Session-based Suggestion of Topics for Geographic Exploratory Search

Noemi Mauro
Computer Science Dept., University of Turin
Torino, Italy
noemi.mauro@unito.it

Liliana Ardissono
Computer Science Dept., University of Turin
Torino, Italy
liliana.ardissono@unito.it

ABSTRACT

Exploratory information search can challenge users in the formulation of efficacious search queries. Moreover, complex information spaces, such as those managed by Geographical Information Systems, can disorient people, making it difficult to find relevant data. In order to address these issues, we developed a session-based suggestion model that proposes concepts as a *"you might also be interested in"* function, by taking the user's previous queries into account. Our model can be applied to incrementally generate suggestions in interactive search. It can be used for query expansion, and in general to guide users in the exploration of possibly complex spaces of data categories.

Our model is based on a concept co-occurrence graph that describes how frequently concepts are searched together in search sessions. Starting from an ontological domain representation, we generated the graph by analyzing the query log of a major search engine. Moreover, we identified clusters of ontology concepts which frequently co-occur in the sessions of the log via community detection on the graph. The evaluation of our model provided satisfactory accuracy results.

CCS Concepts

•Information systems → Information retrieval; Query log analysis; Query suggestion; Query reformulation;

Author Keywords

Geographical Information Retrieval; Session-based Concept Suggestion; Query Expansion.

INTRODUCTION

During an information search task, various issues challenge the specification of efficacious queries. Firstly, as discussed by Belkin in [8], users are often unable to use the appropriate keywords because they are looking for data they do not know about. Secondly, complex information spaces, organized in several topic categories, can overload and disorient people, preventing them from finding the information they need; e.g., see [36]. In order to address these issues, several systems present lists, or hierarchies, of information categories to search for. However, in this way they expose users to possibly large sets of options to choose from.

Our work focuses on exploratory search, which is affected by the above issues because the users' information goals are ill-defined; see [29, 47]. We aim at helping users orientate themselves by suggesting relevant topics (concepts) to be explored, given their observed search behavior. More specifically, we aim at guiding data exploration by proposing a small set of concepts that the user might be interested in, given the search context, as a *"you might also be interested in"* function that helps her/him complete the search. We adopt an associative information retrieval model [15]: our hypothesis is that, by analyzing the first query(ies) of a search session, and by taking into account which types of data are often searched together by people, the system can help the user identify further topics he or she might be interested in, by suggesting terms for query expansion. E.g., if the user looks for kindergartens in a town, he or she might also be interested in play and sports areas, as well as in other data related with children activities. Our work aims at supporting the suggestion of such concepts. Specifically, we aim at answering the following research question:

H1: Can the data about the concepts frequently searched together by people within a search session be exploited to help the user explore the portions of an information space relevant to her/his information needs?

This is particularly useful in Geographic Information Retrieval [24, 6], where several queries are aimed at finding the available items, per category, within a geographical area. Guiding the user towards the exploration of information, such as in the above examples, enables her or him to quickly generate custom maps reflecting individual information needs.

In order to answer our research question, we developed a session-based concept suggestion model which, starting from the queries submitted by the user, proposes a set of possibly relevant data categories, complementing the types of information that he or she has focused on. Our model is based on an ontological representation of the information space, and on a Natural Language approach to the interpretation of search queries that identifies the referred concepts in a flexible way, by considering synonyms and by applying Word Sense Disambiguation to resolve the meaning of words. For the development of our model, we analyzed the log of a major search

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
IUI 2018, March 7–11, 2018, Tokyo, Japan.
Copyright is held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 978-1-4503-4945-1/18/03 ...$15.00.
https://doi.org/10.1145/3172944.3172957

AnonID	Query	QueryTime	ItemRank	ClickURL
67910	las vegas sports teams	2006-04-30 18:14:57	1.0	http://www.vegas.com
67910	las vegas transportation	2006-04-30 18:19:59	1.0	http://www.vegas.com
67910	mccarran international airport	2006-04-30 18:22:30	3.0	http://en.wikipedia.org
67910	mccarran international airport	2006-04-30 18:22:30	1.0	http://www.mccarran.com
67910	hub airports in the united states	2006-04-30 18:25:28	9.0	http://www.airportcodes.us
67910	hub airports in the united states	2006-04-30 18:25:28	9.0	http://www.airportcodes.us
67910	black las vegas itineraries	2006-04-30 18:30:03		
67910	educational facilities in las vegas	2006-04-30 18:30:53		
67910	medical facilities in las vegas nv	2006-04-30 18:31:44	1.0	http://lasvegas.citysearch.com
67910	medical facilities in las vegas nv	2006-04-30 18:31:44	3.0	http://www.lasvegasnevada.gov
67910	medical facilities in las vegas nv	2006-04-30 18:31:44	9.0	http://www.lasvegasrelocating.com
67910	unique architecture in las vegas nv	2006-04-30 18:35:02	10.0	http://www.guggenheim.org
67910	unique architecture in las vegas nv	2006-04-30 18:35:02	2.0	http://travel.yahoo.com
67910	architecture in las vegas nv	2006-04-30 18:40:28	4.0	http://lasvegas.citysearch.com
67910	architecture in las vegas nv	2006-04-30 18:40:28	2.0	http://www.library.unlv.edu
67910	architecture in las vegas nv	2006-04-30 18:40:28	3.0	http://local.yahoo.com
67910	religious sites in lasvegas	2006-04-30 18:50:35	1.0	http://www.lasvegas.worldweb.com
67910	religious sites in lasvegas	2006-04-30 18:50:35	2.0	http://www.lasvegas.worldweb.com
67910	religious sites in lasvegas	2006-04-30 18:56:09	12.0	http://travel2.nytimes.com
67910	religious sites in lasvegas	2006-04-30 18:56:09	12.0	http://travel2.nytimes.com

Table 1. Sample session from the AOL log.

engine and we generated a weighted co-occurrence graph that represents how often concepts are searched together in the sessions of the log. Then, we extracted the clusters of concepts that most frequently co-occur by applying a community detection algorithm to the graph. Those clusters are the basis for query expansion, starting from the concepts referred by the user in the observed part of the search session. We defined a few heuristics for recommending relevant types of information and we tested them on the log, achieving satisfactory accuracy results.

The main contributions of our work are: (i) a session-based query expansion model that suggests complementary concepts for satisfying the user's information needs in geographic exploratory search, and (ii) evaluation results of the model.

The following section positions our work in the related research. Section DATASET describes the dataset we used for our experiments. Section IDENTIFYING CLUSTERS OF FREQUENTLY OCCURRING CONCEPTS describes the approach we adopted to identify the co-occurrences of concepts in search sessions. Section SESSION-BASED QUERY EXPANSION proposes some query expansion strategies and the subsequent section evaluates the strategies. The last section concludes the paper and outlines some future work.

RELATED WORK

Various semantic information retrieval models employ concept networks to identify the meaning of queries and propose query expansions aimed at finding information, regardless of the terminology used by the user. E.g., both [35] and [19] use a local thesaurus inferred from the source pool of documents to identify the concepts referred by the queries. Moreover, [28] shows that the integration of different types of thesauri (linguistic, domain specific, etc.) improves the performance of query expansion techniques. The co-occurrence of word phrases in documents is also mined in [22] and [20] to automatically generate associative conceptual spaces, and in [25] and [1] for term suggestion based on the documents returned by search engines. Differently, Wang et al. [44] classify queries in patterns according to their syntactic components and match them to a knowledge base to generate the answers. Finally, a knowledge-based approach is adopted in [43] for concept interpretation, and in [13] to enhance information retrieval in the Semantic Web. On a different perspective, [32] proposes domain-specific ontologies for the interpretation of queries, assuming that users find it easier to specify what they want to do, rather than the concepts they are interested in.

Similarly, we use an ontology, and linguistic information, to interpret search queries at the conceptual level. However, we offer a *"you can also be interested in"* function to propose complementary concepts, i.e., topics, for expanding queries in order to satisfy the user's information needs, in a serendipitous way. For this purpose, we propose an associative information retrieval model [15] based on the observation of concepts co-occurrence in search sessions. Our work is related to the contextual query suggestion model by Cao et al. [11], who suggest queries on the basis of the context provided by the user's recent search history. However, we mine *ontology concepts* from a linguistic interpretation of search queries, while the concepts defined in [11] are clusters of queries associated to similar sets of click results selected by users.

Our model also relates to session-based term suggestion approaches such as the one by Huang et al. [21], and with term suggestion models used for web site advertisement, e.g., see Chen et al. [12]. However, it differs from those works because we look for *concept co-occurrence*, which abstracts from the specific words used to refer to concepts, while they observe *term co-occurrence* for query expansion.

Some recent work on information filtering and in recommender systems attempts to acquire relations among information items from the observation of users' behavior, and is complementary to our work. E.g., Google search engine manages the Knowledge Graph [16] to relate facts, concepts and entities depending on their co-occurrence in queries. Moreover, Oramas et al. [33] use a knowledge graph for personalized item recommendation in the music domain. Furthermore, CoSeNa [10] employs keyword co-occurrence in the corpus of documents to be retrieved, and ontological knowledge about the domain concepts, to support the exploration of text collections using a keywords-by-concepts graph. Our work differs from the above mentioned ones because we exploit a knowledge graph to predict further concepts that the user might be interested in, i.e., we suggest topics, not individual items.

Working at the conceptual level, our work is also complementary to the research on exploration vs. exploitation in information retrieval, which provides models to recognize the type of search that the user is performing, and/or to adapt the search results accordingly; e.g., [34, 5, 31].

DATASET

We defined our concept suggestion model, and evaluated its accuracy, by using as a dataset the AOL query log.[1] The log is composed of 10 files; however, we excluded one of them because it seemed to be corrupted by queries likely performed by a bot: that file included a session spawning over 6 days and included more than 20000 queries, or click through, submitted by the same user.

Each line of the AOL log represents either a query, or a click-through event (in which the user clicks on one of the search results of the previously submitted query). The lines contain the following fields:

- *AnonID*: ID that represents a user in an anonymous way.
- *Query*: search query submitted by the user.
- *QueryTime*: date and hour when the query was submitted.
- *ItemRank*: given the list of results (links to web pages) returned by the search engine, this field represents the relative position of the link on which the user has clicked. If the user has not selected any search results, this field is empty.
- *ClickURL*: URL of the search result selected by the user. Similar to field *ItemRank*, this field can be empty.

Table 1 shows a session from the AOL log. It can be noticed that the user looks for various types of information, concerning sports, transportation, medical facilities, and others. This substantiates the need for a suggestion model that helps her/him by proposing different topics for exploration.

Since we aim at developing an intelligent search support function for Geographical Information Systems, we selected an example system for our experiments, i.e., OnToMap [3, 4], and we pre-processed the AOL log to work on a smaller dataset, focused on the search sessions relevant to geographical information. The ultimate goal is to train a query expansion model suitable for improving information search in OnToMap.

[1]We retrieved the AOL query log in June 2016 from http://www.cim.mcgill.ca/dudek/206/Logs/AOL-user-ct-collection/. Currently, the log is available at https://archive.org/details/AOL_search_data_leak_2006.

rdfs:label	"Kindergarten"@en
rdfs:comment	"An educational institution for young children, usually before they go to primary school."@en
:keywords	"child"@en, "educational"@en, "young"@en
:lemma	"kindergarten"@en
:synonyms	"0th_grade"@en, "all-day_kindergarten"@en, "didactics"@en, "childcare"@en, "educability"@en, "education_program"@en, "education_system"@en, "junior_kindergarten"@en, "kindergarden"@en, "nursery"@en, "pre-primary"@en, "pre-school"@en, "preschoolar"@en

Table 2. Ontology definition of concept "Kintergarden". Values are tagged with the reference language (@en, for English).

In order to build the dataset of our experiments, we first identified the search sessions from the log: we aggregated the queries performed by the same user according to their temporal proximity, following the widely applied rule that two consecutive queries belong to different sessions if the time interval between them exceeds half an hour; see [46]. Then, we selected the sessions including at least one query that refers to the concepts of the OnToMap ontology.[2] For this purpose, our main task was the identification of the concepts referred by the terms of the queries, which we carried out using the approach described in [2, 30]. Concept identification is a particularly important task: as reported in [44], the analysis of the log of a major search engine proved that about the 62% of the queries contain at least one conceptual class term.

Reference Ontology

The OnToMap ontology describes concepts by merging linguistic, encyclopedic and technical knowledge, focusing on concepts that are typical of Participatory Processes; see [2, 42]. Each concept has a textual description, and a set of lemmatized synonyms and keywords extracted from the description. E.g., Table 2 shows the specification of concept "Kintergarden". It reports the concept name (rdfs:label), its textual description (rdfs:comment), a few lemmatized keywords (:keywords), the lemma of the concept (:lemma), and a set of lemmatized synonyms of the terms occurring in the description (:synonyms).

Creation of the Dataset for the Experiments

We identified the AOL queries relevant to our experiments by matching the knowledge about concepts (i.e., the lemmatized synonyms and keywords extracted from the textual description

[2]We did not consider any broader ontologies, such as WordNet [48] or ProBase [49], because we aimed at obtaining a focused dataset to be analyzed.

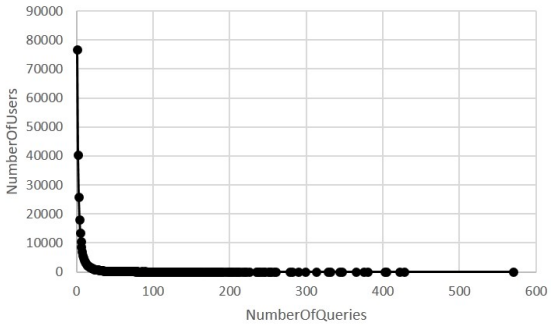

Figure 1. Distribution of the number of queries per user.

Minimum number of queries per user	1
Maximum number of queries per user	428
Mean number of queries per user	6.38
Median number of queries per user	3
Standard Deviation	11.37

Table 3. Measures about the distribution of the queries per user.

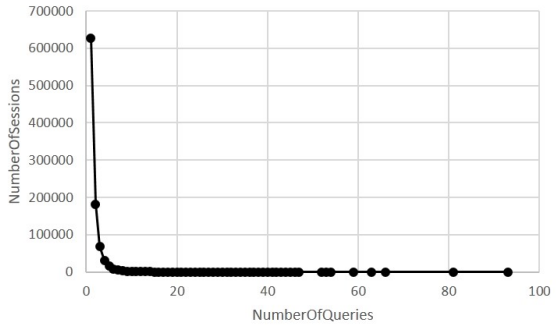

Figure 2. Distribution of queries in sessions.

Minimum number of queries per session	1
Maximum number of queries per session	93
Mean number of queries per session	1.67
Median number of queries per session	1
Standard Deviation	1.45

Table 4. Measures about the length of the search sessions.

of the ontology concepts) to the lemmatized words that compose the queries. If there was at least one match between a concept and a query, we considered the query as relevant, and we included the session to which it belonged in our dataset.

Each query can refer to one or more ontology concepts: the mean number of concepts per query of AOL-reduced is 1.502. This is due to the following reasons:

- The query might refer to independent concepts, each one identified in an unambiguous way. For instance, in a sample query like "public school and transportation in New York" the identified concepts would be *School* and *Local Public Transportation*. Notice that, similar to the findings reported in [7], most queries of the AOL log are short and refer to a single concept, as in the examples of Table 1.
- Because of ambiguity issues, more than one concept could be identified. For instance, given query "missouri child support" from the AOL log, the concepts identified from word "child" are *Childcare Service*, *Play Area* and *Kindergarten*, both including term "child" in their descriptions. As the query is short, it is difficult to understand which topic is the most probable one. Therefore, we assume that the query matches all concepts, with uncertainty.
  In the analysis of the log, we could not exploit the search results selected by users to disambiguate the queries, because we found out that, in most cases, they refer to the root pages of large web sites (e.g., see the URLs in Table 1), or they are obsolete (e.g., the link reported as a search result of query "missouri child support", http://www.dss.state.mo.us). Thus, they help in few cases.

Characteristics of the Dataset

The dataset obtained after the pre-processing phase, henceforth denoted as *AOL-reduced*, is composed of 1581817 queries submitted by 247868 users. The chart in Figure 1 shows the distribution of users with respect to the number of queries they specified. It can be noticed that the distribution follows a Power Law: most users submitted a very small set of queries, whereas a small group of users submitted several queries. As shown in Table 3, the mean number of queries per user is 6, but the distribution differs from this value for about 11 queries (standard deviation).

The dataset includes 945945 sessions, on which we performed two types of analyses:

- Firstly, we analyzed the distribution of sessions w.r.t. the number of queries they include. Table 4 shows the measures related to the length of sessions considering the number of queries. The mean length of a session is 1.67. The distribution follows a Power Law, as shown in Figure 2.
- Secondly, we analyzed the distribution of sessions with respect to users in order to compute the mean number of sessions that the users started; see Figure 3. We found out that users engaged in a mean number of about 3.82 sessions (median = 2). Moreover, the distribution follows a Power Law, with several people engaging in few search sessions, and a small number of users starting a larger number of sessions.

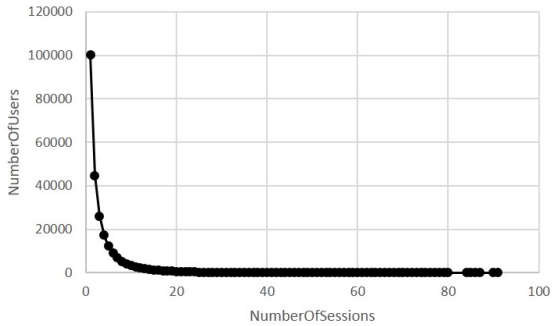

Figure 3. Distribution of sessions per users.

IDENTIFYING CLUSTERS OF FREQUENTLY CO-OCCURRING CONCEPTS

As described in the introduction, starting from the interpretation of the search queries that the user submits, we aim at suggesting further concepts which could be relevant to her/his needs. The idea is to provide an adaptive list of pointers to the types of information provided by the system, which are useful in the context of the user's search. We did this by identifying, from the concepts referred by the queries of the AOL-reduced dataset, other concepts that are related to the former from the viewpoint of commonly shared interests.

Step 1: Creation of the Concept Co-occurrence Graph

Starting from the AOL-reduced dataset, we built a graph that represents concepts co-occurrence in search sessions. By co-occurrence we mean the fact that two or more concepts are referred by the queries belonging to the same session. The graph is composed as follows:

- Each node represents a concept;
- Each edge represents the co-occurrence weight of the two connected concepts. This weight is computed by summing up the evidence of co-occurrence observed in all the sessions of the dataset.

The idea is that, each time two concepts are identified within the same session, the weight of the edge connecting them must increase, in order to capture and reinforce the hypothesis that people frequently search these concepts together.

Given two concepts c_i and c_j, the weight of the edge that connects them is defined as:

$$w_{c_i c_j} = \sum_{S=1}^{nSessions} Freq_{S_{ij}} \quad (1)$$

where $Freq_{S_{ij}}$ represents the contribution provided by session S to the co-occurrence frequency of c_i and c_j.

The contribution of the sessions to the weights of the co-occurrence graph (CG) is represented by a local weighted graph created by interpreting the queries of S. Specifically, $Freq_{S_{ij}}$ is obtained by considering the evidence of co-occurrence provided by the queries that compose S. However, within a session, we avoid summing up the evidence provided by multiple occurrences of the same concepts, because they could derive from a query reformulation [37] or from the repetition of the queries in click-through events; see Table 1.

Given $S = \{Q_1, \ldots, Q_n\}$, $Freq_{S_{ij}}$ is thus computed as follows:

$$Freq_{S_{ij}} = Max_{k=1}^{n}(Freq_{ij_{Q_k}}, ev_{ij_{Q_{k-1}}}) \quad (2)$$

where $Freq_{ij_{Q_k}}$ is the co-occurrence evidence of concepts c_i and c_j provided by query Q_k, and $ev_{ij_{Q_{k-1}}}$ is the one estimated during the interpretation of queries $Q_1, \ldots, Q_{k-1}$.

Finally, the contribution of a query Q is computed as follows:

- If Q contains n terms ($n >= 0$), each one identifying a non-ambiguous concept: $T_1 \Rightarrow c_1, \quad T_2 \Rightarrow c_2, \quad \ldots, \quad T_n \Rightarrow c_n$, then, Q generates nodes $c_1, \ldots, c_n$ of the local graph (if they do not exist yet). For all pairs of concepts c_a and c_b referred by Q, $Freq_{ab_Q} = 1$. Moreover, $Freq_{vw_Q} = 1$ for all the edges connecting c_a, or c_b, to some non-ambiguous concept of the local graph. $Freq_{vw_Q} = 0$ for the other edges of the graph.

  For instance, suppose that $Q =$"public school and transportation in New York" is the first query of a session. The terms and concepts referred by Q are:

  - school $\Rightarrow$ *School* - denoted as concept x;
  - transportation $\Rightarrow$ *Local Public Transportation* - denoted as concept y.

  Then, the local graph is composed of the x and y nodes. Moreover, $Freq_{xy_Q} = 1$.

- If Q contains a term t that refers to m concepts $\{c_1, \ldots, c_m\}$, the interpretation is ambiguous. Therefore, the evidence brought by t to the concepts is $\frac{1}{m}$, in order to take into account the possible interpretations of the query, and spread the weight to the ambiguous concepts. Thus, for each pair c_a, c_b in $\{c_1, \ldots, c_m\}$, $Freq_{ab_Q} = \frac{1}{m}$. $Freq_{vw_Q} = 1/m$ for all the edges connecting c_a, or c_b, to the other concepts of the local graph. Finally, $Freq_{vw_Q} = 0$ for all the edges of the local graph that are not outgoing arcs of c_a or c_b.

  For example, if $Q =$ "missouri child support" is the first query of a session:

  - The concepts identified from word "child" are: *Child-care Services* - denoted as x, *Play Areas* - y, and *Kindergartens* - z.
  - Then, the local graph is composed of the x, y and z nodes. Moreover, $Freq_{xy_Q} = Freq_{xz_Q} = Freq_{yz_Q} = \frac{1}{3}$.

We now sketch the generation of the local co-occurrence graph for a sample search session S. We recall that the contribution of multiple sessions to the overall co-occurrence graph (GC) is incremental: i.e., we sum up the weights provided by the individual sessions. Let's suppose that the terms of S refer to the following concepts:
$Q_1 : t1 \Rightarrow c_1$
$Q_2 : t2 \Rightarrow c_2$
$Q_3 : t3 \Rightarrow c_3, c_4$
$Q_4 : t4 \Rightarrow c_2$
$Q_5 : t5 \Rightarrow c_3$

Figure 4 shows the local co-occurrence graph generated by S.

- Q_1 adds node c_1 to the local graph.
- Q_2 adds c_2 and assigns a weight = 1 to the edge between c_1 and c_2, because there is no ambiguity:
  $Max(Freq_{c_1 c_2 Q_2}, ev_{c_1 c_2 Q_1}) = Max(1, 0) = 1$.
- Q_3 adds c_3 and c_4. Moreover, it assigns a weight = $\frac{1}{2}$ to the edge connecting c_3 and c_4 (the ambiguity concerns two concepts). Furthermore, it assigns the same weight to the edges connecting c_3, and c_4, to the concepts of the previous queries (i.e., the edges between c_3 and c_1, c_3 and c_2, c_4 and c_1, c_4 and c_2).

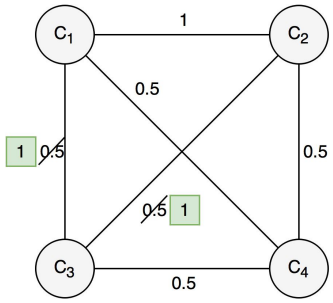

Figure 4. Construction of the graph describing the concepts co-occurrence frequency for a session.

- Q_4 does not add any nodes, nor does it modify the weights in the local graph because the evidence of c_2 doesn't solve the ambiguity between c_3 and c_4:
  $Max(Freq_{c_2c_3Q_4}, ev_{c_2c_3Q_3}) = Max(0.5, 0.5) = 0.5$
  $Max(Freq_{c_2c_4Q_4}, ev_{c_2c_4Q_3}) = Max(0.5, 0.5) = 0.5$
  $Max(Freq_{c_1c_2Q_4}, ev_{c_1c_2Q_3}) = Max(1, 1) = 1.$

- Q_5 does not add any nodes, but it solves the ambiguity between c_3 and c_1, and c_3 and c_2, respectively, because it provides a non ambiguous evidence of c_3. Thus, the respective weights of the local graph are updated:
  $Max(Freq_{c_1c_3Q_5}, ev_{c_1c_3Q_4}) = Max(1, 0.5) = 1.$
  $Max(Freq_{c_2c_3Q_5}, ev_{c_2c_3Q_4}) = Max(1, 0.5) = 1.$
  Notice that we maintain the ambiguity between c_3 and c_4 (in a conservative approach) because, in general, we cannot assume that Q_5 is a reformulation of Q_3. Thus, we do not use it to exclude c_4 from the interpretation of Q_3.

Step 2: Pruning the Graph

The co-occurrence graph of the AOL-reduced dataset is strongly connected: almost all of the ontology concepts are linked to each other by an edge whose weight is > 0. We thus decided to analyze the distribution of weights in the graph, in order to understand the strength of the correlation between concepts. Figure 5 shows this distribution: the x-axis represents the edges, and the y-axis represents their weights, which take values in [0, 68707.5].

The distribution highlights the fact that there is a large number of weakly connected concepts, i.e., of candidate items for query expansion. As, different from keyword suggestion in search engines, we aim at proposing few concepts, we decided to delete the edges having low weight, assuming that they represent weak associations between concepts, and that they capture less commonly shared interests.

Figure 5. Distribution of the weight of edges in the co-occurrence graph.

We selected a threshold for pruning the graph in order to optimize the prediction accuracy of the resulting concept co-occurrence clusters; i.e., the degree of matching between the set of concepts composing the clusters and the concepts identified in the sessions of AOL-reduced. The selected threshold is 2200. Details about the evaluation of accuracy follow.

Step 3: Creation of the Concept Co-occurrence Clusters

Starting from a co-occurrence graph (*CG*) pruned with a threshold, the clusters are created by applying a community detection algorithm which identifies the sets of strictly correlated concepts. Our hypothesis is that they correspond to sets of concepts that are frequently searched together.

For the identification of clusters we analyzed various algorithms. We selected COPRA [18], which works on weighted graphs and detects overlapping communities. In this way, a concept can belong to different clusters at the same time. Our idea was that, within a session, the user might focus on more than one set of highly correlated concepts, and that, starting from the same query, (s)he might explore different paths of the information space.

The COPRA algorithm, applied to the concept co-occurrence graph pruned with threshold 2200, returned 23 clusters having a minimum and maximum number of concepts equal to 1 and 6, respectively.

Three sample clusters are:

1. {Play Area, Sport Area, Kindergarten, Law Enforcement, Hospital}.

2. {National Park, Provincial Park, Regional Park, Urban Park, Furnished Green}.

3. {Play Area, Sport Area, Library, Childcare service, Law Enforcement, School}.

Step 4: Validation of the Clusters

We tested the clusters returned by COPRA on different versions of the pruned graph (using different weight thresholds) in order to estimate how close they reflected user behavior in the AOL-reduced dataset. We considered the AOL-reduced sessions as the ground-truth and we computed standard accuracy measures for the evaluation.

We treated the clusters as unordered sets and we did not care about the observed order of exploration of concepts in the sessions. The reason for this is the fact that, after each search query, we plan to suggest concepts as a set of selectable items, in a multi-choice box. This set is a projection on the information categories managed by the system, based on the search context.

For each threshold, we tested the accuracy of the clusters by applying 10-fold cross-validation, after having randomly

distributed the sessions of the dataset on folders. We used 90% of the sessions as learning set and 10% as test set.

Given a session $S = \{c_{s_1}, \ldots, c_{s_n}\}$, and a cluster $CL = \{c_{cl_1}, \ldots, c_{cl_m}\}$, we evaluated the following measures:

- Precision, describing the rate of concepts of *CL* that also occur in *S*. These concepts represent correct predictions if *CL* is used for query expansion:

$$precision = \frac{|\{c_{s_1}, \ldots, c_{s_n}\} \cap \{c_{cl_1}, \ldots, c_{cl_m}\}|}{|\{c_{cl_1}, \ldots, c_{cl_m}\}|} \quad (3)$$

- Recall, representing the rate of relevant concepts of *S* which *CL* contains, and thus can suggest, if used for query expansion:

$$recall = \frac{|\{c_{s_1}, \ldots, c_{s_n}\} \cap \{c_{cl_1}, \ldots, c_{cl_m}\}|}{|\{c_{s_1}, \ldots, c_{s_n}\}|} \quad (4)$$

- F1 score, computed as:

$$2 * \frac{precision * recall}{precision + recall} \quad (5)$$

Moreover, we considered two types of evaluation:

- Eval1: accuracy of the clusters that best represent search behavior within a session. The aim of this test was to measure the optimal adherence of the clusters to users behavior, assuming to be able to identify the best cluster from the observed queries. In this case, for each session *S*, we computed the precision, recall and F1 score of all the clusters. Then, we selected the accuracy of the best performing one as the representative value for *S*. Finally, we computed the mean accuracy of such best clusters.
- Eval2: mean accuracy of the clusters that include at least one concept referred in the search session. In this case, we aimed at testing the accuracy of a broader set of clusters, related to search sessions in a looser way. As above, we computed the mean precision, recall and F1 score.

Table 5 reports the accuracy values we obtained using the threshold that maximizes the mean F1 ($threshold = 2200$) for pruning the concept co-occurrence graph:

- The precision values show that, during the search sessions, users refer to a substantial portion of the concepts included in the clusters. However, the clusters contain some concepts that are not explored: by using all the clusters that have at least one concept in common with the session, we obtain a precision of about 0.626.
- The recall shows that the suggested clusters largely cover the search sessions: in both evaluations, the recall is about 0.794. Thus, the clusters help propose relevant concepts.

	Eval1	**Eval2**
Precision	0.659	0.626
Recall	0.794	0.794
F1	0.720	0.700

Table 5. Evaluation of the clusters accuracy.

Overall, we can say that, by selecting clusters that have some concepts in common with those referred in the observed portion of a search session, we have good chances to suggest types of information that the user will be interested in. These results helped us in the identification of the query expansion strategies described in the following section.

In the computation of these measures, we did not take into account the occurrence of ambiguities in the interpretation of search queries. Basically, we considered all the concepts referred (either ambiguously or non ambiguously) in the queries as concepts belonging to the search sessions. This might introduce some noise in the evaluation results (e.g., it could increment the number of concepts relevant to the search sessions, reducing the precision results). However, we could not retrieve more precise information about the search interests of the AOL users because, as already discussed, most queries of the dataset are short, and the dataset provides little information to disambiguate them.

SESSION-BASED QUERY EXPANSION

Let's consider a search session $S = \{Q_1, \ldots, Q_i, \ldots, Q_n\}$, and suppose that we have observed the first i queries of *S*. We denote the sequence $\{Q_1, \ldots, Q_i\}$ as $S@i$, and the set of concepts identified by interpreting the queries of $S@i$ as:
$C@i = \{c1, \ldots, \{c_{k_1}, \ldots, c_{k_m}\}, \ldots, \{c_{t_1}, \ldots, c_{t_s}\}, \ldots, c_n\}$.
In $C@i$, the ambiguities in concept identification are represented by including the tuples of ambiguous concepts in subsets of $C@i$. For instance, in the above example, $\{c_{k_1}, \ldots, c_{k_m}\}$ and, respectively, $\{c_{t_1}, \ldots, c_{t_s}\}$ are ambiguous. In the following, we will refer to the tuples of ambiguous concepts as *ambiguity sets*.

We consider three concepts suggestion strategies, which differ from each other in the method for selecting the clusters to be used for query expansion. We assume that they could be applied immediately after the interpretation of the first query of a session, or incrementally, in order to support interactive information search.

- **SLACK:**

  1. Select a set of clusters $\{CL_{x_1}, ..., CL_{x_y}\}$ that contain *at least* one concept c such that $c \in C@i$; i.e., c has been referenced in the observed portion of the search session, $S@i$, either ambiguously or unambiguously.
  2. Propose the concepts that belong to at least one of the selected clusters, and the user has not yet explored. For each cluster CL_{x_j}, we denote the set of concepts included in CL_{x_j} as $Sugg_{x_j}$. The set of concepts to be suggested is computed as follows:
  $Sugg@i = \{Sugg_{x_1} \cup \cdots \cup Sugg_{x_y}\} - C@i$.

  For instance, given $C@i = \{c_1, \{c_3, c_4\}\}$, and two selected clusters, $CL_1 = \{c_1, c_7\}$ and $CL_2 = \{c_2, c_3, c_5, c_8\}$, $Sugg@i = \{c_2, c_5, c_7, c_8\}$.

- **SLACK-selective:** Same as SLACK, but only the clusters that best match $C@i$ are used to compute $Sugg@i$. For our experiments, we selected the best matching cluster. The evaluation of the concept suggestion strategies, described

Indicator	SLACK	SLACK-selective	STRICT
Min number of candidate clusters	0	0	0
Max number of candidate clusters	**2**	N=1	2
Mean number of candidate clusters	**1.192**	1	1.188
Min number of suggested concepts	0	0	0
Max number of suggested concepts	**7**	5	**7**
Mean number of suggested concepts	**2.688**	2.309	2.681
Precision	0.614	**0.621**	0.615
Recall	**0.791**	0.783	0.790
F1	**0.692**	**0.692**	**0.692**
Success rate (%)	**51.5**	51.2	**51.5**

Table 6. Statistical measures of concept suggestion strategies applied to the first query of the search sessions ($S@1$).

Indicator	SLACK	SLACK-selective	STRICT
Min number of candidate clusters	0	0	0
Max number of candidate clusters	**2**	N=1	2
Mean number of candidate clusters	**1.223**	1	1.192
Min number of suggested concepts	0	0	0
Max number of suggested concepts	**7**	5	**7**
Mean number of suggested concepts	**2.861**	2.446	2.799
Precision	0.594	**0.598**	0.596
Recall	**0.819**	0.801	0.818
F1	**0.689**	0.685	**0.689**
Success rate (%)	**39.7**	39.2	39.6

Table 7. Statistical measures of concept suggestion strategies applied to the first two queries of the search sessions ($S@2$).

later on, provided satisfactory accuracy results for this setting of the strategy.

We compute the degree of matching between a cluster CL and $C@i$ as the cardinality of $CL \cap C@i$ (i.e., the number of concepts they have in common), taking the ambiguity in $C@i$ into account. Specifically:

- Each concept c occurring both in CL and, as a non-ambiguous element, in $C@i$, contributes to the computation of the degree of matching with a value = 1.
- Each concept c that occurs in CL, and is part of an ambiguity set AMB in $C@i$, contributes to the computation with $\frac{1}{|AMB|}$; i.e., the cardinality of the ambiguity set mitigates its contribution to the computation of the degree of matching.

For instance, given $C@i = \{c_1, \{c_3, c_4\}, c_5\}$, the degree of matching with cluster $CL_1 = \{c_1, c_7\}$ is 1 and the one with $CL_2 = \{c_2, c_3, c_5, c_8\}$ is 1.5.

- **STRICT:**
  1. Select the set of clusters $\{CL_{x_1}, \ldots, CL_{x_y}\}$ containing *all* of the concepts c such that $c \in C@i$. These are the clusters covering the whole portion of the search session observed so far (regardless of the fact that the concepts are ambiguous or not).
  2. As above, the set of concepts suggested for query expansion is: $Sugg@i = \{Sugg_{x_1} \cup \cdots \cup Sugg_{x_y}\} - C@i$.

  We did not propose a STRICT-selective strategy because STRICT generates very small candidate sets.

EVALUATION OF CONCEPT SUGGESTION STRATEGIES

We tested the accuracy of our strategies by applying 10-fold cross-validation on the AOL-reduced dataset, after having randomly distributed the sessions on folders. For the evaluation, we compared the concepts suggested by the strategies to those explored by users in the search sessions (ground-truth). Specifically, for each search session, we tested the strategies after having interpreted the first query alone ($S@1$), the first two queries ($S@2$), and so forth. In each case, we computed the mean precision, recall and F1 score of the strategies by comparing the concepts of the suggested clusters to the concepts referred in the remainder of the sessions.

Evaluation Results

Table 6 summarizes the results that we obtained in the suggestion of concepts immediately after having interpreted the first query of the search sessions ($S@1$):

- SLACK: the minimum, mean, and maximum number of candidate clusters identified by this strategy are 0, 1.192, and 2, respectively. The mean number of concepts suggested for query expansion is about 2.688. This number goes up to 7 in some sessions.
- SLACK-selective with number of best matching clusters N=1. The minimum, mean and maximum number of selected clusters are 0, 1 and N=1, respectively. The strategy suggests a lower mean number of concepts: about 2.309, and it proposes at most 5 concepts.
- The STRICT strategy proposes a maximum of 2 candidate clusters per session, with about 1.188 clusters in average, and it generates at most 7 concept suggestions, with a mean value of about 2.681 concepts.

The three strategies have almost the same accuracy (see Table 6): the F1 measure can be approximated to 0.692 in all the cases. However, they differ from each other if we look at finer-grained measures:

- Precision: SLACK-selective outperforms the other two, i.e., it suggest the highest number of concepts that users are observed to explore in the remainder of the search session. The lower precision of SLACK can be explained with the fact that it selects a superset of the clusters picked by

SLACK-selective, and thus can introduce more noise in the suggestions. Similarly, STRICT can select more than one cluster (covering all the concepts referred in the first query), thus incrementing noise in some cases.

- Recall: the best strategy is SLACK, which outperforms SLACK-selective because it selects a larger pool of clusters and thus has better chances to guess the concepts explored by users. Interestingly, however, STRICT achieves almost the same recall, having a stricter cluster selection policy. We explain this finding with the fact that, by selecting all the clusters that cover the observed query, the strategy has good chances to guess the user's exploration path, among the possible ones (remember that the concept co-occurrence clusters mirrored common search behavior in a fairly accurate way).

- Success rate: as proposed by Huang et al. in [21], this metric measures the percentage of sessions for which our strategies propose at least one concept that is referred in the portion of the search session to be observed. Looking at the table, we can see that all of the strategies are successful in more than 50% of cases. However, the best performing ones are SLACK and STRICT. In other words, SLACK is as good as STRICT if we aim at suggesting at least one relevant concept.
  We compared our strategies to the Huang et al.'s term suggestion model, which is the most similar one among those available in the literature. As that model achieves 23% success rate, we can say that our approach definitely outperforms it.

Table 7 shows the evaluation results after the interpretation of the first two queries of the search sessions. It can be seen that, for all the strategies, the minimum, mean and maximum number of selected clusters are constant. Moreover, the minimum, mean and maximum number of suggested concepts are constant or increase.
Regarding accuracy, we notice that for all the three strategies, the recall increases, while the precision and F1 scores decrease to about 0.6, but the SLACK-selective strategy is the most precise one. SLACK has the best recall and STRICT a rather similar one; as result, both strategies have the best F1 value. The higher recall of these strategies suggests that the selection of a single cluster, best representing the observed portion of search session, is not general enough to cover all the possible developments of the search session. Thus, a less restrictive approach, that makes it possible to consider more than one development of a search session (as the one adopted in both SLACK and STRICT) achieves better results.

We measured the trend for longer portions of search sessions and we discovered that it is consistent, i.e., the precision decreases, and the recall increases, when more queries are interpreted. See Figures 6 and 7, which report the values of these measures after the interpretation of 1 to 5 queries.[3] However, the three strategies behave slightly differently:

[3]We did not consider any longer sequences because very few search sessions are composed of more than 5 queries.

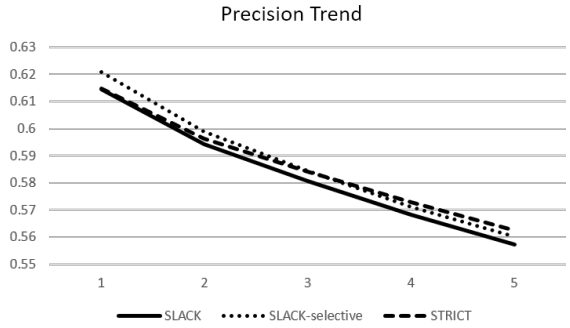

Figure 6. Precision of the concept suggestion strategies in function of the number of interpreted queries ($S@i$).

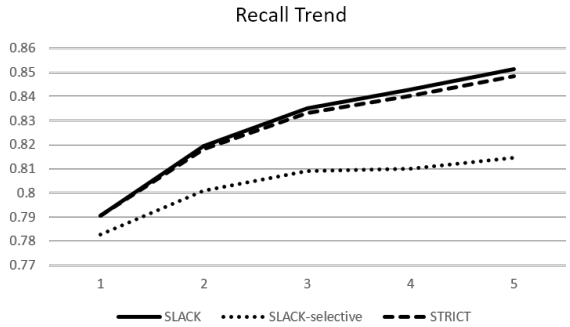

Figure 7. Recall of the concept suggestion strategies in function of the number of interpreted queries ($S@i$)

- SLACK-selective is the most precise strategy when observing the first two queries, and is always better than SLACK, but it is outperformed by STRICT after the observation of 3 queries. We explain this with the fact that, when more information about the search context is available, STRICT is able to select the clusters that best represent the user's search interests, and thus to mirror her/his behavior.

- Regarding recall, we believe that it increases when more information about a search session is available because, if more concepts are identified from the search queries, there are more chances to select good candidates and thus to propose relevant concepts. However, the improvement of this accuracy measure varies depending on the selectivity of the strategy:

  - SLACK selects the highest number of clusters, and thus suggest a larger set of concepts. This increases the chances that the suggested concepts are identified in the session.
  - SLACK-selective chooses a single cluster that best matches the search session, at the expense of recall.
  - STRICT selects the clusters matching the whole portion of session observed so far. These are the best representative of the session and thus have very good changes to predict correct concepts, being used together for concept prediction.

Discussion

The results of our experiments enable us to positively answer our research question:

H1: Can the data about the concepts frequently searched together by people within a search session be exploited to help the user explore the portions of an information space relevant to her/his information needs?

The results show that our concept suggestion strategies have high recall, i.e., they are able to suggest most of the concepts that the users explore in the AOL-reduced dataset. However, their accuracy differs:

- The SLACK-selective strategy is the best one from the viewpoint of precision, at least when the first queries of a session are analyzed in order to suggest concepts for query expansion.
- The SLACK and the STRICT ones achieve the best recall.

Looking at these findings, we believe that the decision of which strategy best suits concept suggestion depends on the way how we want to implement this function:

- If we can suggest a larger number of concepts, then the SLACK strategy is the best one, because (i) it achieves better recall and (ii) the mean number of suggested concepts is just a bit higher than that of the other two.
- Differently, if we aim at precision, SLACK-selective is the preferred one at the beginning of a search session, while STRICT is more suitable for concept suggestion after the first queries.

For our current work, SLACK is the most convenient strategy because it is fairly precise and, at the same time, it suggests more relevant concepts than the other ones. Specifically, we aim at providing the user with an overview of the sets of concepts to choose from, given the search context provided by the previous queries. As the user will be enabled to choose relevant concepts in a multiple-choice box, the presence of very few irrelevant concepts, as is the case of SLACK, is not a problem.

It should be noticed that other strategies might be considered. For instance, the past search history of the individual user might be analyzed to identify the concepts he or she most frequently searched for, assuming that these represent interesting types of information. Then, this might be used to select candidate clusters for query expansion. In this paper, we left these strategies apart for two main reasons:

- Firstly, as discussed by Greenstein et al. in [17], session-based recommendation is of primary importance because users frequently interact with systems in an anonymous way, or there might be privacy or technical reasons for avoiding the tracking of their identities. Therefore, we decided to focus on this function before exploring other scenarios.
- Secondly, the introduction of historic information about user behavior deserves a separate research; e.g., see [45, 40, 39]. Specifically, [9] showed that historic information about the search behavior of an individual user helps re-ranking of search results at the beginning of a search session, while local information about the session is more effective later on, because it provides information about the actual context of search. As we deal with concept suggestion, these results cannot be immediately transferred to our case. We will thus carry out this type of analysis in a broader perspective of ontology-based personalized search support and ontology-based user models; e.g., see [41, 23, 38, 14, 27, 26].

CONCLUSION

We described a session-based concept suggestion model that supports information search by proposing concepts for query expansion, given the context provided by the observed portions of the search sessions. We focus on presenting types of information that the user might also be interested in, given the data he or she is searching for, at the conceptual level. We aim at supporting the exploration of information spaces including rather different types of data, such as those managed by Geographical Information Systems.

Our model is based on an analysis of search behavior, that we carried out by studying the AOL query log. Specifically, our model is based on (i) the identification of clusters of concepts that frequently co-occur in search sessions, and (ii) the definition of strategies that select the clusters for concept suggestion by taking the search context that emerges from the first queries of the sessions. The evaluation of our model has provided satisfactory results about its prediction accuracy.

Having performed off-line experiments, we could not test the serendipity of the concept suggestions, as we relied on data about the types of information that users autonomously explored in the search sessions. Our future work includes (i) analyzing users' past behavior for improving concept suggestion, and (ii) studying the impact of our model on the serendipity of suggestions by testing it in an online experiment with users.

ACKNOWLEDGMENTS

We thank the anonymous reviewers for their fruitful comments and suggestions. This work is partially funded by project MIMOSA (MultIModal Ontology-driven query system for the heterogeneous data of a SmArtcity, "Progetto di Ateneo Torino_call2014_L2_157", 2015-17) and by projects "Ricerca Locale" and "Ricerca Autofinanziata" of the University of Turin.

REFERENCES

1. M. Akaishi, K. Satoh, and Y. Tanaka. 2004. *An Associative Information Retrieval Based on the Dependency of Term Co-occurrence*. Springer Berlin Heidelberg, Berlin, Heidelberg, 195–206.
2. L. Ardissono, M. Lucenteforte, N. Mauro, A. Savoca, A. Voghera, and L. La Riccia. 2016. Exploration of Cultural Heritage Information via Textual Search Queries. In *MobileHCI '16 Proceedings of the 18th Int. Conf. on Human-Computer Interaction with Mobile Devices and Services Adjunct*. ACM, 992–1001.
3. L. Ardissono, M. Lucenteforte, N. Mauro, A. Savoca, A. Voghera, and L. La Riccia. 2017a. Semantic

Interpretation of Search Queries for Personalization. In *Proc. of UMAP 2017 Adjunct*. ACM, 101–102.

4. L. Ardissono, M. Lucenteforte, N. Mauro, A. Savoca, A. Voghera, and L. La Riccia. 2017b. OnToMap - Semantic Community Maps for knowledge sharing. In *Proceedings of Hypertext 2017*. ACM, 317–318.
5. L. Athukorala, A. Medlar, A. Oulasvirta, G. Jacucci, and D. Glowacka. 2016. Beyond relevance: adapting exploration/exploitation in information retrieval. In *Proc. of the 21st Int. Conf. on Intelligent User Interfaces (IUI'16)*. ACM, Sonoma, California, USA, 359–369.
6. A. Ballatore, W. Kuhn, M. Hegarty, and E. Parsons. 2016. Special issue introduction: Spatial approaches to information search. *Spatial Cognition & Computation* 16, 4 (2016), 245–254.
7. S.M. Beitzel, E.C. Jensen, O. Frieder, D.D. Lewis, A. Chowdhury, and A. Kolcz. 2005. Improving automatic query classification via semi-supervised learning. In *Proc. of the Fifth IEEE Int. Conf. on Data Mining (ICDM '05)*. IEEE Computer Society, Washington, DC, USA, 42–49.
8. N.J. Belkin. 1980. Anomalous states of knowledge as a basis for information retrieval. *Canadian Journal of Information Science* 5 (1980), 133–143.
9. P.N. Bennett, R.W. White, W. Chu, S.T. Dumais, P. Bailey, F. Borisyuk, and X. Cui. 2012. Modeling the impact of short- and long-term behavior on search personalization. In *Proc. of the 35th Int. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR '12)*. ACM, New York, NY, USA, 185–194.
10. K.S. Candan, M. Cataldi, L. Di Caro, M.L. Sapino, and C. Schifanella. 2009. CoSeNa: a context-based search and navigation system. In *MEDES '09 Int. Conf. on Management of Emergent Digital EcoSystems*. Chia Laguna, Italy, Art. 33.
11. H. Cao, D. Jiang, J. Pei, Q. He, Z. Liao, E. Chen, and H. Li. 2008. Context-aware query suggestion by mining click-through and session data. In *Proc. of the 14th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD '08)*. ACM, New York, NY, USA, 875–883.
12. Y. Chen, G-R Xue, and Y. Yu. 2008. Advertising Keyword Suggestion Based on Concept Hierarchy. In *Proc. of the 2008 Int. Conf. on Web Search and Data Mining (WSDM '08)*. ACM, New York, NY, USA, 251–260.
13. M. Fernández, I. Cantador, V. López, D. Vallet, P. Castells, and E. Motta. 2011. Semantically enhanced information retrieval: an ontology-based approach. *Web Semantics: Science, Services and Agents on the World Wide Web* 9, 4 (2011), 434–452.
14. S. Gauch, J. Chaffee, and A. Pretschner. 2003. Ontology-based personalized search and browsing. *Web Intelli. and Agent Sys.* 1, 3-4 (2003), 219–234.
15. V.E. Giuliano and P.E. Jones. 1962. *Linear Associative Information Retrieval*. Technical Report ESD-TR-62-294. ASTIA.
16. Google. 2017. Knowledge Graph. https://www.google.com/intl/it_it/insidesearch/features/search/knowledge.html.
17. A. Greenstein-Messica, L. Rokach, and M. Friedman. 2017. Session-based recommendations using item embedding. In *Proc. of the 22nd Int. Conf. on Intelligent User Interfaces (IUI '17)*. ACM, New York, NY, USA, 629–633.
18. S. Gregory. 2010. Finding overlapping communities in networks by label propagation. *New Journal of Physics* 12, 103018 (2010).
19. F.A. Grootjem and T.P. van der Weide. 2006. Conceptual query expansion. *Data & Knowledge Engineering* 56 (2006), 174–193.
20. O. Hoeber, XD. Yang, and Y. Yao. 2005. Conceptual query expansion. In *Advances in Web Intelligence. AWIC 2005, Lecture Notes in Computer Science, vol 3528*. Lodz, Poland, 190–196.
21. C.-K. Huang, L.-F. Chien, and Y.-J. Oyang. 2003. Relevant term suggestion in interactive web search based on contextual in formation in query session logs. *Journal of the American Society for Information Science and Technology* 54, 7 (2003), 638–649.
22. J. Van Den Berg and M. Schuemie. 1999. Information retrieval systems using an associative conceptual space. In *In Proc. of the 7th European Symposium on Artificial Neural Networks (ESANN'99)*. 351–356.
23. X. Jiang and A.-H. Tan. 2009. Learning and inferencing in user ontology for personalized Semantic Web search. *Information Sciences* 179 (2009), 2794–2808.
24. C.B. Jones and R.S. Purves. 2008. Geographical Information Retrieval. *Int. Journal of Geographical Information Science* 22, 3 (2008), 219–228.
25. A. Joshi and R. Motwani. 2006. Keyword Generation for Search Engine Advertising. In *6th IEEE Int. Conf. on Data Mining - Workshops (ICDMW'06)*. ACM, New York, NY, USA, 490–496.
26. K.W-T Leung and Dik D.L. Lee. 2010. Deriving concept-based user profiles from search engine logs. *IEEE Trans. on Knowl. and Data Eng.* 22, 7 (July 2010), 969–982.
27. F. Liu, C. Yu, and W. Meng. 2002. Personalized web search by mapping user query to categories. In *Proc. CIKM'02*. McLean, Virginia, USA.
28. R. Mandala, T. Tokunaga, and H. Tanaka. 1999. Combining multiple evidence from different types of thesaurus for query expansion. In *Proc. of the 22nd Annual Int. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR '99)*. ACM, New York, NY, USA, 191–197.

29. G. Marchiorini. 2006. Exploratory search: from finding to understanding. *Commun. ACM* 49, 4 (2006), 41–46.

30. N. Mauro and L. Ardissono. 2017. Concept-aware Geographic Information Retrieval. In *Proceedings of 2017 IEEE/WIC/ACM Int. Conf. on Web Intelligence (WI)*. ACM, Leipzig, Germany.

31. A. Medlar, J. Pyykkö, and D. Glowacka. 2017. Towards fine-grained adaptation of exploration/exploitation in Information Retrieval. In *Proc. of the 22nd Int. Conf. on Intelligent User Interfaces (IUI '17)*. ACM, New York, NY, USA, 623–627.

32. M. Molina and S. Bayarri. 2011. A multinational SDI-based system to facilitate disaster risk management in the Andean community. *Computers & Geosciences* 18 (2011), 1501–1510.

33. S. Oramas, V.C. Ostuni, T. Di Noia, X. Serra, and E. Di Sciascio. 2016. Sound and music recommendation with knowledge graphs. *ACM Transactions on Intelligent Systems and Technology* 8, 2 (2016), Art. 21.

34. R. Porrini, M. Palmonari, and G. Vizzari. 2014. Composite match autocompletion (COMMA): a semantic result-oriented autocompletion technique for e-marketplaces. *Web Intelligence and Agent Systems: an international journal* 12 (2014), 35–49.

35. Y. Qiu and H.P Frei. 1993. Concept based query expansion. In *Proc. of the 16th Annual Int. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR '93)*. ACM, New York, NY, USA, 160–169.

36. L. Ratzan. 2004. *Understanding information systems: what they do and why we need them*. American Library Association, Chicago, USA.

37. S.Y. Rieh and H. I Xie. 2006. Analysis of multiple query reformulations on the web: the interactive information retrieval context. *Information processing and management* 42 (2006), 751–768.

38. A. Sieg, B. Mobasher, and R. Burke. 2010. Ontology-based collaborative recommendation. *Computing* (2010).

39. B. Smyth, E. Balfe, J. Freine, P. Briggs, M. Coyle, and O. Boydell. 2005. Exploiting query repetition and regularity in an adaptive community-based web search engine. *USER MODELING AND USER-ADAPTED INTERACTION - The Journal of Personalization Research* 14, 5 (2005), 383–423.

40. D. Sontag, K. Collins-Thompson, P.N. Bennett, R.W. White, S. Dumais, and B. Billerbeck. 2012. Probabilistic Models for Personalizing Web Search. In *Proc. of the 5th ACM Int. Conf. on Web Search and Data Mining (WSDM '12)*. ACM, New York, NY, USA, 433–442.

41. J. Teevan, S.T. Dumais, and E. Horvitz. 2005. Personalizing Search via Automated Analysis of Interests and Activities. In *Proc. of the 28th Annual Int. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR '05)*. ACM, New York, NY, USA, 449–456.

42. A. Voghera, R. Crivello, L.Ardissono, M. Lucenteforte, A. Savoca, and L. La Riccia. 2016. Production of spatial representations through collaborative mapping. An experiment. In *Proc. of 9th Int. Conf. on Innovation in Urban and Regional Planning (INPUT 2016)*. 356–361.

43. Y. Wang, H. Huang, and C. Feng. 2017. Query expansion based on a feedback concept model for microblog retrieval. In *Proc. of the 26th Int. Conf. on World Wide Web (WWW '17)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 559–568.

44. Y. Wang, H. Li, H. Wang, and K.Q. Zhu. 2012. Concept-based web search. In *Conceptual Modeling. ER 2012. Lecture Notes in Computer Science, vol 7532*, P. Atzeni and D. Cheung (Eds.). Springer, Berlin, Heidelberg, 449–462.

45. R.W. White, P.N. Bennett, and S.T. Dumais. 2010. Predicting short-term interests using activity-based search context. In *Proc. of the 19th ACM Int. Conf. on Information and Knowledge Management (CIKM '10)*. ACM, New York, NY, USA, 1009–1018.

46. R.W. White, M. Bilenko, and S. Cucerzan. 2007. Studying the use of popular destinations to enhance web search interaction. In *Proc. of the 30th Annual Int. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR '07)*. ACM, New York, NY, USA, 159–166.

47. R.W. White and R.A. Roth. 2006. Supporting exploratory search. *Commun. ACM* 49, 4 (2006), 36–39.

48. WordNet. 2017. WordNet - a lexical database for English. https://wordnet.princeton.edu/.

49. W. Wu, H. Li, H. Wang, and K.Q Zhu. 2012. Probase: A Probabilistic Taxonomy for Text Understanding. In *Proc. of the 2012 ACM SIGMOD Int. Conf. on Management of Data*. ACM, New York, NY, USA, 481–492.