IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received April 18, 2022, accepted May 18, 2022, date of publication May 23, 2022, date of current version June 1, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3177597

An Explanatory Study on User Behavior in Discovering Aggregated Multimedia Web Content

ABDUR REHMAN KHAN[1], **UMER RASHID**[1], **AND NAVEED AHMED**[2], **(Member, IEEE)**
[1]Department of Computer Sciences, Quaid-i-Azam University, Islamabad 45320, Pakistan
[2]College of Computer and Information Sciences, Prince Sultan University, Riyadh 12435, Saudi Arabia

Corresponding author: Abdur Rehman Khan (arkhan@cs.qau.edu.pk)

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Department of Computer Sciences, Quaid-i-Azam University, Islamabad, Pakistan.

ABSTRACT The recent advancements in the web allow users to generate multimedia content, resulting in multimedia information proliferation. Existing search engines provide access to multimedia content via a disjoint assembly of media-specific results called verticals. However, this decentralized assembly of media contents requires manual aggregation and synthesizing efforts at the user's end, hindering the information exploration process and subsequently may cause cognitive overload, hence, demanding innovative tools to discover multimedia content. The researchers have devised numerous state-of-the-art approaches; however, analysis to confirm the efficacy has little emphasis. This study investigates users' complex multimedia information-seeking behavior over state-of-the-art web search systems to unveil the user's information-seeking issues. Our research employs between-subjects study and post hoc analysis strategies to analyze participants' information-seeking characteristics. The study design adopted statistical hypothesis testing to consolidate previous user behavioral studies, confirm existing strategies, and present recommended practices for future general-purpose web search engines. The participants were assigned Google and an advanced discovery search system using the same multimedia dataset to ensure the obtained results' credibility. The primary behavioral parameters include search efforts, multimedia content exploration, search user interface (SUI), information management and presentation, and user cognition. This study uncovers several inadequacies of the search engines in meeting users' complex discovery needs, including 29.6% less user engagement, 43% system and searching dissatisfaction, and 32% less knowledge acquisition with 63.9% increased clicking effort on traditional search engines. The results confirmed previous user studies and suggest novel research recommendations statistically significant in multimedia information exploration-related endeavors.

INDEX TERMS User behavior analysis, information discovery, information seeking, multimedia information, visualization, search engines.

I. INTRODUCTION

In the 21st century, the internet connects millions of devices that generate an enormous amount of multimedia data [1], [2]. The web has accumulated zettabytes of heterogeneous information on the web that creates new challenges in managing the immense amount of multimedia data and therefore is, transforming the web into a giant multimedia graph [3]. Contrarily, users are demanding more from multimedia services to suffice their information needs [2]. However, managing information retrieval tasks in the large-scale multimedia system require considerable processing and storage resources [2]. Users, on the other hand, envisage seeking multimedia information to discover desired information and develop complex intellectual skills such as knowledge acquisition, comprehension, analysis, synthesis, and evaluation [4]. The overall goal is to collect, assemble and make sense of the information, resulting in discovery-oriented seeking behavior [5].

Nowadays, multimedia content is accessible via the web, which also motivates the users to seek the desired information in different media formats such as text, audio, image, and video, via web [6]. Nearly 40%−50% of users are engaged in the dynamic and unplanned nature of web

The associate editor coordinating the review of this manuscript and approving it for publication was Arif Ur Rahman.

multimedia searches to explore information [7] and 78% of the users use the internet with the intent of discovering desired information in addition to factual retrieval [8], [9]. In 71% of web usage scenarios, the user needs a method that allows them to assemble the information from multiple sources in a synthesized form to meet their information needs [10]. Google[1] serves over 2 million search queries per minute, and 25 petabytes are processed daily. Nearly 38% of the issued queries represent exploration and discovery intent. However, the enormity of the heterogeneous data demands efficient and effective management tools and techniques [6].

The prevalent issues present in the existing web search engines are related to the dispersal of relevant multimedia information in various verticals, ranked positions, and pages. [11]. As a result, completing complex exploratory search tasks requires more queries, heterogeneous results, and more sources than less complex tasks on the existing web search engines [12]. In this scenario, information discovery can only be enabled by allowing the user to comprehend, synthesize, and acquire the relevant information [12]. Going beyond known-item search has become a must and demands more effective innovating tools that facilitate discovery [3]. The existing literature concerns analyzing and solving the user challenges in information discovery in specific domains or expertise groups [5], [13].

The emphasis on analyzing the actual user information seeking and composition behavior on the advanced exploratory search systems and existing search engines during the dynamic information search scenarios is inadequate. It is partly due to the lack of generalized approaches that can be instantiated over a real dataset and evaluated via general audiences. As a result, the existing techniques either lack comparable results or provide comparative empirical results via generalization of the existing state-of-the-art approaches [5], [14]. There exists the need to qualitatively understand the characteristics of discovery search tasks [15], including open-ended and multi-faceted questions having many possible answers, often incorporating information exploration and lookup activities [15]. It is reported as a highly complex problem bridging the different areas of information seeking, interactive information retrieval, and SUI design [15].

This research is an extension of our previous work that mainly concentrated on theoretical and architectural formations of lookup, exploratory, and discovery search systems [3], [16]. The research presented in this paper aims to create a comparative information-seeking environment to analyze users' post hoc information-seeking behavior on the proposed exploratory and discovery search system and the existing search engines. The existing search engines present media-specific search results, called verticals, via the aggregation technique. This research uses the same in the proposed exploratory and discovery search system as a foundation. The proposed system exploits user discovery via data mining (clustering, semantic analysis, & summarization) techniques and employs information exploration and discovery fostering parameters.

We performed a detailed user behavioral analysis on the proposed and baseline systems using a between-subjects experimental design. Our behavioral study aimed to provide qualitative insights about the users during their information lookup, exploration, and discovery needs. We used the Google search engine as a baseline in the comparative analysis instead of generalizing the existing state-of-the-art search engine. The contributions of this research are threefold. At first, we present a comparative user study based on the real baseline and advanced exploratory and discovery search systems that analyze information-seeking behavior in information lookup, exploration, and discovery search scenarios. Secondly, we made an effort to confirm previous user studies in this context. Thirdly, we qualitatively establish better standards having statistical significance to enhance information lookup, exploration, and discovery activities via user modeling.

The rest of the discussion is organized as follows. Section II provides the background of the relatively new information discovery search paradigm and establish research hypotheses. Section III outlines the problems & motivation behind this research. Section IV briefly discusses the system setup & preliminaries. Sections V and VI explain the study method and experimentation results, respectively. Section VII briefly discusses the findings and obtained results. Finally, Section VIII formally concludes the paper and provides future research directions.

II. BACKGROUND & HYPOTHESIS

Historically, information lookup was the most primitive search activity instantiated to seek answers to convergent search tasks since information on the web mainly consisted of textual content, and users could only receive/browse them via search engines [17]–[19]. Consequently, the search tasks had precisely defined exploration criteria emphasized (what to look for) to satisfy information needs [20]. The invention of web 2.0 (also known as the social web) allowed the users to publish multimedia content [11]. As a result, the users' information needs are becoming broader, with search tasks having a discovery intent, often starting with divergent questions of "what, how, and why" [19]. Marchionini consolidated the mentioned prevailing trend under the category of exploratory search [18]. The overall objective is to explore the information beyond visible horizons, leading to knowledge acquisition, comprehension, interpretation, comparison, and aggregation [19], [21]. The exploration requires re-modeling the conventional information organization and presentation techniques [11].

Nowadays, web users are also demanding more from the search systems since the multimedia content is proliferating and consists of multiple modalities of information (i.e., text,

[1] https://www.google.com/

audio, image, and video) [3]. The exploration of multimedia information requires the discovery of relevant and desired information. It creates challenges in the information integration and synthesis [16]. Kerne *et al.* firstly addressed the discovery needs of the web users [19]. Information discovery is formally defined as a process that encapsulates the activity of finding, collecting relevant information, and developing understanding and relationships among them [19]. It creates a dynamic information-seeking environment requiring the assimilation of multimedia information from diverse sources having many different valid solutions and, therefore, creating non-linear search patterns often supported by interactive SUIs [3], [11], [22]. Various past researches relied on the use of visual elements such as charts, maps, and graphs [23]–[25].

Recently, increased computational power enabled researchers to deploy complex graph-based dynamic visualizations in more user-friendly SUIs [26]. The result of such interfaces, however, lacks a thorough comparison with existing state-of-the-art approaches [11], [27]. The aggregation of heterogeneous information is also shown to support complex search patterns [14], [22], [27]. Common information aggregation techniques include synthesizing the disjoint media-specific search result verticals [11], [14], [28], predicting the best modality of the search results in response to the user queries [27], and defining optimal positions [29]. However, these studies lack a detailed comparative analysis of the users' information-seeking behavior over the existing search systems. The following sections investigate key information discovery fostering parameters discussed above and formulate hypotheses to confirm the efficacy.

A. DISCOVERY SEARCH TASKS

At present, discovery search tasks are complex, and web surfing is beyond simply finding an answer to a specific question [4]. In particular, a discovery search initiates when a user wants to discover a domain, acquire/expand knowledge, or learn new topics [30]. Most recent literature has made an effort to investigate specific parameters of such tasks using easily measurable user behavior [31]. In less complex search tasks, due to high information seeking confidence, users issue well-articulated queries [31]. The mentioned scenario may allow a search engine to retrieve the most relevant and precise result sets [22]. However, the unclear information needs with multiple exploration targets make the discovery search tasks open-ended [32]. In such cases, the capability of search systems does not primarily depend on complex algorithms [3]. The overall interaction design between the user and the system must be considered holistically symbiotic [20], [33]. Therefore, to confirm this behavioral observation, we hypothesize:

H1: The lookup-based search engine usage complexity is higher during the discovery search tasks than in the discovery-oriented search system.

B. DATA MODEL

The existing commercial web search engines such as Google, Yahoo,[2] and Bing,[3] acknowledge the increasing information needs and to provide the first-aid, often blend media-specific search results in the first page of the web search results [11], [30]. However, the linear organization and disjoint assembly of the multimedia search results hinder the information exploration [5]. Moreover, users reported a widespread concern that they are not informed diversely by search engines, causing difficulties in exploration [34]. This inadequacy lies in the way information is internally structured in a data model [22]. To confirm the inadequacy of the linear list in standard web search engines in exploratory and discovery search tasks, we hypothesize that:

H2: Organization of the information in the linear list increases perceived discovery task difficulty compared to non-linear information organization.

C. INFORMATION INTEGRATION

Discovery and exploratory search tasks require the synthesis of information from a variety of vertical information sources [27], [30]. The present search engines provide access to diverse multimedia sources using disjoint verticals (web, image, video, news, etc.) [16]. Quick access to these disjoint sources is often offered via partially blending the subset of the disjoint list on the traditional web results page [35]. The researchers focused on presenting verticals simultaneously on a single screen [14], [36]. However, such techniques still suffered from the widely criticized hoax of linear lists due to the lack of comparable verticals aggregation techniques instantiated on a real-dataset [3], [35], [37]. The search engines usually use the predictive method to find the optimal position of the vertical to the given user query to enhance the effectiveness of the vertical blending [29]. We hypothesize that:

H3: The partial blending of the verticals on the traditional web search results page provides less complex search task completion satisfaction than fully-blended verticals aggregation.

D. INFORMATION SYNTHESIS

The users often encounter difficulty in synthesizing the multimedia information, especially in the case when relevant contents are retrieved from disjoint sources and scattered in various disjoint multimedia verticals, search positions, and pages [3]. For instance, a topic within the search results space, e.g., FIFA in Mexico, may have the relevant image, video, web, and news search results [11]. Due to the displacement of the search results related to a topic over various places, users often encounter hindrance in synthesizing the required information [16]. As a result, users

[2] https://www.yahoo.com/
[3] https://www.bing.com/

may acquire less knowledge and subsequently issue more well-articulated queries to seek the relevant information [38]. Hence, we hypothesize that:

H4: Lack of information synthesis leads to lower knowledge acquisition, increased number of queries, and query length.

E. NON-LINEAR JOURNEY

Information exploration and discovery tasks are dynamic; finding a relevant piece of information can trigger the need for additional browsing and search for unanticipated reasons [4], [12]. It forms non-linear information-seeking patterns [4], [11], [39]. In this scenario, a user needs to exploit the multiple information organization structures, such as the ability to explore an item holistically and browse more interesting results related to a focused item, to enhance sensemaking [40]. Moreover, the linear list of the search results solely is inadequate to support this non-linear behavior in discovery search tasks [22]. To confirm this behavior, we hypothesize that:

H5: Dynamic search tasks require the exploitation of multiple information organization structures compared to unified presentation styles.

F. ENGAGEMENT VIA SUI

The use of SUIs is known to enhance user experience and cognitive engagement [38]. An appealing user interface and interactive visualizations detain the user's interest in the task and keep them captivated to seek the information for a more extended period [41]. As a result, the user can browse more interesting content [36]. Whereas, if the user interface cannot visually appeal to a user, then the result introduces a quicker disinterest in the user's task [11]. Hence, we hypothesize that:

H6: Appealing visualizations and search user interfaces enhance users' interest, increase the number of contents browsed, and task exploration time.

G. AGGREGATED PRESENTATION

The optimum presentation of aggregated multimedia search results is a focus of many researchers [21], [29], [42], [43]. Mainly, a predictive method is utilized to measure the optimum positioning of the aggregated media-specific verticals and placed either horizontally (grid-based layout) [14] or vertically (linear lists) [29]. The goal of verticals aggregation is to reduce user efforts in discovering interesting multimedia artifacts [14], [42], [44]. The grid-based aggregation of the verticals enable increased exploration [14], [42]. Conversely, linear list-based verticals aggregation helps in the detailed examination of information [42]. The discovery search tasks, on the other hand, are dynamic, and the goal of an information seeker switches between holistic examination of the information in detailed analysis [34]. Nevertheless, a study on the combination of vertical and horizontal presentation of aggregated disjoint verticals in discovery search scenarios is still to be studied. Hence, we hypothesize that:

H7: The horizontal and vertical presentation combination reduces search efforts when discovering multimedia artifacts.

III. PROBLEMS & MOTIVATION

Information discovery is an emerging search and exploration paradigm [3], [45]. It is a challenging task to pose an exploratory search leading towards discovery [46]. The previous studies measured the effectiveness of advanced proposed search engines by comparing them with a generalization of traditional search engines [45]. However, to the best of our knowledge, there exists a lack of detailed and statistically significant study that analyzes the users' behavior over an advanced search system and the present state-of-the-art search engine on different types and user categories. There exists a need to expand the areas of information retrieval on the web and user studies for comprehensive identification, and finding of the users' actual search needs [47], [48]. Approximately every three out of four usability studies do not assess the outcome of the users' interaction, therefore, leaving unsupported any broad claims about usability [48], [49]. There exists a weak relation between the usage patterns and quality-in-use [48], [49].

Additionally, the users' satisfaction with the search interfaces is often inadequately measured since most studies reinvent the questionnaires while ignoring the validated questionnaires readily available in literature [23], [49]–[51]. We identified and analyzed various main parameters in the information systems, including information organization, presentation, aggregation, synthesis, and visual appearance. We determined the outcome of each parameter individually and collectively on the users. Specifically, in this research, we mainly:

- Studied the effect of various information system parameters in the discovery search scenario.
- Statistically analyzed and discussed the outcome of each parameter using standard usability scales.
- Confirmed the parameters analyzed in previous research and the newly recommended discovery practices.

IV. SYSTEM SETUP & PRELIMINARIES

This section discusses a brief system walk-through to provide a generic system overview. Subsequently, detailed system instantiation, including the data model and the inherited organization of information, are explained. Finally, an interactive user interface explains the interconnection of each discussed component.

A. SYSTEM WALKTHROUGH EXAMPLE

The web search engines present the results as a disjoint assembly of the verticals (Figure 1(b)). The approach is simple lookup-based in which a user knows what to look for in exploration [22]. To enhance the efficacy of information lookup, the most common popular commercial web search

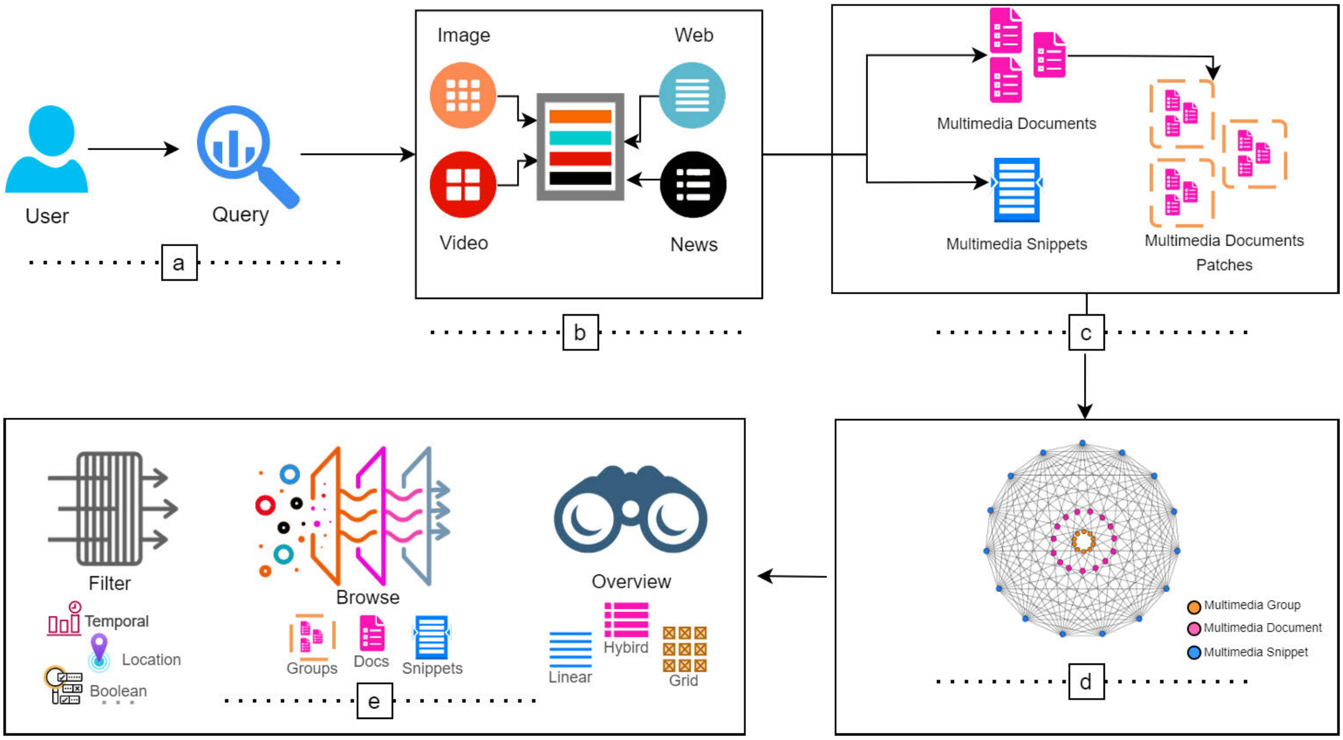

FIGURE 1. **Approach overview. (a) User query instantiation, (b) Verticals aggregation, (c) Search results encoding, (d) Non-linear graph exploration, and (e) Discovery mechanism.**

engines such as Baidu,[4] Bing, Google, Yahoo!, Yandex[5] etc., blend top few search results assembled from the different verticals into the linear ranked list of expected results. The integration of the verticals is primarily partial-blended because the relationships among the retrieved search results are ignored [11]. This strategy may suffice in simple lookup searches; however, it is inadequate to support complex information exploration and discovery tasks [15]. Users often have difficulty in the expression of their information needs [15]. Such tasks require synthesized and topically diverse search results [22]. It challenges the traditional display of vertical search results in disjoint sets [11]. In the advanced search system, a user usually initiates a query (Figure 1(a)). In response to the user query, search result verticals are retrieved and aggregated (Figure 1(b)). The search results are then synthesized into various coherent topics (Figure 1(c)). The synthesized form refers to multimedia documents, document groups, and snippets. The subsequent sections explain the search results synthesis in detail.

B. DATA-MODEL

The dataset and data model is based on the instantiation of the approach presented in literature [3]. Two well-known search results data models are employed, categorized into linear and non-linear. The former to organize search results, often in a ranked list (Figure 1(b)). The traditional commercial search engines, such as Google, Yahoo!, Bing, etc., are based on linear data models. The latter usually includes mash-like graphs (Figure 1(d)). The non-linear data models are mainly employed in advanced semantics-based exploratory search engines [52]. We employed a combination of the linear and non-linear data models to compare each data model's efficiency in existing state-of-the-art approaches. The linear and non-linear data models are instantiated dynamically on a real dataset retrieved from the Google (baseline) search engine. The meta-information associated with each search result includes title, date, URL, and description (snippet). This setup ensured the accurate and unbiased retrieval of the search results on the baseline and the instantiated discovery search system.

C. SEARCH RESULTS ORGANIZATION

The linear data model reorganizes the disjoint image, news, web, and video search result lists by text semantics mining and ranks the search result according to the semantic similarity to the given user query. As a result, instead of retrieval of a single type of content, multiple types of content are retrieved if they are semantically similar. Traversal of linearly organized information is a non-trivial task. Hence, the information is conceptually organized in the form of multimedia documents and groups [3]. The multimedia documents are created by performing clustering on the semantics extracted from the search results to aggregate semantically similar search results on a particular topic.

[4]http://www.baidu.com/
[5]https://yandex.com/

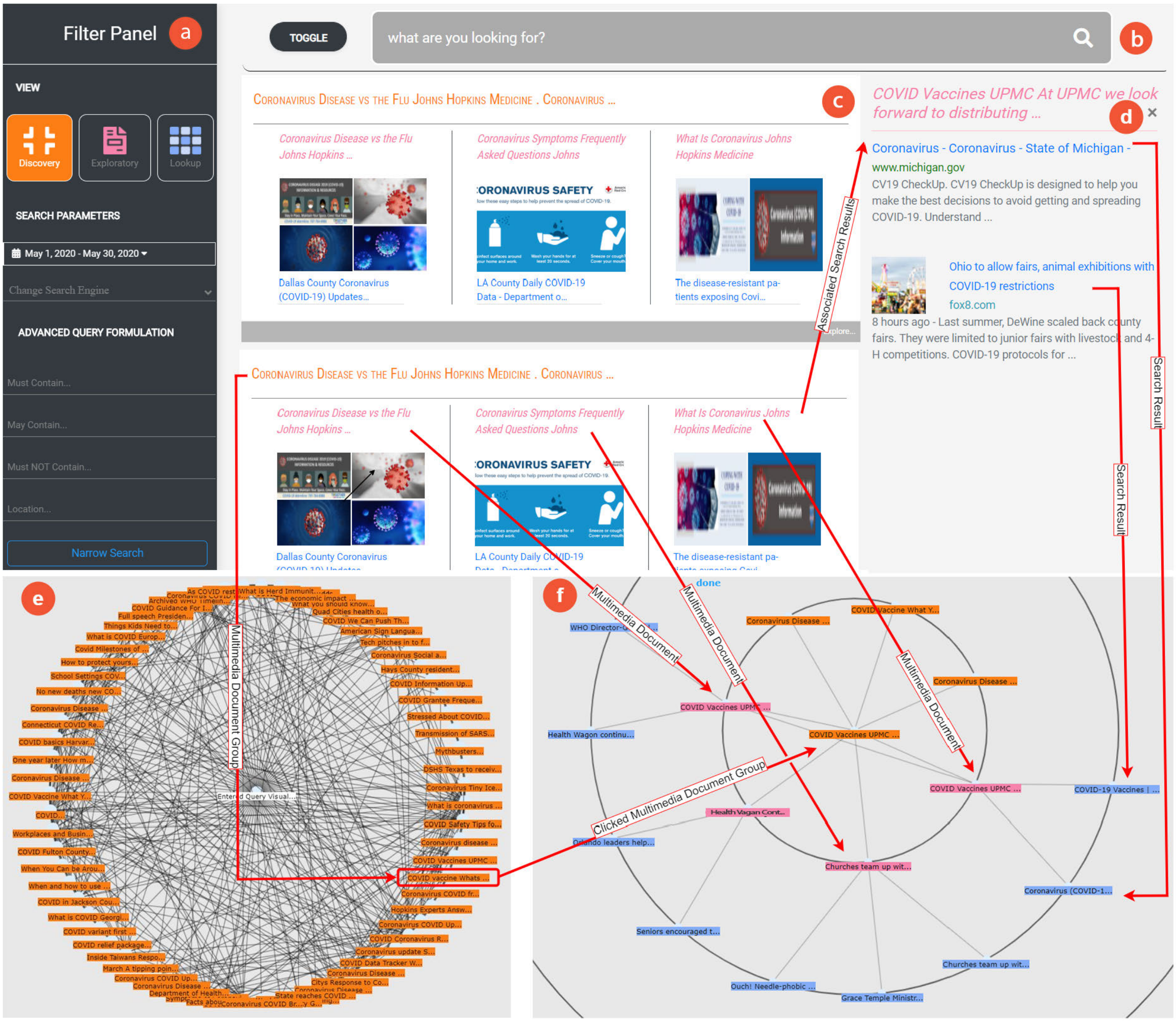

FIGURE 2. (Top) discovery SUI with (bottom) conceptual data-model. (a) Filtration, (b) Query formulation, (c) Search results browsing, and (d) Overview panel, with (e) underlying conceptual data model, and (f) associated contents.

Likewise, multimedia documents are clustered to aggregate broader topic categories, called multimedia document groups (Figure 1(c)). Each multimedia document and group is summarized and labeled using the summary of the content present inside. It may instantiate a complex non-linear graph to aid topical and categorical traversal within a search results space (Figure 1(d)).

D. SEARCH USER INTERFACES

The SUIs should primitively support overview, browsing, and filtration of results (Figure 2(c)). In a conventional discovery search interaction, a user initially gains an overview of the search space and browse the contents in a non-linear manner. This allows the user to narrow-down the information of interest, and subsequently, may issue well-articulated filters to obtain the desired information. To support three distinct types of information presentation and visualization scenarios, first, the discovery SUI present categorically organized multimedia documents (Figure 2) to provide a quick overview of the search results space. The second SUI for information exploration provide browsing and exploration of conceptually similar aggregated search results in multimedia documents (Figure 3(top)). The third is the lookup SUI that presents precise search results in a traditional linear list, instantiated via a linear data model (Figure 3(bottom)). All SUIs include standard interaction panels; filtration (by views, date, keywords, and location) panel (Figure 2(a)), query formulation panel (Figure 2(b)), search results browsing panel (Figure 2(c)) and overview panel (Figure 2(d)). The conceptualization of the information presentation using data-model is presented (Figure 2(e)). The detailed non-linear browsing mechanism supported by underlying complex non-linear graph is also presented (Figure 2(f)). These multiple conceptualizations of the search results are necessary to enhance user cognitive abilities [14].

Furthermore, we established standard coloring in the filtration panel to minimize user efforts in learning the system and maximize user experience [53]. Throughout the system usage, the multimedia document groups were represented by the orange color, the pink color represented multimedia documents, and the traditionally used blue color represented the snippets. These SUIs were instantiated on a non-linear data model to support non-linear information exploration [3]. The links between different content types represent containment relationships (e.g., a multimedia document is part of a multimedia document group). The links among the same content type represent high betweenness similarity. The non-linear links facilitate search results' traversal using topical or similar browsing mode. Moreover, a user can dynamically change the organization via the search results view in the filtration panel. The baseline SUI includes the Google search results page.

V. METHOD

The method discusses the proposed study design, users' detailed background information, evaluation environment, and apparatuses. Furthermore, the evaluation procedure, including evaluation sequence, study tasks design, and evaluating instruments, is explained.

A. STUDY DESIGN

We performed the usability analysis based on the between-subjects study design. The reason for choosing this study design is twofold. First, this study design allowed us to comprehend the user behavior in advanced approaches and existing search engines disjointly. In the between-subjects study, a user can only use one system throughout the experiment session. Since exploratory and discovery search tasks impose high anxiety and low confidence on the users [20], we made an effort to foster the same environment by using the between-subjects study design. Since users can gain confidence after completing a task, they were limited to completing a task only on one system to maintain the credibility of the results. Therefore, a between-subjects study ensures that no prior biases of the users affect the experimental results. Secondly, exploratory experiment sessions are long in duration. A typical session lasts between 30 to 45 minutes. The between-subjects design hence situates less burden on the users.

B. USERS

According to the previous research, a minimum sample size of 5 is required to uncover 80% of the usability issues [54], [55]. A sample size of 12 is required to obtain statistically significant results to cover and analyze the performance metrics and success rate [54], [55]. Recent usability studies employ between 9 to 24 number of users [11], [56], [57]. We recruited a total of 44 participants in the evaluation process. The participants were invited from all around the world via social media advertisements. Only the participants fulfilling the experiment criteria, including access to a computer and the ability to experiment in a calm environment, were recruited without age, education, or professional limits. Afterward, the participants were divided into groups and further categorized as professional (working), student (studying), and domestic (homeworker) users. The users contain 26 males and 18 females with an average age of 31 years (SD=12.11). 32%−41% of the participants were students, 41%−54% were the professionals, and 14%−18% were the domestic participants. All consented to be part of this evaluation, and they did not receive any compensation for their participation. The users were further distributed into two groups. The first group was assigned the baseline system, and the second group was given the proposed system. Table 1 shows overall demographic statistics of the participants.

TABLE 1. Users demographic background.

		Proposed System (N=22)	Baseline System (N=22)
Gender	Male	13 (59.10%)	13 (59.10%)
	Female	9 (40.90%)	9 (40.90%)
Age	18-35	16 (72.72%)	17 (72.72%)
	36-55	5 (22.72%)	5 (22.72%)
	>55	1 (4.55%)	1 (4.55%)

C. APPARATUS

We deployed our tool on an Intel(R) Core(TM) i5-9400F CPU @ 2.90 GHz equipped workstation having 16 GB RAM, 250 GB Samsung Evo SSD, and 64-bit Linux operating system. The workstation also incorporated GeForce RTX 2060 GPU. The workstation was connected to a 19'' widescreen LCD monitor having a resolution of 1366 x 786 pixels. The interaction with our tool was provided using the Mozilla Firefox web browser, a standard QWERTY layout keyboard, and an office mouse, delivered in a calm office-like environment. For the participants residing in remote locations, access to this workstation was provided via remote desktop software such as Teamviwer[6] and Anydesk.[7] The participants were interviewed before the experimentation process. The objective was to ensure the complete integrity of the apparatus used in this research.

D. PROCEDURE

In the beginning, a demographic questionnaire was collected from the participants. A 5-minute demonstration of the experimental procedure was given to the participants. Subsequently, the participants were asked to select a least familiar discovery search task incorporating lookup and exploratory activity and complete a pre-task performing questionnaire to measure their interest in the task, perceived task difficulty,

[6]https://www.teamviewer.com/
[7]https://anydesk.com/

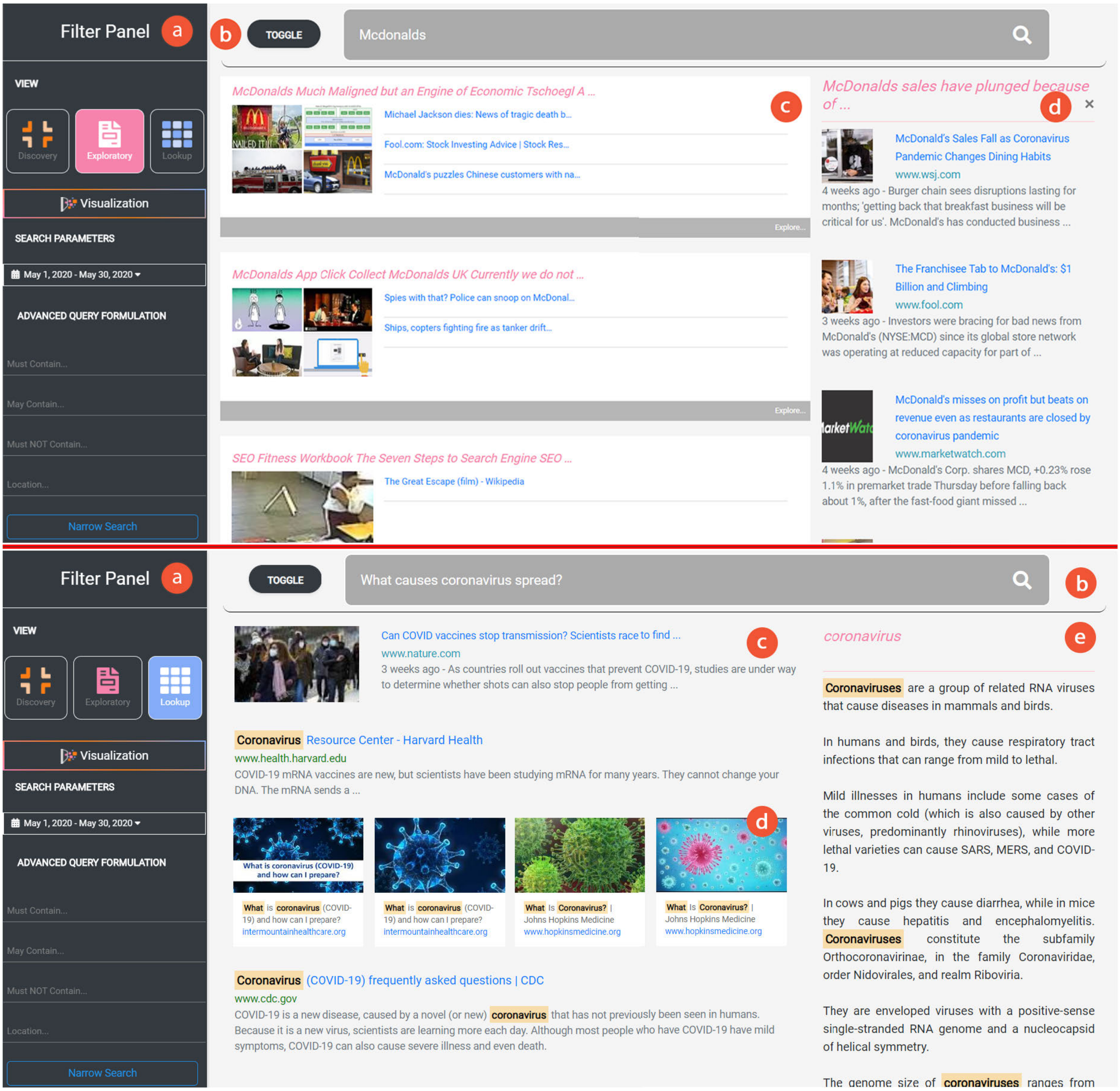

FIGURE 3. **(Top) exploratory and (bottom) lookup SUI with (a) Filtration, (b) Query formulation, (c) Search results browsing, and (d) Overview panel.**

and present knowledge. Then the users were given either a baseline search system or an advanced exploratory search system. The Google search engine was used as a baseline because of the web users' first preference [16]. In the case of the advanced exploratory search system, we used the same presented in Section IV to evaluate the users. Afterward, the users were given free time to explore their chosen search task; however, the software imposed an internal timer of 15 minutes for completing a search task to ensure consistency in the results. The users were not informed about this task completion time limit to imitate their realistic search behavior and eliminate emotional factors (stress, urgency, panic, etc.) affecting the results. Afterward, participants were presented with post-task performing questionnaires to measure qualitative parameters such as searching efforts, cognitive gain, system usability, and satisfaction. The procedure was recorded using a freely available screen and mouse recording software for detailed behavioral analysis.

E. TASKS

Previous literature has shown a lack of appropriate user behavioral analysis on "report like" exploratory and discovery search tasks such as locating all the information on some topic given with open-ended conditions, despite these being commonplace in practice [38]. We designed three discovery search tasks based on the pattern suggested in the literature [38]. These tasks were designed to incorporate lookup and exploratory information seeking. Throughout the evaluation procedure, the user had to select one of the following tasks based on their slightest prior familiarity with it to ensure the user's natural exploratory behavior without previous knowledge biases affecting the process. The tasks are as follows.

- Task-1: BMW is an automobile company. Explore the information about BMW, including when BMW was established and where it is located, the current affairs, and recent advancements of BMW.
- Task-2: Piano is a musical instrument. Explore the information about Pianos, such as when and who invented the Piano and recent trends and advancements in Piano.
- Task-3: NASA is an independent agency of America. Explore the information about NASA, such as when NASA was established, its abbreviation, and recent trends and advancements in NASA.

F. INSTRUMENTS

To measure the cognitive user state, including task difficulty, knowledge acquisition, user interest, and search experience satisfaction before and after task completion, we used a 5-points Likert scale from the research study [13]. System satisfaction was measured using the standard After Scenario Questionnaire (ASQ) [58]. For the demographic purpose, we used the literacy scale from a previously presented study [59]. User behavior is recorded using a free screen recorder software and behavioral tracking JavaScript software. The browser cookies were cleared for each participant, and countries' internet proxies were employed to eliminate the underlying bias of search results retrieved concerning the past browsing history and location (if present).

VI. RESULTS

This section briefly discusses the obtained evaluation results and statistically analyzes their significance. Based on the statistical significance, we unveil the outcome of each observed hypothesis.

A. SEARCH TASKS COMPLEXITY

The discovery search tasks are complex since they incorporate exploratory and lookup activities. We measured the complexity using the ASQ. The ASQ consists of three questions asked at the end of the experiment from the users. These include user satisfaction with the system supporting the users with the information, the amount of time taken to complete the task, and search scenario satisfaction after finishing the search task. The higher scores denote less task difficulty perceived, and the lower scores denote higher task difficulty perceived. The users on the advanced search system (Group-I) could perform complex search tasks with significantly less difficulty (on average 94.42% ASQ score). Alternatively, the users on the baseline system (Group-II) were moderately challenged (on average 49.63% ASQ score). Therefore, users felt 44.75% less complex search task complexity on the advanced discovery search system. Hence, we confirm (H1, $\alpha \leq .01$).

B. USERS COGNITIVE STATE

We measured the users' prior web search skills, computer literacy (e.g., ability to use computers), and digital literacy (e.g., ability to use digital technologies) using the self-assessment computing skills questionnaire given in literature [59]. Overall, most participants had "Good" web search skills and digital and computer literacy. Users were primarily confident in their web search skills on the baseline system compared to the advanced search system. This user behavior confirms the study presented in [60], indicating that the first time using a system can be difficult for the users.

Similarly, we measured the user knowledge, interest, and perceived difficulty before and after performing the search task using the scale adopted from the study [13]. We also measured the user search task completion satisfaction and calculated the p-value against each reading using a 2-tailed t-test, displayed in table 3. On average, the participants felt 1.6% more complex search task completion difficulty in the linear organization of the information. Hence, we accept (H2, $\alpha \leq .01$). On average, participants were 43.7% more satisfied after complex task completion on the fully-blended aggregated than the partially blended aggregated approaches. It leads to acceptance of (H3, $\alpha \leq .01$). Similarly, participants gained 32.2% more knowledge and perceived 1.6% less difficulty with the advanced discovery search system despite indicating low web searching and digital literacy skills before performing the search tasks.

C. USERS SEARCH EFFORTS & QUERY STATISTICS

The more clicks and the number of issued queries translated to more effort to explore the search results. The increased query length denotes that users need to specify detailed information to get the relevant results. On average, the users of the advanced search system issued 31.9% fewer queries with 64% fewer keywords to look up, explore, and discover relevant information. Similarly, an enormous difference was noted in the discovery of visual search results. The users, on average, found the relevant visual content with ten times less click effort on the advanced search system. The users discovered textual content with 63.9% less click effort on the advanced search system for textual content. This leads to acceptance of (H4, Knowledge Acquisition $\alpha \leq .01$, Query Length $\alpha \leq .05$, Query # $\alpha \leq .05$). Likewise, the search efforts were found to be overall three times less on the advanced search system ($p \leq .05$) except for the visual results ($p \leq .4$). Hence, we weakly accept (H7).

TABLE 2. Users skill background for (I) Proposed system and (II) Baseline System.

Group	Background	Very Poor			Poor			Acceptable			Good			Very Good		
		P	S	D	P	S	D	P	S	D	P	S	D	P	S	D
I	Web Search Skill	1 (8%)	1 (14%)	0 (0%)	2 (17%)	0 (0%)	1 (33%)	0 (0%)	2 (29%)	0 (0%)	4 (33%)	3 (43%)	1 (33%)	5 (42%)	1 (14%)	1 (33%)
	Computer Literacy	1 (8%)	0 (0%)	0 (0%)	1 (8.3%)	0 (0%)	1 (33%)	0 (0%)	2 (29%)	0 (0%)	3 (25%)	4 (57%)	1 (33%)	7 (58%)	1 (14%)	1 (33%)
	Digital Literacy	2 (17%)	2 (29%)	0 (0%)	0 (0%)	0 (0%)	1 (33%)	3 (25%)	1 (14%)	0 (0%)	4 (33%)	3 (43%)	1 (33%)	3 (25%)	1 (14%)	1 (33%)
II	Web Search Skill	0 (0%)	1 (11%)	0 (0%)	2 (22%)	1 (11.1%)	2 (50%)	4 (44%)	3 (33%)	1 (25%)	3 (33%)	2 (22%)	0 (0%)	0 (0%)	2 (22%)	1 (25%)
	Computer Literacy	0 (0%)	1 (11%)	0 (0%)	2 (22%)	2 (22.2%)	2 (50%)	3 (33%)	1 (11%)	2 (50%)	3 (33%)	2 (22%)	0 (0%)	1 (11%)	3 (33%)	0 (0%)
	Digital Literacy	0 (0%)	0 (0%)	0 (0%)	3 (33%)	2 (22.2%)	1 (25%)	3 (33%)	1 (11%)	2 (50%)	3 (33%)	3 (33%)	1 (25%)	0 (0%)	3 (33%)	0 (0%)

TABLE 3. Participants cognitive state before and after performing the search task on (I) Proposed system, and (II) Baseline system.

		Group	Professional	Student	Domestic	Overall	Sign. (2-Tailed)
Knowledge	Before	I	28%	29%	40%	32%	.001
		II	42%	36%	35%	38%	.000
	After	I	77%	77%	80%	78%	.000
		II	51%	59%	60%	57%	.000
Interest	Before	I	75%	71%	87%	78%	.000
		II	76%	64%	55%	65%	.001
	After	I	77%	89%	80%	82%	.000
		II	53%	64%	40%	52%	.002
Difficulty	Before	I	65%	43%	67%	58%	.002
		II	60%	51%	70%	60%	.001
	After	I	40%	54%	40%	45%	.001
		II	47%	58%	40%	48%	.001
Satisfaction	After	I	90%	89%	93%	91%	.000
		II	51%	40%	50%	47%	.000

TABLE 4. Users search efforts statistics testing based on average number of clicks issued.

	Group	Professional	Student	Domestic	Sig. (2-Tailed)
Query #	I	2.44	1.89	1.25	.012*
	II	1.42	1.14	1.67	.015**
Query Length	I	3.66	3.87	1.75	.005*
	II	1.85	2.14	1.67	.005*
Visual Snippets	I	0.89	2.89	1.25	.112***
	II	0.42	0	0	.423
Textual Snippets	I	1.67	3	2	.031**
	II	1.17	1.57	1.33	.007*

*Significant at $p \leq .01$
**Significant at $p \leq .05$
***Significant at $p \leq .1$

D. INFORMATION ORGANIZATION & MANAGEMENT

Discovery search tasks are complex and require the exploitation of multiple information organizations and presentations. This aspect is tested statistically using a chi-square test based on the navigation pattern of the users. According to the chi-square analysis (table 5), there exists statistically significant strong association between the discovery search tasks and exploitation of multiple information organizations, such as multimedia document, document groups and visualization ($\chi^2(2, N = 22) = 5.867$, $p = .05$, Cramer's $V = .516$). However, there exists statistically insignificant and no association between the discovery search tasks and disjoint presentation of the information, e.g. video, news, image and web verticals ($\chi^2(2, N = 22) = 7.961$, $p = .99$, Cramer's $V = .025$). It can be due to the linear presentation of search results in web search engines, causing information space disorientation for the user [61]. As a result, the experimentation results could not find significant patterns of browsing in linearly ranked verticals-based search engines such as Google. Hence, we weakly accept (H5).

TABLE 5. Validation of discovery search tasks exploitation in (Group-I) multiple information organizations and (Group-II) multiple presentations.

	Group-I			Group-II		
	Value	df	Sign. (2-tailed)	Value	df	Sign.(2-tailed)
Pearson Chi-Square	5.867	2	0.053*	0.014	2	0.993
Likelihood Ratio	7.961	2	0.019*	0.014	2	0.993
Phi	0.516	-	0.053*	0.025	-	0.993
Cramer's V	0.516	-	0.053*	0.025	-	0.993

*Significant at $p \leq .05$

TABLE 6. The UES scores for (Group-I) advanced search system and (Group-II) baseline system.

	Group	Aesthetics	Focused Attention	Perceived Usability	Reward
Professional	I	80.56%	88.33%	94.44%	89.44%
	II	58.52%	60.74%	60.74%	51.11%
Student	I	80.95%	80.95%	83.81%	84.76%
	II	56.19%	66.67%	51.43%	51.43%
Domestic	I	84.44%	88.89%	93.33%	95.56%
	II	51.11%	68.89%	57.78%	55.56%

All results are significant at $p \leq .01$

E. SEARCH USER INTERFACES

The effects of SUIs are measured using quantitative and qualitative measures. The quantitative measures include measuring user engagement using the new and improved short form of the User Engagement Scale (UES-SF). Similarly, User interest is measured using the scale given in literature [13]. The qualitative measures include the empirical results calculated via user behavioral interaction with the system. It includes the number of contents browsed, clicked, and time spent in information exploration. The higher quantitative and qualitative results denote the capability of SUIs to engage the user in the task at hand cognitively.

Overall, on average, users were 29.6% more engaged ($p \leq .01$), 29.7% more interested ($p \leq .01$) on the advanced search system. They explored 94.38% more visual ($p \leq .05$), 313.6% more textual ($p \leq .05$) contents and spent 40.23% longer time ($p \leq .05$) on the advanced search system. The obtained UES results are summarized in table 6. Hence, we accept (H6, $\alpha \leq .05$).

VII. FINDINGS & DISCUSSION

This section provides deep insights into the obtained hypothetical outcomes and explains qualitative users' experiences. The insights provided in this section also contribute toward establishing a standard guideline for future search engines.

TABLE 7. The qualitative content exploration and time spent on the task scores for (Group-I) advanced search system and (Group-II) baseline system.

	Group	Professional	Student	Domestic	Sign. (2-Tailed)
Visual Snippets	I	40.58	24	51.53	.040**
	II	23.33	22.89	13.5	.025**
Textual Snippets	I	43.75	37.57	53.67	.011*
	II	13.22	10.67	8.75	.040**
Time Spent	I	296	513	324	.031**
	II	231	272	305	.006*

*Significant at p $\leq$ .01
**Significant at p $\leq$ .05

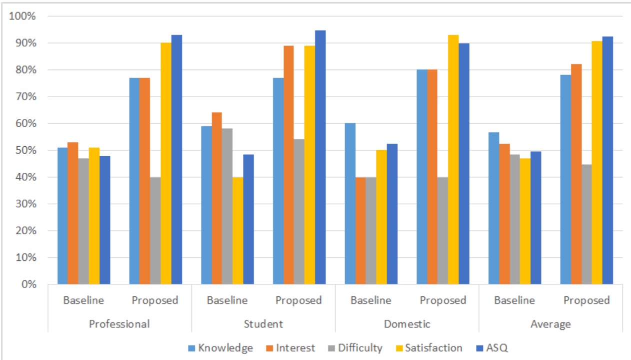

FIGURE 4. Holistic comparison of users cognitive states in proposed and baseline system.

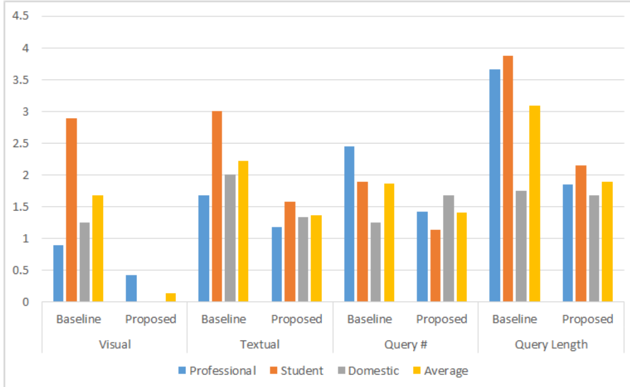

FIGURE 5. Holistic comparison of users searching effort in proposed and baseline system.

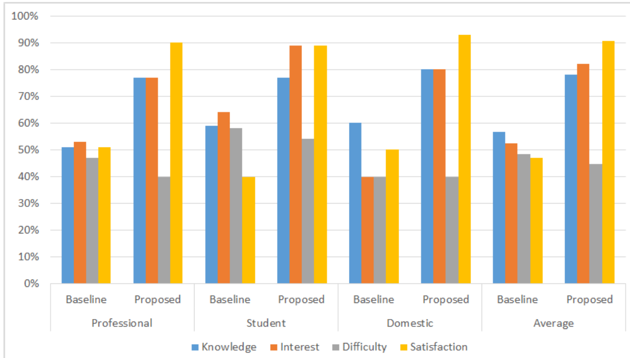

FIGURE 6. Holistic comparison of UES and ASQ scale in proposed and baseline system.

A. TRADITIONAL SEARCH ENGINES INHIBITION

At present, the search engines are the sole gateways to access the information on the web, and a significant number of web users prefer Google search engine [62], [63]. Our demographic collection recorded the users' preferred search source on the web. All chose Google as their primary source for media-specific search results in ranked disjoint verticals. This approach is best optimized for lookup searches that incorporate clearly defined search criteria. However, for exploratory and discovery search tasks, the search criteria are vague and require topical browsing of the search results (Figure 4). In such search tasks, the main objective is to synthesize the information and gain knowledge. Our finding confirms the exploratory study conducted by the authors in [45], indicating that users explore information with fewer efforts when visual cues along with textual content are synthesized. The users spent extra effort on the baseline system in synthesizing the information from various disjoint verticals; as a result, users considered the traditional lookup search engines difficult (H1) and less satisfactory in performing the complex search tasks (H3). Consequently, the overall perceived task difficulty is higher in the existing search engine (H2). It adds to the significance of the previous study reported in literature [64], indicating that existing search engines are inadequate to foster exploration and discovery of multimedia information.

B. RE-ENGINEERING INFORMATION

The disjoint placement of the multimedia search results induces information disorientation in the users during topical exploratory searches. For instance, a topic like "NASA Moon Landing" may contain news, video, and image search results. These dispersed search results require manual assembling efforts and makes efficient knowledge acquisition challenging. As a result, users issue more queries and clicks to obtain relevant search results (Figure 5). Users acquire 32% less knowledge with extra assembling efforts (H4). Hence, providing aggregated access at various granularity levels is the most effective method in exploratory and discovery-oriented search tasks (H5). It also confirms a previous study in [65], concluding that categorical organization of the search results helps exploration and discovery of the results without necessarily adding more SUI complexity than standard Google-like interfaces.

C. SUI IMPLICATIONS

The SUI is the most neglected aspect of the information exploration and discovery search systems. An aesthetically pleasing SUI is a primitive aspect of human-centered system design (Figure 6). A cognitively incompatible SUI leads to lower user engagement and limits the exploration of the information. As a result, the previous exploratory search systems achieve at most the "B" grade in overall usability results [11], [30]. A consistent and aesthetically pleasing

TABLE 8. Users perceived search tasks complexity using (ASQ) scores.

	Group-I				Group-II			
	P	S	D	Sign. (2-Tailed)	P	S	D	Sign. (2-Tailed)
Overall, I am satisfied with the support information when completing the tasks	96.67%	90.00%	90.00%	.001	44.44%	46.67%	57.78%	.005
Overall, I am satisfied with the ease of completing the tasks in this scenario	88.57%	94.29%	85.71%	.001	44.44%	48.89%	44.44%	.000
Overall, I am satisfied with the amount of time it took to complete the tasks	93.33%	100.00%	93.33%	.001	55.00%	50.00%	55.00%	.006

TABLE 9. Summary of studied variables, parameters, hypotheses, result and statistical significance.

S#	Variables	Parameters	Hypothesis	Result	Significance
1	Usage Complexity	After Scenario Questionnaire	H1	Accept	0.01
2	Perceived Task Difficulty	Perceived Task Difficulty Questionnaire	H2	Accept	0.01
3	Task Completion satisfaction	Task Satisfaction Questionnaire	H3	Accept	0.01
4	User cognition	Knowledge Acquisition Questionnaire, Queries Length, Number of Queries Issued	H4	Accept	0.05
5	Information Organization	Content Modality, Content Granularity	H5	Weakly Accept	0.05 & 0.99
6	User Interface	User Interest, Exploration Time Contents Browsed	H6	Accept	0.05
7	Search Efforts	# of Clicks Issued	H7	Weakly Accept	0.05 & 0.4

SUI increases the users focused attention, perceived usability, reward, and exploratory activities (H5 & H6). However, an aesthetically pleasing user interface must be accompanied by a practical layout of information presentation.

The grid layout can accumulate multiple types of information effectively. As evident from (Figure 2 and Figure 3), each textual information is endorsed by visual information in a grid, therefore enhances users' comprehensive. It allows a user to rapidly deduce the relevancy of required information without explicitly clicking on the source, which reduces the search efforts (Table 4). Therefore the grid layout should be preferred when rendering multimedia information to enhance the information overview. The linear layout is more suitable for detailed information scrutiny. Therefore, adopting an effective information presentation style according to the task goal eases information exploration and leads to lower search efforts (H7).

D. FREE-FORM COMMENTS

At the end of the evaluation session, the users were allowed to provide free-form comments to express their feedback. On the positive side, based on the obtained comments for the advanced search system, most users preferred the multiple visualizations of the search results (e.g., multimedia snippets, documents, and document groups) and find semantically aggregated search results space comparatively easier to comprehend and explore. Similarly, the users stated that the appealing SUI and increased visibility of all the possible filtration options on the screen added to a pleasant search experience. Users find the traditional linear lists simple to use in the baseline system.

As a negative aspect, users of the baseline system find that the linearly ranked search results list excessively narrows the topic in the displayed information. Due to the invisibility of complete filtration controls in traditional web search engines, users repeatedly reformulated queries to get diverse results. For the advanced search system, the only reported drawback indicated by the users was increased time required to effectively use the complex system due to lack of prior familiarity. These general observations provided by the users further support the observed hypothesis and confirms the inadequacy of existing traditional web search engines to meet the user's multifaceted information needs [3], [30], [52].

Prospective search engines can incorporate multiple data visualization and views to enhance user cognitive abilities and maximize the controlling interaction points (such as filtration options, visually aiding material, and consistent coloring conventions) in the SUI. According to the users' feedback, most of the users were attracted to the look and feel of the advanced search system; however, we believe that the users implicitly learned the consistent coloring convention, which added symmetry to their learning and enhanced the overall user experience [53]. Therefore, considering these minor visual cues while designing the SUI can enhance users' experience and system learn-ability.

E. USER CATEGORICAL BEHAVIOR

Most of the previous research performed usability analysis without user categorization and presented the results as a single unit [14]. A few of the most recent studies categorized the users into general groups, such as employees, undergrad and graduate students, and homemakers [11], [30], [66]. Therefore, we performed the user categorization to further measure the quality of user interaction by the three major holistic user categories (e.g., professionals, students, and homemakers) without creating excessive experimental

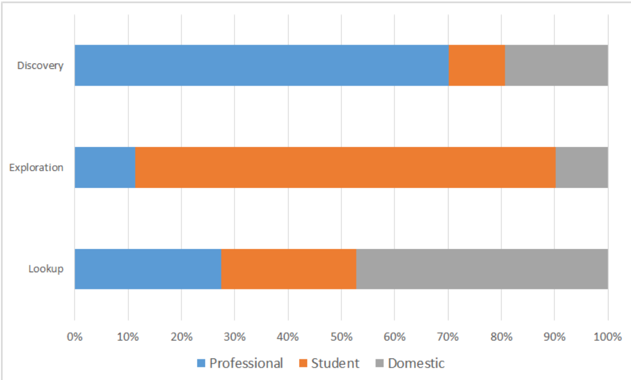

FIGURE 7. Categorized information seeking behavior of the participants using time usage.

complexity and providing group-wise insights about the results obtained [49]. All the achieved results between the different categories of the users (e.g., professionals, students, and homemakers) demonstrated almost similar outcomes. The only exception was the student users. The students felt a slight increase in the post-task difficulty. It can be due to behavioral characteristics of the students, such as willingness to trade results quality for effort, time spent in searching, and natural tendency to face difficulties in locating information and resources [67].

Another interesting finding was observed in the holistic information-seeking behavior of the participants in terms of usage time (Figure 7). The professionals, students, and domestic participants spent most on the discovery, exploratory, and lookup interfaces, respectively. It indicates that professionals' information needs are creative, requiring the discovery of new information and gaining a quick overview of the information. The students demonstrated a preference for the topical exploration of the information. The domestic users mainly were females. As a result, they utilized the most traditional (linear lookup list) means of searching. This domestic behavior confirms a previous study that indicated risk-averse behavior of the female user group [68].

VIII. CONCLUSION & FUTURE WORK

This research briefly studied the exploratory and discovery search characteristics and their influences on the users in an actual scenario using a real-dataset on the existing state-of-the-art (Google) search engine. We recruited 44 participants, mainly categorized as professionals, students, and domestic/homemakers groups, divided into the between-subjects design. Various user behavioral factors such as users' cognitive state, including knowledge acquisition, interest, system and task difficulty, and satisfaction were measured. We confirmed previous studies indicating users' increased difficulty (H1, +44.75%) during information exploration and discovery on traditional web search engines. The widely criticized linear list organization of information also showed a vital role in increasing user searching specifically (H2, +1.6%) and decreased task completion satisfaction (H3, −43.7%). In contrast, disjoint multimedia aggregated information and non-linear information exploration data model increased the user cognitive abilities, including perceived knowledge (+32.2%), and allowed the users to gain required information with 31.9% fewer queries issued 64% short-typed queries (H4).

We also confirmed that the efficacy of the search engines that supported the exploitation of multiple information organization structures (H5) augmented with appealing SUIs, and aggregated the relevant information semantically (H6) reduced the search efforts up to three times (H7). Therefore, users demand the exploitation of multimedia aggregated information that can be browsed non-linearly and explored in-depth when required. Hence, an exploratory search engine should allow the exploitation of multiple visualization views to stimulate the users' cognitive abilities and semantically aggregate relevant information to prevent information disorientation. Overall, we confirmed the inadequacy of the existing search engines in discovery search scenarios qualitatively with statistical significance to serve as a recommended guideline for future research.

This study mainly employed a between-subjects design to ensure no previous trailing biases of the user affected the evaluation process. However, it remains an open question to empirically quantify the biases induced in a within-subjects design and to what extent it affects the credibility of the obtained results. Therefore, in the future, we look forward to investigating the effects of between-subjects and within-subjects design in discovery search scenarios. We are also interested in performing in-depth categorical information-seeking behavior of the users based on extended holistic taxonomic categories.

ACKNOWLEDGMENT

The authors would like to acknowledge the support of Prince Sultan University for paying the Article Processing Charges (APC) of this publication. They would also like to thank Prince Sultan University for their support.

CODE & DATA AVAILABILITY

The code [69] and data [70] associated with this article is publicly published and available for reproducibility at DOI 10.24433/CO.8476348.v1 & 10.21227/xzb9-xm80 respectively.

REFERENCES

[1] D. Flores-Martin, J. Berrocal, J. García-Alonso, C. Canal, and M. Juan Murillo, "Enabling the interconnection of smart devices through semantic web techniques," in *Web Engineering*, M. Bakaev, F. Frasincar, and I.-Y. Ko, Eds. Cham, Switzerland: Springer, 2019, pp. 534–537.

[2] S. Mahmoudi and M. A. Belarbi, "Large scale multimedia management: Recent challenges," *Information*, vol. 13, no. 1, p. 28, Jan. 2022.

[3] A. R. Khan, U. Rashid, K. Saleem, and A. Ahmed, "An architecture for non-linear discovery of aggregated multimedia document web search results," *PeerJ Comput. Sci.*, vol. 7, p. e449, Apr. 2021.

[4] E. Palagi, F. Gandon, A. Giboin, and R. Troncy, "A survey of definitions and models of exploratory search," in *Proc. ACM Workshop Explor. Search Interact. Data Analytics*, Mar. 2017, pp. 3–8.

[5] T. Ruotsalo, J. Peltonen, M. J. A. Eugster, D. Głowacka, P. Floréen, P. Myllymäki, G. Jacucci, and S. Kaski, ''Interactive intent modeling for exploratory search,'' *ACM Trans. Inf. Syst.*, vol. 36, no. 4, pp. 1–46, Oct. 2018.

[6] S. Pouyanfar, Y. Yang, S.-C. Chen, M.-L. Shyu, and S. S. Iyengar, ''Multimedia big data analytics: A survey,'' *ACM Comput. Surv.*, vol. 51, no. 1, pp. 1–34, Jan. 2018.

[7] L. C. J. Tseng, D. Tjondronegoro, and A. Spink, ''Analysing web multimedia query reformulation behaviour,'' in *Proc. 14th Australas. Document Comput. Symp.*, 2009, pp. 118–125.

[8] S. S.-L. Tan and N. Goonawardene, ''Internet health information seeking and the patient-physician relationship: A systematic review,'' *J. Med. Internet Res.*, vol. 19, no. 1, p. e9, Jan. 2017.

[9] A. Kopliku, F. Damak, K. Pinel-Sauvagnat, and M. Boughanem, ''Interest and evaluation of aggregated search,'' in *Proc. IEEE/WIC/ACM Int. Conferences Web Intell. Intell. Agent Technol.*, Aug. 2011, pp. 154–161.

[10] E. Koh, B. Dworaczyk, J. Albea, R. Hill, H. Choi, and D. Caruso, ''Combinformation: A mixed-initiative system for representing collections as compositions of image and text surrogates,'' in *Proc. 6th ACM/IEEE-CS Joint Conf. Digital Libraries (JCDL)*, Jun. 2006, pp. 11–20.

[11] U. Rashid and M. A. Bhatti, ''A framework to explore results in multiple media information aggregated search,'' *Multimedia Tools Appl.*, vol. 76, no. 24, pp. 25787–25826, Dec. 2017.

[12] T. Ruotsalo, J. Peltonen, M. J. A. Eugster, D. Glowacka, A. Reijonen, G. Jacucci, P. Myllymäki, and S. Kaski, ''SciNet: Interactive intent modeling for information discovery,'' in *Proc. 38th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, 2015, pp. 1043–1044.

[13] J. Mao, Y. Liu, N. Kando, M. Zhang, and S. Ma, ''How does domain expertise affect users' search interaction and outcome in exploratory search?'' *ACM Trans. Inf. Syst.*, vol. 36, no. 4, pp. 1–30, 2018.

[14] M. Taramigkou, D. Apostolou, and G. Mentzas, ''Supporting creativity through the interactive exploratory search paradigm,'' *Int. J. Hum.–Comput. Interact.*, vol. 33, no. 2, pp. 94–114, Feb. 2017.

[15] A. Hassan Awadallah, R. W. White, P. Pantel, S. T. Dumais, and Y.-M. Wang, ''Supporting complex search tasks,'' in *Proc. 23rd ACM Int. Conf. Conf. Inf. Knowl. Manage.*, Nov. 2014, pp. 841–844.

[16] A. R. Khan and U. Rashid, ''A relational aggregated disjoint multimedia search results approach using semantics,'' in *Proc. Int. Conf. Artif. Intell. (ICAI)*, Apr. 2021, pp. 62–67.

[17] A. Kerne and S. M. Smith, ''The information discovery framework,'' in *Proc. 5th Conf. Designing Interact. Syst., Processes, Practices, Methods, Techn.*, 2004, pp. 357–360.

[18] G. Marchionini, ''Exploratory search: From finding to understanding,'' *Commun. ACM*, vol. 49, no. 4, pp. 41–46, 2006.

[19] A. Kerne, S. M. Smith, E. Koh, H. Choi, and R. Graeber, ''An experimental method for measuring the emergence of new ideas in information discovery,'' *Int. J. Hum.-Comput. Interact.*, vol. 24, no. 5, pp. 460–477, Jun. 2008.

[20] T. Russell-Rose and T. Tate, *Designing the Search Experience: The Information Architecture of Discovery*, 1st ed. San Francisco, CA, USA: Morgan Kaufmann, 2012.

[21] L. Turpin, D. Kelly, and J. Arguello, ''To blend or not to blend?: Perceptual speed, visual memory and aggregated search,'' in *Proc. 39th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, Jul. 2016, pp. 1021–1024.

[22] V. Tablan, K. Bontcheva, I. Roberts, and H. Cunningham, ''Mimir: An open-source semantic search framework for interactive information seeking and discovery,'' *J. Web Semantics*, vol. 30, pp. 52–68, 2015.

[23] M. Dörk, S. Carpendale, C. Collins, and C. Williamson, ''Visgets: Coordinated visualizations for web-based information exploration and discovery,'' *IEEE Trans. Vis. Comput. Graphics*, vol. 14, no. 6, pp. 1205–1212, Nov./Dec. 2008.

[24] A. Khalili, P. van Andel, P. van den Besselaar, and K. A. de Graaf, ''Fostering serendipitous knowledge discovery using an adaptive multigraph-based faceted browser,'' in *Proc. Knowl. Capture Conf.*, Dec. 2017, p. 15.

[25] P. Lisena, R. Troncy, K. Todorov, and M. Achichi, ''Modeling the complexity of music metadata in semantic graphs for exploration and discovery,'' in *Proc. 4th Int. Workshop Digit. Libraries Musicol.*, Oct. 2017, pp. 17–24.

[26] N. Kanjanakuha, P. Janecek, and C. Techawut, ''The comprehensibility assessment of visualization of semantic data representation (VSDR) reflecting user capability of knowledge exploration and discovery,'' in *Proc. 7th Int. Conf. Comput. Commun. Manage.*, Jul. 2019, pp. 195–199.

[27] M. Taramigkou, D. Apostolou, and G. Mentzas, ''Leveraging exploratory search with personality traits and interactional context,'' *Inf. Process. Manage.*, vol. 54, no. 4, pp. 609–629, Jul. 2018.

[28] S. Achsas and E. H. Nfaoui, ''An analysis study of vertical selection task in aggregated search,'' *Proc. Comput. Sci.*, vol. 148, pp. 171–180, Jan. 2019.

[29] D. Bakrola and S. Gandhi, ''Enhancing web search results using aggregated search,'' in *Proc. Int. Conf. ICT Sustain. Develop.* (Advances in Intelligent Systems and Computing), vol. 409, S. Satapathy, A. Joshi, N. Modi, and N. Pathak, Eds. Singapore: Springer, 2016, pp. 675–688. [Online]. Available: https://link.springer.com/chapter/10.1007/978-981-10-0135_65, doi: 10.1007/978-981-10-0135-2_65.

[30] U. Rashid, K. Saleem, and A. Ahmed, ''MIRRE approach: Nonlinear and multimodal exploration of MIR aggregated search results,'' *Multimedia Tools Appl.*, vol. 80, pp. 20217–20253, Mar. 2021.

[31] K. Athukorala, D. Głowacka, G. Jacucci, A. Oulasvirta, and J. Vreeken, ''Is exploratory search different? A comparison of information search behavior for exploratory and lookup tasks,'' *J. Assoc. Inf. Sci. Technol.*, vol. 67, no. 11, pp. 2635–2651, Nov. 2016.

[32] S. Alspaugh, N. Zokaei, A. Liu, C. Jin, and M. A. Hearst, ''Futzing and moseying: Interviews with professional data analysts on exploration practices,'' *IEEE Trans. Vis. Comput. Graphics*, vol. 25, no. 1, pp. 22–31, Jan. 2019.

[33] D. Glowacka, T. Ruotsalo, K. Konyushkova, K. Athukorala, S. Kaski, and G. Jacucci, ''SciNet: A system for browsing scientific literature through keyword manipulation,'' in *Proc. Companion Publication Int. Conf. Intell. Interfaces Companion*, 2013, pp. 61–62.

[34] M. Steiner, M. Magin, B. Stark, and S. Geiß, ''Seek and you shall find? A content analysis on the diversity of five search engines' results on political queries,'' *Inf., Commun. Soc.*, vol. 25, no. 2, pp. 217–241, 2020.

[35] S. Achsas and E. H. Nfaoui, ''Improving relational aggregated search from big data sources using stacked autoencoders,'' *Cognit. Syst. Res.*, vol. 51, pp. 61–71, Oct. 2018.

[36] K. Klouche, T. Ruotsalo, D. Cabral, S. Andolina, A. Bellucci, and G. Jacucci, ''Designing for exploratory search on touch devices,'' in *Proc. 33rd Annu. ACM Conf. Hum. Factors Comput. Syst.*, New York, NY, USA, Apr. 2015, pp. 4189–4198.

[37] S. Achsas, ''Improving relational aggregated search from big data sources using deep learning,'' in *Proc. Intell. Syst. Comput. Vis. (ISCV)*, Apr. 2017, pp. 1–6.

[38] P. H. Cleverley, S. Burnett, and L. Muir, ''Exploratory information searching in the enterprise: A study of user satisfaction and task perfo rmance,'' *J. Assoc. Inf. Sci. Technol.*, vol. 68, no. 1, pp. 77–96, Jan. 2017.

[39] P. Pirolli and S. Card, ''Information foraging,'' *Psychol. Rev.*, vol. 106, no. 4, p. 643, 1999.

[40] C. D. Sciascio, V. Sabol, and E. E. Veas, ''Rank as you go: User-driven exploration of search results,'' in *Proc. 21st Int. Conf. Intell. Interfaces*, 2016, pp. 118–129.

[41] C. L. Paul, J. Chang, A. Endert, N. Cramer, D. Gillen, S. Hampton, R. Burtner, R. Perko, and K. A. Cook, ''TexTonic: Interactive visualization for exploration and discovery of very large text collections,'' *Inf. Vis.*, vol. 18, no. 3, pp. 339–356, Jul. 2019.

[42] M. Bron, J. van Gorp, F. Nack, L. B. Baltussen, and M. de Rijke, ''Aggregated search interface preferences in multi-session search tasks,'' in *Proc. 36th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, Jul. 2013, pp. 123–132.

[43] X. Li, Y. Liu, R. Cai, and S. Ma, ''Investigation of user search behavior while facing heterogeneous search services,'' in *Proc. 10th ACM Int. Conf. Web Search Data Mining*, Feb. 2017, pp. 161–170.

[44] L. Punchoojit and N. Hongwarittorrn, ''Usability studies on mobile user interface design patterns: A systematic literature review,'' *Adv. Hum.-Comput. Interact.*, vol. 2017, pp. 1–22, 2017.

[45] S. Lu, T. Mei, J. Wang, J. Zhang, Z. Wang, and S. Li, ''Exploratory product image search with circle-to-search interaction,'' *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 7, pp. 1190–1202, Jul. 2015.

[46] M. A. Sharaf and H. Ehsan, ''Efficient query refinement for view recommendation in visual data exploration,'' *IEEE Access*, vol. 9, pp. 76461–76478, 2021.

[47] A. Rorissa and X. Yuan, ''Visualizing and mapping the intellectual structure of information retrieval,'' *Inf. Process. Manage.*, vol. 48, no. 1, pp. 120–135, Jan. 2012.

[48] D. Yuniarto, M. Suryadi, E. Firmansyah, D. Herdiana, A. Subiyakto, and A. B. A. Rahman, ''Integrating the readiness and usability models for assessing the information system use,'' in *Proc. 6th Int. Conf. Cyber IT Service Manage. (CITSM)*, Aug. 2018, pp. 1–6.

[49] K. Hornbæk, "Current practice in measuring usability: Challenges to usability studies and research," *Int. J. Hum.-Comput. Stud.*, vol. 64, no. 2, pp. 79–102, Feb. 2006.
[50] M. Dunaiski, G. J. Greene, and B. Fischer, "Exploratory search of academic publication and citation data using interactive tag cloud visualizations," *Scientometrics*, vol. 110, no. 3, pp. 1539–1571, Mar. 2017.
[51] R. C. Basole, A. Srinivasan, H. Park, and S. Patel, "ecoxight: Discovery, exploration, and analysis of business ecosystems using interactive visu alization," *ACM Trans. Manage. Inf. Syst.*, vol. 9, no. 2, pp. 1–26, 2018.
[52] U. Rashid, M. Viviani, and G. Pasi, "A graph-based approach for visualizing and exploring a multimedia search result space," *Inf. Sci.*, vols. 370–371, pp. 303–322, Nov. 2016.
[53] P. Khast, "Ux/Ui design process for a peer to peer financial platform," M.S. thesis, Dept. Inf. Technol., Helsinki Metropolia Univ. Appl. Sci., Helsinki, Finland, 2017. [Online]. Available: https://www.theseus.fi/handle/10024/122685
[54] R. Alroobaea and P. J. Mayhew, "How many participants are really enough for usability studies?" in *Proc. Sci. Inf. Conf.*, Aug. 2014, pp. 48–56.
[55] C. W. Turner, J. Nielsen, and J. R. Lewis, "Current issues in the determination of usability test sample size: How many users is enough," in *Proc. Usability Professionals' Assoc. Conf.*, 2002, pp. 1–16.
[56] E. Hoque, O. Hoeber, and M. Gong, "CIDER: Concept-based image diversification, exploration, and retrieval," *Inf. Process. Manage.*, vol. 49, no. 5, pp. 1122–1138, Sep. 2013.
[57] N. Marie, F. Gandon, D. Legrand, and M. Ribière, "Exploratory search on the top of dbpedia chapters with the discovery hub application," in *The Semantic Web: ESWC 2013 Satellite Events*, P. Cimiano, M. Fernández, V. Lopez, S. Schlobach, and J. Völker, Eds. Berlin, Germany: Springer, 2013, pp. 184–188.
[58] J. R. Lewis, "IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use," *Int. J. Hum.-Comput. Interact.*, vol. 7, no. 1, pp. 57–78, Jan. 1995.
[59] J.-B. Son, S.-S. Park, and M. Park, "Digital literacy of language learners in two different contexts," *JALT CALL J.*, vol. 13, no. 2, pp. 77–96, Aug. 2017.
[60] F. Gustafsson, "An explorative design study of a booking system: Evaluating the usability and experience of a user interface for novice admin users," M.S. thesis, School Elect. Eng. Comput. Sci. (EECS), KTH, Stockholm, Sweden, 2019.
[61] Z. Yang, N. Peng, and C. Xue, "Exploratory search behaviors in linear list visualization," in *Proc. Int. Conf. Converg. Digit. World-Quo Vadis (ICCDW)*, Feb. 2020, pp. 1–6.
[62] V. Shrivastava, "A methodical study of web crawler," *J. Eng. Res. Appl.*, vol. 8, no. 11, pp. 1–8, 2018.
[63] S. Ali and S. Gul, "Search engine effectiveness using query classification: A study," *Online Inf. Rev.*, vol. 40, no. 4, pp. 515–528, Aug. 2016.
[64] D. Tjondronegoro and A. Spink, "Web search engine multimedia functionality," *Inf. Process. Manage.*, vol. 44, no. 1, pp. 340–357, Jan. 2008.
[65] B. Kules and B. Shneiderman, "Users can change their web search tactics: Design guidelines for categorized overviews," *Inf. Process. Manage.*, vol. 44, no. 2, pp. 463–484, Mar. 2008.
[66] M. A. M. Nieto, P. D. V. Mora, J. De La Calleja Mora, M. T. Vidal, E. L. Dominguez, D. A. Diaz, and I. E. B. Patino, "Web service to retrieve and semantically enrich datasets for theses from open educational repositories," *IEEE Access*, vol. 8, pp. 171933–171944, 2020.
[67] J. Griffiths and P. Brophy, "Student searching behavior and the web: Use of academic resources and Google," *Library Trends*, vol. 53, pp. 539–554, May 2005.
[68] A. Shekhar and N. Marsden, "Cognitive walkthrough of a learning management system with gendered personas," in *Proc. 4th Conf. Gender IT (GenderIT)*, 2018, pp. 191–198.
[69] A. R. Khan. (Feb. 2022). *Multimedia Aggregated Information Discovery Software*. [Online]. Available: https://www.codeocean.com/
[70] A. R. Khan, U. Rashid, and N. Ahmed, "Multimedia aggregated information discovery data," *IEEE Dataport*, Mar. 2022. [Online]. Available: https://ieee-dataport.org/documents/multimedia-aggregated-information-discovery-data, doi: 10.21227/xzb9-xm80.

ABDUR REHMAN KHAN received the Distinguished B.S. degree (Hons.) in computer science from the National College of Business Administration and Economics, Lahore, and the M.Phil. degree in computer science from Quaid-i-Azam University, Islamabad, Pakistan. He possesses diverse working experience in private, semi-government, and government institutes, as an Android Developer, Graphic Designer, Research Assistant, and IT Professional. His research interests include multimedia information retrieval, user-centered computing, AI, search engines, UI/UX, web, and data science.

UMER RASHID received the B.S. degree in computer sciences from the University of Lahore, Pakistan, in 2005, and the M.Phil. and Ph.D. degrees in computer sciences from Quaid-i-Azam University, Islamabad, Pakistan, in 2008 and 2017, respectively. He is currently an Assistant Professor of computer sciences at Quaid-i-Azam University. He has teaching and research experience in international organizations. His work is published in international journals and conferences. He is also the author of several book chapters. He also reviewed for several international journals and conferences. His research interests include user-centered computing, multimedia information retrieval, and multimedia technology.

NAVEED AHMED (Member, IEEE) graduated in computer science from the Department of Computer Science, University of Peshawar, Pakistan, in 2007. He received the Ph.D. degree in computer science from the Center for Communication System Research (CCSR), University of Surrey, U.K., in 2013. He is currently working as an Associate Professor at the College of Computer and Information Sciences, Prince Sultan University, Saudi Arabia. Prior to his current position, he remained an Assistant Professor at the Department of Computer Science, University of Peshawar. He has over 45 publications that include international reputed journals, conferences, and workshops. He worked in the area of cyber security, privacy, blockchain technology, and penetration testing. He has managed three research and development grants related to blockchain, transport system for emergency vehicles, and platoon management. He remained reviewers of various research proposals/grants funded by the U.K. Research and Innovation (UKRI). He also served on the program committee for various national conferences and workshops. He remained reviewer of various IEEE, Elsevier, Springer, and MPDI journals. He has the honor of serving as the Guest Editor for a special issue in the IEEE Internet of Things Journal.