Journal of King Saud University – Computer and Information Sciences (2018) 30, 206–222

King Saud University

Journal of King Saud University – Computer and Information Sciences

www.ksu.edu.sa
www.sciencedirect.com

A Secured Cognitive Agent based Multi-strategic Intelligent Search System

Neha Gulati *,[1], **Atul Garg**

M.M. Institute of Computer Technology & Business Management, MM University, Mullana, Ambala, Haryana, India

Received 30 January 2016; revised 6 June 2016; accepted 28 June 2016
Available online 6 July 2016

KEYWORDS

BDI model;
Cognitive Agent;
Emotion;
Information retrieval;
Intelligent search;
Search Engine

Abstract Search Engine (SE) is the most preferred information retrieval tool ubiquitously used. In spite of vast scale involvement of users in SE's, their limited capabilities to understand the user/searcher context and emotions places high cognitive, perceptual and learning load on the user to maintain the search momentum. In this regard, the present work discusses a Cognitive Agent (CA) based approach to support the user in Web-based search process. The work suggests a framework called Secured Cognitive Agent based Multi-strategic Intelligent Search System (CAbMsISS) to assist the user in search process. It helps to reduce the contextual and emotional mismatch between the SE's and user. After implementation of the proposed framework, performance analysis shows that CAbMsISS framework improves Query Retrieval Time (QRT) and effectiveness for retrieving relevant results as compared to Present Search Engine (PSE). Supplementary to this, it also provides search suggestions when user accesses a resource previously tagged with negative emotions. Overall, the goal of the system is to enhance the search experience for keeping the user motivated. The framework provides suggestions through the search log that tracks the queries searched, resources accessed and emotions experienced during the search. The implemented framework also considers user security.

Abbreviations: BDI, Belief Desire Intention; CA, Cognitive Agent; CAbMsISS, Cognitive Agent based Multi-strategic Intelligent Search System; CIMM, Cognitive Information Mapping Model; COGSEMO, Cognitive Search Engine based on Emotions; MsIS, Multi-strategic Intelligent Search; OTP, One Time Password; PSE, Present Search Engine; Q, Query; QRT, Query Retrieval Time; R, Resource; RDJI, Retrieved Documents Judged Irrelevant; RDJR, Retrieved Documents Judged Relevant; SE, Search Engine; SQ, Semantic Query; Webcam, Web Camera

* Corresponding author.
E-mail addresses: nehagulati_pu@rediffmail.com (N. Gulati), atulgarg1411@gmail.com (A. Garg).
[1] University Business School, Panjab University, Chandigarh, India.

Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

http://dx.doi.org/10.1016/j.jksuci.2016.06.005

1. Introduction

The size of the Web is growing at a very fast pace with billions of websites available today. Majority of users use SE's as the main tool to retrieve information from this huge repository (Rangaswamy et al., 2009). The scalability of the Web and everyday growing number of users poses formidable challenge to the SE's. The users submit their information needs to the SE's in the form of query consisting of a sequence of 2–3 keywords (Barr et al., 2008; Guo et al., 2009). In the present scenario, there is gap between the user's needs and results retrieved from the SE's. The reason attributed for this is the inability of the SE's to understand the user context for the query (Gulati and Garg, 2015a). As, the SE's do not understand the relations among concepts embedded into semantic annotations (Lamberti et al., 2009), therefore, understanding the user context presents another major challenge. This leads to inappropriate understanding of the query and thereby results in irrelevant retrieval i.e. retrieving from varying ranges of the query (Alafif and Sasi, 2012). Thus, the user needs to reframe the query multiple times to retrieve the relevant results.

Emotions are pervasive in online environment and certainly affect the user engagement in search activity, motivation to persist search in the face of irrelevant retrieval and desire to stay on the task. The method used to perform search and alternatives chosen are greatly influenced by affective states (emotions) experienced during search process (Flavian-Blanco et al., 2011; Kim, 2008). In addition, emotions that arise during search also act as "affective filters" for the user to evaluate the relevance of information (Nahl, 2005). The ill-defined tasks may put more strain on the user emotions (Kim, 2008). Emotion control has a significant effect on search behavior and users with different levels of emotion control tend to search Web differently (Kim, 2008). Overall, the search process ends with positive emotions if the user feels good and enjoys the search. In contrast, negative emotions generated from negative task-related feelings decrease the performance. The feelings of anxiety, fear and nervousness generated during the user interaction with computers influence the user's performance on computer related tasks and other affective processes (Liaw and Huang, 2006; Yee et al., 2004). Users with lower emotion control are more likely influenced by negative feelings spawned while working on complex task; and their performance might suffer. They are less likely to handle pressure well, which might make them stressed and distracted, therefore making more frequent, hasty and inefficient search moves. The work thereby suggests providing a mechanism to decrease cognitive and affective burden on the user to prevent stress and getting distracted during the search process.

Keeping this in view and the problem of irrelevant retrieval, there is a need to keep the user motivated during the search process. For this, the authors have suggested to include user emotions as a crucial parameter in the design of a Multi-strategic Intelligent Search System based on user context and emotions. Since, a CA as defined by (Lawniczak and Di Stefano, 2010) performs the cognitive acts of perceiving, reasoning, judging, responding and learning in a way similar to human beings, it is suggested to design a Cognitive Agent based Multi-strategic Intelligent Search System. Belief, Desire and Intention as described by (Rao and George, 1995) are the three components that constitute the "brain" of the agent in the proposed work.

2. Research contributions

Contributions of this research fall into the following categories:

- Clear knowledge of problems of the SE's from the aspect to achieve user satisfaction during search process.
- Framework called Secured Cognitive Agent based Multi-strategic Intelligent Search System (CAbMsISS) to perform search based on the user's cognitive style, context, emotions and timestamp.
- Implementation of the Secured CAbMsISS framework called COGSEMO to improve user satisfaction in terms of QRT, Precision and Proportion of RDJI.
- Analyzing the results retrieved and comparison of PSE: Google with Secured CAbMsISS.

The organization of the paper is as follows: Section 3 gives an overview of unsolved problems of SE's and the need for MsIS. In Section 4, we present our proposal for a Secured CAbMsISS framework. Different sub-sections in this section describe the functionality of the proposed framework. Section 5 elucidates the implementation of Secured CAbMsISS framework (COGSEMO) along with the experimental setup. Graphs and tables present the performance evaluation of the implemented framework in Section 6 followed by conclusion and future work in Section 7.

3. Problem formulation

The incapability of the SE's in understanding the searcher's diverse needs, choices and feelings experienced during the search process yields superfluous and useless information. Fig. 1 shows the result set obtained from PSE: Google (as on 20th Jan 2016, 11:26 am) for the query (Q1) = ns3. The user's context for Q1 was to fetch content related to nonstructural protein 3 (ns3) found in Hepatitis C virus. But, none among the top 10 results related to the protein ns3. Instead, the entire result set was associated with the theme of computer science. In this case, Proportion of RDJR = 0 i.e. Precision = 0. The SE's cannot determine the user context (motive/interest area) while searching (Gulati and Garg, 2015a). Instead, they are dependent on the searcher for formulating the query using all the required keywords.

The query (Q1) was restated as Q2 = ns3 + protein for retrieving the relevant content. The result set retrieved by PSE: Google (as on 20th Jan 2016, 11:35 am) for Q2 is presented in Fig. 2. The resources (results) obtained are consecutively numbered as R1, R2 and so on for elaboration. The following actions were performed on these resources: R1 – read (click the resource to open it), R2 – read and save (store it), R3 – read and found frustrating and R8 – read and save.

The user can manually save any resource for future using the following feature of the Web browser:

- adding the resource to the favorites (Browser – Internet Explorer),
- creating bookmarks (Browser-Mozilla Firefox, Google Chrome and Apple Safari).

This aspect relates to the browser and not to the SE, hence if the user shifts to another Web browser, then the favorites or

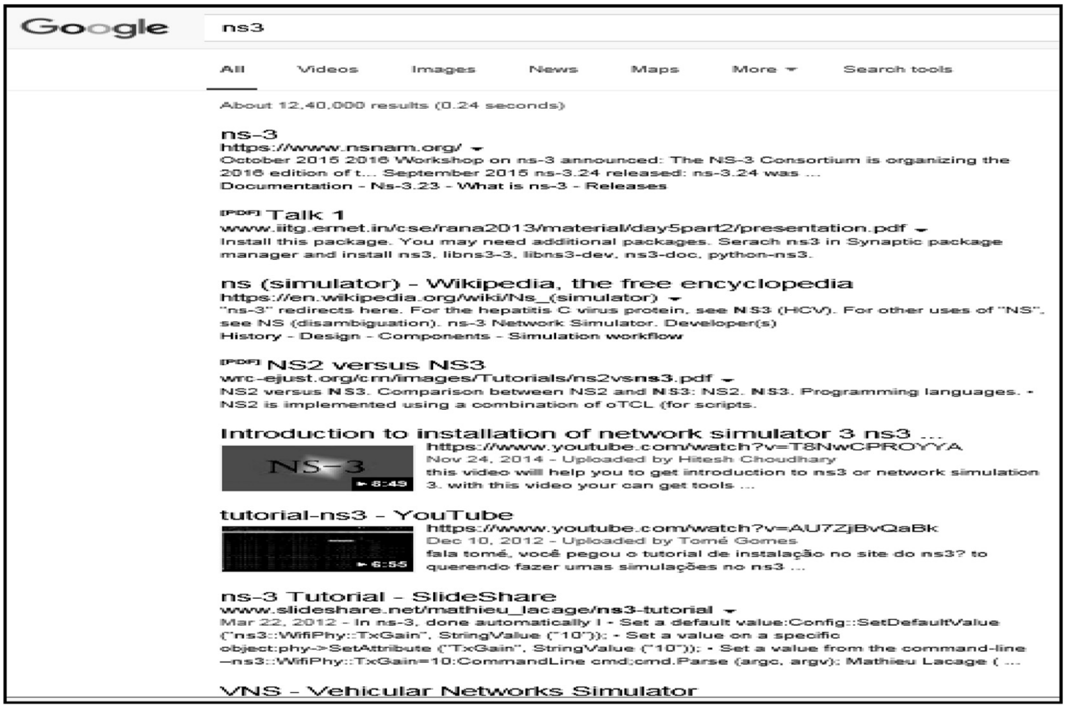


Figure 1 Results retrieved from PSE for query (Q1) = ns3.

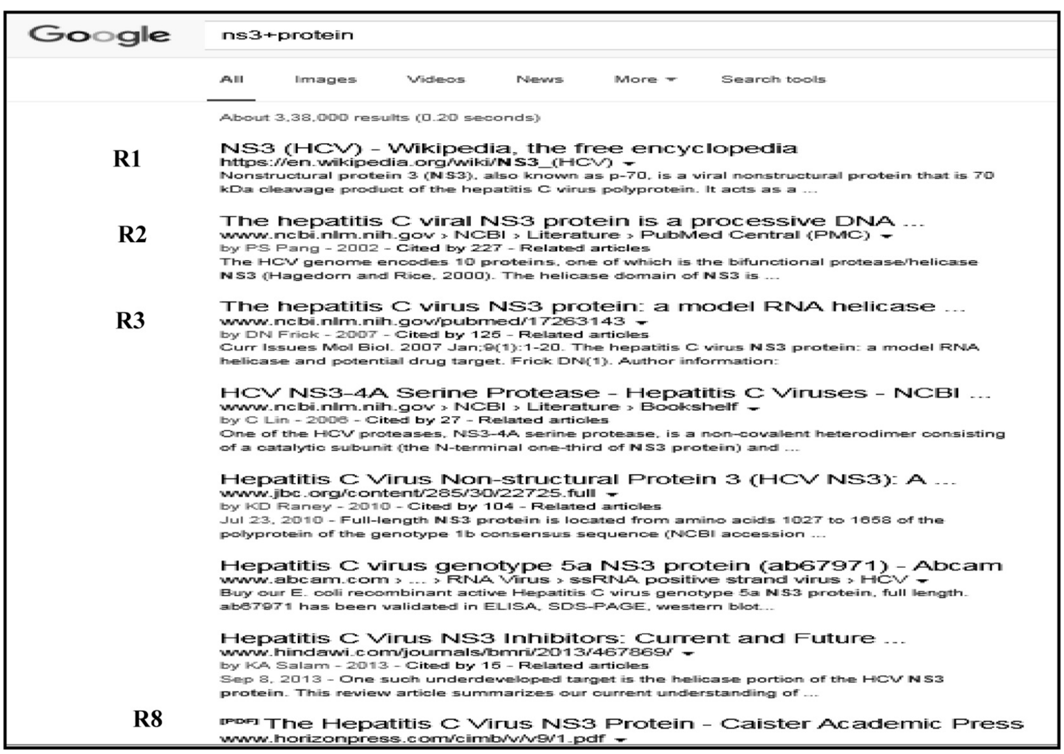


Figure 2 Results retrieved from PSE for query (Q2) = ns3 + protein.

bookmarks needs to be exported to the new browser. Thus, demanding action from the user.

Another concern is the SE's inability to notify regarding a formerly accessed resource associated with negative emotions or stated as useless, thus avoiding its recurring access. Moreover, the SE's also do not list the queries searched along with the corresponding resources accessed (Gulati and Garg, 2015b). Thus, placing high cognitive, perceptual and learning burden on the searcher to memorize the queries along with the context and resources accessed.

Emotions impact learning, engagement and achievement in online environment (Artino, 2012). While searching, the searcher might encounter some feelings (like anxiety, joy etc.). In spite of the effect of human emotions on human–computer interaction

(Charlton, 2009), the human affective element (emotions) has been ignored by the SE's. No concern has been given to the emotions felt during searching and hence, the SE's do not relate the emotions with the post-search user contentment. There exists no mechanism to associate emotions with the queries executed and resources accessed. This wastes user's time in repetitive access of some resource formerly established as not useful or frustrating. It has paved the way in the present work to consider emotions as a critical factor in revealing Web-based search experience.

The above-discussed issues provide ample support for considering adaptivity in the SE's. Adapting the SE to satisfy the user's needs efficiently entails modeling the user behavior. The user behavior refers to the way user processes and organizes information i.e. user's cognitive style (Belk et al., 2013). It emphasizes on the strengths and preferences during information processing. The user's cognitive style correlates with the performance in Web-based environment (Wang et al., 2006; Tsianos et al., 2008). In addition to this (Kinley et al., 2010) also elicited the relationship between the Web user searching behavior and cognitive style. Supplementing this association, navigation metrics have been suggested and utilized for grouping the users on the basis of the navigation behavior (linear and non-linear) (Belk et al., 2013). This provides strong foot for modeling the user in the present work based on cognitive style, context, emotions and timestamp.

4. Proposed framework

The researchers suggest performing the Web-based search by utilizing the user's cognitive style, context, emotions and timestamp. The proposed search called MsIS is illustrated in Fig. 3. User modeling is the key component of the proposed work. The user's static data are collected through the registration form filled during the process of "New User Registration". The dynamic data (Belk et al., 2013) are gathered using behavioral analysis, self-reports and image recognition.

- Behavioral data refer to the user's navigational conduct (queries initiated, Web-pages accessed, activities performed and timestamp).
- Self-reports (Flavian-Blanco et al., 2011) are based on emoticons (emotions self-described by the user).
- Web Camera (Webcam) captures the user image. Image recognition through clmtrackr[2] helps to extract the emotions experienced by the user during Web-based search. The clmtrackr is a java script library used to fit facial models to faces in images and video.

The user guided modeling and dynamic user modeling has been employed to prepare the user's searching model. The user's Web-browsing actions facilitate in identifying the regularities in the keywords used, path followed to access the content (resource) and emotions felt. Clustering the queries and resources based on emotions, timestamp etc. guides the reasoning, judging and decision making process.

The design of MsIS inspired by BDI model (Rajendran and Iyakutti, 2009) of CA is based on five-layer architecture of CA (Lawniczak and Di Stefano, 2010) as described below:

- Perceptual Layer is responsible for perceiving the following information from the environment:
  - user through the Webcam during authentication,
  - queries initiated,
  - Web-resources accessed and
  - emotions generated during searching.

The judging block extracts the features, identifies them and provides estimate.

- Reasoning Layer also referred to as thinking layer forms the decisions by utilizing the information obtained from Perceptual Layer and existing knowledge. The decisions are then passed to the Judging Layer.
- Judging Layer not only accepts the information from Reasoning Layer, but also conveys the processed information to Reasoning Layer for enhancing the decisions.
- Response Layer receives the content from Judging Layer, applies the rules based on the situation and further instructs Perceptual Layer to issue the response.
- Learning Layer is responsible for monitoring the actions performed and thus, deriving new knowledge. It is responsible for updating the existing knowledge based on which the Reasoning Layer functions.

Besides this, CIMM discussed in (Gulati and Garg, 2015b) provides the foundation for Secured CAbMsISS (Fig. 3). The proposed framework illustrated in Fig. 3 performs Web-based search on the basis of CIMM (cognitive style and emotions), context and timestamp. The high-level view of Secured CAbMsISS as shown in Fig. 4 consists of three sub-systems namely: User Entry, User Authenticate and Secured Search. Two-layered security integrated in the framework CAbMsISS provides a secured search environment to the user.

4.1. User Entry

This sub-system (labeled as 1) in Fig. 4 ensures the security in the system. It permits either registering as new user (sub-module Register of module Entry) or entering the login identification for the existing user (sub-module Login of module Entry).

4.1.1. Register

The module Entry initiates the sub-module Register for the new user and generates a registration form to be filled by the user. Filling the registration form is a prerequisite to initiate the Web-based search process using Secured CAbMsISS framework. The registration form consists of the following fields: Name, E-Mail, Date of Birth, Sex, Qualification, Organization of the user, Work Field, Interest and Address. These fields guide the system to provide adaptive search results. The field E-Mail augments security in the system. This module generates and sends an OTP to the user E-Mail for authenticating the user during new registration.

If OTP entered by the user

= OTP generated by the system, then the user is authenticated.

Finally, the user image is captured through Webcam and the user details including image are stored in the Database (Fig. 4). Algorithm: Register illustrates the rationale for this sub-module.

[2] It is an open source distributed under the MIT license. It is available at: https://github.com/auduno/clmtrackr.

Algorithm: Register [New User Registration]

1. Display Registration Form consisting of fields: Name, E-Mail, Date of Birth, Sex, Qualification, User's Organization, Work Field, Interest and Address.
2. Generate OTP and send it to User E-Mail.
3. Allow User to enter OTP.
4. If (OTP entered matched), then : /* User is Authenticated */
   a. Activate Webcam to capture User image.
   /* Image stored is used for User Recognition */
   b. Authenticate = True.
   c. Store User details and image in the Database.

   Else: /* OTP not matched, User not Authenticated */
   a. Authenticate = False.
   b. Goto step 2.

   [End of If structure]

[End of Algorithm Register]

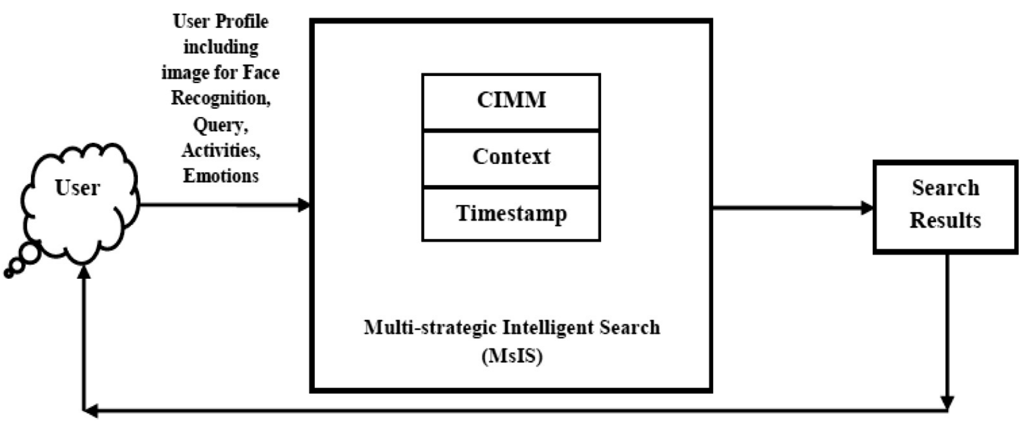


Figure 3 Framework – Secured CAbMsISS.

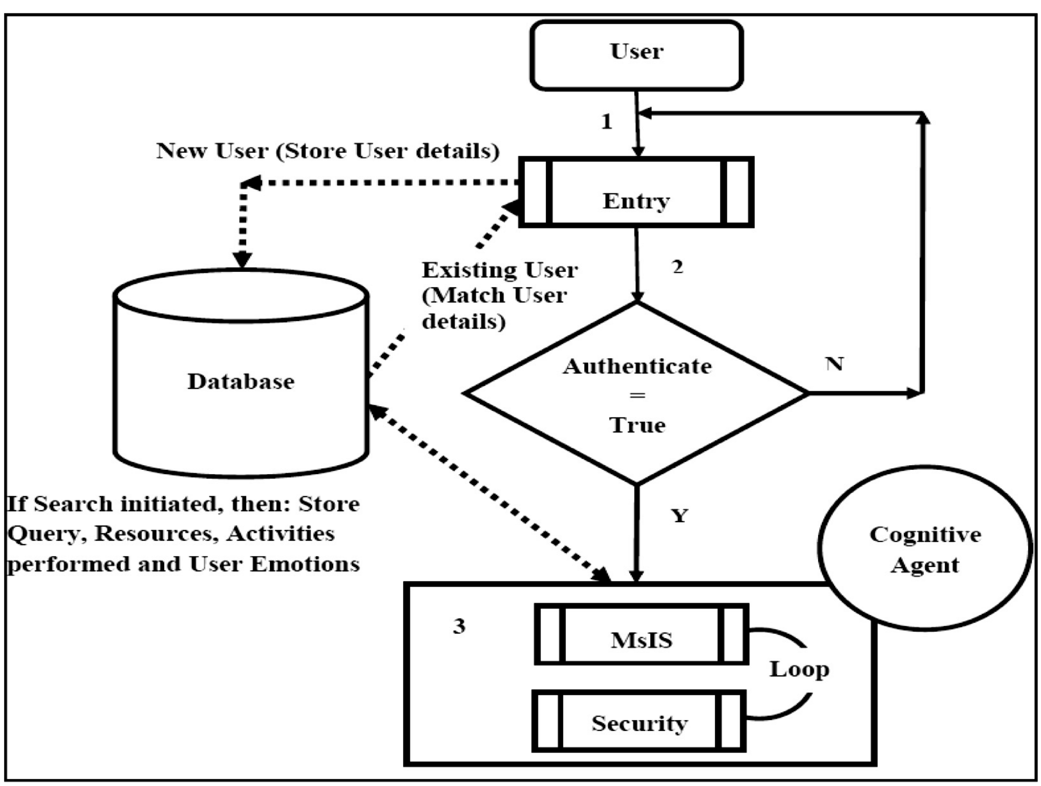


Figure 4 High-level view of Secured CAbMsISS.

4.1.2. Login

The Login sub-module is initiated by the module Entry for the exiting user. It accepts the E-Mail from the existing user and searches it in the Database. If the E-Mail entered by the user is found in the Database i.e. the user already exists, then Webcam captures the user image. The captured image is then matched with the user image stored in the Database. If it matches, then the user is authenticated. In case, the user image does not match, then this module generates and sends an OTP to the user E-Mail for authentication. On entering the correct OTP, the user is permitted to access the system. The access is denied if the:

- user E-Mail does not match with the E-Mail stored in the Database or
- OTP does not match.

Algorithm: Login highlights the user authentication process.

Algorithm: Login [Existing user]

```
1.  Enter User E-Mail.
2.  /* Match E-Mail with the Database */
    If (E-Mail matched), then:
        a. Activate Webcam to capture User image for User Identification.  /* Face Recognition */
        b. If (User Image matched / Identified) then:  /* User is Authenticated */
                              Authenticate = True.
           Else:       /* Image not matched, User is not Authenticated */
                               i. Authenticate = False.
                              ii. Display "User not Identified".
                             iii. Generate OTP and send it to User E-Mail.
                              iv. Allow User to enter OTP.
                               v. If (OTP entered matched), then : /* User is Authenticated */
                                      Activate Webcam to capture User image.
                                      /* Image stored is used for User Recognition */
                                      Authenticate = True.
                                      Store User image in the Database.
                                  Else:  /* OTP not matched, User not Authenticated */
                                      Authenticate = False.
                                      Goto step iii.
                                  [End of If structure]
           [End of If structure]
    Else:  /* E-Mail not matched */
        Authenticate = False.
        Display : "E-Mail not Matched".
        Goto step 1.
    [End of If structure]
[End of Algorithm Login]
```

4.2. User Authenticate

After successfully crossing the security check, this sub-system (labeled as 2) authenticates the user and provides the access to search by transferring the control to the sub-system called Secured Search. The unauthorized user is denied the access and control is transferred to User Entry.

4.3. Secured Search

The sub-system labeled as 3 presents the searching interface to the authenticated user. The modules MsIS and Security (Fig. 4) control the functionality of this sub-system.

4.3.1. MsIS

It executes search on the basis of user cognitive style, query context, emotions and timestamp. CIMM presented in Gulati and Garg (2015b) enhances the present work by enabling the MsIS module to capture the user search activities, emotions and timestamp.

4.3.1.1. CIMM (Gulati and Garg, 2015b). The authors in the present work have also thoroughly discussed the functionality of the modules (Search Activity, Capture, Store, Decision Making and Search Suggestions) presented in the high-level view of CIMM (Gulati and Garg, 2015b). Fig. 5 illustrates the steps performed to capture the searcher's cognitive style, context, emotions and timestamp. Since, according to (Flavian-Blanco et al., 2011), the emotional outcomes of an online search process can be influenced by different structures of perceptions, initial affective states and emotions felt during searching, thus CIMM is instructed to capture the emotions during the search process.

The different terms used to illustrate the working of modules Search Activity and Capture are as follows:

R – Resource, RL – Resource List (list of resources previously accessed), SE – Search Engine, Q – Query, QL – Query List (list of previously searched queries), W – Window. The module Capture is called when the user inputs a query or selects an existing query from QL. It is also activated when some R is selected. The algorithm Capture(Q) acquires the information associated with the search process.

The following actions and sub-modules describe the working of the module Capture:

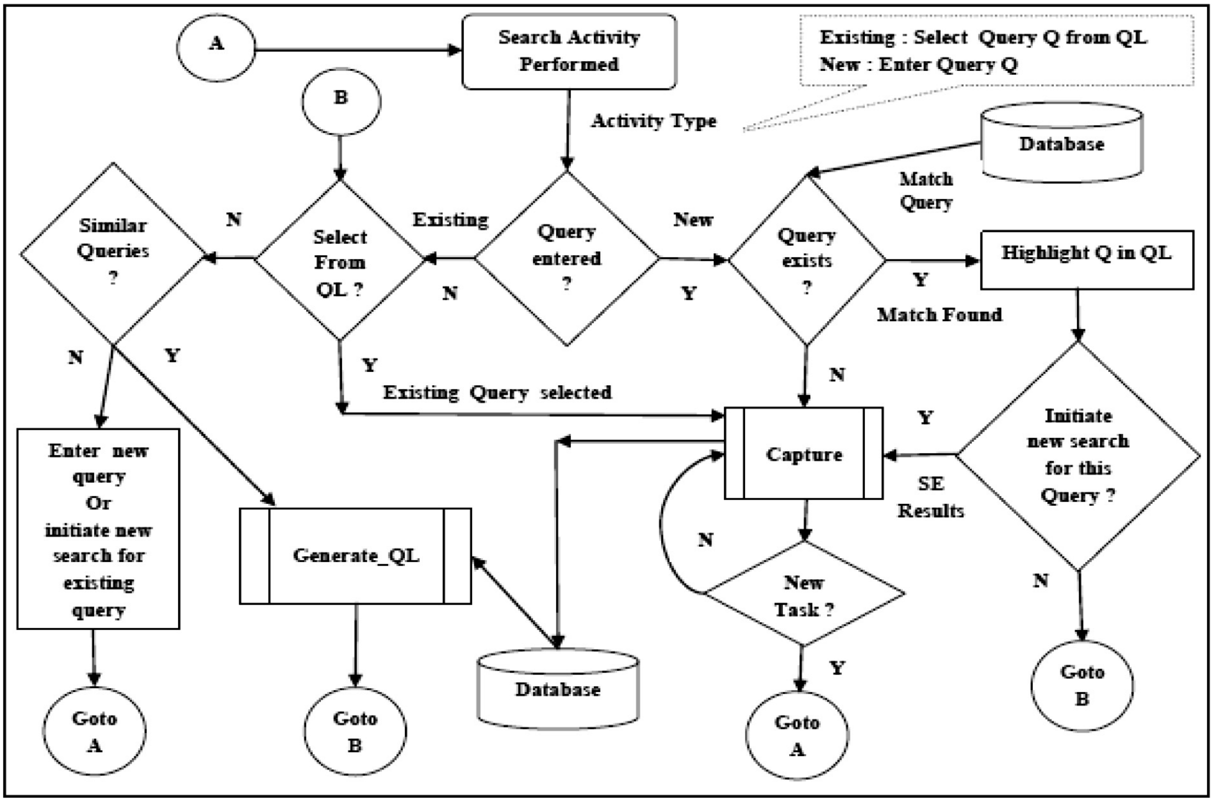


Figure 5 Functionality of module Search Activity of CIMM.

- ACT{Q} (Action performed on Q): If action performed = Click, then call module Record_keywordlog (Q).
- Module Record_keywordlog(Q): It stores the details of Q including query, logstamp and emotions captured through Webcam in the Database. It also displays RL corresponding to Q. In case Q is selected from QL, then predefined RL is displayed, otherwise SE generates a new RL for Q.
- ACR{RL} (Action performed on RL): If action performed = Click, then Resource R is selected from RL and module Record_searchlog(R) is called.
- Module Record_searchlog(R): It captures and stores the details of the resource R accessed, logstamp, activity performed on resource and user emotions.
- ACT{W.R[Current]} (Action performed on window W of current resource R): If action performed = Close, then the user has closed the current resource R. Decrease the existing number of opened resources by 1. The module Record_searchlog(R[Current]) now stores the current status of the closed resource R[Current] and the window for R [Current] is closed.
- Module Check_window(i, current): Checks the current resource window that is presently activated by the user.
- ACT{W.RL} (Action performed on the window of RL): If action performed = Close, i.e. the user no longer wishes to select an existing resource for the query Q. If (ACT{RL} ! = Click and ACT{W.RL} = Close), then no action is performed by the user for Q. No resource is selected for Q from RL. In this case, it allows the user to again start the search process for some task (same query using SE or new query).

Algorithm : Module Check_window(i, current)

```
Repeat for (k = 1 ,2 ,….., i)
[i is the total no. of resource windows opened by the user at any time]
{
    If ACT{W.R[k]} == Click
    {
            current = k.
    }
    [End of If structure]
    break.
}
```

/* End of Algorithm Module Check_window */

The logic used to describe working of the module Capture is illustrated using Algorithm: Capture(Q).

Algorithm: Capture(Q)

It obtains the information about the user's search process. There are two types of activities: New (user enters Q) and Existing (user selects Q from QL).

```
1. U: Repeat while ACT{Q} = Click                         [Q is entered or selected from QL]
   {
     a) Call Record_keywordlog(Q)                         [store details of Q in Database]
     b) Display RL       [RL is predefined for existing Q][RL depends on SE for new Q]
     c) i = 0 , j = 0               [temporary variables, i is global and j is local]
     d) V : Repeat while ACT{RL} = Click                  [resource R is selected from RL]
        {
      i.     j = 1                                 [check if any R is selected for Q or not]
      ii.    Copy the resource selected in R[++i]
                                                   [i is first incremented and then used]
      iii.   R[i].activity = read                  [if R is selected then user wants to read R]
      iv.    current = i        [current is global variable that specifies the current resource
                                 used by the user among the multiple R opened by user]
      v.     Call Record_searchlog(R[current])
                                                   [store details of R[current] in Database]
      vi.    Call Check_window(i, current)         [check  current window used by user]
      vii.   Repeat while ACT{W.R[current]} != Close
             {
               If (R[current].activity != read), then:    [user wants to save / print R]
               {
                 Call Record_searchlog(R[current])    [store details of new activity]
                 R.activity[current] = read
               }
               If (ACT{RL} == Click), then :  Goto V
               Call Check_window(i, current)
             }   /* end of step vii loop */
      viii.  R.activity[current] = close,  i = i-1
                                  [if any resource R is closed by user then, number
                                   of resources currently opened are decreased by 1]
      ix.    Call Record_searchlog(R[current])
                              [store details of the current R which is  closed by the user]
      x.     Close W.R[current]
      xi.    Call Check_window(i, current)
        }  /* end of step d loop */
     e) If ( (j == 0) and (ACT{W.RL} == Close, then :
                        [no R is selected for Q and user presses the close button for RL]
        {
          Call Record_keywordlog(Q)                       [store details of Q in Database]
        }
   }  /* end of step 1 loop */
2. If new task is initiated , then:          [either Q is issued or existing Q is selected]
      Goto U
   Else
      Stop
/* End of Algorithm Capture */
```

The data gathered by the module Capture is stored in the Database through the module Store. The Conceptual ER Model of the Database (Fig. 6) consists of five entities (webuser, webuserphoto, keywordlog, searchlog and emoticon) associated by the relationship "has".

The Cognitive Mapping (Gulati and Garg, 2015b) for the user search process is based on BDI model. Guided by the data stored in the Database and Cognitive Mapping, the Decision Making module analyzes the data from the perspective of query, resource, emotion and timestamp. The analyzed data help in providing suggestions to the user regarding the Query (Q) and Resource (R).

Query (Q)

- Q was searched earlier or not?
- If yes, then how many times it was searched?
- All Q's beginning with root keyword of Q.
- Timestamp of Q (this helps to analyze if the user has some fixed specific time to search Q).
- Emotions associated with Q.
- Categorize all Q's on the basis of positive and negative emotions.
- Whether any R was accessed for Q or not?

Resource (R)

- Various activities performed on R.
- Time spent on R.
- Emotions associated with R
- Classify useful and useless R.
- Categorize all R's tagged with positive and negative emotions.

4.3.1.2. Context. Various fields used in the registration process help to provide the user's context for the query.

4.3.1.3. Timestamp. The system's decision for executing the query at a specific timestamp is fulfilled using the CA architecture (Lawniczak and Di Stefano, 2010).

4.3.2. Security

The security is activated through different layers defined in the structure of the CA (Lawniczak and Di Stefano, 2010). The Webcam and image recognition system (clmtrackr) enables security.

5. Implementation and Experimental Setup

The implementation of the proposed framework called Secured CAbMsISS (Fig. 7) is referred to as Cognitive Search Engine based on Emotions (COGSEMO). The structural design and working uses the following technologies:

- Page Layout – HTML 5 + JADE (Template Engine)
- Client Interactivity – Java Script, AJAX and Jquery
- Database – Structured Query Language (My SQL)
- Server Site Processing and Database Handling – PHP

The implemented system is executed using:

- OS: Windows 7 Home Premium (64- bit)
- CPU – Intel Core i5-2430 M Processor 2.40 GHz
- Memory – 4 GB
- Hard Disk Drive – 500 GB

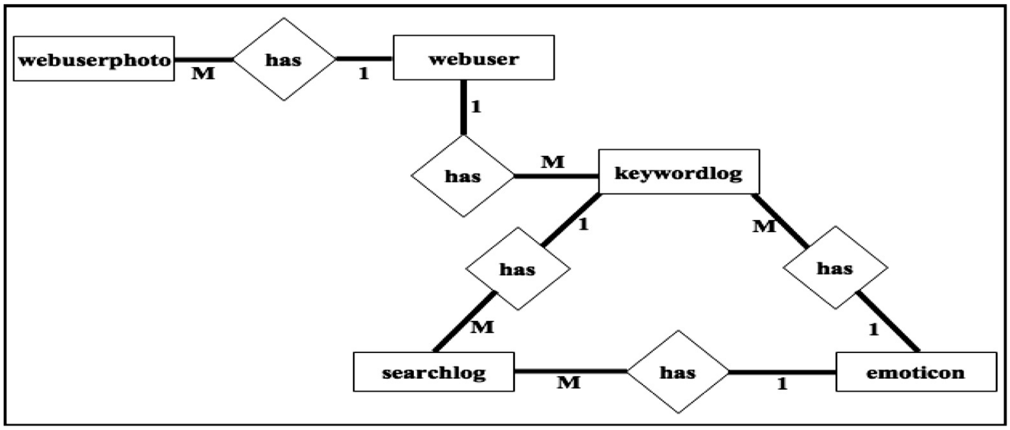


Figure 6 Conceptual ER Model Diagram of the Database.

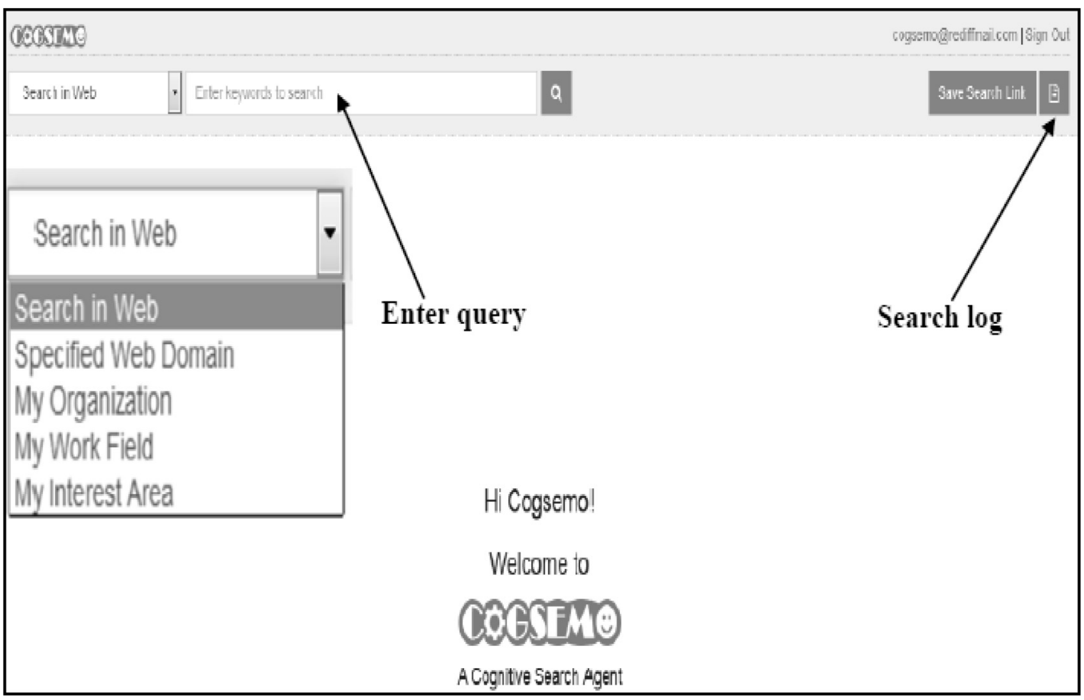


Figure 7 COGSEMO – Implementation of Secured CAbMsISS framework.

COGSEMO facilitates secured and adaptive Web-based search by building the user model based on the cognitive style, context, emotions and timestamp. The security element permits the access for Web-based search only to the authenticated user. This prevents unauthorized access to the search log. After successful login, the user can either search for a new query by typing the keywords in the text box (Fig. 7) or search from the list of queries available in the search log as shown in Fig. 7.

5.1. Search Options

COGSEMO assists the user during search by displaying a drop down box with some specific options as shown in Fig. 7. The values for the fields: My Organization, My Work Field and My Interest Area are provided by the user during the registration process. These fields help to specify the user's query context. It helps to improvise the relevancy of the retrieved results by performing the search based on the user context. COGSEMO also permits the user to perform the search from a specific domain as discussed in Section 5.1.1.

5.1.1. Specified Web Domain

Fig. 8 illustrates the option provided for Domain specific search. COGSEMO also generates the list of domains previously searched.

For the query = Education and Domain = ieee.org (Fig. 8), the result set obtained by COGSEMO is presented in Fig. 9.

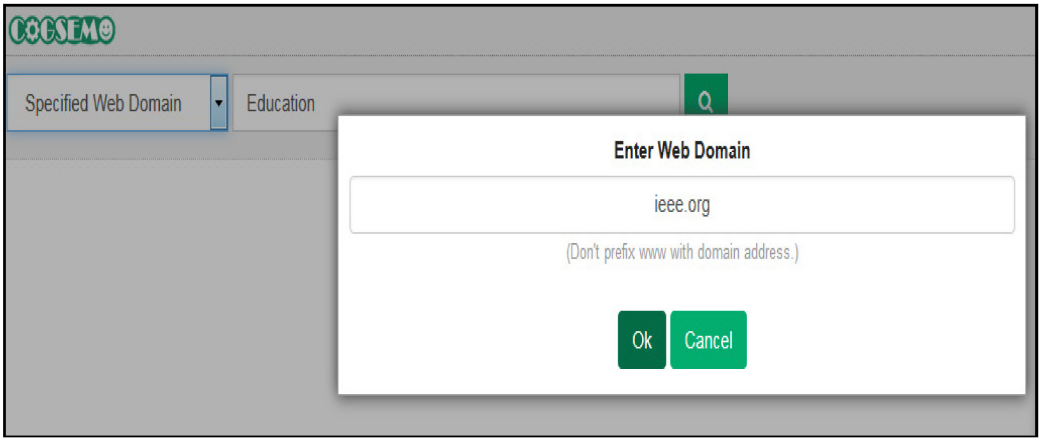


Figure 8 COGSEMO – Domain Specific Search.

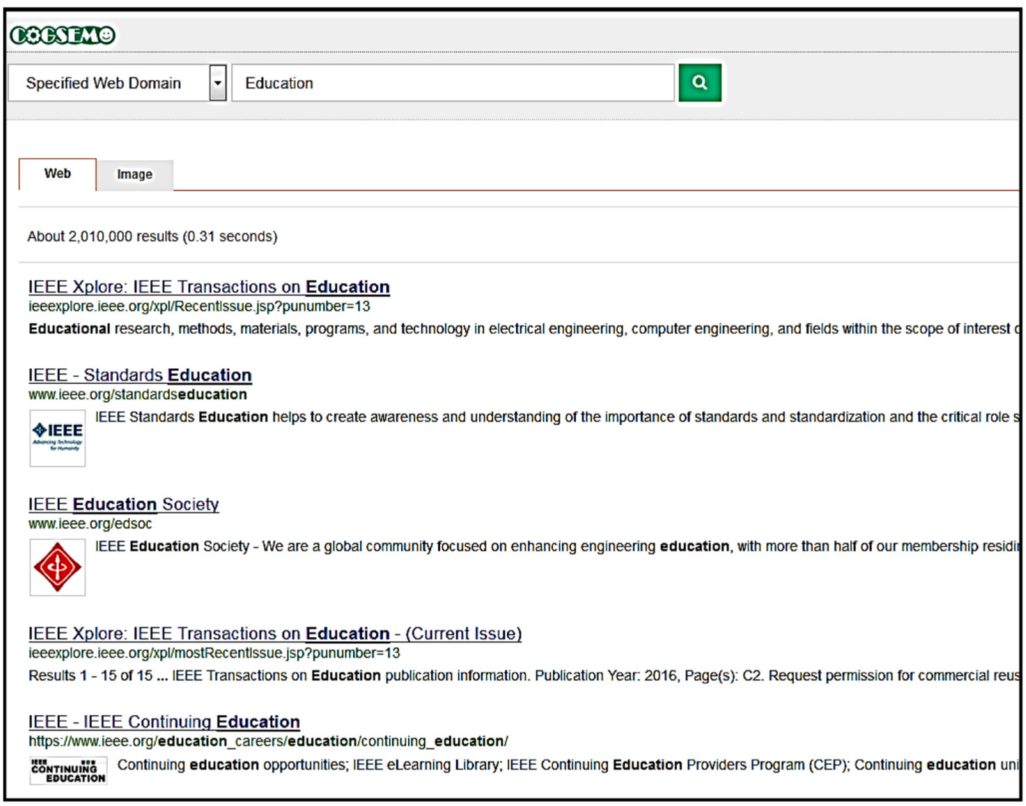


Figure 9 COGSEMO – Result set for query = Education and Domain = ieee.org.

5.1.2. My Organization

On selecting the option: My Organization for the query (Q), COGSEMO converts the query (Q) to Semantic Query (SQ) and sends SQ to PSE for retrieving the relevant results. Google Search Engine referred to as PSE has been used to retrieve the search results. Fig. 10 highlights the results obtained after executing the query = Senate with user context = My Organization ("Panjab University"). This eliminates the need to repeatedly specify the user context in the query through the keywords. Once specified, the context can be used repetitively. COGSEMO also facilitates updating the values of the fields entered in the registration form by clicking on the E-Mail shown in Fig. 11.

5.2. Emotions

The impact of human emotions (stress, anxiety or frustrations) on computer performance in various fields of human–computer interaction has been stated (Charlton, 2009; Rozell and Gardner, 2000). Moreover, the emotional outcomes of a search process can have important implications for the potential actions that online users perform on the Web (Flavian-Blanco

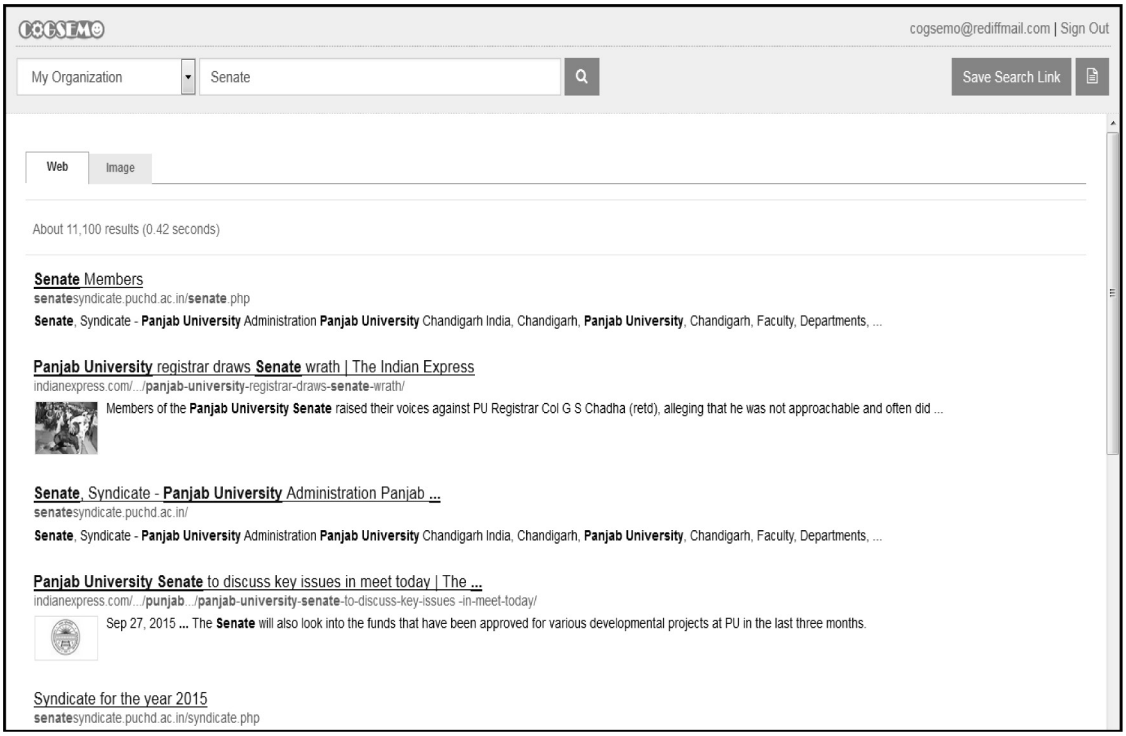


Figure 10 COGSEMO – Result set for query = Senate searched with context = My Organization.

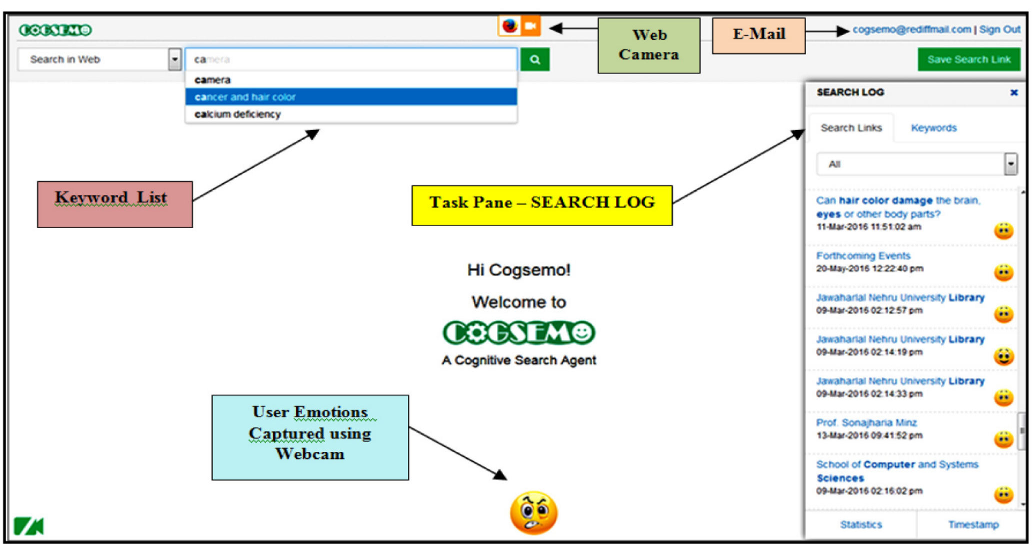


Figure 11 COGSEMO – Structure.

et al., 2011). In line with this, the present implemented system provides active assistance to the user while accessing a resource tagged with negative emotions. COGSEMO captures the user image through Webcam as shown in Fig. 11 and analyzes it to extract the emotions using clmtrackr.

The emotions can be extracted:

- automatically through Webcam and clmtrackr.
- manually by explicit specification through the emoticon. Fig. 12 enumerates the emoticons that can be selected by the user to express the emotion felt.

After extracting the user emotion through clmtrackr or emoticon, the corresponding emotion is tagged (associated) with the query initiated or resource accessed. This provides assistance to the user in restricting the access to a resource or query previously tagged with negative emotions or found useless.

COGSEMO guides the user by providing the list of queries searched, resources accessed and emotions felt using a task pane called **SEARCH LOG** (Fig. 11).

- **Option: Search Links** specifies the resources accessed, time of search and emotion captured (Fig. 11).

Figure 12 COGSEMO – Emotions presented to the user.

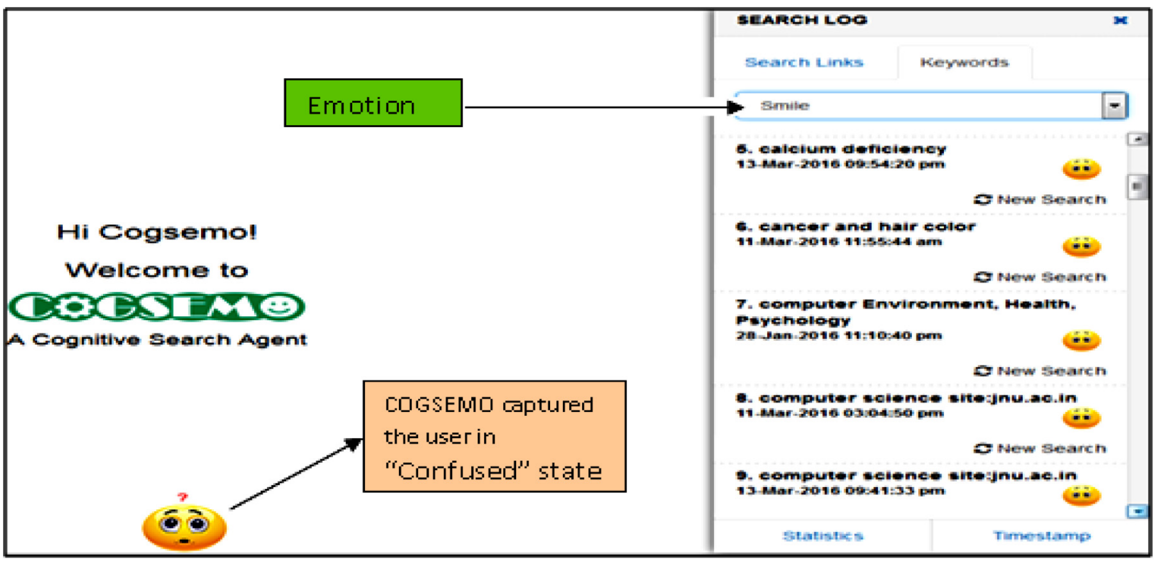


Figure 13 COGSEMO – Queries tagged with user emotion = Smile.

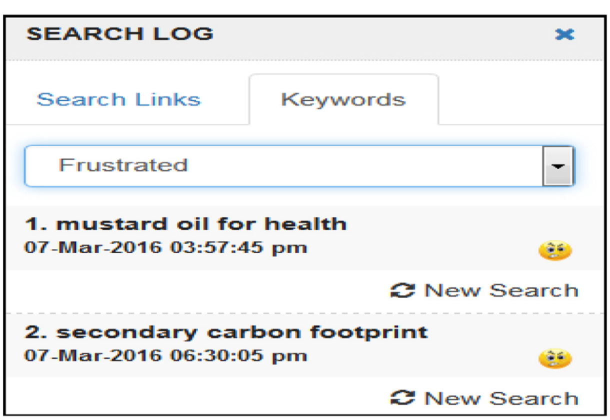


Figure 14 COGSEMO – Queries tagged with user emotion = Frustration.

- **Option: Keywords** enumerates the queries initiated. It also categorizes the queries and resources according to the emotion tagged with them.

Fig. 13 displays the queries (keywords) associated with **user emotion = Smile.** Fig. 14 lists the queries tagged with **user emotion = Frustrated**. The feature prevents accessing a resource previously tagged with negative emotions. Any previously conducted search can be initiated from the **Button: New Search** (Fig. 13 and 14).

As illustrated in Fig. 15, COGSEMO displays the queries initiated and corresponding resources accessed (highlighted in blue color) in the SEARCH LOG task pane. On executing the **Button: New Search, a message is displayed regarding the emotion experienced majority of time for the corresponding query and resource.**

Emotion Statistics generated by COGSEMO for the queries are presented in Fig. 16. The symbols + and – depict the % age of the number of times positive and negative emotions were experienced corresponding to the query.

5.3. Timestamp

The system provides the list of queries and resources accessed corresponding to different timestamps (Morning, Afternoon, Evening and Night) interpreted by COGSEMO (Fig. 17). It

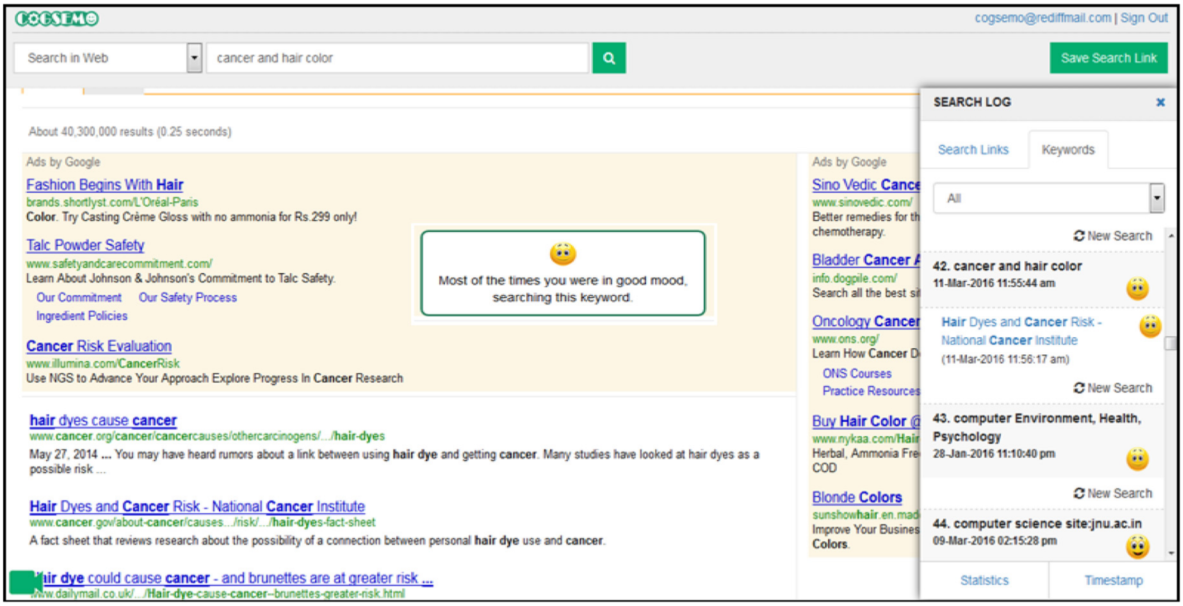


Figure 15 COGSEMO – Result set for query (No. 42) executed using Button: New Search.

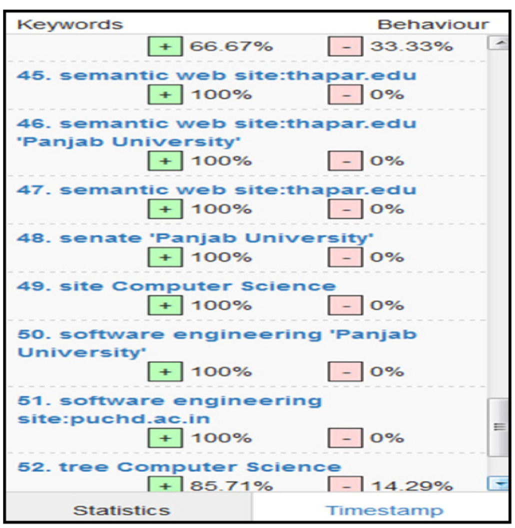


Figure 16 COGSEMO – Emotion Statistics.

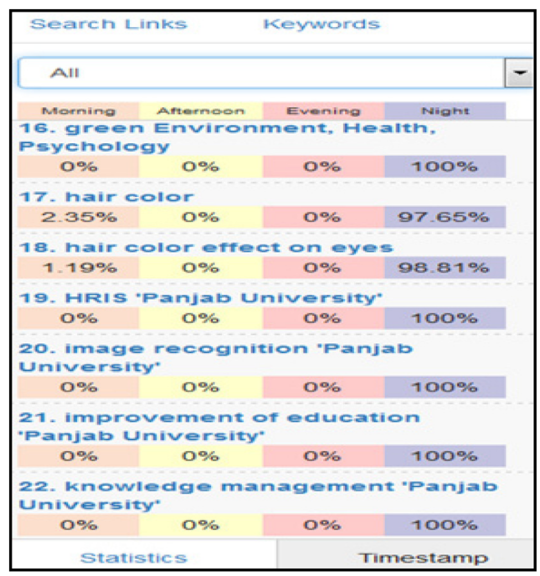


Figure 17 Queries and Timestamp.

shows the % age of the number of times a query has been accessed in a particular timestamp. Query No. 17 (hair color) shown in Fig. 17 has been accessed 2.35% in Morning and 97.65% at Night.

The retrieval effectiveness of the Secured CAbMsISS (implemented as COGSEMO) has been tested during the study. The paper presents comparative analysis of the results retrieved from PSE: Google and CAbMsISS. QRT i.e. time taken to execute the query was observed for both the approaches with Document Cut-off Value (DCV) = 10. Based on Judgmental Sampling, ten queries (N = 10) were selected for the experimental study. All the 10 queries were executed for both the approaches (PSE and CAbMsISS) and the obtained QRT (in min:sec.millisec) was recorded. Each query was re-executed for both the approaches and mean value of QRT observations obtained per query for each of the approach was calculated and analyzed. The authors also analyzed the retrieved result sets to find the Proportion of RDJR i.e. Precision and Proportion of RDJI. Descriptive Statistics: Mean was used to calculate the mean values. Finally, Inferential Statistics: Independent *t*-test was used to measure the significance of difference between the various retrieval effectiveness measures (QRT, Precision, RDJI) for PSE and CAbMsISS.

6. Experimental results and comparative analysis

The effectiveness of CAbMsISS and its comparative analysis with PSE: Google is based on the following parameters:

- QRT
- Precision – Proportion of RDJR (No. of RDJR/Total No. of documents retrieved)
- Proportion of RDJI

Table 1 Hypotheses used for comparative analysis.

Sr. No.	Hypothesis	Test
1	H_0 (1): There is no significant difference in the Mean QRT for PSE and CAbMsISS	Independent *t*-test
2	H_0 (2): There is no significant difference in the Proportion of RDJR for PSE and CAbMsISS **or** There is no significant difference in the Precision for PSE and CAbMsISS	Independent *t*-test
3	H_0 (3): There is no significant difference in the Proportion of RDJI for PSE and CAbMsISS	Independent *t*-test

Table 1 presents the hypotheses formulated for the above-mentioned parameters. Tabulated data are presented and analyzed using MS-Excel software.

6.1. QRT

It refers to the time taken by the SE to display the first result set starting from the time the query is issued.

Null Hypothesis H_0 (1): QRT (PSE) = QRT (CAbMsISS)

Table 2 presents the calculated mean of the QRT obtained after executing the 10 queries for both PSE and CAbMsISS. Fig. 18 presents the comparison of the Mean QRT (PSE) and Mean QRT (CAbMsISS).

Table 2 Mean QRT for PSE and CAbMsISS.

Mean QRT (min: s·ms)

Sr. No.	PSE			CAbMsISS		
	Query No.	QRT (PSE)	Mean QRT (PSE)	Query No.	QRT (CAbMsISS)	Mean QRT (CAbMsISS)
1	1a	0.21 0.39	0.30	1b	0.16 0.18	0.17
2	2a	0.57 0.61	0.59	2b	0.36 0.31	0.34
3	3a	0.53 0.5	0.52	3b	0.42 0.26	0.34
4	4a	0.55 0.33	0.44	4b	0.47 0.26	0.37
5	5a	0.44 0.4	0.42	5b	0.37 0.33	0.35
6	6a	0.7 0.62	0.66	6b	0.31 0.26	0.29
7	7a	0.5 0.36	0.43	7b	0.36 0.28	0.32
8	8a	0.49 0.45	0.47	8b	0.34 0.3	0.32
9	9a	0.59 0.38	0.49	9b	0.38 0.33	0.36
10	10a	0.37 0.66	0.52	10b	0.2 0.18	0.19

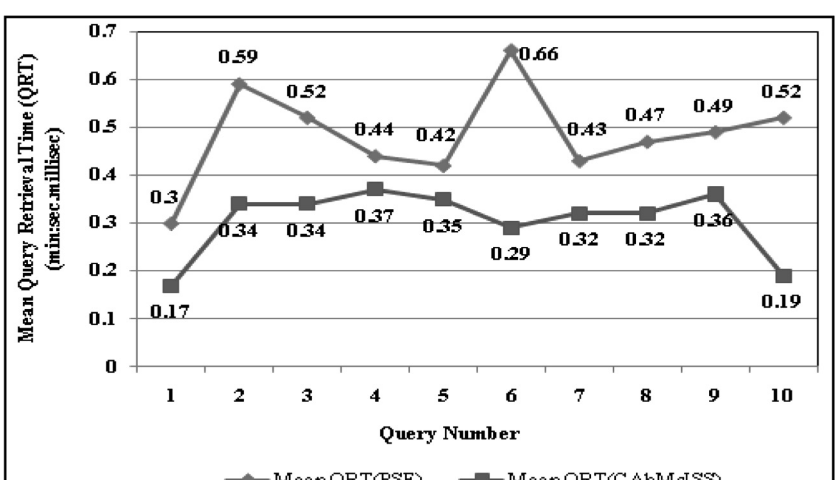


Figure 18 Comparison of mean QRT (PSE) and mean QRT (CAbMsISS).

Independent t-test was used to find if there exists any significant difference in the Mean QRT (PSE) and Mean QRT (CAbMsISS).

Findings: Table 3 presents the values obtained after applying t-test on Mean QRT (PSE) and Mean QRT (CAbMsISS) shown in Table 2.

Calculated t-value = 4.729 and tabulated t-value (for two-tail test) = 2.101.

Since calculated t-value > tabulated t-value, hence Null Hypothesis H_0 (1) is rejected.

This depicts significant difference in the Mean QRT (PSE) and Mean QRT (CAbMsISS).

Proved: Mean QRT (PSE) ≠ Mean QRT (CAbMsISS).

Moreover, Mean QRT (PSE) = 0.4825 > Mean QRT (CAbMsISS) = 0.303, hence CAbMsISS consumes less time for retrieving the query results in comparison to PSE.

Thus, CAbMsISS is more efficient in terms of QRT.

6.2. Precision

It refers to the proportion of Retrieved Documents Judged Relevant (RDJR) on query execution.

Null Hypothesis H_0 (2): Proportion of RDJR (PSE) = Proportion of RDJR (CAbMsISS) i.e. Precision (PSE) = Precision (CAbMsISS)

Table 4 presents the Mean Document Retrieval Results obtained after executing the 10 queries for PSE and CAbMsISS. Independent t-test was used to find if there exists any significant difference in the Precision (PSE) and Precision (CAbMsISS).

Findings: Table 5 presents the values obtained after applying t-test on Precision values shown in column 4 and column 10 of Table 4.

Calculated t-value = −5.743 = ABS (−5.743) = 5.743 and tabulated t-value (for two-tail test) = 2.101.

Since calculated t-value > tabulated t-value, hence Null Hypothesis H_0 (2) is rejected.

This depicts significant difference in the Precision (PSE) and Precision (CAbMsISS).

Proved: Precision (PSE) ≠ Precision (CAbMsISS).

Moreover, Mean Precision (PSE) = 0.01 < Mean Precision (CAbMsISS) = 0.66, so less number of retrieved documents were judged as relevant by PSE in comparison to CAbMsISS.

Thus, CAbMsISS is more efficient in terms of Precision.

Fig. 19 shows the comparison of Precision values of the 10 Queries (Q1 to Q10) for PSE and CAbMsISS. The graph depicts that CAbMsISS shows more Precision as compared to PSE. This proves that the CAbMsISS is more efficient in presenting relevant results to the user.

Table 3 t-Test for mean QRT.

t-Test: two-sample assuming equal variances

	PSE	CAbMsISS
Mean	0.4825	0.303
Variance	0.009679167	0.004723333
Observations	10	10
Hypothesized mean difference	0	
Df	18	
t stat	4.729829776	
$P(T \leqslant t)$ one-tail	8.35832E−05	
t critical one-tail	1.734063592	
$P(T \leqslant t)$ two-tail	0.000167166	
t critical two-tail	2.100922037	

Table 5 t-Test on precision.

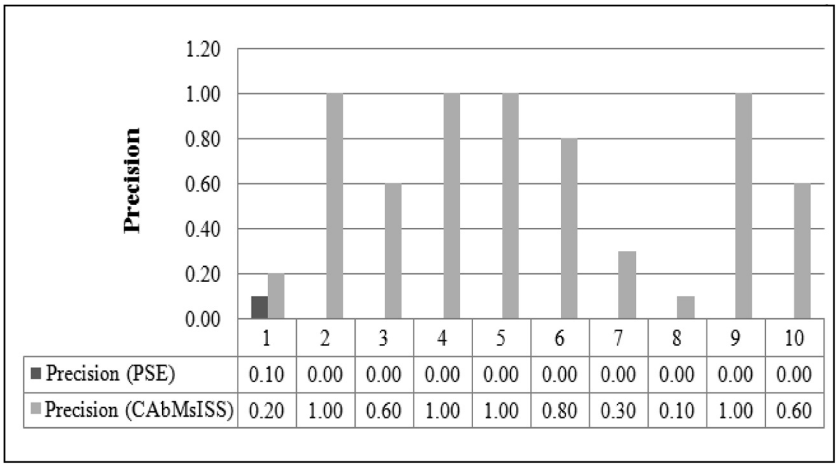
t-Test: two-sample assuming equal variances

	PSE	CAbMsISS
Mean	0.01	0.66
Variance	0.001	0.127111111
Observations	10	10
Hypothesized mean difference	0	
Df	18	
t stat	−5.742750625	
$P(T \leqslant t)$ one-tail	9.58379E−06	
t critical one-tail	1.734063592	
$P(T \leqslant t)$ two-tail	1.91676E−05	
t critical two-tail	2.100922037	

Table 4 Mean document retrieval results.

Mean – document retrieval results

PSE						CAbMsISS					
Query No.	Total Retrieved	No. of RDJR	Precision	No. of RDJI	Proportion of Retrieved Documents Judged Irrelevant	Query No.	Total Retrieved	No. of RDJR	Precision	No. of RDJI	Proportion of Retrieved Documents Judged Irrelevant
1a	10	1	0.10	9	0.90	1b	10	2	0.20	8	0.80
2a	10	0	0.00	10	1.00	2b	10	10	1.00	0	0.00
3a	10	0	0.00	10	1.00	3b	10	6	0.60	4	0.40
4a	10	0	0.00	10	1.00	4b	10	10	1.00	0	0.00
5a	10	0	0.00	10	1.00	5b	10	10	1.00	0	0.00
6a	10	0	0.00	10	1.00	6b	10	8	0.80	2	0.20
7a	10	0	0.00	10	1.00	7b	10	3	0.30	7	0.70
8a	10	0	0.00	10	1.00	8b	10	1	0.10	9	0.90
9a	10	0	0.00	10	1.00	9b	10	10	1.00	0	0.00
10a	10	0	0.00	10	1.00	10b	10	6	0.60	4	0.40



	1	2	3	4	5	6	7	8	9	10
Precision (PSE)	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Precision (CAbMsISS)	0.20	1.00	0.60	1.00	1.00	0.80	0.30	0.10	1.00	0.60

Figure 19 Comparison of precision values for PSE and CAbMsISS.

6.3. Proportion of Retrieved Documents Judged Irrelevant (RDJI)

Null Hypothesis H_0 (3): Proportion of RDJI (PSE) = Proportion of RDJI (CAbMsISS)

Apply independent *t*-test on data shown in column 6 and column 12 of Table 4.

Findings: calculated *t*-value = 5.743 and tabulated *t*-value (for two-tail test) = 2.101.

Since calculated *t*-value > tabulated *t*-value, hence Null Hypothesis H_0 (3) is rejected.

This depicts significant difference in the Proportion of RDJI (PSE) and Proportion of RDJI CAbMsISS.

Proved: Proportion of RDJI (PSE) ≠ Proportion of RDJI (CAbMsISS).

Mean Proportion of RDJI (PSE) = 0.99 > Mean Proportion of RDJI (CAbMsISS) = 0.34, so more number of irrelevant documents were retrieved by PSE as compared to CAbMsISS.

Thus, CAbMsISS is more efficient in terms of RDJI.

7. Conclusion and future work

The implemented framework (COGSEMO) performs MsIS based on the user cognitive style, context, timestamp and emotions experienced during searching. The fields specified in the registration form liberate the user from repeatedly specifying all the keywords to establish context for the query each time search is initiated. Moreover, as the system attaches emotions, this prevents the user to access queries and resources previously tagged with negative emotions. The emotion statistics can be used to determine positive and negative elements of the search process. This can be used to enhance the search experience. The system also presents suggestions based on the emotions and timestamp.

Moreover, the performance evaluation of PSE: Google and CAbMsISS based on QRT, Precision and RDJR suggests that CAbMsISS outperforms PSE in all the above-mentioned retrieval effectiveness measures. The analysis of CAbMsISS with PSE (Google) highlights the efficiency of CAbMsISS in terms of less QRT, higher Precision and decreased retrieval of irrelevant results. Hence, this illustrates the success of the proposed and implemented framework for CAbMsISS.

The comparative study of the implemented framework with other SE's should also be performed. Moreover, the implemented framework must be tested with vast data set. The authors suggest to use enhanced approach for image processing to map the user emotions.

References

Alafif, T., Sasi, S., 2012. Domain and Range Identifier Module for Semantic Web Search Engines. In: 2012 International Conference on Data Science & Engineering (ICDSE). IEEE, pp. 8–12. http://dx.doi.org/10.1109/ICDSE.2012.6281904.

Artino, A.R., 2012. Emotions in online learning environments: introduction to the special issue. Internet High. Educ. 15, 137–140, <http://dx.doi.org/doi:10.1016/j.iheduc.2012.04.001>.

Barr, C., Jones, R., Regelson, M., 2008. The linguistic structure of english web-search queries. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing – EMNLP '08. Association for Computational Linguistics, Morristown, NJ, USA. http://dx.doi.org/10.3115/1613715.1613848.

Belk, M., Papatheocharous, E., Germanakos, P., Samaras, G., 2013. Modeling users on the World Wide Web based on cognitive factors, navigation behavior and clustering techniques. J. Syst. Softw. 86 (12), 2995–3012. http://dx.doi.org/10.1016/j.jss.2013.04.029.

Charlton, J.P., 2009. The determinants and expression of computer-related anger. Comput. Human Behav. 25 (6), 1213–1221. http://dx.doi.org/10.1016/j.chb.2009.07.001.

Flavian-Blanco, C., Gurrea-Sarasa, R., Orus-Sanclemente, C., 2011. Analyzing the emotional outcomes of the online search behavior with search engines. Comput. Human Behav. 27 (1), 540–551. http://dx.doi.org/10.1016/j.chb.2010.10.002.

Gulati, N., Garg, A., 2015a. A proposed framework to optimize the query by filtering noise using Semantic information processing. In: 2015 International Conference on Green Computing and Internet of Things (ICGCIoT). IEEE, pp. 1331–1338. http://dx.doi.org/10.1109/ICGCIoT.2015.7380673.

Gulati, N., Garg, A., 2015b. Cognitive information mapping model for web-based search. Int. J. Appl. Res. Inf. Technol. Comput. 6 (3), 144–153. http://dx.doi.org/10.5958/0975-8089.2015.00018.4.

Guo, J., Xu, G., Cheng, X., Li, H., 2009. Named entity recognition in query. In: Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval – SIGIR '09. ACM Press, New York, USA, pp. 267–274. http://dx.doi.org/10.1145/1571941.1571989.

https://github.com/auduno/clmtrackr.

Kim, K.-S., 2008. Effects of emotion control and task on Web searching behavior. Inf. Process. Manage. 44 (1), 373–385. http://dx.doi.org/10.1016/j.ipm.2006.11.008.

Kinley, K., Tjondronegoro, D., Partridge, H., 2010. Web searching interaction model based on user cognitive styles. In: Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction – OZCHI '10. ACM Press, New York, USA, pp. 340–343. http://dx.doi.org/10.1145/1952222.1952296.

Lamberti, F., Sanna, A., Demartini, C., 2009. A relation-based page rank algorithm for semantic web search engines. IEEE Trans. Knowl. Data Eng. 21 (1), 123–136. http://dx.doi.org/10.1109/TKDE.2008.113.

Lawniczak, A.T., Di Stefano, B.N., 2010. Computational intelligence based architecture for cognitive agents. Proc. Comput. Sci. 1 (1), 2227–2235. http://dx.doi.org/10.1016/j.procs.2010.04.249.

Liaw, S.-S., Huang, H.-M., 2006. Information retrieval from the World Wide Web: a user-focused approach based on individual experience with search engines. Comput. Human Behav. 22 (3), 501–517. http://dx.doi.org/10.1016/j.chb.2004.10.007.

Nahl, D., 2005. Measuring the affective information environment of web searchers. Proc. Am. Soc. Inf. Sci. Technol. 41, 191–197. http://dx.doi.org/10.1002/meet.1450410122.

Rajendran, B., Iyakutti, K., 2009. Cognitive Agents for Understanding the Complexities Involved in Web-based Knowledge-gathering Tasks. In: 2009 IEEE International Conference on Systems, Man and Cybernetics. IEEE, pp. 1310–1315. http://dx.doi.org/10.1109/ICSMC.2009.5346237.

Rangaswamy, A., Giles, C.L., Seres, S., 2009. A strategic perspective on search engines: thought candies for practitioners and researchers. J. Interact. Mark. 23 (1), 49–60. http://dx.doi.org/10.1016/j.intmar.2008.10.006.

Rao, A.S., George, M.P., 1995. BDI agents : from theory to practice. Proc. first Int. Conf. multiagent Syst. ICMAS95, 312–319, <https://www.aaai.org/Papers/ICMAS/1995/ICMAS95-042.pdf>.

Rozell, E.J., Gardner III, W.L., 2000. Cognitive, motivation, and affective processes associated with computer-related performance: a path analysis. Comput. Human Behav. 16, 199–222, <https://www.deepdyve.com/lp/elsevier/cognitive-motivation-and-affective-processes-associated-with-computer-46MRcw65Ng>.

Tsianos, N., Lekkas, Z., Germanakos, P., Mourlas, C., Samaras, G., 2008. User-Centric Profiling on the Basis of Cognitive and Emotional Characteristics: An Empirical Study. In: Nejdl, W., Judy, K., Pearl, P., Eelco, H. (Eds.), Adaptive Hypermedia and Adaptive Web-Based Systems. Springer, Berlin Heidelberg, pp. 214–223. http://dx.doi.org/10.1007/978-3-540-70987-9_24, 5149.

Wang, K.H., Wang, T.H., Wang, W.L., Huang, S.C., 2006. Learning styles and formative assessment strategy: enhancing student achievement in Web-based learning. J. Comput. Assist. Learn. 22 (3), 207–217. http://dx.doi.org/10.1111/j.1365-2729.2006.00166.x.

Yee, P.L., Hsieh-Yee, I., Pierce, G.R., Grome, R., Schantz, L., 2004. Self-evaluative intrusive thoughts impede successful searching on the Internet. Comput. Human Behav. 20, 85–101. http://dx.doi.org/10.1016/S0747-5632(03)00042-6.