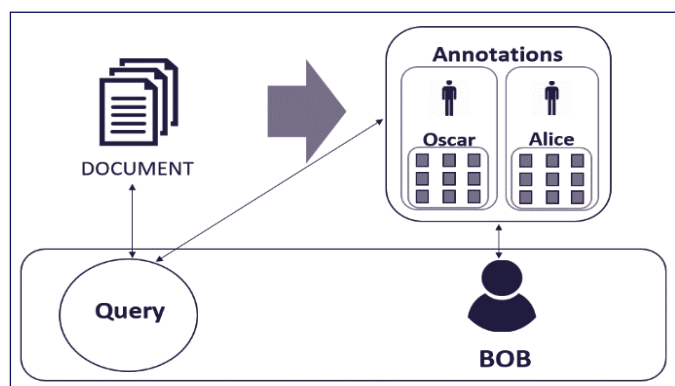
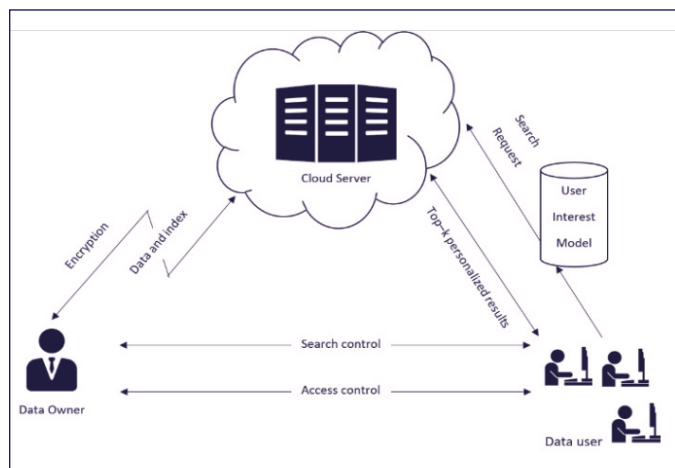


Fig. 2. Architecture of the search over encrypted cloud data [3]



Exploring effectively and efficiently for encrypted data has become extremely important. In any case, most existing strategies on the scrambled hunt over cloud information pursue the model of “one size fits all” and neglect the purpose of personalized search. The majority of them bolster just the accurate catchphrase look. Zhangjie Fu [3] proposed a paper that handles the issue of Personalized Multi-keyword Ranked Search over Encrypted data (PRSE) while guaranteeing insurance in distributed computing and gave two PRSE plans dependent on secure internal items to meet diverse pursuit expectations. The common technique used in information retrieval was query extension, which could extend original query terms before submitting the search request. Using both question augmentation and custom positioning, an accessible encryption plot was structured. Using WordNet semantic metaphysics, the maker created a model of customer interest for an individual customer by inspecting the history of the customer and getting the component to demonstrate client intrigue unmistakably. In the basic PRSE scheme, the author built an accessible sub-record for each archive in the document collection. The initial stage was to delineate catchphrases separated from the archive into the watchword set to produce an information vector and after that encode the vector. To improve look effectiveness and to scale well for huge information sources, the analyst utilized MDB-tree to build an access list to address the top-k issue. Fig. 2 shows the search architecture over encrypted cloud data. The model of user interest is based on the search history of the user, using WordNet, Documents the get entry to query keywords and key phrases linked with them. Different get entry to frequencies of key phrases can constitute the distinct relevance of the consumer's attitude as a keyword priority. The user interest framework must conduct a request reformulation meeting the keyword priority of the query word. The encrypted search query is sent to the cloud via the search control system. Once the approved user receives the search application, the cloud server will perform the assigned search procedure over the index and return appropriate encrypted files to the user.

The experiment of searchable encryption took place on Enron Datasets. The generated results were compared with MRSE [14]. In order to reduce the duplicity from the dataset, MySQL database is used which was provided by Jitesh Keyword were extracted from each email, (TF X IDF) was calculated and arranged in descending order. Dummy keywords were used to reduce the lack of reliability, privacy and precise management technique. After experimental verification, around 75% of the users were satisfied with the system's performance. The sub-index in MRSE was a binary vector and sub-index in PRSE was a decimal vector. Because of the difference of index structure, the storage of PRSE scheme was twice that of MRSE. Query encryption operation consumes more time in all the schemes. So, by increasing the number of Schematic Dictionary Keywords increased the price. In terms of search, the efficiency of PRSE\_1 and PRSE\_2 was more than MRSE when user requested more relevant documents.

P. Palleti, [2] used a probabilistic query extension to develop a custom web search. The internet includes large amounts of

records, and the quest engine advises professional methods to help navigate the net to get the appropriate facts. On the other side, most popular search engines provide query outcomes without the intent of the user behind the request. The custom web search system aims to provide users with appropriate outputs by taking into account user interests. In this strategy, the writers created a custom Web search system applied to the proxy by generating user profiles using collaborative filtering to change user interests. A user profile includes probabilistic correlations between the terms of the request and conditions used to deliver custom search outcomes. Experimental findings indicate that this suggested Web search system is highly efficient and effective.

Kenneth Wei- Ting Leung [4] proposed the paper that examines the customized social pursuit dependent on the client's social associations; looked results were re-positioned by their relationship with people in the client's interpersonal organization. Informal organization – based customized social hunt approaches the inclinations of different individuals, who should have “comparable” interests as the searcher, give the great expectation to the searcher's inclinations and consequently help with uncovering the query items that may especially fulfill the searcher's needs. Personalization required the capacity of demonstrating the client's inclinations and interests which were finished by following and accumulating the client's cooperation with the framework. These aggregations include user's past inquiries [15], navigate examination [16, 17] and eye following [17] amid the hunt session. The paper portrayed the adequacy of the informal community for personalization: 1) Familiarity based system that speaks to unequivocal commonality associations; 2) Similarity-based system which mirrors the general population like the client by their social exercises; 3) Overall systems that give the two connections types. The idea of social pursuit depicted the inquiry procedure over “social” information gathered from the web, for example, social bookmarking framework, wikis, online journals, gatherings, Social Network Sites (SNSs). The inclinations of client-related individuals are used to help with getting the client's inclinations which are the fundamental thought behind affiliated sifting techniques [4]. The paper also compared the social search based personalization with topic-based personalization [18]. The location feature vector and content feature vector are extracted to represent the records:

#### Location Feature Vector

If location concept  $x_i$  is in a web snippet  $d_k$ , increases the value of location feature vector  $\phi_L(q, d_k)$ . For other location concepts  $x_j$  related to the concept  $l_i$  in the ontology of the location, they are increased according to the following equation in the location feature vector  $\phi_x(q, d_k)$ :

$$\forall x_i \in d_i ; \phi_x(q, d_k)[x_i] = \phi_x(q, d_k)[x_j] + ancestor(x_i, x_j) + descendant(x_i, x_j) + sibling(x_i, x_j) \quad (1)$$

#### Content Feature Vector

If content concept  $y_i$  is in a web snippet  $s_k$ , the content feature vector  $\phi_y(q, d_k)$  increases its value. For other content concepts  $y_j$  that are related to the concept  $c_i$  in the ontology of the content, they are increased according to the following equation in the content feature vector  $\phi_y(q, d_k)$ :

$$\forall y_i \in s_k ; \phi_y(q, d_k)[y_i] = \phi_y(q, d_k)[y_j] + sim_R(y_i, y_j) + ancestor(y_i, y_j) + descendant(y_i, y_j) + sibling(y_i, y_j) \quad (2)$$

The experiment for location preferred personalized search was done using two mining algorithms namely Joachims and spyNB. These two algorithms are applied on different number of users. The result of first part of experiment showed that the higher location entropy  $H_L(q)$  and content entropy,  $H_c(q)$  gives better personalized efficiency because higher-content or location entropy queries are more ambiguous about content or location data. The improvement in average rank for queries with low content entropies,  $H_c(q, u)$  and location entropies,  $H_L(q, u)$  benefited more from personalization because in the search results users concentrated on a few particular topics. The average rank improvement on sorted personalization efficiency on queries with higher extracted location entropy,  $e_L(q, u)$  and extracted content entropy,  $e_c(q, u)$  provides better results. Joachims(content) performed better than Joachin(location) but the combination of both gave 110% improvement. On the other hand, spyNB generated more accurate preferences compared to Joachin's method.

David Carmel [5] introduced a paper that offers a customization approach to web search captures the client's interest and preferences by mining search results in the form of notions and clicking through. Content concepts and location concepts are formed to build ontologies ontology-based, multi-faceted (OMF) profiles to obtain the content and territorial interests of the user. The paper characterizes the adequacy dependent on entropies and uses it to adjust the angles of content and location. SVM is prepared to change a customized positioning capacity for re-positioning. By catching the client's advantages, a customized inquiry middleware adjusts the list items got from general web search tools to the client's inclinations through customized re-positioning of the indexed lists. Content entropies and location entropies measure the decent variety of substance and area data from the indexed lists of a question. Customized adequacy consolidates the client's substance and area inclinations for re-positioning the query items. Two inclination mining calculations to be specific, Joachims Method and SpyNB methods were utilized for the extraction of the client inclinations. As a reporting function, a collection of content concepts and a set of location concepts are extracted from RSVM [19] search results. Utilizing navigate information as input, RSVM finds a linear ranking function.

Social Network and Discovery (*SaND*) retrieves  $N(x)$  - a ranked users list associated with user  $x$ , and  $T(x)$  - a ranked list of associated terms. During the search session, these lists are used to customize the search results for all user queries as a user profile. Given the client's profile,

$$P(x) = (N(x), T(x)) \quad (3)$$

The search results are re-ranked as follows:

$$S_p(k, z | P(x)) = \alpha S_{np}(k, z) + (1-\alpha) [\beta \sum w(x, y) \cdot w(y, z) + (1-\beta) \sum w(x, t) \cdot w(t, z)] \quad (4)$$

$S_p(k, z | P(x))$  – Custom entity z score to query k provided client x' profile.

$S_{np}(k, z)$  – Non-personalized score for SaND of z to k.

$w(x, y)$  and  $w(x, t)$  – Relationship strength of user y and time period t to x, as given with the aid of the client's profile.

$w(y, z)$  and  $w(t, z)$  – relationship strength among y and t to entity z, as given by means of SaND.

$\alpha$  – Relative customization score weight relative to the original non-custom ranking.

$\beta$  – Relative weight between people and configuration words.

The experiment performed off-line study using bookmarks as personal queries. The query was submitted to SaND and the search result were re-ranked using client's profile. Measuring average accuracy (AP) and reciprocal rank (RR) is used to evaluate the information. As a consequence, the Similarity-based strategy's MAP (mean-AP) was three times greater than non-personalized search.

Fang Lui [6] introduced a technique in his paper in order to learn client profiles from client's search histories which were used to improve recovery adequacy. Client profile and a general profile are joined to outline client inquiry into a lot of classifications that speaks to the client's pursuit goals and fill in as a setting to disambiguate the words in the client's inquiry. The paper proposed a two-advance approach to improve recuperation sufficiency. In the underlying advance, the structure normally amasses, for each client, a ton of arrangements for every request set up together by the client in view of the history of his / her concern. The scheme used the classes match scheme to extend the application to guide the internet glance in the second step. The procedure demonstrated and concluded the client's pursuit history, build a client profile dependent history of the searched content and generate a general profile. The framework assembles proper classes for every client question dependent on the client's profile and the general profile. It improved the web seek adequacy by utilizing classifications as a setting for each inquiry. The development of two profiles was finished by utilizing four diverse learning approaches, in particular, k-Nearest Neighbor (kNN), the pseudo-LLSF (pLLSF) approach, the Linear Least Squares Fit (LLSF) approach and (Rocchio) or versatile Rocchio (aRocchio).

Mapping of user query to a set of categories is done by using:

$$1. \text{ Only client profile: } \quad Sim(k, x) = Sim(k, x^u) \quad (5)$$

$$2. \text{ Only general profile: } \quad Sim(k, x) = Sim(k, x^g) \quad (6)$$

$$3. \text{ Combined method 1: } \quad Sim(k, x) = (Sim(k, x^u) + Sim(k, x^g))/2 \quad (7)$$

$$4. \text{ Combined method 2: } \quad Sim(k, x) = 1 - (1 - Sim(k, x^u)) * (1 - Sim(k, x^g)) \quad (8)$$

$$5. \quad \text{ Combined method 3: } \quad Sim(k, x) = \max(Sim(k, x^u), Sim(k, x^g)) \quad (9)$$

where, k is query, x is category,  $x^u$  and  $x^g$  are client profile vectors and general profile vectors, respectively. And  $Sim(k, x)$  is the similarity between query and the category.

The experiment took place in two phases after collecting seven datasets from seven different users. In first phase, client submitted the queries to the search engine. Search record consist of query, set of associated categories and document list. Every query was presented in three distinct methods in the second stage. The effectiveness of pLLSF, KNN and bRocchio is comparable and has been given 76.8% to 97.5% accuracy. After combining client profile and general profile are more accurate than one. General accuracy is as follows:

General Profile < Client Profile < Combination of general and client profiles

Customized seek is an approach to improve the precision of the web look. Successful customized seek requires assembling and accumulating user data which lifts up the major concern of privacy. Xuehua Shen [7] defines four levels of privacy protection in his paper. The levels include Pseudo identification, Group identification, No identification, and No personal information. The paper also explained the analysis of software architecture which includes server-side personalization, customer side personalization, and customer server agreeable personalization. The experiment demonstrated that customer side personalization rules the current server-side customized look in looking after security. Fig. 3 shows the server-side personalization architecture. The personal information is located on the search engine side. The search engine creates and updates user profiles either by explicit input from the user or by implicitly gathering the search history of the user. Fig. 4 shows the Customer side customization (CSP). On the user's laptop, personal data will be gathered. The CSP search agent performs query expansion to generate a brand new question earlier than sending the software to the hunt engine. The agent may additionally re-rank the quest consequences to match the consumer preferences inferred after acquiring the quest results from the quest engine. Fig. 5 Shows collaborative customization of the Client-Server. Customer-side user profile is still stored, but customization also involves the server. At the time of request, the client extracts contextual statistics from the user profile and sends them together with the request to the search engine.

Fig. 3. Server- side personalization [7]

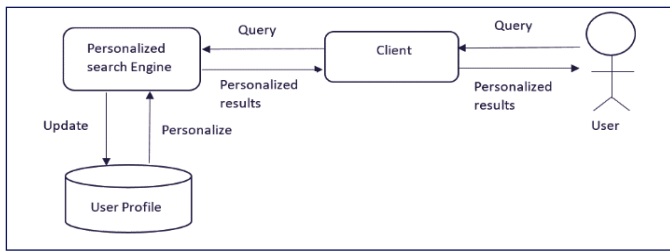


Fig. 4. Client- side Personalization [7]

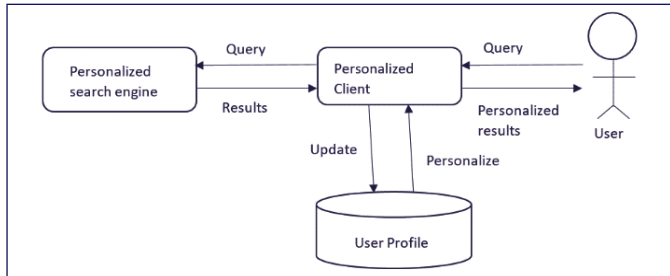


Fig. 5. Client- Server Collaborative Personalization [7]

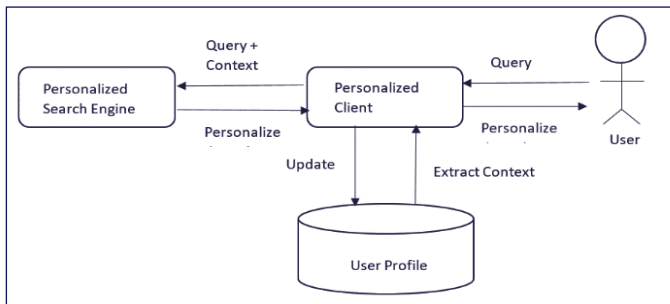
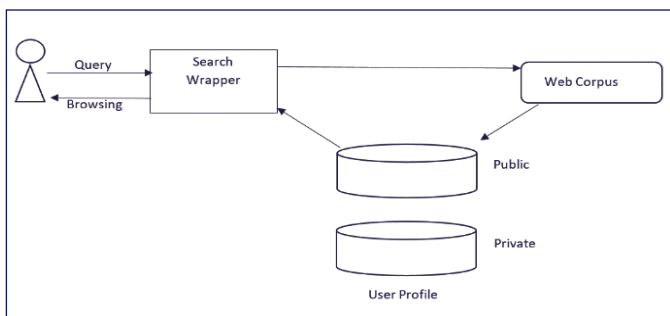


Fig. 6. Privacy enhancing system architecture [8]

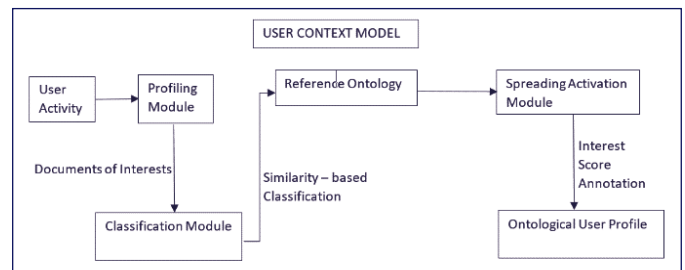


The proportion of information on the web is increasing persistently and it has turned out to be strenuous for the web crawler to discover information that persuades the user’s needs. Customized web seek is an approach to upgrade look quality by altering query items for the user with explicit data. Be that as it may, users feel uneasy by revealing private data to internet searchers. To take care of the issue, Yabo Xu [8] displayed a route for users to naturally assemble rich user profiles in his paper. These profiles embody the user’s interests into a various leveled association as indicated by individual interests. Mindetail and expRatio are the two parameters proposed for security insurance. The outcome demonstrated the improved user’s inquiry quality when contrasted with standard MSN (Microsoft Network) rankings [3]. Fig. 6 shows the privacy-

enhancing system architecture. The algorithm automatically creates a hierarchical user profile representing the implicit private interests of the user. At a greater level, general interests are placed, particular interests are placed at a lesser level. This is performed according to the privacy environments of the user. To integrate a partial user profile with the outcomes obtained, a search wrapper is developed from the search engine on the server side.

Ahu sieg [9] supplied a custom search approach that covered modeling the user’s heritage as ontological profiles by means of implicitly assigning derived interest scores to contemporary ontological thoughts. Via assigning interest ratings to provide thoughts in a website ontology, the approach produces ontological consumer profiles (OUPs). These profiles are retained and updated as annotated specializations of a pre-existing ontology of the reference area. A spreading activation set of rules is used to preserve and an increasing number of replacing the hobby scores based totally on the user’s non-stop conduct. The experimental evaluation is designed to deal with three specific issues: (1) Is the ontological profile coverage of the results of the situation for the individual mind? (2) Do modifications within the outcomes of interest replicate the user’s interest in a selected subject matter well? (3) Can the semantic proof furnished by way of the ontological profiles be used to effectively re-rank the consequences of the net seek for gift the consumer with a customized view? The experimental findings imply that re-ranking the search effects in an OUP primarily based on interest outcomes and linguistic proof efficaciously offers the user with a tailor-made perspective of the outcomes of the search by way of putting the consequences toward the pinnacle while they are maximum relevant to the patron. Fig. 7 represents the OUP version of context model. The OUP is an example of the ontology of reference. For every concept in the consumer profile, an interest score is annotated. The consumer interacts with the system by means of selecting and viewing new files, updating the OUP and altering the annotation by means of dispensing activation for present thoughts. For this reason, based totally on the consumer’s ongoing conduct, the user context is incrementally maintained and up to date.

Fig. 7. Ontological User Profile as the Context Model [9]



Alexander Pretschner [10] explored ways in which user interests can be incorporated in the search system to enhance the search outcomes in his article. The article proposed methods to model the interests of a user into the profile of a user and showed how these profiles can be implemented to retrieve and filter data more effectively. A system that “watches over a user’s shoulder” while surfing the internet is being created. Over time,

via analyzing surfed websites, a personal profile is developed to define their content material and companion that content with the period and time of the file. Whilst web sites on sure topics are repeatedly visited, this is taken as a demonstration of the person's hobby in this difficulty. Except for browsing, there's no need for personal interaction with the system. These profiles can, therefore, be used to enhance the efficiency of the search. The experiment evaluates the usability of these profiles and examined two applications to verify the usability of advanced profiles, namely re-ranking and filtering search results. The experimental outcome showed that the performance was re-ranked by 8 percent and filtered by 9-15 percent of irrelevant papers and 6-12 percent of appropriate papers.

Zhengyu Zhu [11] suggested internet search model-based query extension. This article proposes an innovative query extension strategy. It relies upon on a customized net search business enterprise's representation. The unconventional gadget is solved at the user gadget as a middleware connecting a person to an internet search engine. It is able to implicitly search the fave of the user after which creates the user profile automatically. The recommended technique generates more personalized extension sentences when the user enters question keywords, after which these words are passed at the side of the question keywords to a recognized seek engine which includes Baidu or Google. These extension phrases can make it easier for a user to retrieve search engine data based on their implicit objectives of the search. The novel Web search portrayal can generate a standard custom search engine, particularly through custom query creation the search engine can provide separate search results for separate clients entering the relevant keywords. The experimental findings show the effect and use of the suggested

job on a search engine's personalized data service.

Xuwei Pan [12] suggested adaptive custom Web context-based search to improve the efficiency of information retrieval. The latest net search/metasearch engines are intended to serve all users in extraordinary conditions, regardless of any person's unique requirements. Net seek customization is for each user to conduct retrieval integrating their hobbies. In this technique, the authors proposed a state-of-the-art adaptive personalized method primarily based on context to comply the quest consequences in line with every user's need for suitable statistics on awesome occasions with moderate user attempts. Following the context-based adaptive personalized seek investigation process, three widespread strategies are given for the execution of this technique, that is net aid semantic indexing, obtaining and matching semantic similarities among internet resources and user context, modeling and user context. Experimental observations demonstrate that most users carry out the adaptive custom search system and that the strategy to customizing web search is efficient.

B. Smyth [13] suggested a Community-based customization strategy for web search. Researchers can influence the underlying understanding generated in search groups by collecting community-level search habits of users-the queries they enter and the outcomes they choose. They can use this information to build a relevant model throughout conventional web search to promote community-relevant outcomes. This paper makes a specialty of the cooperative web search technique that promotes the idea that community search behaviors can generate and share precious kinds of search information of data enables normal search engine outputs to be adapted.

S. No.	Paper Title	Description	Approach	Efficiency
1.	SoPRa: A New Social Personalized Ranking Function for Improving Web Search.	Presented a fresh document ranking function considering the internet social context to improve the Information Retrieval (IR) model.	Context Model	Medium
2.	Enabling Personalized Search over encrypted outsourced data with efficiency improvement	Proposed a paper that handles the issue of Personalized Multi-keyword Ranked Search over Encrypted data (PRSE).	Ontological Model	High
3.	Personalized web search using Probabilistic Query Expansion	Created a custom Web search system applied to the proxy by generating user profiles using collaborative filtering to change user interests.	User Profile Model	High
4.	Personalized Web Search with Location Preferences	Proposed a method that examines the customized social pursuit dependent on the client's social associations.	Ontological Model	Medium
5.	Personalized social search based on the user's social network	Introduced a method that offers a customization approach to capture the client's interest and preferences by mining search results in the form of notions and clicking through.	Ontological Model	High
6.	Personalized web search for improving retrieval effectiveness	Introduced a technique in order to learn client profiles from client's search histories which are used to improve recovery adequacy.	Ontological Model	High
7.	Context-based Adaptive personalized web search for improving information retrieval effectiveness	Proposed a personalized method primarily based on context to comply the quest consequences in line with every user's need for suitable statistics.	Context Model	Low
8.	Privacy Protection in Personalized search	The paper explained the analysis of software architecture which includes server-side personalization, customer side personalization, and customer server agreeable personalization.	Context Model	Low

9.	Privacy- enhancing personalized web search	Displayed a route for users to naturally assemble rich user profiles in his paper to overcome the feeling of <u>uneasiness of the users by revealing their private data.</u>	Ontological Model	High
10.	Learning ontology-based user profiles: A semantic approach to personalized web search	Proposed a custom search approach that covered the user's history as ontological profiles by means of implicitly assigning derived interest scores to contemporary ontological thoughts.	Ontological Model	High
11.	Ontology based Personalized Search	The article proposed methods to model the interests of a user into the profile of a user and showed how these profiles can be implemented to retrieve and filter data <u>more effectively.</u>	Ontological Model	Low
12.	Query expansion based on a personalized web search model	Proposes an innovative query extension strategy that relies upon on a customized net search business enterprise's representation.	User Profile Model	Medium

#### 4. CONCLUSION

In this paper, we addressed a number of personalized methods on personalized information retrieval done by using user's search history, social connection, location, bookmarks and privacy. A deep analysis of different methodologies and their experimental results has been done. A social personalized ranking function gives the intuition about the webpages to which the user is connected that helps in knowing the interest of the user. Another method suggested two mining algorithms namely Joachims and spyNB in which results showed that spyNB is better than Joachims. The combination of both the algorithms gave 110% improvement. A similarity-based strategy was introduced in an experiment which have performed off-line study using bookmarks as personalized queries and concluded that the introduced strategy was three times greater than non-personalized search. Another experiment used three algorithms namely pLLSF, KNN and bRocchio for the generation of client profile and general profile. The result showed that the combination of both the profiles gave more accurate results than using one at a time. Our evaluation suggests that every method give much better result than the non- personalized methods. Since each method has its own pros and cons but every technique of retrieving personalized information is effective. We will next evaluate privacy protection in personalized information retrieval.

#### REFERENCES

1. Mohamed Reda, Hakim Hacid, Mokrane Bouzeghoub: *SoPRA: A New Social Personalized Ranking Function for Improving We Search*. In: *ACM SIGIR*, August (2013).
2. Pallavi Palleti, Harish Karnick, Pabitra Mitra: *Personalized web search using Probabilistic Query Expansion*. In: *IEEE/WIC/ACM International Conference of Web Intelligence and Intelligent Agent Technology* (2007).
3. Zhangjie Fu, Kui Ren: *Enabling Personalized Search over encrypted outsourced data with efficiency improvement*. In: *IEEE Transaction on parallel and distributed systems*, vol. 27, No. 9, September (2016).
4. Kenneth Wei-Ting Leung, Dik Lun Lee, Wang-Chien Lee: *Personalized Web Search with Location Preferences*. In: *ICDE Conference* (2010).
5. David Carmel, Naama Zwerdling, Ido Guy, Shila Ofek-Koifman, Nadav Har'el, Inbal Ronen, Erel Uzeil, Sivan Yogev, Sergey Chernov: *Personalized social search based on the user's social network*. In: *CIKM' 09*, November (2009).
6. Fang Liu, Clement Yu: *Personalized web search for improving retrieval effectiveness*. In: *IEEE Transaction on knowledge and data engineering* vol. 16, No. 1, January (2004).
7. Xuehua Shen, Bin Tan, ChengXiang Zhai: *Privacy Protection in Personalized search*. In: *ACM SIGIR Forum*, vol.41, No.1, June (2007).
8. Yabo Xu, Benyu Zhang, Zheng Chen, Ke Wang: *Privacy-enhancing Personalized web search*. In: *International World Wide Web Conference (WWW 2003)*, Canada, May (2007).
9. Ahu Sieg, Bamshad Mobasher, Robin Burke: *Learning ontology based user profiles: A semantic approach to personalized web search*. In: *IEEE Intelligent Informatics Bulletin*, vol.8 No. 1, November (2007).
10. Alexander Pretschner, Susan Gauch: *Ontology based Personalized Search*. Partially supported by National Science Foundation CAREER Award Number 97-03307.
11. Zhengyu ZHU, Jingqiu XU, Xiang REN, Yunyan TIAN Lipei LI: *Query expansion based on a personalized web search model*. In: *IEEE Third International Conference on semantics, Knowledge and Grid* (2007).
12. Xuwei Pan, Zhengcheng Wang, Xinjian Gu: *Context-based Adaptive personalized web search for improving information retrieval effectiveness*. In: *IEEE* (2007).
13. Barry Smyth: *A Community-based approach to personalizing web search*. In: *The IEEE computer society*, August (2007).
14. Fang Liu, Clement Yu, Weiyi Meng: *Personalized web search by mapping user queries to categories*. In: *CIKM' 02*, November (2002).
15. B.Tan, X. Shen, C.Zhai: *Mining long term search history to improve search accuracy*. In: *ACM press*, pages 718-723, (2006).

16. Zhicheng Dou, Ruihua Song, Ji-Rong Wen: *A large scale evaluation and analysis of personalized search strategies*. In: *International World Wide Web Conference Committee*, May (2007).
17. T. Joachims, L. Granka, B.Pang, H. Hembrooke, G. Gay: *Accurately interpreting clickthrough data as implicit feedback*. In: *SIGIR*, pages 154-161, (2005).
18. S. Xu, S. Bao, B. Fei, Z. Su, Y. Yu: *Exploring folksonomy for personalized search*. In: *ACM SIGIR*, pages 155-162, (2008).
19. Filip Radlinski, Susan Dumais: *Improving personalized web search using result diversification*. In: *ACM SIGIR'06*, August (2006).
20. <https://whatis.techtarget.com/definition/personalized-search>
21. Nathaneal Ramesh, J.Andrews: *Personalized search engine using social networking activity*. In: *Indian Journal of science and technology*, vol8 (4), 301-306, February (2015).
22. Gabriella Pasi: *Issues in Personalizing Information Retrieval*. In: *IEEE Intelligent Informatics Bulletin*, vol.11 No. 1, December (2010).
23. Arnau Erola, Jordi- Castella-Roca, Alexandre Viejo, josep M.Mateo- Sanz: *Exploiting Social networks to provide privacy in personalized web search*. In: *The Journal of system and software*, May (2011).
24. Jia Hu: *Personalized web search by using learned user profiles in re-ranking*. A thesis, Florida Institute of technology, May (2008).
25. Pau Valles Fradera: *Personalizing web search and Crawling from Clickstream Data*. Master thesis, Universitat Politecnica De Catalunya, January (2009).
26. Mariam Daoud, Lynda Tamine-Lechani, Bilal Chebaro, Mohand Boughanem: *A session based personalized search using an ontological user profile" in Proc of SAC' 09*, March (2009).
27. Adan Ortiz-Cordova, Bernard J. Jansen: *Classifying web search queries to identify high revenue generating customers*. In: *Journal of the American Society for Information Science and Technology*, January (2012).
28. Kenneth Wai-Ting Leung, Wilfred Ng, Dik Lun Lee: *Personalized concept-based Clustering of Search engine queries*. In: *IEEE Transactions on Knowledge and Data Engineering*, vol.20, No. 11, November (2008).
29. Jae-Wook Ahn, Peter Brusilovsky Daqing He, Jonathan Grady, Qi Li: *Personalized web exploration with task models*. In: *International World Wide Web Conference Committee*, April (2008).
30. Orland Hoerber: *Web information retrieval systems: The future of web search*. In: *IEEE/WIC/ACM International Conference of Web Intelligence and Intelligent Agent Technology*, (2008).
31. Zhicheng Dou, Ruihua Song, Ji-Rong Wen, Xiaojie Yuan: *Evaluating the effectiveness of personalized web search*. In: *IEEE Transactions on Knowledge and Data Engineering*, vol.21, No. 8, August (2009).
32. Cyrus Shahabi, Yi-Shin Chen: *Web information personalization: Challenges and Approaches*. In: *Springer-Verlag Berlin Heidelberg* (2003).
33. Fang Qiu, Junghoo Cho: *Automatic identification of User interest for personalized search*. In: *International World Wide Web Conference Committee*, May (2008).
34. Steve Lawrence: *Context in web search*. In: *IEEE Computer Society Technical Committee on Data Engineering*, (2000).
35. S. Gauch, M. Speretta, A. Chandramouli, A. Micarelli: *User profile for personalized information access*. *The adaptive*, pp. 54-89 (2008).
36. T.Havelivali: *Topic Sensitive PageRank*. In: *12<sup>th</sup> International World Wide Web Conference (WWW 2003)*, Hawaii (2002).
37. N. Cao, C. Wang, M. Li, K. Ren, W. J. Lou: *Privacy preserving multi-keyword ranked search over encrypted cloud data*. In: *IEEE INFOCOM*, pp. 829-837, (2011).
38. J.Anil, K.F. Bharathi: *Improving the efficiency of personalized web search using scoring function*. In: *International Journal or advanced research in computer science engineering and technology (IJARCET)*, vol. 5, issue 10, October (2016).
39. Miss Jayashri Divakar, Prof. D. R. Patil: *Personalized and effective search over encrypted data on cloud: A survey*. In: *Open Access International Journal of Science and Technology (OAIJST)*, vol. 2, special issue, December (2017).
40. T. Joachims: *Optimizing search engines using clickthrough data*. In: *ACM SIGKDD Conference*, (200)

#### AUTHORS

**Ms. Sakshi Singh**, Research Scholar, Ambedkar Institute of Advanced Communication Technologies And Research, Krishna Nagar Road Chacha Nahru Bal Chikitsalaya, Geeta Colony, New Delhi - 110 031  
Email: 010sakshisingh@gmail.com

**Suresh Kumar**, Associate Professor, Ambedkar Institute of Advanced Communication Technologies And Research,, Krishna Nagar Road Chacha Nahru Bal Chikitsalaya, Geeta Colony, New Delhi - 110 031  
Email: sureshpoonia@aiactr.ac.in