

Article

Understanding Query Combination Behavior in Exploratory Searches

Pengfei Li ¹, Yin Zhang ^{2,*} and Bin Zhang ²

¹ School of Computer Science and Engineering, Northeastern University, No.195 Chuangxin Road, Hunnan District, Shenyang 110167, China; leepengfei.neu@gmail.com

² Software College, Northeastern University, No.195 Chuangxin Road, Hunnan District, Shenyang 110167, China; zhangbin@mail.neu.edu.cn

* Correspondence: zhangyin@mail.neu.edu.cn

Abstract: In exploratory search, users sometimes combine two or more issued queries into new queries. We present such a kind of search behavior as query combination behavior. We find that the queries after combination usually can better meet users' information needs. We also observe that users combine queries for different motivations, which leads to different types of query combination behaviors. Previous work on understanding user exploratory search behaviors has focused on how people reformulate queries, but not on how and why they combine queries. Being able to answer these questions is important for exploring how users search and learn during information retrieval processes and further developing support to assist searchers. In this paper, we first describe a two-layer hierarchical structure for understanding the space of query combination behavior types. We manually classify query combination behavior sessions from AOL and Sogou search engines and explain the relationship from combining queries to success. We then characterize some key aspects of this behavior and propose a classifier that can automatically classify types of query combination behavior using behavioral features. Finally, we summarize our findings and show how search engines can better assist searchers.

Keywords: query combination behavior; exploratory search; user study

Citation: Li, P.; Zhang, Y.; Zhang, B. Understanding Query Combination Behavior in Exploratory Searches. *Appl. Sci.* **2022**, *12*, 706. <https://doi.org/10.3390/app12020706>

Academic Editor:
Agostino Forestiero

Received: 6 December 2021
Accepted: 10 January 2022
Published: 11 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Exploratory search is a special type of information seeking [1] carried out by search engine users who are not familiar with their search domains, who are motivated by complex information needs, or who encounter indexes of information which are inadequate [2–4]. Under such conditions, users have to spend more time searching for related documents, obtaining new knowledge from them, and planning for new queries by synthesizing the information they gathered during the search process [5,6]. Exploratory search studies focus on diversifying user search behaviors to help users learn and explore the unknown [7–9].

During the process of analyzing user behavior in exploratory searches, there is a wealth of search behavior where users combine two or more issued queries into new queries. In this work, we present such a type of search behavior as query combination behavior (QCB). Here, we consider two examples to illustrate this behavior. Both examples come from the AOL data set (released by America Online, a collection of real query logs that is based on 650,000 real users). Figure 1a shows a search session where the user is seeking information on dogs. We can see that, in this session, the user issued content-related queries (all queries pertaining to dogs) and clicked on relevant results. The descriptions of information needs are becoming more and more accurate (“pitbull” is a kind of dog, but the query “pitbull” yields results on a variety of topics, such as the singer

“Pitbull” and the nickname “Pitbull”; therefore, the query “pitbull dogs” is a more accurate description of the information needs compared with the query “pitbull”, thereby reducing redundant topics), providing some evidence that the user is exploring during information seeking. The user performed QCB by combining the query “dogs” and the query “pitbull” into the following query: “pitbull dogs”. Figure 1b shows another search session, where the user was seeking musical resources. The query “everything I want the world” and the query “a pleasing sense of happiness” are both lyrics of a song named “Top of the world”, which was evidence that the user was exploring information pertaining to this song. The user performed QCB by combining the query “everything I want the world” and the query “a pleasing sense of happiness” into the following query: “everything I want the world a pleasing sense of happiness for me”.

00:59:14 **Queried** dogs
 Clicked <http://dogs.about.com> <http://www.thetypesofdogs.com>
 01:06:40 **Queried** pitbull
 Clicked <http://www.pitbullmusic.com>
 01:06:57 **Queried** pitbull dogs
 Clicked <http://www.pitbull411.com/>
 01:21:32 END OF SESSION
 (a) A search session where the user is seeking information on dogs
 10:43:55 **Queried** everything I want the world
 Clicked None
 10:45:08 **Queried** a pleasing sense of happiness
 Clicked None
 10:45:23 **Queried** everything I want the world a pleasing sense of happiness for
 me
 Clicked <https://www.azlyrics.com/lyrics/acdc/topoftheworld.html/>
 11:06:07 END OF SESSION
 (b) A search session where the user is seeking information on music

Figure 1. Examples of QCB sessions.

We observed that users can meet their information needs by combining queries where they analyzed and synthesized the information they have learned. For example, in Figure 1a, the user performed QCB after he or she: (1) analyzed the search results (found the query “pitbull” yields results on a variety of topics); (2) synthesized the information needs “pitbull” with the topic “dogs”. We also notice that the inter-query time between “pitbull” and “pitbull dogs” is relatively short (the click on the web page is a music website which has nothing to do with dogs), and the inter-query time between “pitbull dogs” and the end of the session is the longest, providing some evidence that the query “pitbull dogs” can better meet the user’s information need.

Based on this observation, the users who performed QCB have to search and learn during their information retrieval processes. Therefore, users who perform QCB not only have sufficient searching abilities but also sufficient learning abilities. The current literature on users’ searching behaviors has figured out that users would be able to search better and possibly learn more with their learning skills (e.g., understanding, analysis, application, or synthesis) improved [10,11]. These studies emphasize the importance of conceptualizing learning as search outcomes [12].

Searching as learning (SAL) studies bring together two disconnected research areas: information retrieval and pedagogy [13]. From the perspective of searching as learning, searching is conceptualized as a process in which searchers engage in various search activities for learning [10,14]. During this process, people critically analyze information, bringing pieces of information together to create something new, through evaluating and using information [10]. This aligns with our previous observation from QCB sessions, where users analyze search results and synthesize the information they learned.

Therefore, being able to understand this kind of exploratory search behavior is important for understanding how users analyze search results and how they bring pieces of information together; therefore, this can further help users find information more easily by designing a range of applications, including recommendation systems and feature engineering, etc. However, to the best of our knowledge, this is the first study that characterizes this typical type of search behavior in exploratory searches. We thus address this gap in the literature with the research presented in this paper.

We first focus on exploring the space of QCB types to answer how and why users combine queries, and to explain the relationship between combining queries and success. We then characterize some key aspects of this behavior and present a methodology that uses machine learning techniques to autonomously classify QCB sessions.

Through our analysis, this paper proposes a two-layer hierarchical structure of QCB, discusses the role QCB play in exploratory search, characterizes QCB features, and uses these insights to develop machine learning models capable of accurately distinguishing QCB sessions. Specifically, we make the following four research contributions with the work presented in this paper:

- We define QCB in search sessions and present a two-layer hierarchical structure of QCB.
- We propose the role that QCB plays in exploratory searches.
- We characterize differences in the search behavior, which are associated with different types of QCB in the hierarchy.
- We build classified models, based on these differences, to automatically classify QCB for a given set of behavioral data.

2. Related Work

Characterizing the query behavior of exploratory search users has been the subject of previous studies from different perspectives. The line of work most related to ours is query reformulation, because QCB can be seen as a typical kind of query reformulation in exploratory searches.

Query reformulation has been widely studied to reduce the gap between queries and information needs [15,16]. By reformulating user queries into appropriate new forms, it enables users to retrieve more relevant information [17]. Therefore, studies of query reformulation are particularly important in exploratory search, as users may lack the necessary domain knowledge to describe their information needs [3,18,19].

Some research studied query reformulation from the perspective of reformulation strategies. Huang et al. [20] focus on the taxonomy developed by analyzing search logs and summarizing query reformulation strategies from prior works. These strategies include adding words, removing words, word reorder, word substitution, etc. Boldi et al. [21] classify user query reformulations into four categories, as follows: generalization, specialization, error correction, and parallel move. They built a model for classifying query reformulations. Bosung Kim et al. [22] proposed that the studies on query reformulation can be divided into three categories, as follows: query reduction, substitution, and query expansion. While these studies serve as insights into query re-formulation, these approaches all focus on superficial lexical aspects of reformulation [20].

Web search involves abundant human interactions, and these activity footprints can reflect the real information needs of the user [23–25]. A number of recent studies have shown that log-based user behaviors provide rich context for performing query reformulation [26]. To better understand user reformulations, recently, more studies have focused on analyzing users' reformulating behaviors [3,15]. Therefore, search logs are widely used to analyze user reformulating behavior because they record information about the full interaction between users and search engines [27].

Medlar et al. [3] investigated the use of query reformulation in exploratory search and explore the effect of query reformulation on user behavior. They found that query

reformulation suggestions are important in the decision-making process of whether or not to terminate the current phase of their information seeking. Jiang et al. [28] proposed an exploitation of heterogeneous network embedding to learn the semantics of queries and term the dependencies of queries in search logs. They figured out that homomorphic reformulations embeddings that precisely capture both syntactic and semantic reformulations can essentially benefit query reformulation suggestion. Sloan et al. [29] explored the relationship between terms in adjacent queries to aid in query reformulation. They indicated that a significant number of added terms in query reformulations can be sourced from the terms that the user was exposed to in the previous impression.

This work extends existing research in the following ways: no previous study on exploratory search examines how users combine queries and why they perform QCB. We summarize a hierarchy of this kind of search behavior and employ log analysis to provide new insight into combination success. We characterize differences in the search behavior associated with different types of QCB in the hierarchy and develop a classifier based on these differences to automatically classify QCB for a given search log.

3. A hierarchy of Query Combination Behavior

Different types of QCB were observed in user search sessions at the beginning of this work. To illustrate the differences between QCB types, let us review the examples in Figure 1. In Figure 1a, the user combined queries to find “pitbull” in the topic “dogs” where “pitbull” is a subtopic of “dogs” (pitbull is a breed of dog). Compared with the example in Figure 1b, however, there is no subtopic relevance between the query “everything I want the world” and the query “a pleasing sense of happiness”.

To explore the space of QCB types, we begin in this section by defining QCB in search sessions for the purposes of this study. We then describe the process by which we gradually obtained the hierarchy of QCB. Based on the QCB hierarchy, we manually classified a variety of QCB sessions taken from AOL search logs and Sogou search logs. Through the judgment of success for QCB sessions, we propose the role QCB plays in exploratory search processes.

3.1. Definitions

To explore the space of QCB types, we begin in this section by defining QCB in search sessions. Through our definitions, one would be able to identify QCB sessions for any given search session. We apply the following definitions for this purpose:

Definition: Query combination behavior (QCB)—a type of search behavior where users combine two or more issued queries into new queries during the information retrieval process.

Definition: Combined Queries—denoted as CBQ, that is, the newly created queries during QCB. For example, the query “pitbull dogs” in Figure 1a and the query “everything I want the world a pleasing sense of happiness for me” are both CBQ.

Definition: Component Queries—denoted as CPQ, are the queries which consist of CBQ during QCB. For example, the query “pitbull” and the query “dogs” in Figure 1a are both the CPQ of the CBQ “pitbull dogs”.

Definition: Query combination behavior sessions—those search sessions where QCB is performed. For example, the search session in Figure 1a,b.

3.2. A Hierarchy of QCB

To explore the space of QCB types, in this section we develop a hierarchy and categorization scheme for this behavior. For this purpose, we first brainstormed and listed candidates of QCB type as extensively as possible with our experiences. The list served as a prototype of our hierarchy. We then gradually revised the prototype. Five exploratory search researchers took part in this task. Researchers were asked to collaboratively judge 100 QCB sessions by examining the queries, the search results pages, and the clicked-on

web pages. These QCB sessions were identified and randomly picked from either AOL search logs [30] or Sogou search logs [31].

Categories were revised, or new categories were added, if a QCB session fell out of the prototype hierarchy. For example, the previous QCB type “accurate description” was removed as unrepresentative (almost all QCB will fall into this category, providing evidence that this is a common feature of QCB).

If categories were modified or new categories were added, another 100 QCB sessions were presented for judging. The judge–revise–judge cycle was repeated until all 100 randomly picked QCB sessions were classified into our hierarchy. As the prototype hierarchy of QCB types was gradually modified, some QCB types were split into sub-types, and finally a two-layer hierarchy of QCB naturally emerged. Figure 2 shows the hierarchy.

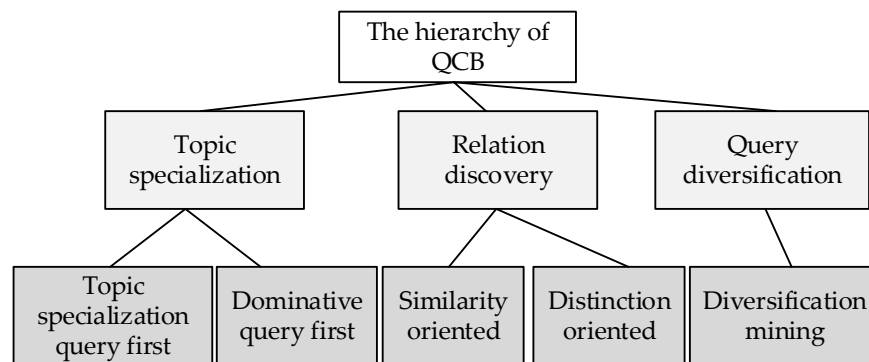


Figure 2. The hierarchy of QCB.

The first layer in Figure 2 is the category layer, which contains three QCB categories: the topic specialization category, the relation discovery category, and the query diversification category. The second layer is the pattern layer. For example, the topic specialization category can be divided into topic specialization query first pattern and dominative query first pattern. To better understand the space of QCB types, we illustrate the hierarchy in the remainder of this section. Note that, to better illustrate the examples of QCB, we omit queries other than CPQ and CBQ in the following examples.

(1) Topic specialization

The topic specialization category of QCB demonstrates a requirement by the user to limit the search results for CPQ into specific topics, represented by other CPQs. CPQs in this category can be divided into two types:

Dominative queries are a particular type of CPQ, the search results for which cover the objects that users are seeking.

Topic specialization queries are also a particular type of CPQ, the search results for which correspond to the topic that the user has in mind.

To be considered as being under the topic specialization category of QCB, some CPQs must correspond to the search object that the user already has in mind, while the others correspond to the topic of the objects which he or she wants. We found that search results for the dominative query are usually across different topics. Users exhibited this category of QCB after they learned that their information needs could be met by synthesizing topics and subjects.

Two patterns naturally emerged according to whether the time sequence, dominative query, or topic specialization query came first. We found that these patterns correspond to two types of search goals: exploring the topics of their search objects or exploring the search objects in their topics. We compared the patterns in this QCB category as follows:

- Dominative query first: This pattern describes users’ issue dominative queries in advance of the topic specialization queries. Users who exhibited this pattern of QCB were exploring the topics of their information needs. For example, the QCB session

such as <“online games”, “pinball”, “pinball online games”>, where the user was exploring to find online games, is judged as a dominative query first pattern. In this session, the CPQ “online games” contains the subject that the user wants in mind, which is considered as a dominative query, and another CPQ, “pinball”, contains the topic information, which expresses that the type of online game he or she searched for was “pinball”.

- Topic specialization query first: This pattern describes users’ issue dominative queries in advance of the topic specialization queries. Users who exhibited this pattern of QCB were exploring search objects in topics. The most common example of this type of QCB pattern was the search session in Figure 1a. In this session, the CPQ “dogs” was a topic specialization query, and the component query “pitbull” was a dominative query. The user’s information need was to find the information of “pitbull” in topic “dogs”.

(2) Relation discovery

QCB which falls into this category demonstrates a desire by the user to find the potential relations between CPQs. We found that users exhibit this category of QCB after analyzing the search results yielded by CPQs and learning the relations between CPQs, having discovered the objects that they were seeking. We also found that CPQs in these QCB sessions are semantically similar. According to the kinds of relations between CPQs, two patterns emerged:

- Similarity-oriented pattern: The search objects of users who exhibit this pattern are seeking similarities between CPQs. The commonest example of this pattern is the QCB in Figure 1b. As we discussed above, the user was exploring to find musical resources about the song “Top of the world”, and the similarity relation between the two CPQs is the name of the song—this is evidence that the user was exploring to find the similarities of the CPQs by combining the CPQs.
- Distinction-oriented pattern: The search objects of users who exhibit this pattern are seeking distinctions between CPQs. For example, a search session including <“estrogen”, “progestin”, “estrogen and progestin”>, where the user was exploring to find distinctions between “estrogen” and “progestin”, would be included in this pattern. In this session, neither the CPQ “estrogen” nor the CPQ “progestin” was the search object that the user had in mind. In fact, what the user really wanted was to explore the distinctions between the two CPQs.

(3) Query diversification

QCB which falls into this category demonstrates a desire by the user to discover potential information which could not be found by issuing CBQ. Note that users who perform this QCB category are not exploring relationships between CPQs but discovering the new meaning of the CBQ. Therefore, CBQ in this pattern usually contains meanings that cannot be found in the search results for each CPQ. There is only one pattern in this category of QCB, as follows:

- Diversification mining: The most common example of this type of QCB pattern may be the search session <“black”, “sheep”, “black sheep”>. The combined query “black sheep” contains a new meaning, which is not present in “black” and “sheep”.

3.3. Manual QCB Classification

To better understand QCB, in this section we develop a system to manually classify QCB sessions. Once this was achieved, we were able to characterize features to build models which can automatically classify QCB types.

In this paper, we extracted QCB sessions from the AOL and the Sogou search log data sets. We chose these two data sets for the following four main considerations: firstly, search log data are extremely private, so collecting search logs may raise privacy issues, thus leading to few publicly available query log data sets. Secondly, the query logs we can collect is limited; however, both AOL and Sogou data sets have sufficient search logs

available for analyses and for training machine learning algorithms (the AOL search log data set is released by America Online with search logs of 650,000 real users [32], and the Sogou search log data set—founded by Chinese commercial search engine Sogou—consists of real query logs of a given month [33]). Third, search logs in these two data sets are collected from real information needs, which make them more able to reflect the user's real exploratory process [32]. Fourth, both AOL and Sogou data are publicly available, so that our research findings could be reproduced. These factors also make AOL and Sogou data sets the most used data sets in information retrieval.

For each search log, we used 30 min for identifying time gap sessions of our query log [34]. To find sessions in which users may show exploratory behavior, we performed the following:

- Filter short sessions: We first filtered sessions with less than three queries since these cannot be exploratory search sessions [35,36]. Similar criteria have been used in previous work, e.g., [34].
- Filter no click sessions: We also removed search sessions with no click, since these cannot reflect user behavioral characteristics.
- Normalize plural nouns in sessions: We replaced the plural nouns in QCB sessions with their corresponding singular nouns.

We applied these criteria to find sessions in which users are likely to show exploring [34]. We found that more than 5% of the search sessions show QCB (145,783 QCB sessions in 2,667,714 AOL search sessions and 78,927 QCB sessions in 1,511,916 Sogou search sessions), which illustrates QCB as a common exploratory search behavior. In addition, 98% of these QCB sessions (219,227 QCB sessions) have two CPQs. Due to the large scale of the data set, we sampled 6000 QCB sessions (4000 sessions are from the AOL data set and the others are from the Sogou data set). We believe that in many cases, QCB types can be deduced from looking at user behaviors which are available to the search engine [37]. Therefore, judges classified these QCB sessions based on behaviors, such as: (1) queries; (2) search results; (3) clicked-on web pages; (4) dwell times; (5) further searches after CBQ.

Previous work on behavioral analysis has shown that dwell times are associated with the queries, which can yield satisfying information for searchers [38,39]. For the issued query in a search session, the time stamps of this query and the following query can represent the dwell time of the query. For instance, at time t_0 , the user issues a query, and at time t_1 , the user issues the next query. Given the two timestamps, we calculate the dwell time of the query as $t_1 - t_0$. Similar criteria have been used in previous work to demarcate search sessions, e.g., [7].

For the experiments, we recruited 12 participants affiliated with the Northeastern University campus. We divided these participants into three groups, and the judges in each group were asked to work together on the manual classification of 2000 QCB sessions.

Classification results showed that 5787 QCB sessions (3876 QCB sessions in the AOL data set and 1911 in the Sogou data set) could be manually classified based on our hierarchy, and 213 QCB sessions (124 QCB sessions in the AOL data set and 89 QCB sessions in the Sogou data set) could not be classified because of insufficient behavior evidence, for example, if all the clicked-on web pages in a QCB session were invalid.

Figure 3 shows the category and pattern distribution in QCB sessions; we found that more than 54% of the QCB sessions fell into topic specialization QCB category (43% of them fell into the topic specialization query first pattern and 57% of them fell into the dominative query first pattern), 36% of the QCB sessions fell into the relation discovery category, and only 10% of the QCB sessions fell into the query diversification category.

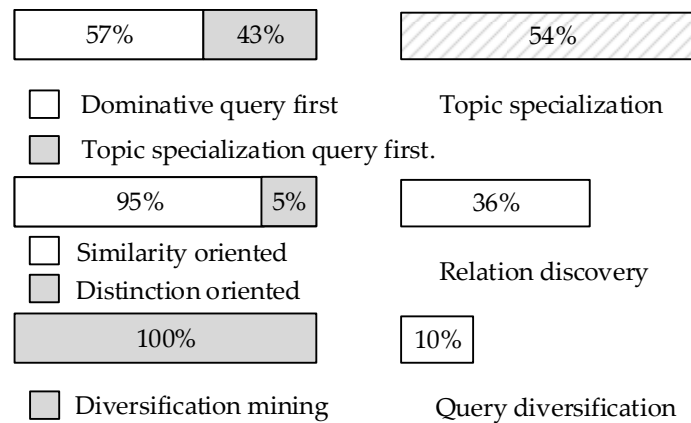


Figure 3. Category and pattern distribution in QCB sessions.

3.4. Successful vs. Unsuccessful QCB

We now turn our attention to understanding the role QCB play in the exploratory search process, to know, for example, whether users can find their required information by performing QCB. This is important because methods should be designed to help users combine queries if most users benefit from QCB. Conversely, if QCB blocks users’ exploratory processes, search engines should avoid users combining queries. For this purpose, we began by identifying the success of QCB, as follows:

- Successful: QCB sessions where users can find their required information by issuing CBQs.
- Unsuccessful: QCB sessions where searchers failed to find the required information by issuing CBQs.

These criteria are inspired by Ahmed Hassan’s study [26]. After labeling, participants were asked to answer the following question, if a QCB was judged as “successful”:

- Do you think CPQ should be issued if the user issues CBQ in advance? Select “Yes” or “No”.

In addition, participants were asked to answer the following question, if a QCB was judged as “unsuccessful”:

- Why do you think the user failed to succeed by issuing CBQ? Describe what you think made this QCB fail in one sentence.

The distribution of “successful” and “unsuccessful” searches is shown in Table 1. The table shows that most of the QCB were successful (more than 86%). We also found that 43% of users completed their search processes after issuing CBQs. In addition, most QCB were “successful” for each of the 3 QCB categories (the lowest was more than 75% for the query diversification category). These findings suggest that users can find their required information by exhibiting QCB.

Table 1. Distribution of “successful” and “unsuccessful” searches.

	Total	Topic Specialization	Relation Discovery	Query Diversification
Successful	86%	88%	86%	75%
Unsuccessful	14%	12%	14%	25%

Figure 4 shows the distribution of the judgments of the question when QCB was judged as “successful”. It is observed from Figure 4 that most (more than 88%) users can find their required information by issuing CBQ without issuing CPQ, providing strong

evidence that prompting users to combine queries can help them to complete their search task quickly.

In addition, Table 2 summarizes the top three mentioned reasons which make QCB “unsuccessful”. About 54% of QCB were judged as “unsuccessful” because users explored the wrong search objects in the topic. This highlights that accurate identification of user information requirements is helpful for these users. We also notice that about 11% of “unsuccessful” searches were judged as such because users issued invalid combinations; this is where users combine queries for unknown reasons, for example, the QCB session <“fire”, “flame”, “fire flame”>, with all kinds of unrelated clicks. We suspect that these users may randomly combine queries together, or forget to delete the previously issued terms when reissuing a query.

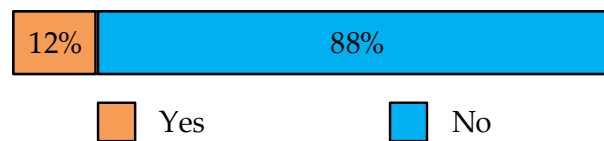


Figure 4. Distribution of “Yes” and “No” in answering the question “Do you think CPQ should be issued if the user issues CBQ in advance?”.

Table 2. Top three mentioned reasons that make QCB fail.

Reason		
Top 1	Explored wrong search object in topic.	54%
Top 2	The search results of CBQ and CPQ are same.	17%
Top 3	Invalid combinations.	11%

4. Characterizing Query Combination Behavior

In this section, we analyze some key characteristics to support automatic classification for QCB sessions, such as query characteristics, click characteristics, and content characteristics.

4.1. Query Characteristics

The most obvious factors to distinguish between QCB types may be the characteristics of queries. Therefore, we examined several different aspects of queries in QCB sessions, as follows:

Average Semantic Distance of Queries: Semantic distance between CPQs may be a factor distinguishing the “relation discovery category” from all other QCB categories. To examine this, we measured the average similarity distance between CPQs in each QCB session for each QCB type.

We measure semantic distance between queries based on CLSM (convolutional latent semantic model), presented by Shen and He [40,41], which can generate a continuous semantic vector representation for each query string. The semantic distance between any two CPQs is calculated as the cosine similarity between their vectors in that semantic space, which is computed as Equation (1), as follows:

$$Sim(CPQ_i, CPQ_j) = \cos \angle(v_i, v_j) = \frac{v_i^T v_j}{\|v_i\| \|v_j\|} \quad (1)$$

v_i and v_j are the semantic vectors of CPQ_i and CPQ_j . A semantic distance value of one indicates QCB with perfect same semantics. On the contrary, a semantic distance value of zero indicates that the semantic between QCB is completely unrelated.

Figure 5 shows the average semantic distances between all CPQ in every QCB session for each QCB pattern. We notice from the figure that the average semantic distance of CPQ seemed to be a factor to distinguish between QCB patterns except for distinguishing between dominative query first pattern and topic specialization query first pattern. This can

be explained by the analysis, where we discussed that the difference between dominative query first pattern and topic specialization query first pattern is the time sequence between dominative queries and topic specialization queries. The differences (other than the difference between dominative query first pattern and topic specialization query first pattern) between QCB patterns were statically significant at the 0.05 level, according to the two-tailed t-test and variance analyses.

Subtask Belonging to Queries: In addition, we also considered the subtask belonging to the queries. Exploratory search tasks often tend to have multiple subtasks associated with them [42,43]. As we discussed, a given CPQ in a QCB session which was judged into the topic specialization category may be more likely to belong to the same subtask.

To examine this, we considered the Bayesian rose tree (BRT) model, proposed by Mehrotra and Yilmaz, which has been proved to be one of the most advantageous techniques for searching subtask clustering [44]. Figure 6 shows the distribution of subtask belonging for each QCB pattern. A value of one indicates that all CPQs in every QCB session fell into the same search subtasks, while a value of zero indicates that each CPQ in every QCB session fell into different search subtasks.

We notice from the figure that CPQs in the topic specialization QCB sessions are more likely in the same subtasks (more than 82%), which aligns with our previous observation. In addition, the CPQ in the relation discovery category is more likely in the same subtasks than that in the query diversification category.

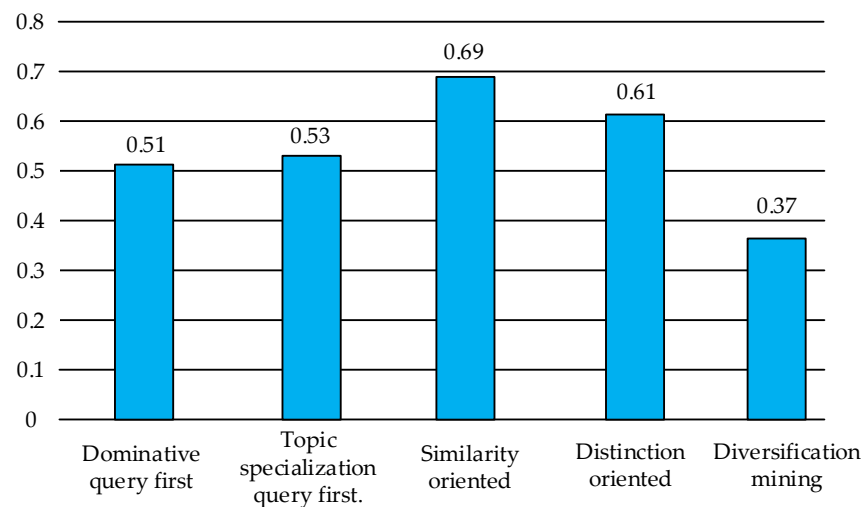


Figure 5. Average semantic distances between all CPQs in every QCB session for each QCB pattern.

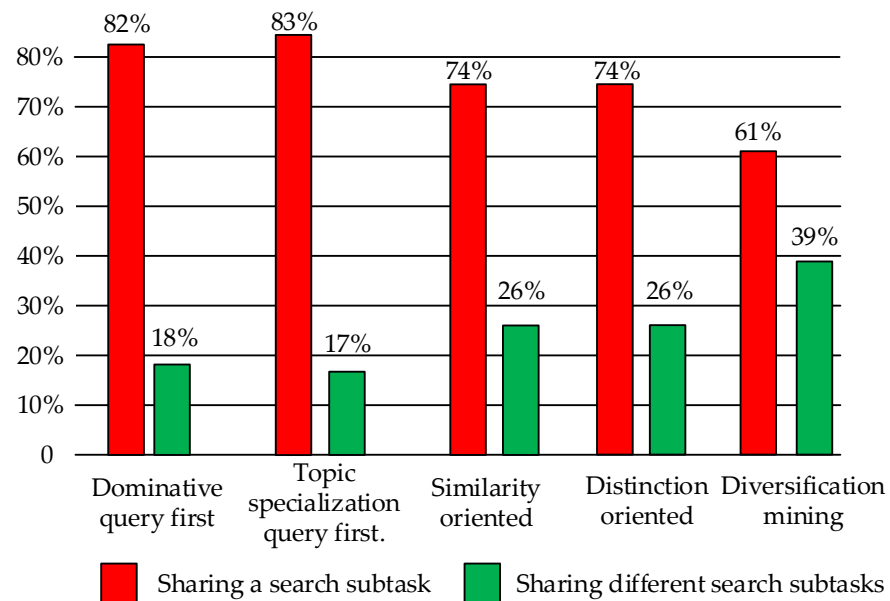


Figure 6. Distribution of subtask belonging for each QCB pattern.

The differences between patterns in each QCB category were not statistically significant at the 0.05 level according to the two-tailed *t*-test and the variance analysis, suggesting that this was unlikely to be a distinguishing factor. However, the differences between categories were statistically significant at the 0.05 level according to a two-tailed *t*-test and the variance analysis, suggesting that this characteristic could be able to distinguish between QCB categories.

4.2. Click Characteristics

Users who performed different QCB patterns were likely to have different exploratory features, and clicks could provide additional information about users' exploration processes. Therefore, click characteristics may be another factor to distinguish between QCB patterns. We examined this as follows:

Number of Clicks: A large number of clicks after issuing a query may indicate that the exploratory process is complex, or the returned results cannot meet the user's information needs. We considered that these features may also contribute to distinguishing QCB patterns. We examined this by computing three aspects of the average number of clicks, as follows:

- The average number of clicks after issuing a CPQ for each QCB pattern.
- The average number of clicks after issuing a CBQ for each QCB pattern.
- The average number of clicks after issuing the other queries in every QCB session for each pattern.

Figure 7 shows the distribution of the three aspects. While QCB patterns in the same category show similar distributions, we only display the distribution for QCB categories. It is interesting to note that the average numbers of clicks after CPQ in each QCB category show significant differences. The average number of clicks after CPQ in the topic specialization category is obviously higher than those in the other categories. This aligns with our previous work, where we discussed that the search results returned by both dominative queries and topic specialization queries contained all the users' information needs. Users who performed this QCB pattern have to explore more information from the search results for the CPQ. On the contrary, CPQs in the other QCB pattern will not be the users' search objects. The average number of clicks after a CPQ in the relation discovery category

is clearly less than those in the other categories. The differences between an average number of clicks after a given CPQ are statistically significant at the 0.05 level, according to a two-tailed t-test, suggesting that this is likely to be a distinguishing factor for QCB categories.

We also notice from the figure that the average number of clicks after issuing CBQ in the “query diversification category” of QCB sessions is significantly less than those in the other category of QCB sessions. This may be because users who performed this QCB category were less likely to be judged as “successful” in their searches. The differences between an average number of clicks after a given CBQ are statistically significant at the 0.05 level according to a two-tailed t-test, suggesting that this is likely to be a distinguishing factor of query diversification QCB sessions.

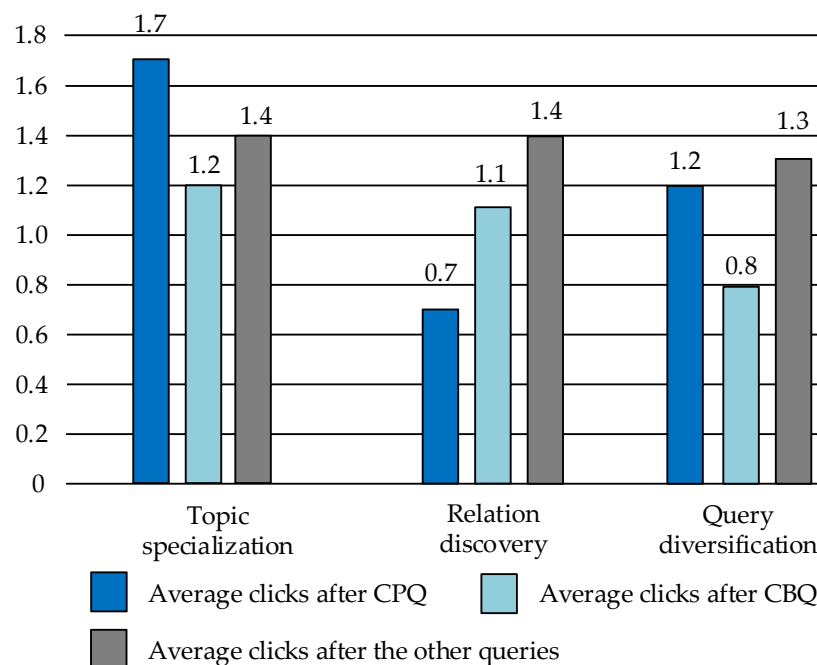


Figure 7. Distribution of three aspects of the average number of clicks in QCB sessions for each QCB category.

4.3. Content Characteristics

During the information retrieval process, characterizing search outcomes is the most direct way to access users [10]. According to Vakkari [13], the search outcome, typically text content, is an indicator of learning. In this paper, we conceptualize content characteristics as learning characteristics where we could identify the learning characteristics of QCB users during the information retrieval process.

An increasing interest in the outcome of searching has led to a better understanding of the structures of learning from the outcome. A typical outcome analyzed is a web page text, based on search results. Compared with click characteristics, content characteristics (such as web pages, search engine results pages) might provide more implicit features [45]. Therefore, we further analyzed the contents to mine users’ learning characteristics.

Learn from Contents: In the process of manual classification, we observed that some CBQ can be directly found on search engine results pages returned by issuing CPQs or on web pages clicked after issuing CPQ. We apply the following definition for this learning behavior:

Definition: Content Inspire Behaviors Sessions (CIB) are QCB sessions where the CBQ can be directly found on search engine results pages returned by issuing CPQ or on web pages clicked after issuing CPQ.

CIB is a typical searching as learning behavior, where users learned CBQ in contents of search outcomes. In QCB sessions that performed CIB, CBQs are a learning outcome for QCB search sessions that performed CIB. Therefore, we can mine users’ learning characteristics by analyzing CIB.

Figure 8 shows the distribution of CIB vs. no-CIB for each QCB pattern. We notice from the figure that patterns in topic specialization QCB category show similar distribution (show more than 71% of the CIB). It is also interesting that patterns in relation to the discovery of QCB category show opposite distribution. Distinction-oriented pattern QCB sessions have significantly less likely to show CIB. This might be because the difference between CPQs is latent. All differences, except the difference between patterns in the topic specialization category, reported in Figure 8 are statically significant at the 0.05 level, according to two-tailed t-test and the variance analysis.

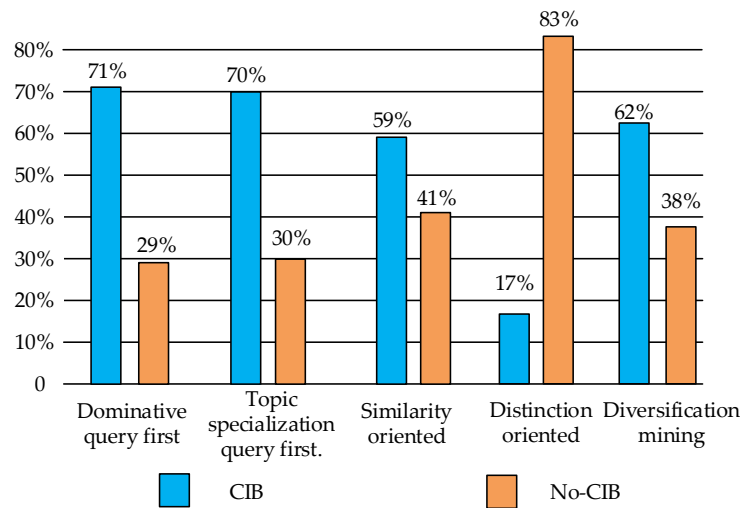


Figure 8. Distribution of CIB vs. no-CIB for each QCB pattern.

5. Automatic QCB Classification

We successfully completed the task of manual classification for QCB sessions and have characterized a series of key aspects of QCB sessions that are significant in distinguishing between QCB patterns. In this section, we try to utilize them as features to build models for automatic QCB pattern classification.

5.1. Features

In addition to the features mentioned, we found some features in existing works which may correlate with our study for classification [46–50], for example, topic features, which have been provided to be a factor to differ from search behaviors [51]. In addition, we also utilize some common features in exploratory search studies [52], such as session length (the number of web search engine queries issued in a session) [53]. A summary of the features used for automatically classifying QCB sessions is shown in Table 3.

Table 3. Features used for automatic QCB classification.

QCB Features	
Name	Description
Query Features	
Number of queries	Number of queries in QCB session.
Semantic similarity of queries	Average semantic similarity of queries in QCB session.
Subtask belonging of queries	Subtask belonging of CPQ in QCB session.

Distance between queries	Number of queries between the first CPQ and the last CPQ in QCB session.
End of session	Whether the CBQ is the last query in QCB session.
Time between queries	Average time between CPQ.
Similarity queries	(1) Average Jaccard coeff between the term sets of CPQ; (2) Average Jaccard coeff between the term sets of CBQ and CBQ.
Edit distance between queries	Average norm edit distance between query strings.
Inter-query time	(1) Average inter-query time per CPQ; (2) inter-query time of CBQ.
Query length	Query length in number of terms.
Embedding distance *	Average cosine distance between embedding vectors of the CBQ and CPQ.
Click Features	
Number of clicks	(1) Average number of clicks per CPQ;
	(2) Number of clicks after CBQ;
	(3) Number of clicks in QCB session;
	(4) Number of CPQ with no click;
	(5) Number of CBQ with no click.
Number of repeating clicks	Same click after CBQ and any CPQ.
Content Features	
Learn from contents	Number of combined queries which can be directly found on clicked on web pages or on SERP.
Similarity of SERP	(1) Average SERP Similarity returned by CPQ;
	(2) Average SERP Similarity between CBQ and each CPQ.
Content semantic similarity	(1) Average content semantic similarity of web pages per clicks after CPQ;
	(2) Average content semantic similarity of web pages between CPQ;
	(3) Average content semantic similarity of web pages between CPQ and CBQ.
Other Features	
Total dwell time	Total dwell time in QCB session.
Topic *	Binary variable for every visited URL topic.
Topic Entropy *	Topic distribution entropy.

“*” means these features are inspired by previous studies on exploratory search behaviors.

5.2. Classification for QCB Types

After identifying the features, we first described an automatic QCB classification model, and then we described our experimental setup and the results for the classification experiments. We conducted experiments using the QCB sessions described, which had approximately 6000 QCB sessions and 23,871 queries. These QCB sessions were manually classified and labeled. To ensure the validity of features in each QCB session, we began by performing standard text normalization where we removed leading spaces, trailing spaces, and stop words.

We used 10-fold cross-validation for all experiments and we trained our classifier to classify QCB based on the gradient boosting tree (GBT) model [54,55]. We adopted the GBT model because it can handle heterogeneous features and has a good prediction power [48]. We trained each QCB pattern as one class, and the GBT algorithm is summarized in Algorithm 1.

Algorithm 1 Gradient boosting tree (GBT)

- 1: Initialize $F_0(x) = \arg \min_a \sum_{i=1}^N L(y_i, a)$
- 2: **for** $m=1,2,\dots,M$ **do**
- 3: **for** $i=1,2,\dots,N$ **do**
- 4: $\tilde{y}_{im} = -\left[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)}\right]_{F(x)=F_{m-1}(x)}$

```

5:   end for
6:    $a_m = \arg \min_a \sum_{i=1}^N (\hat{y}_{im} - T(x_i, a))^2$ 
7:    $\mu_m = \arg \min_{\mu} \sum_{i=1}^N L(y_i, F_{m-1}(x_i) + \mu T(x_i, a))$ 
8:    $F_m(x) = F_{m-1}(x) + \sum_{m=1}^M \mu_m T(S, a_m)$ 
9: end for
10:  $F(x) = \sum_{m=1}^M \mu_m T(S, a_m)$ 

```

The input of Algorithm 1 is the training data set $S = \{(x_1, y_1), (x_2, y_1), \dots, (x_N, y_N)\}$, and the sample size is N . $x_i \in x \subset R^n$, where x is the feature set we proposed in Table 3. $y_i \in y \subset R$, where y is the QCB pattern set. The output is the optimal mapping function $F(x)$, which minimizes the loss function $L(y, F(x))$ by ensemble weak learners (CATR decision tree).

The GBT model combines M individual CATR decision trees $T(S, a_1), T(S, a_2), \dots, T(S, a_M)$, where $T(S, a_i)$ is the i th CATR decision tree. Therefore, the function $F(x)$ can be expressed as Equation (2), as follows:

$$F(x) = \sum_{m=1}^M \mu_m T(S, a_m) \tag{2}$$

For the m th CATR decision tree $T(S, a_m)$, μ_m is the weight for this classifier, and a_m determines the structure of the tree. The parameters μ_m and a_m are determined iteratively, which can ensure that the loss function $L(y, F(x))$ is minimized. μ_m and a_m can be determined as Equation (3), as follows:

$$(\mu_m, a_m) = \arg \min_{\mu, a} \sum_{i=1}^N L(y_i, F_{m-1}(x_i) + \mu T(x_i, a)) \tag{3}$$

The GBT algorithm initialized $F_0(x) = \arg \min_a \sum_{i=1}^N L(y_i, a)$ at the beginning of Algorithm 1; however, it is difficult to directly solve μ_m and a_m . Therefore, we estimate μ_m and a_m with a simple two-step procedure [56,57]. The GBT algorithm first estimates a_m as Equation (4), as follows:

$$a_m = \arg \min_a \sum_{i=1}^N (\hat{y}_{im} - T(x_i, a))^2 \tag{4}$$

After the a_m estimated, the output of the CATR decision tree is constant, so that the weight μ_m can be straightforwardly estimated using a line search on the loss function [58,59].

To study the performances of our approach, we also experimented with other classifiers, such as multiple additive regression trees (MART) [50], SVM, and C4.5 [60].

For every classifier, we use the area under the curve (AUC) of ROC, accuracy, and the F1 score to evaluate these models' performance. Table 4 presents the results of the QCB pattern classification experiment. Experimental results show that the GBT model yielded better performances than the other baseline classifiers. From the results, we can see that, using the features, the gradient boosting tree can effectively classify QCB sessions. The AUC reaches 79.23 in the QCB classification.

Table 4. Performance of QCB sessions classification.

	AUC	Accuracy	F1
Gradient Boosting Tree	79.23 *	77.19 *	78.81 *
Logistic Regression	72.27	70.05	72.41
SVM	73.42	72.23	73.37
C4.5	72.84	71.61	72.27

* indicates statistical significance at $p < 0.05$ using paired t-tests compared to baseline.

Table 5 shows the classification power of GBT for each QCB pattern. The results for the classification experiments demonstrate that GBT shows more performances in the classification of patterns in the topic specialization category and relation discovery category. Indeed, as can be seen from Table 5, the relative proportions of AUC in the diversification mining of the QCB pattern are much less than the other patterns.

Table 5. Performance of QCB sessions' classification.

	AUC	Accuracy	F1
Dominative query first	81.21	77.47	78.94
Topic specialization query first	81.34	77.21	70.11
Similarity-oriented pattern	82.15	77.91	79.59
Distinction-oriented pattern	79.89	77.14	78.21
Diversification mining	77.67	74.38	76.02

6. Discussion and Conclusions

Exploratory search users have to seek with more effort to meet their information needs [61]. Therefore, exploratory search studies aim to help users find their required information with minimal efforts [62].

QCB is a special type of query reformulation behavior. Query reformulation studies focus on studying the users who modify their initial or previous queries, based on the judgment of search results. Previous work mainly focuses on (i) exploring the inspiration of search results to users [29]; (ii) exploring the regularity of query reformulation in historical data [63]. However, previous studies neglected to explore the relationship between the issued queries and the reformulate queries. We believe that exploring these relationships is helpful for understanding how users synthesize information. Our work is differentiated from previous studies because we successfully explored a typical kind of relationship where users combine issued queries into a new query; furthermore, we found the regular patterns of how users synthesize information.

Our first finding was that we introduced a kind of novel exploratory search behavior where users combine two or more previously issued queries into new queries. On both the AOL and the Sogou data sets, we observed that users can easily find their required information by combining queries where they analyzed and synthesized the information they have learned. However, to the best of our knowledge, no studies have focused on this typical exploratory search behavior. Better understanding of this kind of exploratory search behavior is important for exploring how users search and learn during information retrieval processes and can further help people find information more easily.

We list our main contributes as follows: (1) through analysis on millions of search logs, we have explored the space of QCB types and have described a two-layer hierarchical structure to better understand the nature of QCB; (2) through manual classification of QCB sessions, we have proposed that users can find their required information easily by performing QCB, and we further found that prompting users to perform QCB can also help them find their required information easily; (3) we characterized aspects of how users combine queries, where we found some statically significant differences to support the manual classification of QCB patterns; (4) based on these characteristics, we developed a classifier to accurately predict QCB types, with a view to helping systems understand what web searchers really want.

Our research supports further analysis of different QCB patterns. Ideally, a search engine would interpret the behavioral signals that indicate QCB, accurately classify QCB types, and provide personalized help to searchers to help them attain task success. For example, through the following possibilities:

Enhance understanding of the constructs of learning in searching: Soo Young Rieh and Kevyn Collins-Thompson [10] proposed that providing users with a high-quality re-

sult list is not enough—it is important to distinguish different kinds of constructs of learning in searching. As we demonstrated, users perform QCB after learning search outcomes, analyzing search results, and bringing pieces of information together. Different search patterns can reflect users' different learning characteristics. Therefore, different QCB categories and patterns can reflect different kinds of constructs of learning in searching and can help search engines to provide more personalized learning strategies.

Enhance query recommendation strategies: As we demonstrated, users who explore the topics of their information needs, or who explore their information needs within topics, are highly likely to perform QCB. Learning the relationship between users' search goals and topics allows search systems to generate hints of queries to prompt QCB (e.g., add a topic term of pivotal queries), and propose them in real-time, as people are searching.

Overall, a better understanding QCB and accurate prediction of QCB patterns can help searchers and search engines reduce searching efforts. Although our findings are promising, there are still some limitations of this study. In future work, we want to expand on several ideas touched upon by this work. (i) Firstly, we must enhance our classifier—human-labeled data limited our prediction capability; therefore, we plan to explore methods that use unlabeled data to improve our classification efficiency and to explore high-level features by deep learning methods, developing a more accurate classifier for QCB patterns. (ii) Secondly, we plan to develop a further understanding of when users have needed to combine queries, and further develop QCB pattern-oriented recommendation systems that support exploratory search users to synthesize information and meet their information needs quickly.

Author Contributions: Conceptualization, P.L., Y.Z., and B.Z.; methodology, P.L., Y.Z., and B.Z.; software, P.L. and Y.Z.; validation, P.L., and Y.Z.; formal analysis, P.L., Y.Z., and B.Z.; investigation P.L.; resources, P.L., Y.Z., and B.Z.; investigation, P.L. and Y.Z.; resources, P.L. and Y.Z.; data curation, P.L., Y.Z., and B.Z.; writing—original draft preparation, P.L. and Y.Z.; writing—review and editing, P.L. ; supervision, P.L., Y.Z., and B.Z.; project administration, P.L., Y.Z., and B.Z.; funding acquisition, Y.Z.. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the the National Natural Science Foundation of China (Foundation No. 61572116, 61572117, and 61502089).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Carevic, Z.; Lusky, M.; van Hoek, W.; Mayr, P. Investigating exploratory search activities based on the stratagem level in digital libraries. *Int. J. Digit. Libr.* **2018**, *19*, 231–251.
2. Han, S.; Yue, Z.; He, D. Understanding and Supporting Cross-Device Web Search for Exploratory Tasks with Mobile Touch Interactions. *ACM Trans. Inf. Syst.* **2015**, *33*, 1–34.
3. Medlar, A.; Li, J.; Głowacka, D. Query Suggestions as Summarization in Exploratory Search. In Proceedings of the 2021 Conference on Human Information Interaction and Retrieval (CHIIR '21), Canberra, Australia, 14–19 March 2021; Association for Computing Machinery: New York, NY, USA, 2021; pp. 119–128. <https://doi.org/10.1145/3406522.3446020>.
4. Kropotov, I.; Medlar, A.; Glowacka, D. Exploratory Search of GANs with Contextual Bandits. In Proceedings of the 30th ACM International Conference on Information & Knowledge Management (CIKM '21), Gold Coast, QLD, Australia, 1–5 November 2021; Association for Computing Machinery: New York, NY, USA, 2021; pp. 3157–3161. <https://doi.org/10.1145/3459637.3482103>.
5. Crescenzi, A.; Li, Y.; Zhang, Y.; Capra, R. Towards Better Support for Exploratory Search through an Investigation of Notes-to-self and Notes-to-share. In Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '19), Paris, France, 21–25 July 2019, Association for Computing Machinery: New York, NY, USA, 2019; pp. 1093–1096. <https://doi.org/10.1145/3331184.3331309>.

6. Ward, A.R.; Capra, R. OrgBox: Supporting Cognitive and Metacognitive Activities during Exploratory Search. In Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '21), Montreal, QC, Canada, 11–15 July 2021; Association for Computing Machinery: New York, NY, USA, 2021; pp. 2570–2574. <https://doi.org/10.1145/3404835.3462790>.
7. Choi, D. A Study of Information Seeking Behavior Using Physical and Online Explorations. In Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval (SIGIR '16), Pisa, Italy, 17–21 July 2016; p. 1163.
8. Lissandrini, M.; Mottin, D.; Palpanas, T.; Velegrakis, Y. Example-based Search: A New Frontier for Exploratory Search. In Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '19), Paris, France, 21–25 July 2019; Association for Computing Machinery, New York, NY, USA, 2019; pp. 1411–1412. <https://doi.org/10.1145/3331184.3331387>.
9. Anarfi, R.; Kwapong, B.; Fletcher, K.K. Towards a Reinforcement Learning-based Exploratory Search for Mashup Tag Recommendation. In Proceedings of the 2021 IEEE International Conference on Smart Data Services (SMDs), Chicago, IL, USA, 5–10 September 2021; pp. 8–17. <https://doi.org/10.1109/SMDs53860.2021.00012>.
10. Rieh, S.Y.; Collins-Thompson, K.; Hansen, P.; Lee, H.J. Towards searching as a learning process. *J. Inf. Sci.* **2016**, *42*, 19–34.
11. Savolainen, R. Berrypicking and information foraging: Comparison of two theoretical frameworks for studying exploratory search. *J. Inf. Sci.* **2018**, *44*, 580–593.
12. Hansen, P.; Rieh, S.Y. Editorial: Recent advances on searching as learning: An introduction to the special issue. *J. Inf. Sci.* **2016**, *42*, 3–6.
13. Vakkari, P. Searching as learning: A systematization based on literature. *J. Inf. Sci.* **2016**, *42*, 7–18.
14. Howard, P.N.; Massanari, A. Learning to Search and Searching to Learn: Income, Education, and Experience Online. *J. Comput.-Mediat. Commun.* **2010**, *12*, 846–865.
15. Du, J.T.; Arif, A.S.M.; Hansen, P. Collaborative query reformulation in tourism information search. *Online Inf. Rev.* **2019**, *43*, 1115–1135.
16. Zhang, X. Improving personalised query reformulation with embeddings. *J. Inf. Sci.* **2021**. <https://doi.org/10.1177/0165551520968698> (accessed on 15 December, 2021).
17. Sankhavara, J.; Dave, R.; Dave, B.; Majumder, P. Query specific graph-based query reformulation using UMLS for clinical information access. *J. Biomed. Inform.* **2020**, *108*, 103493.
18. Crescenzi, A.; Ward, A.R.; Li, Y.; Capra, R. Supporting Metacognition during Exploratory Search with the OrgBox. In Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '21), Montreal, QC, Canada, 11–15 July 2021; Association for Computing Machinery: New York, NY, USA, 2021; pp. 1197–1207. <https://doi.org/10.1145/3404835.3462955>.
19. Ruotsalo, T.; Peltonen, J.; Eugster, M.J.; Głowacka, D.; Floréen, P.; Myllymäki, P.; Jacucci, G.; Kaski, S. Interactive Intent Modeling for Exploratory Search. *ACM Trans-Actions Inf. Syst.* **2018**, *36*, 44.1–44.46.
20. Huang, J.; Efthimiadis, E.N. Analyzing and evaluating query reformulation strategies in web search logs. In Proceedings of the 18th ACM conference on Information and Knowledge Management, Hong Kong, China, 2–6 November 2009; pp. 77–86.
21. Boldi, P.; Bonchi, F.; Castillo, C.; Vigna, S. Query reformulation mining: Models, patterns, and applications. *Inf. Retr.* **2011**, *14*, 257–289.
22. Bosung Kim, Hyewon Choi, Haeun Yu, and Youngjoong Ko. 2021. Query Reformulation for Descriptive Queries of Jargon Words Using a Knowledge Graph based on a Dictionary. In Proceedings of the 30th ACM International Conference on Information & Knowledge Management (CIKM '21), Gold Coast, QLD, Australia, 1–5 November 2021; Association for Computing Machinery: New York, NY, USA, 2021; pp. 854–862. <https://doi.org/10.1145/3459637.3482382>.
23. Wicaksono, A.F.; Moffat, A. Modeling search and session effectiveness. *Inf. Process. Manag.* **2021**, *58*, 102601.
24. Forestiero, A.; Mastroianni, C.; Spezzano, G. A Multi Agent Approach for the Construction of a Peer-to-Peer Information System in Grids. In Proceedings of the 2005 Conference on Self-Organization and Autonomic Informatics, Glasgow, UK, 16 May 2005.
25. Forestiero, A. Metaheuristic algorithm for anomaly detection in Internet of Things leveraging on a neural-driven multiagent system. *Knowl.-Based Syst.* **2021**, *228*, 107241. <https://doi.org/10.1016/j.knosys.2021.107241>.
26. Hassan, A.; Jones, R.; Klinkner, K.L. Beyond DCG: User behavior as a predictor of a successful search. In Proceedings of the third ACM International Conference on Web Search and Data Mining (WSDM '10), New York, NY, USA, 4–6 February 2010.
27. Chen, J.; Mao, J.; Liu, Y.; Zhang, F.; Zhang, M.; Ma, S. Towards a Better Understanding of Query Reformulation Behavior in Web Search. In Proceedings of the Web Conference 2021 (WWW '21), Ljubljana, Slovenia, 19–23 April 2021; Association for Computing Machinery: New York, NY, USA, 2021; pp. 743–755. <https://doi.org/10.1145/3442381.3450127>.
28. Jiang, J.Y.; Wang, W. RIN: Reformulation Inference Network for Context-Aware Query Suggestion. In Proceedings of the 27th ACM International Conference on Information and Knowledge Management (CIKM '18), Torino, Italy, 22–26 October 2018; Association for Computing Machinery, New York, NY, USA, 2018; pp. 197–206. <https://doi.org/10.1145/3269206.3271808>.
29. Sloan, M.; Yang, H.; Wang, J. A term-based methodology for query reformulation understanding. *Inf. Retr. J.* **2015**, *18*, 145–165.
30. Hong, Y.; Vaidya, J.; Lu, H.; Karras, P.; Goel, S. Collaborative Search Log Sanitization: Toward Differential Privacy and Boosted Utility. *IEEE Trans. Dependable Secur. Comput.* **2015**, *12*, 504–518.

31. Zhang, Z.; Yang, M.; Li, S.; Qi, H.; Song, C. Sogou Query Log Analysis: A Case Study for Collaborative Recommendation or Personalized IR. In Proceedings of the 2009 International Conference on Asian Language Processing, Singapore, 7–9 December 2009; IEEE Computer Society: Washington, DC, USA, 2009; pp. 304–307.
32. Chen, W.; Cai, F.; Chen, H.; de Rijke, M. Hierarchical neural query suggestion with an attention mechanism. *Inf. Processing Manag.* **2019**, *57*, 102040.
33. Yi, D.; Zhang, Y.; Wei, B. Query Subtopic Mining via Subtractive Initialization of Non-negative Sparse Latent Semantic Analysis. *J. Inf. Sci. Eng.* **2016**, *32*, 1161–1181.
34. Choi, D.; Matni, Z.; Shah, C. Switching sources: A study of people's exploratory search behavior on social media and the web. *Proc. Assoc. Inf. Sci. Technol.* **2016**, *52*, 1–10.
35. Backhausen, D.T. Adaptive IR for exploratory search support. In Proceedings of the Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '12), Portland, OR, USA, 12–16 August 2012.
36. Singer, G.; Norbistrath, U.; Vainikko, E.; Kikkas, H.; Lewandowski, D. Search-logger analyzing exploratory search tasks. In Proceedings of the 2011 ACM Symposium on Applied Computing (SIGIR '11), TaiChung, Taiwan, 21–24 March 2011.
37. Liu, J.; Wang, Y.; Mandal, S.; Shah, C. Exploring the immediate and short-term effects of peer advice and cognitive authority on Web search behavior. *Inf. Processing Manag.* **2019**, *56*, 1010–1025.
38. Athukorala, K.; Głowacka, D.; Jacucci, G.; Oulasvirta, A.; Vreeken, J. Is exploratory search different? A comparison of information search behavior for exploratory and lookup tasks. *J. Assoc. Inf. Sci. Technol.* **2016**, *67*, 2635–2651.
39. Seki, Y.; Yoshida, M. Analysis of User Dwell Time by Category in News Application. In Proceedings of the 2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI), Santiago, Chile, 3–6 December 2018.
40. Shen, Y.; He, X.; Gao, J.; Deng, L.; Mesnil, G. Learning Semantic Representations using Convolutional Neural Network for Web Search. In Proceedings of the 23rd International World Wide Web Conference (WWW '14), Seoul, Korea, 7–11 April 2014; pp. 373–374.
41. Shen, Y.; He, X.; Gao, J.; Deng, L.; Mesnil, G. A Latent Semantic Model with Convolutional-Pooling Structure for Information Retrieval. In Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management (CIKM '14), Shanghai, China, 3–7 November 2014; Association for Computing Machinery, New York, NY, USA, 2014; pp. 101–110. <https://doi.org/10.1145/2661829.2661935>.
42. Kulahcioglu, T.; Fradkin, D.; Palanivelu, S. Incorporating Task Analysis in the Design of a Tool for a Complex and Exploratory Search Task. In Proceedings of the Conference on Human Information Interaction and Retrieval, Oslo, Norway, 7–11 March 2017; pp. 373–376.
43. Sarrafzadeh, B. Supporting Exploratory Search Tasks through Alternative Representations of Information. UWSpace. 2020. Available online: <http://hdl.handle.net/10012/15854> (accessed on 15 October, 2021).
44. Mehrotra, R.; Yilmaz, E. Extracting Hierarchies of Search Tasks & Subtasks via a Bayesian Nonparametric Approach. In Proceedings of the The 40th International ACM SIGIR Conference on Research and Development in Information Retrieval, Tokyo, Japan, 7–11 August 2017.
45. Hassan, A.; White, R.W.; Dumais, S.T.; Wang, Y.M. Struggling or Exploring? Disambiguating Search Sessions. In Proceedings of the WSDM 2014: Seventh ACM International Conference on Web Search and Data Mining, New York, NY, USA, 24–28 February 2014.
46. Rose, D.E.; Levinson, D. Understanding User Goals in Web Search. In Proceedings of the WWW04: The 2004 World Wide Web Conference (in conjunction with ACM Conference on Electronic Commerce [EC'04]), New York, NY, USA, 17–20 May 2004; pp. 13–19.
47. Mao, J.; Liu, Y.; Luan, H.; Zhang, M.; Ma, S.; Luo, H.; Zhang, Y. Understanding and Predicting Usefulness Judgment in Web Search. In Proceedings of the SIGIR '17: The 40th International ACM SIGIR Conference on Research and Development in Information Retrieval Shinjuku, Tokyo, Japan, 7–11 August 2017; pp. 1169–1172.
48. Di Nunzio, G.M.; Faggioli, G. A Study of a Gain Based Approach for Query Aspects in Recall Oriented Tasks. *Appl. Sci.* **2021**, *11*, 9075. <https://doi.org/10.3390/app11199075>.
49. Hisada, S.; Murayama, T.; Tsubouchi, K.; Fujita, S.; Yada, S.; Wakamiya, S.; Aramaki, E. Surveillance of early stage COVID-19 clusters using search query logs and mobile device-based location information. *Sci. Rep.* **2020**, *10*, 18680.
50. Wang, L.; Zhang, J.; Chen, G.; Qiao, D. Identifying comparable entities with indirectly associative relations and word embeddings from web search logs. *Decis. Support Syst.* **2020**, *141*, 113465.
51. Qu, P.; Liu, C.; Lai, M. The effect of task type and topic familiarity on information search behaviors. In Proceedings of the IliX 2010: Information Interaction in Context Symposium, New Brunswick, NJ, USA, 18–21 August 2010.
52. Maxwell, D.; Azzopardi, L.; Järvelin, K.; Keskkustalo, H. Searching and Stopping: An Analysis of Stopping Rules and Strategies. In Proceedings of the CIKM '15: 24th ACM International Conference on Information and Knowledge Management, Melbourne, Australia, 13–18 October 2015.
53. White, R.W.; Dumais, S.T.; Teevan, J. Characterizing the Influence of Domain Expertise on Web Search Behavior. In Proceedings of the WSDM '09: Second ACM International Conference on Web Search and Web Data Mining, Barcelona, Spain, 9–12 February 2009.
54. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232.
55. Müller, F.; Schug, D.; Hallen, P.; Grahe, J.; Schulz, V. Gradient Tree Boosting-based Positioning Method for Monolithic Scintillator Crystals in Positron Emission Tomography. *IEEE Trans. Radiat. Plasma Med. Sci.* **2018**, *2*, 411–421.

56. Yamagishi, J.; Kawai, H.; Kobayashi, T. Phone duration modeling using gradient tree boosting. *Speech Commun.* **2008**, *50*, 405–415.
57. Liu, Z.; Zhang, M.; Liu, F.; Zhang, B. Fault Diagnosis on the Braking System of Heavy-haul Train based on Multi-dimensional Feature Fusion and GBDT Enhanced Classification. *IEEE Trans. Ind. Inform.* **2020**, *17*, 41–51.
58. Deng, S.; Wang, C.; Wang, M.; Sun, Z. A gradient boosting decision tree approach for insider trading identification: An empirical model evaluation of China stock market. *Appl. Soft Comput.* **2019**, *83*, 105652.
59. Sagi, O.; Rokach, L. Approximating XGBoost with an interpretable decision tree. *Inf. Sci.* **2021**, *572*, 522–542.
60. Ruggieri, S. Efficient C4.5 [classification algorithm]. *IEEE Trans. Knowl. Data Eng.* **2002**, *14*, 438–444.
61. Gao, H.; Jiang, T. A Review of Research Studies on Exploratory Search. *J. Libr. Sci. China* **2013**, *39*, 36–47.
62. Alan Medlar and Dorota Glowacka. 2017. Using Topic Models to Assess Document Relevance in Exploratory Search User Studies. In Proceedings of the 2017 Conference on Conference Human Information Interaction and Retrieval (CHIIR '17), Oslo, Norway, 7–11 March, 2017; Association for Computing Machinery, New York, NY, USA, 313–316. DOI:<https://doi.org/10.1145/3020165.3022141>.
63. Cao, K.; Chen, C.; Baltes, S.; Treude, C.; Chen, X. Automated Query Reformulation for Efficient Search Based on Query Logs from Stack Overflow. In Proceedings of the 2021 IEEE/ACM 43rd International Conference on Software Engineering (ICSE), Madrid, Spain, 25–28 May 2021; pp. 1273–1285. <https://doi.org/10.1109/ICSE43902.2021.00116>.