DOI: 10.1111/cgf.13403

Eurographics Conference on Visualization (EuroVis) 2018
J. Heer, H. Leitte, and T. Ropinski
(Guest Editors)

Volume 37 (2018), Number 3

Visualizing Expanded Query Results

Michael Mazurek[1] and Manuela Waldner[1]

[1]TU Wien, Austria

Abstract

When performing queries in web search engines, users often face difficulties choosing appropriate query terms. Search engines therefore usually suggest a list of expanded versions of the user query to disambiguate it or to resolve potential term mismatches. However, it has been shown that users find it difficult to choose an expanded query from such a list. In this paper, we describe the adoption of set-based text visualization techniques to visualize how query expansions enrich the result space of a given user query and how the result sets relate to each other. Our system uses a linguistic approach to expand queries and topic modeling to extract the most informative terms from the results of these queries. In a user study, we compare a common text list of query expansion suggestions to three set-based text visualization techniques adopted for visualizing expanded query results – namely, Compact Euler Diagrams, Parallel Tag Clouds, and a List View – to resolve ambiguous queries using interactive query expansion. Our results show that text visualization techniques do not increase retrieval efficiency, precision, or recall. Overall, users rate Parallel Tag Clouds visualizing key terms of the expanded query space lowest. Based on the results, we derive recommendations for visualizations of query expansion results, text visualization techniques in general, and discuss alternative use cases of set-based text visualization techniques in the context of web search.

CCS Concepts

•***Information systems*** → *Search interfaces;* •***Human-centered computing*** → *Empirical studies in visualization;*

1. Introduction

Users normally use very few key terms to formulate queries for web search engines. Depending on the broadness of the query, users issue an average number of three words per query [PBW07]. However, the users' information needs can hardly be specified accurately in such a short query. It is estimated that around around 16% of online queries are ambiguous [SLW*07]. This means that the users choose query terms that can have multiple meanings, such as "java" or "apple", and therefore reveal a set of incoherent documents [CTZC02]. Term mismatches, on the other hand, occur when indexers and users employ a different vocabulary to describe the same phenomenon [FLGD87]. An example for such a term mismatch is the usage of "lawyer" instead of "attorney", which leads to different sets of top-ranked results in web search engines.

The information retrieval community resolves ambiguous queries and term mismatches by expanding the user queries with additional terms that are related to the key terms given by the user. These expansion terms may either be statistically related to the user's query terms (for instance, because they co-occur in the same documents) or lexically related (for instance, by consulting a thesaurus) [Voo94]. Expansion terms may also be selected based on the user's personal search history [CFN07] or based on the user's relevance feedback about retrieved documents [SB97].

Query expansion has become a standard feature in nowadays' search engines. Most search engines provide query expansion suggestions from which users can choose interactively (*interactive query expansion* IQE), such as "Google Suggest" [Goo17b]. PubMed automatically expands the user's query by mapping the query terms to concepts of the curated "Medical Subject Headings" vocabulary map [LKW09] (*automatic query expansion* AQE). For general web search, AQE is considered too unstable [CR12]. IQE is generally more effective than AQE, but users find it hard to choose the best IQE terms. In an experiment [Rut03], users reported that, even though they understood the semantic relationships of suggested IQE terms, they could not infer which would attract more relevant documents. Ruthven [Rut03] therefore suggests to provide users with more information than just the expansion terms to facilitate discrimination of good from poor query expansion suggestions.

Indeed, there have been attempts to visualize how suggested query expansions relate to the user's original query [FFW91, HYY05, KTZ*07]. However, while these examples visualize how the *expanded query terms* relate to each other, our goal was to visualize how the *results* of these expanded queries relate to each other. This way, users can judge if the resulting documents of a query are relevant for their information needs. Our hypothesis is that users can decide faster and more accurately if expanding the query is beneficial for their search task, and which expansion terms are most appropriate, when seeing a visual summary of the results of these queries. To test this hypothesis, we selected and partially extended three set-based text visualization techniques to visualize the relationship between expanded queries and their associated text-based results, namely: Compact Euler Diagrams [RD10], Parallel Tag Clouds [CVW09], and a List View [SGL08]. In a user study, where users had to conduct ambiguous queries, we compared these three visualizations to a baseline, showing a text list of query expansion suggestions. Our contributions are:

1. The selection, adoption, and extension of set-based text visualization techniques to visualize query expansion results.
2. The implementation of these visual query expansion techniques as interactive extension for the Google search engine.
3. The results of our user study, showing that users selected expanded queries most quickly and without significant precision loss from the text list.

We reflect on the results and suggest directions for future research.

2. Related Work

The most common way to show query expansion suggestions is a simple text list. An early example was presented by Harman [Har88], whose interface shows three lists of query expansion terms: *feedback terms* from top-ranked result documents, *term variants* of the original query terms, and *related terms* to the original query terms from a thesaurus. Instead of lists, expansion terms, extracted from a thesaurus or a knowledge map, have also been visualized as graphs, where the relation between the user's original query terms and the suggested expansion terms is explicitly encoded [FFW91, HYY05]. Kozanidis et al. [KTZ*07] visualize expansion terms in a tree – from a general root node (e.g., "transportation") to specific leaf nodes (e.g., "station wagon"). Similarly to these works, we use a linguistic approach to choose expansion terms. However, we are interested whether visualizing the query expansion *results* has an influence on the user's selection strategy.

"Visual Query Suggestions" [ZYM*09] provide a list of related key terms to the given query, where each key term is also associated with images. While this work also visualizes query expansion results, our work differs in two aspects: First, we are interested in visualizing the text results of the expanded queries. Second, instead of showing the query expansion suggestions in a linear list, we are comparing visualization techniques differing in the way how they encode overlaps and differences between queries.

There is a plethora of work focusing on the visualization of results of a single web query (see Hearst's survey on search user interfaces for an overview [Hea11]). While some works encode the results based on meta-data [DCCW08, HS17], visualizations showing text-based search results are more relevant for the present study. One way to visualize search results is to encode the retrieved documents' similarities to the user's query terms, either using glyphs attached to the document surrogates [Hea95, HY06] or stacked bars [RTM05, dSSV15]. Another common approach is to cluster the text of the resulting documents into topics and visualize each document's association with these topics as glyphs attached to document surrogates [ISY*12], by spatializing key terms of documents based on topic similarities [PBR17], or by spatializing the document surrogates themselves and color-coding them topic-wise [GNSRP*14]. In our work, we also use topic modeling, but the focus of the visualizations is to reveal the connections between multiple query variations with the topic key terms.

Notably, Sparkler [HHP*01] is a visualization of search results of *multiple* user queries or multiple search engines. Retrieved documents are represented as dots arranged in a circle, where the circle segment represents the associated query and the radius encodes the document's relevance score for the query. This way, users can easily compare which query reveals the most relevant documents. In the VIBE system [OKS*93], users can specify multiple queries and spatially arrange icons representing these queries in the visualization. Document icons are then automatically positioned so that they are close to the most related queries. However, if users issue ambiguous queries or use mismatched query terms, the employed relevance scores are not reliable. We therefore do not visualize document relevances, but key terms of these documents and their associations with multiple linguistic query variations and thereby leave the relevance judgment to the user.

Visualization of text-based information, such as documents retrieved during web search, have become an important research topic within the visualization community (see Kucher and Kerren for a survey [KK15]). Of most interest for our work are visualization techniques comparing the text content of multiple documents or corpora, respectively. In this use case, documents or corpora can be viewed as sets, containing terms as set elements. A goal of set visualization techniques is to facilitate the exploration of relations between sets, for instance to discover overlaps between sets or hierarchical relations [AMA*14].

Visualizations of multiple documents or corpora therefore often use classic set visualization techniques, such as Euler Diagrams. For instance, Riche et al. [RD10] visualize multiple overlapping sets with associated text data with simple rectangles by splitting sets or by duplicating set elements. DiTop View [OSR*14] partitions a 2D plane into three sets, as well as four overlap regions between these sets, and assigns topic glyphs into these seven regions. In RadCloud [BLB*14], terms are not explicitly associated with sets, but are arranged within a circle. Set labels are placed on equidistant circle segments and apply attractive forces onto their associated key terms. ConcentriCloud [LHB*15] uses a similar principle, but explicitly partitions the circle to show predefined set overlaps. However, similarly to classic Venn Diagrams, rendering *possible* set relations leads to poor scalability in terms of the number of sets.

A straight-forward way to visualize the association of elements with sets is to use parallel lists. Stasko et al. [SGL08] use this principle to reveal co-occurring sets of entities in document collections, where vertical lists show the entities, and links between these lists indicate co-occurrences within documents. TheMail [VGD06] and Parallel Tag Clouds [CVW09] also use parallel lists, but combine them with the concept of classic tag clouds, where font size encodes term frequency. Parallel Tag Clouds use links between terms to indicate co-occurrences between corpora. WordBridge [KKEE11] embeds tag clouds into a node-link diagram, where nodes are tag clouds of distinct terms in documents, and links show terms shared

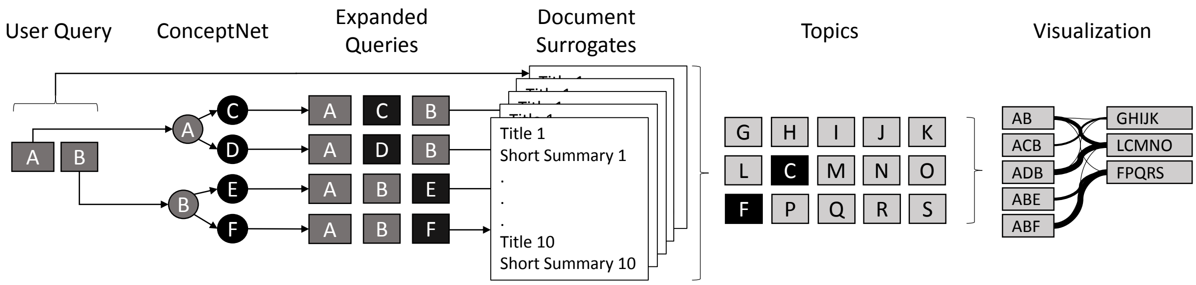


Figure 1: *Pipeline of preprocessing steps to create a visualization of query expansion results: given a user query (here consisting of two terms), n expanded queries are generated from ConceptNet. Each of the n expanded queries, plus the original query, results in up to 10 document surrogates. From these document surrogates, k topics are extracted, and labeled by m key terms each, which serve as input for the visualization. In this example, we use* $n = 4$, $k = 3$, *and* $m = 5$, *and the List View for visualization.*

between pairs of documents. However, this visualization does not explicitly reveal term overlaps between more than two sets.

3. Query Expansion

We use a linguistic approach to provide query expansion suggestions to resolve ambiguous queries and term mismatches. We use the semantic network ConceptNet [LS04], which features more term relations than a classic thesaurus. We parse the users' original query and extract nouns, adjectives, and verbs from the query. Nouns are all converted to singular, and verbs are conjugated.

To disambiguate polysemes (i.e., words with multiple meanings), we find generalizations of each query term in ConceptNet. Therefore, we parse the following directed edges from the query term: UsedFor, HasContext, DefinedAs, isA, RelatedTo, and PartOf. For instance, the term "java" has isA-edges like "programming language" or "an island". To resolve term mismatches, we select English synonyms, irrespective of their edge direction. For instance, "lawyer" has "attorney" and "solicitor" as synonyms in ConceptNet. We obtain a list of expansion terms for each query term and rank them by their edge weight from ConceptNet. From these lists, the highest ranked terms for each query term are selected, so that the number of expansions per query term is balanced (see first three steps in Figure 1).

In contrast to classic IQE, the resulting expanded queries are actually performed in the background to be able to visualize the results. We use Google's Custom Search API [Goo17a] to perform these queries. The API returns the results as document surrogates, each consisting of a title and short text summary (see step "Document Surrogates" in Figure 1).

4. Key Terms Extraction

To visually summarize the text results from the expanded queries, we extract the most representative key terms of the entire expanded query space. The document surrogates, consisting of the titles and short summaries of the results, serve as input for this step. Document surrogates are cleaned by removing stop words, filter symbols, separators, and lowercasing all remaining terms.

For each query's document surrogate list, we count the number of occurrences of nouns, verbs, and adjectives, identified through Part-of-Speech tagging. From this information, we create an $l \times (n+1)$ term-query matrix T, where l is the number of unique key terms extracted from all document surrogates, and n is the number of expanded queries. Each cell of the matrix T contains the weight of a term t in query q, which is computed as follows:

$$\text{tf*iqf}(t,q,Q) = \text{tf}(t,q) \cdot log \frac{n+1}{|\{q \in Q : t \in q\}|} \tag{1}$$

where Q represents all $n+1$ queries, and $\text{tf}(t,q)$ is the number of occurrences of term t in query q. This corresponds to the widespread tf*idf weighting scheme [SJ72].

While prominent key terms could be determined by ranking the term weights, it has been shown that the tf*idf heuristic is not as discriminative as topic modeling [ZYT11]. We therefore use topic modeling, which can be seen as clustering of the matrix T into $k < l$ coherent topics, to obtain discriminative and expressive key terms for our visualizations. For our examples, we used $k \leq (n+1)$, because we assumed that there would be no more truly distinct topics in the document surrogates than query variations. We use nonnegative matrix factorization [CLRP13] to decompose the matrix T into two matrices W and H of a lower rank k. The matrix W is an $l \times k$ matrix, containing scores of each term for each of the k topics. The $k \times (n+1)$ matrix H contains the scores of each query for a given topic. For our visualizations, we pick m key terms with the highest scores in the matrix W from each topic to represent the expanded query space (see "Topics" step in Figure 1). We associate a key term t with a query q if $\text{tf}(t,q) > 0$. The *topic-query association* between a topic j and a query i is given by the score H_{ji} in matrix H. *Term similarities* between two terms t_i and t_j – irrespective of their associated topics – are computed in a document surrogate vector space (DVS), where each document surrogate represents a dimension, and key terms are represented as vectors. The similarities of two terms is then defined by the cosine similarity:

$$cos(\vec{t_i}, \vec{t_j}) = \frac{\vec{t_i} \cdot \vec{t_j}}{|\vec{t_i}| \cdot |\vec{t_j}|} \tag{2}$$

To compute the *term relevance* of a term t, we calculate the cosine similarity $cos(\vec{q_c}, \vec{t})$ between the query vector $\vec{q_c}$, which is the centroid of all query term vectors used for the original user query [QF93], and the term vector $\vec{t}$ in DVS. The same concept is often used for document ranking with respect to a given query [LCS97].

5. Visualization of Expanded Query Results

When performing a web search, the overall goal is to find query terms resulting in the highest possible precision. This means, the fraction of retrieved documents that are relevant for the search task should be maximized. Since users pay most of their attention to the first result page (i.e., usually, the first up to 10 results of a query) [JSS00], precision is often defined as the fraction of relevant document surrogates on the first result page for web search [KT00]. To identify the most promising query during IQE, the user has to perform the following steps: First, she needs to understand if her original query already reveals the desired information. Second, she needs to identify queries that add relevant information that is not contained in her original query. Third, she selects the most appropriate query expansion and performs the query.

We selected three set-based text visualization techniques using different visual encodings to reveal associations between key terms and queries. For our comparison, we considered set-based text visualization techniques described in Section 2 satisfying the following criteria: First, they should explicitly encode term-query relations. Second, they should scale up to at least 10 sets. Third, they should encode intersections of multiple sets. We chose visualization techniques that substantially differ in their way how they encode the information: Compact Euler Diagrams (ComED) [RD10] use a spatial layout to encode associations of key terms to sets, while Parallel Tag Clouds [CVW09] and the List View (similarly as used in JigSaw [SGL08]) use parallel lists in combination with links connecting these lists for visualizing the resulting key terms. However, Parallel Tag Clouds have a column for each query, while our List View is limited to one column listing the queries, and a second column listing the extracted topics. Therefore, the number of rows in the List View is limited to the number of queries $n+1$ and topics k, respectively, while Parallel Tag Clouds list all key terms associated with a query, so they can have up to $k \times m$ rows. In Table 1, we summarize how the key term attributes of the expanded query space are encoded by the different techniques. Below, we describe these visual mappings for each visualization technique and adaptations we made to use the selected techniques for visualizing expanded query results.

Table 1: *Visual encodings of text results and attributes.*

	Term-query association	**Term similarities**	**Term relevance**
ComED	enclosures	proximity	font size
Parallel Tag Clouds	presence in columns	–	font size
List View	links	co-occurrence by topic	–

5.1. Density-Based Compact Euler Diagrams

Euler diagrams are a natural choice to visualize set relationships [SA08]. However, as the number of sets increases, and the contained items become more numerous and large, conventional Euler diagrams and Euler-like diagrams easily become cluttered and hard to read [RD10]. Riche et al. [RD10] therefore introduced a new class of Compact Euler Diagrams (ComED), which splits sets into strict hierarchies, so that every item is associated with exactly one set enclosure. Each set is represented by one or more rectangular enclosures, depending on how often it was split. Rectangles of split sets are nested so that the maximum number of top-level rectangles is limited to the number of sets. Rectangles belonging to one set are visually linked and are assigned the same color.

The splitting algorithm by Riche et al. requires a ranking of the sets, because lower-ranked sets get split more frequently. We therefore order the queries according to the edge weight of the expansion terms in ConceptNet, and always rank the original query first. As a result, the original query will not be split, and all terms associated with the original query will be enclosed by the rectangle of the original query. The lower the edge weight of an expansion term in ConceptNet, the more likely and more often the query set will be split. As a result, lower-ranked queries will only enclose those key terms that are distinct in their resulting document surrogates in their main rectangle. This way, the user can quickly determine if the original query contains relevant key terms for the search topic. If not, the visualization shows which additional key terms the query expansions would yield. Thereby, the font size encodes the term relevance with respect to the original query.

ComEDs easily become cluttered as the number of queries increases, because of the different colors used to encode the sets, as well as the links between split query enclosures. Also, the original ComED does not take term similarities into account. We therefore introduced two variations to ComED for visualizing query expansion results: a density-based visualization of set overlaps and term similarities as attractive forces between individual key terms.

The density-based ComED visualizes only one set association explicitly – namely the one of the original query. The rectangle representing the currently displayed query is assigned a distinct color so that the user can quickly grasp the most important key terms contained in the original query (see light blue box in Figure 2). For the remaining queries, we render the rectangular enclosures, but we drop the links between split sets. This way, the user cannot immediately see which key terms are associated with a particular query, but how a query differs from other queries. All key terms associated with a query can be revealed by interactively hovering set enclosures (see Figure 2). Overlapping sets are indicated through half-transparent rectangular enclosures, so that the density of the enclosure encodes the number of queries in which a key term is contained. For instance, in Figure 2, the key terms “painting” and “ideas” are associated with five queries each. In addition, we only label top-level enclosures with the respective query terms. As a result, query expansions that do not yield distinct key terms are not revealed to the user. In the example of Figure 2, the expanded queries “rock and roll art” and “rock art painting” do not contain any unique key terms, and are therefore not labeled.

To reflect term similarities within the visualization, we construct a graph, where terms are nodes, and term similarities are weighted edges between these nodes. We use a constraint-based force-directed layout with query rectangles as constraints to place similar terms close to each other, while respecting the grouping

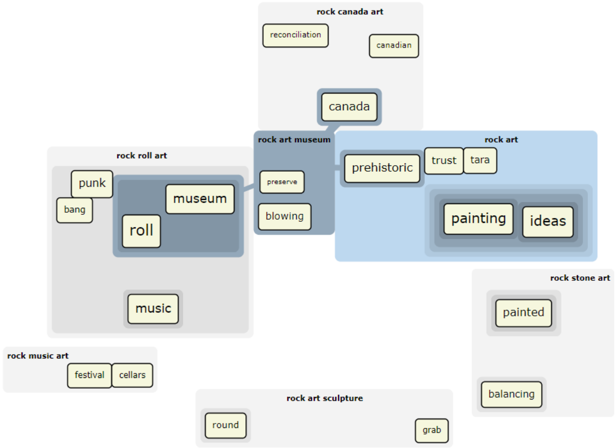


Figure 2: *Density-based Compact Euler Diagram showing results for $n = 9$ expanded queries of "rock art" with 21 key terms (from $k = 7$ topics and $m = 3$ key terms per topic). The user is hovering the expanded query "rock art museum" to reveal the key terms associated with the split query. Two of the query expansion suggestions are not revealed to the user because all their associated key terms are shared with other queries.*

given by the query enclosures. In Figure 2, this causes, for instance, key terms related to rock painting to be placed close together within the original query enclosure. In addition, expanded queries covering similar topics are in close proximity.

5.2. Parallel Tag Clouds

Parallel Tag Clouds (PTC) [CVW09] (Figure 3) combine the concepts of parallel coordinates and tag clouds. Each query is visualized as a column, where key terms associated with the query are listed alphabetically. Key terms occurring in more than one column are visually connected by link stubs. As in density-based ComED, the font size encodes the key term relevance.

Figure 3: *Parallel Tag Clouds showing results for $n = 9$ expanded queries of "rock art" with 21 unique key terms (from $k = 7$ topics and $m = 3$ key terms per topic). The user is hovering the key term "roll" to reveal the queries associated with it.*

Queries are listed from left to right with decreasing weight of their expansion terms. The left-most query always represents the original query. Compared to ComED, users can more easily scan all key terms associated with a particular query, as query sets are not split. However, depending on the similarity of the query results, this visualization can comprise a lot of duplicated key terms.

5.3. List View

A common usage of List Views in the context of text and document visualization is to link co-occurrences of extracted entities in document collections, such as in the JigSaw system [SGL08]. We can apply the same concept to visualize connections between queries and the topics extracted from the resulting document surrogates: one list shows the queries, ranked by their expansion term weights, linked to a second parallel list containing the extracted topics. The edge weights are defined by the topic-query association strengths.

Figure 4 shows the expanded "rock art" query space using the List View. The original query is always on top and highlighted by a black boundary. Hovering over a topic or a query, respectively, highlights the associated links with the adjacent list (Figure 4).

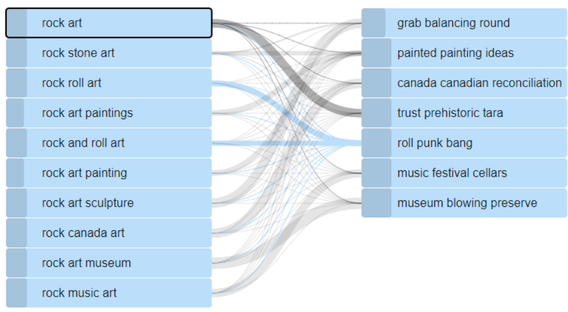


Figure 4: *List View showing results for $n = 9$ expanded queries of "rock art" with $k = 7$ topics and $m = 3$ key terms per topic. The user is hovering the topic "roll punk bang" to reveal the queries associated with it.*

The List View is more compact than the PTC and ComED representations. It also explicitly reveals the topics, as opposed to the other two visualizations, which show the topic key terms independently. This way, it allows users to quickly scan associations between topics and queries. However, without user interaction to highlight the connections, the edges can be hard to discriminate, because most queries have at least a weak association with every topic. In addition, due to the aggregation of key terms into topics, this representation does not support the identification of distinct key terms in queries.

6. Interactive Query Expansion

Visualizations of expanded query results can make use of the void display space usually available in search engines next to the search result list on large monitors, as shown in Figure 5. We allow for interactive query expansion and exploration of the query space through brushing and linking between the currently displayed search results and the visualization. In addition, users can easily select or construct expanded queries using the visualization.

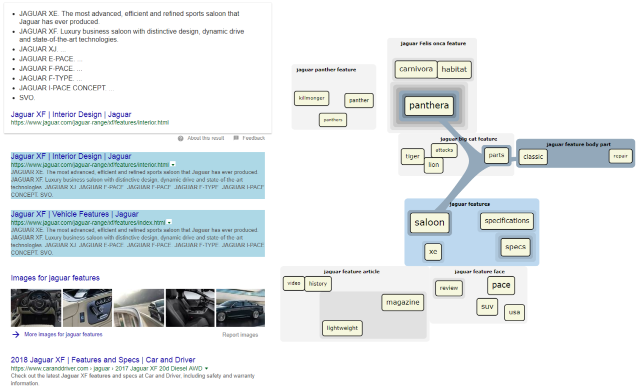


Figure 5: *Brushing and linking between the search result list and ComED of the query "jaguar features": Hovering over the term "saloon" reveals an expanded query also containing the term in ComED and the document surrogates containing the brushed term in the original query.*

As illustrated in Figure 5, brushing over a key term in the visualization highlights all search results in the list containing the selected term. Similarly, brushing over a search result highlights the associated key terms within the visualization. Expanded queries can be conducted by simply clicking their label in the visualization. In addition, by clicking on key terms, they are added to the query text box. By clicking on a topic box in the List View, all topic terms are added to the query box. This way, users can achieve a similar effect as relevance feedback [SB97], where queries are expanded by key terms extracted from result documents rated as relevant by the user.

7. Implementation

To integrate visualizations of expanded query results into a common web search engine, we implemented it as extension for the Chrome web browser. We intercept search requests for the Google search page and parse the queries from its query text box. In a background script, we query the n expansion terms and phrases using ConceptNet's REST API. For the expanded queries, document surrogates are retrieved using Google's Custom Search REST API. Document surrogates of the original query are parsed directly from the web site. Part-of-Speech selection and preprocessing of query terms, as well as of document surrogates, are performed using the JavaScript library compromise.js. For topic modeling, we use the library nmf.js.

We use d3.js to render the visualization on the search engine's result page. For the ComED visualization, we use cola.js to compute the constraint-based force-directed graph layout of the key term nodes.

8. User Study

We conducted a controlled user study to compare users' effectiveness and efficiency when selecting query expansion suggestions using the three different visualization techniques. In addition, we added a baseline condition, listing the query expansion suggestions without a visualization of the query expansion results, similar to query expansion suggestions provided by popular web search engines. Our major goal was to test whether a visualization of query expansion results helps users to quickly and accurately decide whether a query expansion can improve the precision of their search results. We believed that the spatial organization and enclosures used by ComED would make it easiest for the users to see which query expansions lead to distinctive results. We had the following hypotheses:

H1: Using any visualization, users will select queries leading to higher precision and recall than without a visualization.

H2: With density-based ComED, users will decide most quickly which query to select.

H3: Overall, users will prefer a visualization of the query results space over just having a text list of query expansion suggestions.

8.1. Tasks

We asked users to perform a set of ambiguous queries from the web tracks of the Text REtrieval Conference (TREC). These collections of query topics consist of queries logged by commercial search engines, together with verbose sub-topic descriptions [CCS09]. TREC web topics are classified as "faceted" or "ambiguous", where faceted topics are underspecified and can reveal multiple sub-topics. More interesting for us, ambiguous topics comprise those queries that allow for multiple interpretations.

We parsed all ambiguous topics in all TREC web tracks from 2009 to 2014. We only picked those topics that fulfilled two criteria: First, the query should lead to at least five query expansion suggestions using ConceptNet. Second, there should be at least five relevant document surrogates per sub-topic in the entire expanded query result space. The TREC web collections contain many abbreviations or movie names as ambiguous topics, from which many cannot be resolved through ConceptNet. We finally had a selection of 10 topics, each with two different sub-topic descriptions. From the 20 sub-topic descriptions, we slightly re-phrased four to broaden the search scope. Among the 10 topics, there was only a single abbreviation that could be disambiguated to two different sub-topics using ConceptNet ("pvc"). The selected topics consist of one or two query terms. For such short queries, query expansion is more likely to beneficial than for longer, more well-defined queries [Voo94]. Table 2 shows a selection of six sub-topics used for the experiment.

Table 2: *Example sub-topics from TREC ambiguous web queries.*

Query	Sub-Topic Description
kiwi	Find information on kiwi fruit. Find information on kiwi birds.
rock art	Where can I learn about rock painting or buy a rock-painting kit? Find information on cave paintings all around the world.
iron	Find information about iron as an essential nutrient. Find information about the element iron (Fe).

We set the number of query expansions to $n = 9$ so that for each sub-topic, we could guarantee a precision of at least 0.5 for at least one query. This is also a similar number of query expansion suggestions provided by common search engines. However, for "pvc", we only could find six possible expansion terms in ConceptNet. We retrieved the document surrogates in an offline process through the Google Custom Search API prior to the experiment to ensure consistent results across participants. Each document surrogate was manually labeled as relevant or not relevant for the given sub-topic description. From the 20 sub-topics, the highest relevance score could be achieved by the original query in four cases. For the remaining 80% of the sub-topics, one of the expanded queries leads to higher precision.

We manually assigned the 20 individual sub-topics to four task sets, where sub-topics with the same query terms were always assigned to separate task sets. The presentation order of sub-topics within a task set was randomized.

8.2. Apparatus and Procedure

The study was conducted using the Google Chrome web browser on a 27" monitor. Users had to fill in a consent form, followed by a demographic questionnaire, and then they were asked to read a printed task description. The presentation order of the four interface conditions, as well as the assignment of the four task sets to the interfaces, was balanced using a Graeco-Latin Square design. For each sub-topic, we displayed the query, as well as the description (see Table 2), and asked users to read out the description aloud. After pressing the "Query"-button below, the Google page was called with the respective query string. As shown in Figure 5, the visualization was displayed next to the document surrogates. Interaction capabilities, as described in Section 6, were enabled but not used by any participant.

For each interface condition, there was a warm-up task consisting of two queries ("spider" and "jaguar") to get familiar with the interface. After the warm-up task, users performed five sub-topics within a task set, before proceeding to the next interface condition and task set, respectively. After the experiment, we asked users to rate the overall preference of the four interface conditions for solving the tasks on a five-point Likert-scale, as well as to list positive and negative aspects of the interfaces.

8.3. Design

We used a within-subjects design with query expansion interface as independent variable: Compact Euler Diagrams (euler), Parallel Tag Clouds (PTC), List View (lists), and text list without visualization (text). We logged the task completion time (TCT), the selected query, and overall preference ratings from the post-study questionnaire. We measured the task completion time from pressing the "Query"-button to selecting the target query using the provided interface. The number of relevant documents were read from the manually labeled relevance scores for each selected query and sub-topic, respectively (see Section 8.1). For each response, we then computed precision as the number of relevant documents divided by the number of document surrogates in the selected query, and recall as the number of relevant documents of the selected query divided by the overall number of relevant documents of all queries for the sub-topic.

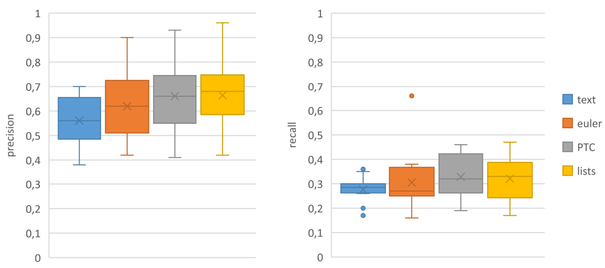
For all three visualization techniques, we set the number of topics k to the number of queries (i.e., 7 for "pvc" and 10 for the remaining tasks). For each topic, we selected the $m = 5$ top-ranked key terms. We empirically chose the number of key terms to ensure inclusion of expressive key terms while keeping visual clutter as low as possible.

8.4. Participants

In the study, 16 users participated (five female, 11 male), aged 25 to 40. All users except one have a background in computer science and use online search engines on a daily basis. 14 out of the 16 reported to be familiar with simple visualization techniques, like bar charts or pie charts. Eight users stated to use dynamic query suggestions to expand queries while typing frequently or very frequently, while only one user never uses this feature. However, the textual query expansion suggestions presented at the search engines' results page is used rarely or never by around 70% of the users.

8.5. Results

Before performing the statistical tests, we removed 18 outlier samples and aggregated all TCT, precision and recall measures per user and visualization by average. To test hypothesis H1, we compared precision and recall of the user-selected queries between the four interface conditions. A repeated measures ANOVA showed that there is no significant difference for precision, and the effect is medium ($F(3,45) = 1.992; p = .129; \eta^2 = .117$). There is also no difference for recall between the conditions and only a small effect ($F(3,45) = .919; p = .439; \eta^2 = .058$). Figure 6 shows box plots of precision and recall per condition. *We therefore have to reject H1: A visualization of expanded query results does not significantly increase precision or recall for expanded query selections.*



(a) (b)

Figure 6: *Precision (a) and recall (b) per condition.*

Mind, however, how expanding the query could significantly improve the overall precision: While the average precision of the original query for the sub-topics was 0.37, users could achieve an average precision of 0.63 across all conditions in our study (Mann-Whitney U test: $Z = -3.111; p = .002$).

To test hypothesis H2, we compared the time between sending the original query and selecting the target query between the four conditions. Since the repeated measures ANOVA showed a large and significant effect ($F(3,45) = 34.923; p < .001; \eta^2 = .70$), we performed pair-wise Bonferroni-correct post-hoc comparisons. These post-hoc comparisons showed that there is a significant difference between text and the three visualization conditions. Contrary to our hypothesis, users selected the target query significantly faster using the common text list (14 seconds on average) than any of the visualizations (32 seconds on average), as visualized in Figure 7. Informally, we could observe that users spent considerable time carefully parsing the visualizations rather than quickly scanning them. *We therefore also have to reject H2: Deciding which query to select is significantly faster without a visualization.*

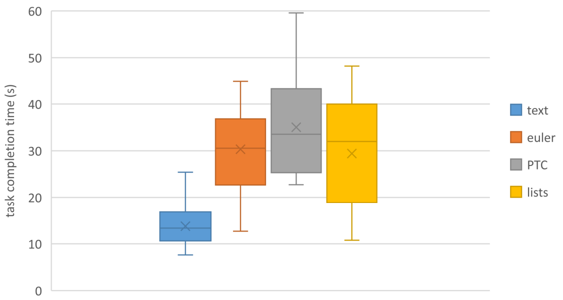


Figure 7: *Task completion time per condition.*

Finally, we compared users' preference ratings of the four conditions to test hypothesis H3. A Friedman test showed a significant difference between the ratings of the four interfaces ($\chi^2(3) = 15.396; p = .002$). Bonferroni-corrected pair-wise Wilcoxon Signed-Rank post-hoc comparisons revealed a significant difference between PTC and text, as well as between PTC and lists. While PTC received an average score of 2.1, text and lists were consistently rated higher on average (3.7), as illustrated in Figure 8. *This also disproves our hypothesis H3: Users do not prefer visualizations to choose query expansion suggestions. They prefer a simple text list and a list view of queries linked to result topics over Parallel Tag Clouds.*

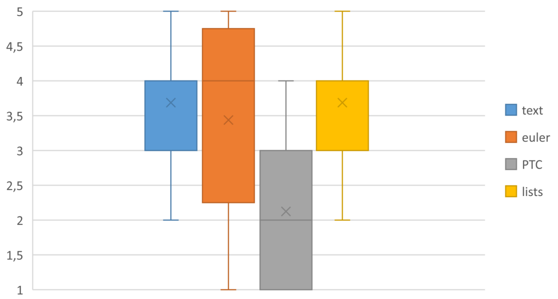


Figure 8: *User preference ratings for solving the task using the four interfaces on a five-point Likert scale.*

To explore reasons for these findings, we performed open coding on the users' feedback given for the post-experiment questionnaire. We grouped utterances into six categories: speed, ease of use, clarity, expressiveness, appearance, and learnability. Furthermore, each utterance was assigned a positive or negative sentiment. PTC received the highest number of negative utterances in the category clarity, where users wrote comments like *"hard to read"* or *"hard to make connections"*. Lists received most positive utterances in the categories appearance and expressiveness, with comments like *"interesting information about overlaps"*. Text received most positive utterances in the categories speed and clarity. For instance, users mentioned that it was *"easy to scan quickly"* or *"fast to find a query"*. On the other hand, text received the highest number of negative utterances in the category expressiveness. Users commented that the interface was *"not very detailed"* or *"I couldn't actually see what each query meant"*.

8.6. Discussion

Our study showed that visualizing expanded query results does not make it easier for users to decide which expansion suggestion to select for query disambiguation. But why does seeing additional text information about expanded query results not improve precision and efficiency? To answer this question, we will first discuss the strengths of the simple text list. While five users mentioned the lack of expressiveness as negative aspect of the text list, this aspect did not have a significantly negative impact on the precision of their selection. In fact, in most cases, the query expansion terms themselves were already descriptive enough for the users to make a decision. One such example is shown in Table 3: The highest precision could be achieved by selecting the query expansion suggestion containing the phrase of the sub-topic description (last row). Indeed, from our 16 users, 14 selected this query.

Table 3: *Query expansion suggestions for the sub-topic "How are premature ventricular contractions treated?".*

Query	Precision
pvc	0.14
pvc polyvinyl chloride	0.0
pvc iv catheter	0.0
pvc extrasystole	0.5
pvc abs	0.0
pvc artificial substance	0.0
pvc premature ventricular contraction	1.0

On the other hand, the second sub-topic description of the query "pvc" was to find information about pipes and fittings, where the highest precision could be achieved using the query "pvc abs". In Figure 9, it can be seen that there is only one topic that contains the term "fitting", which is connected to the best query suggestion. No participant using the text list for this sub-topic selected this query. However, even when having a visualization of the expanded query results, only a third of the users selected the best query.

What was consistently mentioned as negative aspect of all visualization types was the lack of clarity. Users, for instance, criticized that the Parallel Tag Clouds had *"too many words to read"* and found Compact Euler Diagrams *"crowded"*. This negative feedback was especially pronounced for Parallel Tag Clouds, which had the highest number of key terms in the visualization due to duplications. The information gain of showing the key terms of the results apparently did not outweigh the added cost of scanning the

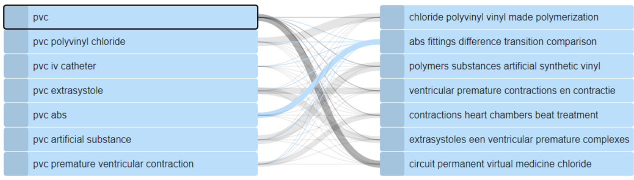


Figure 9: *Selecting the best query expansion ("pvc abs") for the sub-topic "Find information about PVC pipes and fittings" using the List View.*

additional text. To reduce this cost, the number of topics (k) or the number of key terms per topic (m) need to be reduced. However, to ensure expressiveness of the visualization, the expressiveness of key terms has to be high – a topic, which is still undervalued in text visualization research [CMH12]. According to Chuang et al. [CMH12], higher expressiveness can be achieved, for instance, by grouping similar key terms into bigrams.

Another way to reduce the amount of text is to decrease the number of expanded queries (n). However, for some of the sub-topics in our experiment, expansion terms that were ranked rather low according to ConceptNet revealed the highest precision. Reducing the number of query expansions therefore increases the risk of missing relevant query variations.

Alternatively, key terms can be substituted or enhanced by images – similarly as proposed by Zha et al. [ZYM*09] for lists of query expansion suggestions. Figure 10 shows a density-based ComED with a reduced number of key terms, but instead including one image result for each top-level enclosure. To query the images, we used the query terms of the enclosure's associated query label, as well as all enclosed key terms. Mind that the original query in Figure 10 covers multiple sub-topics, while the single image result illustrates only the material sub-topic. Also mind that three images show pipes, because this is one of the most common usages of PVC. However, two of these queries only reveal a small number of documents actually discussing pipes and fittings specifically. This illustrates that picking expressive images for a given document content is also a challenging task.

Notably, there have been previous studies in the context of search result visualization that also could not show an improvement of search performance. Hornbæk et al. [HF99], as well as Reiterer et al. [RTM05], added visualizations to text-based search result lists, showing either a thematic map of the retrieved documents [HF99] or document relevance for the given query terms, for instance as bar chart [RTM05]. In both cases, adding a visualization could not improve retrieval performance. Hornbæk et al. observed that the visualization was sometimes distracting and misinterpreted, and that labels of document clusters were not always understandable.

On the other hand, some visual enhancements of search results have been shown to improve search performance over classic lists of document surrogates: HotMap [HY06] and AspecTiles [ISY*12] visualize document-query associations and document-topic associations, respectively, as simple glyph visualizations attached to document surrogates. Using both interfaces, users were more likely to select relevant documents from the result list. What these two ex-

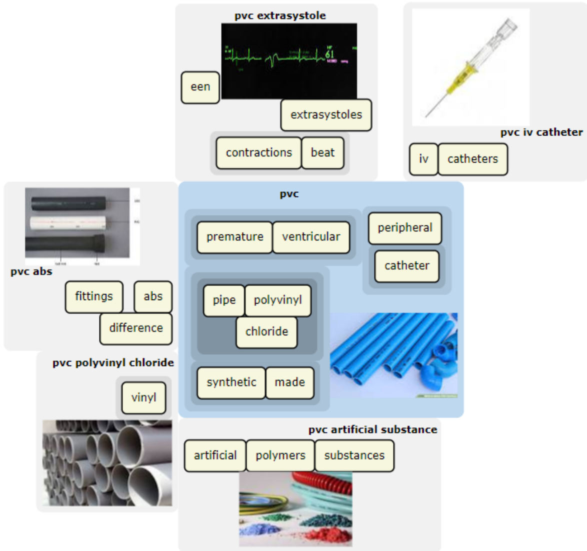


Figure 10: *ComED showing image results for every top-level enclosure with $k = 6$ topics and $m = 4$ key terms per topic.*

amples have in common is that their visualization is reduced to a minimum and is not detached from the document surrogates. Exploring how visualizations can seamlessly integrate into the result list, which is well-known to the users, therefore could be a way to increase their acceptance and their effectiveness. Also, simplifying the visual encodings used for this study, for instance by removing weak edges from the List View or putting more emphasis on the query terms unique to a query rather than shared between queries in the Parallel Tag Clouds, could already have a positive effect.

8.7. Limitations

A limitation of our study is that, due to the repetitive and controlled experiment design similar to previous work in the field [ISY*12], we could not evaluate whether users are more likely to expand a query when provided with a visualization of the expanded query results. Users were rather "forced" to consider all query options, regardless of the interface. Also, we asked users to select a query suggestion, but did not evaluate the "relevance feedback" feature, where users interactively expand their query with document key terms (see Section 6). To evaluate the usage frequency and expansion strategies, longer-term field studies, logging the users' search behaviors, would be necessary.

Furthermore, our study was limited to ambiguous queries from the TREC database. Therefore, we cannot infer whether visualization of expanded query results can increase precision when facing term mismatches, or increase the desired learning effect when performing exploratory search. In addition, the queries of the TREC topics were rather simple. However, since the TREC query topics were mined from real web search engines, we can assume that they are representative for ambiguous queries.

Finally, we used a homogeneous user sample in our study, consisting of knowledge workers using search engines on a daily basis and being used to regularly having to adapt to new user interfaces. Also, the self-reported level of visualization literacy was quite high. We therefore cannot generalize our results to a broader group of users. However, we can assume that efficiency of users not being used to working with a wide range of user interfaces may decrease even more, as the complexity of the interface increases.

9. Conclusions

In this paper we analyzed the usage of set-based text visualization techniques for facilitating the selection of query expansion suggestions. We presented the preprocessing pipeline to construct such visualizations, and how the Compact Euler Diagram representation was modified for clutter reduction and better disambiguation between sub-topics within a single query and across queries. In our study, we showed that visualizations cannot significantly improve the decision quality which query to pick, but require more time to come up with the decision.

However, this does not mean that visualization of expanded query results is generally not beneficial. Query expansion is not only useful to resolve disambiguities, but can also be a valuable tool for resolving term mismatches (Figure 11) – a use case we have not evaluated in the course of our study. This can be especially useful when using domain-specific ontologies rather than a general thesaurus to perform more in-depth exploratory search. Other usage scenarios of set-based text visualization techniques in the context of web search are comparisons of results of different search engines and visualizing a user's personal search history [HHP*01] to support learning during exploratory search.

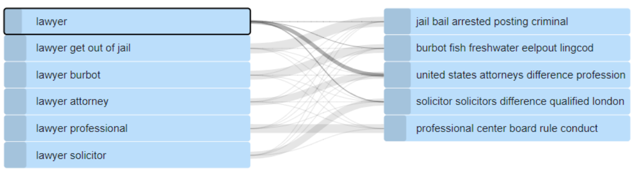


Figure 11: *List View showing results for an expansion of the query "lawyer". The associated key terms reveal that the synonymous term "attorney" is more common in the US, while "solicitor" is more common in the UK. Furthermore, it reveals a polysemy of the term "lawyer", which is also a name for a fish species.*

In our study, users were less efficient when more text information was shown. For the field of text and document visualization research, it will therefore be important to establish guidelines how much text information should be encoded in visualizations to balance the trade-off between visualization effectiveness and expressiveness. Additionally, supplementing or substituting text information with images could increase the effectiveness of a visualization, if users can correctly interpret the image on a single glance. Choosing whether it is more expressive to use text- or image-based labels is therefore also a future topic to be explored.

In addition, user feedback indicates that many users prefer a clean and aligned layout and find non-aligned layouts or non-orthogonal links visually unpleasing – at least when arranged next to a strictly aligned list of search results. Aggregating topic terms into one line of text could have been a reason why users rated the List View higher than Parallel Tag Clouds, which contained a lot of single key terms. This is also in accordance with earlier studies, showing search performance improvements for very simple glyph-based visualizations [HY06, ISY*12], but no improvements for more complex add-on visualizations [HF99, RTM05]. For in-the-wild usage of visualizations, like for visualizing search results, a strong focus on aesthetics principles, such as in the graph drawing community [BRSG07], and reduction to simple visual encodings seem to be important.

Acknowledgments

This work was financed by the Austrian Science Fund (FWF): T 752-N30. We thank Michael Gusenbauer for valuable discussions, as well as Lukas Eibensteiner and Manuel Kapferer for an early implementation of the List View.

References

[AMA*14] ALSALLAKH B., MICALLEF L., AIGNER W., HAUSER H., MIKSCH S., RODGERS P.: Visualizing sets and set-typed data: State-of-the-art and future challenges. In *Eurographics conference on Visualization (EuroVis)–State of The Art Reports* (2014), pp. 1–21. 2

[BLB*14] BURCH M., LOHMANN S., BECK F., RODRIGUEZ N., DI SILVESTRO L., WEISKOPF D.: Radcloud: Visualizing multiple texts with merged word clouds. In *Information Visualisation (IV), 2014 18th International Conference on* (2014), IEEE, pp. 108–113. 2

[BRSG07] BENNETT C., RYALL J., SPALTEHOLZ L., GOOCH A.: The aesthetics of graph visualization. *Computational aesthetics 2007* (2007), 57–64. 10

[CCS09] CLARKE C. L., CRASWELL N., SOBOROFF I.: *Overview of the trec 2009 web track*. Tech. rep., WATERLOO UNIV (ONTARIO), 2009. 6

[CFN07] CHIRITA P.-A., FIRAN C. S., NEJDL W.: Personalized query expansion for the web. In *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval* (2007), ACM, pp. 7–14. 1

[CLRP13] CHOO J., LEE C., REDDY C. K., PARK H.: Utopian: User-driven topic modeling based on interactive nonnegative matrix factorization. *IEEE transactions on visualization and computer graphics 19*, 12 (2013), 1992–2001. 3

[CMH12] CHUANG J., MANNING C. D., HEER J.: "without the clutter of unimportant words": Descriptive keyphrases for text visualization. *ACM Transactions on Computer-Human Interaction (TOCHI) 19*, 3 (2012), 19. 9

[CR12] CARPINETO C., ROMANO G.: A survey of automatic query expansion in information retrieval. *ACM Computing Surveys (CSUR) 44*, 1 (2012), 1. 1

[CTZC02] CRONEN-TOWNSEND S., ZHOU Y., CROFT W. B.: Predicting query performance. In *Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval* (2002), ACM, pp. 299–306. 1

[CVW09] COLLINS C., VIEGAS F. B., WATTENBERG M.: Parallel tag clouds to explore and analyze faceted text corpora. In *Visual Analytics Science and Technology, 2009. VAST 2009. IEEE Symposium on* (2009), IEEE, pp. 91–98. 2, 4, 5

[DCCW08] DÖRK M., CARPENDALE S., COLLINS C., WILLIAMSON

C.: Visgets: Coordinated visualizations for web-based information exploration and discovery. *IEEE Transactions on Visualization and Computer Graphics 14*, 6 (2008). 2

[dSSV15] DI SCIASCIO C., SABOL V., VEAS E.: urank: Visual analytics approach for search result exploration. In *Visual Analytics Science and Technology (VAST), 2015 IEEE Conference on* (2015), IEEE, pp. 217–218. 2

[FFW91] FOWLER R. H., FOWLER W. A., WILSON B. A.: Integrating query thesaurus, and documents through a common visual representation. In *Proceedings of the 14th annual international ACM SIGIR conference on Research and development in information retrieval* (1991), ACM, pp. 142–151. 1, 2

[FLGD87] FURNAS G. W., LANDAUER T. K., GOMEZ L. M., DUMAIS S. T.: The vocabulary problem in human-system communication. *Communications of the ACM 30*, 11 (1987), 964–971. 1

[GNSRP*14] GOMEZ-NIETO E., SAN ROMAN F., PAGLIOSA P., CASACA W., HELOU E. S., DE OLIVEIRA M. C. F., NONATO L. G.: Similarity preserving snippet-based visualization of web search results. *IEEE transactions on visualization and computer graphics 20*, 3 (2014), 457–470. 2

[Goo17a] GOOGLE: Custom Search JSON/Atom API. https://developers.google.com/custom-search/json-api/v1/overview, 2017. [Online; accessed Sep-2017]. 3

[Goo17b] GOOGLE: Search using autocomplete. https://support.google.com/websearch/answer/106230, 2017. [Online; accessed Oct-2017]. 1

[Har88] HARMAN D.: Towards interactive query expansion. In *Proceedings of the 11th annual international ACM SIGIR conference on Research and development in information retrieval* (1988), ACM, pp. 321–331. 2

[Hea95] HEARST M. A.: Tilebars: visualization of term distribution information in full text information access. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (1995), ACM Press/Addison-Wesley Publishing Co., pp. 59–66. 2

[Hea11] HEARST M.: User interfaces for search. *Modern Information Retrieval* (2011), 21–55. 2

[HF99] HORNBÆK K., FRØKJÆR E.: Do thematic maps improve information retrieval? In *Interact* (1999), pp. 179–186. 9, 10

[HHP*01] HAVRE S., HETZLER E., PERRINE K., JURRUS E., MILLER N.: Interactive visualization of multiple query results. In *Proceedings of the IEEE Symposium on Information Visualization 2001 (INFOVIS'01)* (2001), IEEE Computer Society, p. 105. 2, 10

[HS17] HASITSCHKA P., SABOL V.: Visual exploration and analysis of recommender histories: A web-based approach using webgl. In *Proceedings of the 2017 ACM Workshop on Exploratory Search and Interactive Data Analytics* (2017), ACM, pp. 33–40. 2

[HY06] HOEBER O., YANG X. D.: The visual exploration of web search results using hotmap. In *Information Visualization, 2006. IV 2006. Tenth International Conference on* (2006), IEEE, pp. 157–165. 2, 9, 10

[HYY05] HOEBER O., YANG X.-D., YAO Y.: Visualization support for interactive query refinement. In *Web Intelligence, 2005. Proceedings. The 2005 IEEE/WIC/ACM International Conference on* (2005), IEEE, pp. 657–665. 1, 2

[ISY*12] IWATA M., SAKAI T., YAMAMOTO T., CHEN Y., LIU Y., WEN J.-R., NISHIO S.: Aspectiles: Tile-based visualization of diversified web search results. In *Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval* (2012), ACM, pp. 85–94. 2, 9, 10

[JSS00] JANSEN B. J., SPINK A., SARACEVIC T.: Real life, real users, and real needs: a study and analysis of user queries on the web. *Information processing & management 36*, 2 (2000), 207–227. 4

[KK15] KUCHER K., KERREN A.: Text visualization techniques: Taxonomy, visual survey, and community insights. In *Visualization Symposium (PacificVis), 2015 IEEE Pacific* (2015), IEEE, pp. 117–121. 2

[KKEE11] KIM K., KO S., ELMQVIST N., EBERT D. S.: Wordbridge: Using composite tag clouds in node-link diagrams for visualizing content and relations in text corpora. In *System Sciences (HICSS), 2011 44th Hawaii International Conference on* (2011), IEEE, pp. 1–8. 2

[KT00] KOBAYASHI M., TAKEDA K.: Information retrieval on the web. *ACM Computing Surveys (CSUR) 32*, 2 (2000), 144–173. 4

[KTZ*07] KOZANIDIS L., TZEKOU P., ZOTOS N., STAMOU S., CHRISTODOULAKIS D.: Ontology-based adaptive query refinement. In *WEBIST (2)* (2007), pp. 43–50. 1, 2

[LCS97] LEE D. L., CHUANG H., SEAMONS K.: Document ranking and the vector-space model. *IEEE software 14*, 2 (1997), 67–75. 4

[LHB*15] LOHMANN S., HEIMERL F., BOPP F., BURCH M., ERTL T.: Concentri cloud: Word cloud visualization for multiple text documents. In *Information Visualisation (iV), 2015 19th International Conference on* (2015), IEEE, pp. 114–120. 2

[LKW09] LU Z., KIM W., WILBUR W. J.: Evaluation of query expansion using mesh in pubmed. *Information retrieval 12*, 1 (2009), 69–80. 1

[LS04] LIU H., SINGH P.: Conceptnet – a practical commonsense reasoning tool-kit. *BT technology journal 22*, 4 (2004), 211–226. 3

[OKS*93] OLSEN K. A., KORFHAGE R. R., SOCHATS K. M., SPRING M. B., WILLIAMS J. G.: Visualization of a document collection: The vibe system. *Information Processing & Management 29*, 1 (1993), 69–81. 2

[OSR*14] OELKE D., STROBELT H., ROHRDANTZ C., GUREVYCH I., DEUSSEN O.: Comparative exploration of document collections: a visual analytics approach. In *Computer Graphics Forum* (2014), vol. 33, Wiley Online Library, pp. 201–210. 2

[PBR17] PELTONEN J., BELORUSTCEVA K., RUOTSALO T.: Topic-relevance map: Visualization for improving search result comprehension. In *Proceedings of the 22nd International Conference on Intelligent User Interfaces* (2017), ACM, pp. 611–622. 2

[PBW07] PHAN N., BAILEY P., WILKINSON R.: Understanding the relationship of information need specificity to search query length. In *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval* (2007), ACM, pp. 709–710. 1

[QF93] QIU Y., FREI H.-P.: Concept based query expansion. In *Proceedings of the 16th annual international ACM SIGIR conference on Research and development in information retrieval* (1993), ACM, pp. 160–169. 4

[RD10] RICHE N. H., DWYER T.: Untangling euler diagrams. *IEEE Transactions on Visualization and Computer Graphics 16*, 6 (2010), 1090–1099. 2, 4

[RTM05] REITERER H., TULLIUS G., MANN T. M.: Insyder: a content-based visual-information-seeking system for the web. *International Journal on Digital Libraries 5*, 1 (2005), 25–41. 2, 9, 10

[Rut03] RUTHVEN I.: Re-examining the potential effectiveness of interactive query expansion. In *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval* (2003), ACM, pp. 213–220. 1

[SA08] SIMONETTO P., AUBER D.: Visualise undrawable euler diagrams. In *Information Visualisation, 2008. IV'08. 12th International Conference* (2008), IEEE, pp. 594–599. 4

[SB97] SALTON G., BUCKLEY C.: Improving retrieval performance by relevance feedback. *Readings in information retrieval 24*, 5 (1997), 355–363. 1, 6

[SGL08] STASKO J., GÖRG C., LIU Z.: Jigsaw: supporting investigative analysis through interactive visualization. *Information visualization 7*, 2 (2008), 118–132. 2, 4, 5

[SJ72] SPARCK JONES K.: A statistical interpretation of term specificity and its application in retrieval. *Journal of documentation 28*, 1 (1972), 11–21. 3

[SLW*07] SONG R., LUO Z., WEN J.-R., YU Y., HON H.-W.: Identifying ambiguous queries in web search. In *Proceedings of the 16th international conference on World Wide Web* (2007), ACM, pp. 1169–1170. 1

[VGD06] VIÉGAS F. B., GOLDER S., DONATH J.: Visualizing email content: portraying relationships from conversational histories. In *Proceedings of the SIGCHI conference on Human Factors in computing systems* (2006), ACM, pp. 979–988. 2

[Voo94] VOORHEES E. M.: Query expansion using lexical-semantic relations. In *Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval* (1994), Springer-Verlag New York, Inc., pp. 61–69. 1, 6

[ZYM*09] ZHA Z.-J., YANG L., MEI T., WANG M., WANG Z.: Visual query suggestion. In *Proceedings of the 17th ACM international conference on Multimedia* (2009), ACM, pp. 15–24. 2, 9

[ZYT11] ZHANG W., YOSHIDA T., TANG X.: A comparative study of tf* idf, lsi and multi-words for text classification. *Expert Systems with Applications 38*, 3 (2011), 2758–2765. 3