Intentional Query Suggestion: Making User Goals More Explicit During Search

Markus Strohmaier
Graz University of Technology and Know-Center
Inffeldgasse 21a
8010 Graz, Austria
markus.strohmaier@tugraz.at

Mark Kröll
Graz University of Technology
Inffeldgasse 21a
8010 Graz, Austria
mkroell@tugraz.at

Christian Körner
Graz University of Technology
Inffeldgasse 21a
8010 Graz, Austria
christian.koerner@student.tugraz.at

ABSTRACT
The degree to which users' make their search intent explicit can be assumed to represent an upper bound on the level of service that search engines can provide. In a departure from traditional query expansion mechanisms, we introduce *Intentional Query Suggestion* as a novel idea that is attempting to *make users' intent more explicit* during search. In this paper, we present a prototypical algorithm for Intentional Query Suggestion and we discuss corresponding data from comparative experiments with traditional query suggestion mechanisms. Our preliminary results indicate that intentional query suggestions 1) diversify search result sets (i.e. it reduces result set overlap) and 2) have the potential to yield higher click-through rates than traditional query suggestions.

Categories and Subject Descriptors
H.1.2 [**User/Machine Systems**]: Human Factors; H.3.3 [**Information Storage and Retrieval**]: Query Formulation, Search Process, Retrieval Models

General Terms
Algorithms, Human Factors, Experimentation

Keywords
Query Suggestion, User Intent

1. INTRODUCTION
In IR literature, the purpose of query suggestion has often been described as the process of making a user query resemble more closely the documents it is expected to retrieve ([26]). In other words, the goal of query suggestion is commonly understood as maximizing the similarity between query terms and expected documents. The task of a searcher then is to envision the expected documents, and craft queries that reflect their contents.

However, research on query log analysis suggests that many queries exhibit a lack of user understanding about the specific documents users expect to retrieve. Broder [6] found that only ~25% of queries have a clear navigational intent, and up to ~75% of queries need to be understood as informational or transactional queries, meaning they are not directed towards a specific set of expected documents. Recent studies even estimate more drastic ratios [13]. While users crafting informational or transactional search queries often have a high level search intent ("plan a trip to Europe"), in many situations they have no clear idea or knowledge about the specific documents they expect to retrieve. This makes it difficult for users to craft successful queries and makes query suggestion a particularly important and challenging problem.

In this paper we are interested in exploring the following question: What if search engines would, rather than letting users guess arbitrary words from the set of documents they are expected to retrieve, encourage users to tell them their original search intent in a more unambiguous and natural way? In other words, what if search engines would encourage users to make their search intent more explicit (e.g. "buy a car") rather than formulating their query in a rather artificial manner ("car dealership")? In future search interfaces (such as audio search interfaces for cell phones or natural language search interfaces), current mechanisms for query suggestion might become inadequate and natural language search queries might play a more important role. This work is interested in understanding how current search methods would cope with such a development.

For this purpose, we introduce and study a novel approach to query suggestion: *Intentional Query Suggestion* or query suggestion by user intent. While traditional query suggestion often aims to make a query resemble more closely the documents a user is expected to retrieve (which might be unknown to the user), we want to study an alternative: expanding queries to make searchers' intentions more explicit.

To give an example: In traditional query suggestion, a query "car" might receive the following suggestions: "car rental", "car insurance", "enterprise car rental", "car games" (actual suggestions produced by Yahoo.com on Nov 27th 2008). In query suggestion based on explicit user intent, the suggestions could be "buy a car", "rent a car", "sell your car", "repair your car" (see Table 1 for examples). We can speculate that in innovative search interfaces (such as audio search interfaces), such suggestions would be easier to verify with a user than verifying traditional query suggestions (e.g. "Do you want to: buy a car OR sell a car OR ...?").

Table 1: Comparison of suggested queries provided by Yahoo!, MSN and Intentional Query Suggestion.

Initial Query	*Semantic Query Suggestion*[1]	*Semantic Query Suggestion*[2]	*Intentional Query Suggestion*
car	car rental, car insurance, enterprise car rental, car games	used cars, new cars, 2007 new cars, used cars for sale, cars for sale, fast cars, classic cars, car games	buy a car, rent a car, sell your car, repair your car
poker	online poker, poker games, world series of poker, party poker, free poker	free online poker, full tilt poker, free poker games, free poker, poker rules, absolute poker, online poker, poker hands	cheating at poker, learn to play poker, buy poker table, design your own poker chips
house	house plans, white house, house of fraser, columbia house, house of blues, full house	house TV show, houses for sale, houses for rent, house plans, house MD, house fox, haunted houses, Hugh Laurie	insure my house, sell your house, make offer on house, buy house online, build my own house

We are interested in studying the effects of this idea on the search result sets obtained from experiments with a current search engine provider. In particular, we are interested in seeking answers to the following questions: *How do today's search engines deal with queries that contain explicit user goals? How would queries expanded by user intent influence search results and click through?*

This paper introduces Intentional Query Suggestion as a novel type of query suggestion. Specifically, this paper 1) introduces a definition of Intentional Query Suggestion 2) presents a preliminary algorithm to perform intentional query suggestion based on historic query log data and 3) discusses experimental results and potential implications for future research on search interfaces.

2. QUERY SUGGESTION

The general idea of query suggestion is to support the searcher in formulating queries that have a better chance to retrieve relevant documents [21], [3]. Methods offered to expand queries can be divided into two major categories. Global methods employ entire document collections or external sources such as thesauri as corpora for producing suggestions. Local methods reformulate the initial query based on the result set it has retrieved. Relevance feedback represents another query reformulation strategy in which a searcher is involved by marking retrieved documents as relevant or not. Global as well as local methods aim to eventually move the initial query closer to the entire cluster of relevant documents.

2.1 Intentional Query Suggestion

While traditional query suggestion techniques aim at narrowing the gap between the initial query and the set of relevant documents, we seek to approximate the user's intentions behind a query and expand it based on a better understanding of the corresponding information need – thereby aiming to *make user intent more explicit*.

We define *Intentional Query Suggestion* as the incremental process of transforming a query into a new query based on intentional structures found in a given domain, in our case: a search query log. An initial query is replaced by the most probable intentions that underlay the query. To give an example: for the query "playground mat", an Intentional Query Suggestion mechanism might suggest the following 5 user intentions: "buy playground equipment", "build a swing set", "covering dirt in a playground", "buy children plastic slides", "raise money for our playground".

In our case, we extract the proposed intentions from search query logs, but they could potentially be extracted from other knowledge bases containing common human goals as well, such as ConceptNet [19] or others. To the best of our knowledge, the application of explicit search intent [24] to query suggestion represents a novel idea that has not been studied yet.

3. EXPERIMENTAL SETUP

In traditional query suggestion, an initial query formulation is replaced by some other query that refines, disambiguates or clarifies the original query. In our approach, the initial query is replaced by a query that exhibits a higher degree of intentional explicitness, meaning that it makes user intent more explicit [24].

Definition: We define this replacement as *query suggestion based on user intent*. The suggested queries can be considered to represent Intentional Query Suggestions whenever they *1) contain at least one verb and 2) describe a plausible state of affairs that the user may want to achieve or avoid (cf.) in 3) a recognizable way.*

We developed a parametric algorithm that executes the function $f(q) \rightarrow R_{QE} = \{q_{e,1}, q_{e,2} \ldots q_{e,k}\}$, mapping *implicit* intentional queries (length <= 2) to a set of potential *explicit* intentional query suggestions (e.g. "car" → "buy a car", "rent a car", "repair my car").

3.1 Datasets

The MSN Search query log excerpt contains about 15 million queries (from US users) that were sampled over one month in May, 2006. The search query log data is split into two files, one file containing attributes Time, Query, QueryID and ResultCount, the other one attributes QueryID, Query, Time, URL and Position providing click-through data. The queries were modified via the following normalization steps (i) trimming of each query, and (ii) space sequence reduction to one space character. Queries and corresponding click-through data containing adult content were filtered out (and were not taken into account in our study).

A set of ~46.000 explicit intentional queries was extracted from the MSN Search Asset Data Spring 2006 applying the algorithm described in [24]. The resulting set has an estimated precision of 77% of explicit intentional queries (based on the evaluations reported in [25]) and represents our knowledge base for Intentional Query Suggestion. We call this subset of queries the Explicit Intentional Query Dataset from here on.

Our parametric algorithm for Intentional Query Suggestion approximates the searcher's intent by combining two different yet complementary approaches, i.e. *text-based* Intentional Query

[1] Related query suggestion results from Yahoo!

[2] Related query suggestion results from MSN

Suggestion (see Section 3.2) and *neighborhood-based* Intentional Query Suggestion (see Section 3.3). The two approaches can be combined yielding a ranked list of potential intentional query suggestions.

3.2 Text-Based Intentional Query Suggestion

In the text-based approach, the tokens of input queries are textually compared to all query tokens in the Explicit Intentional Query Dataset. We experimented with several text-based similarity measures including Cosine Similarity, Dice Similarity, Jaccard Similarity and Overlap Similarity [11], [3]. Because the similarity measures did not exhibit significant differences, we decided on using Jaccard Similarity throughout our experiments for reasons of simplicity. In text-based intentional query suggestion, we calculate Jaccard Similarity in the following way:

$$S_T(q_A, q_B) = \frac{|q_A \cap q_B|}{|q_A \cup q_B|}$$

where q_A and q_B are the respective token sets representing two queries.

3.3 Neighborhood-Based Intentional Query Suggestion

In addition to Intentional Query Suggestion based on text, we are using a similarity construct based on query log session neighborhood. This has the potential to include behavioural intentional structures in our algorithm. For that purpose, we are conceptualizing query logs as consisting of two types of nodes (a bipartite graph), where nodes of one type correspond to *explicit* intentional queries and nodes of the other type correspond to *implicit* intentional queries. We construct a bipartite graph based on session proximity between these two types of nodes. Thereby, we use neighboring queries to further describe and characterize explicit intentional queries, building characteristic term vectors for explicit intentional queries. In the following, we introduce the parametric algorithm for intentional query expansion in a more formal way.

Table 2: Search query log excerpt illustrating the explicit intentional query $q_{e,1}$ and its neighborhood $N(q_{e,1}, 3)$.

Type	Query	Date
$q_{u,1}$	types of diet pills	2006-05-24 13:34:16
$q_{u,2}$	Lipo6	2006-05-24 13:36:24
$q_{u,3}$	lose 20 pounds in 8 weeks	2006-05-24 13:37:23
$q_{e,1}$	lose weight fast	2006-05-24 13:38:42
$q_{u,4}$	lose weight fast	2006-05-24 13:39:06
$q_{u,5}$	weight loss upplements	2006-05-24 13:39:51
$q_{u,6}$	weight loss supplements	2006-05-24 13:39:56

3.3.1 *Parametric Algorithm*

Let $Q = \{q_1, q_2 \dots q_n\}$ denote the set of n queries in a search query log. Q consists of two disjoint sets $Q_E=\{q_{e,1}, q_{e,2} \dots q_{e,s}\}$ and $Q_U=\{q_{u,1}, q_{u,2} \dots q_{u,t}\}$ so that $Q = Q_E \cup Q_U$ and $s + t = n$. Q_E represents the set of explicit intentional queries, such as "lose weight fast", and Q_U the neighboring implicit intentional queries such as "weight loss supplements" as illustrated in Table 2.

We define the neighborhood of an explicit intentional query q_e as $N(q_e, P_d)$, where the parameter P_d determines the number of queries that are considered before and after the query q_e. The neighborhood $N(q_e, P_d)$ contains $2 * P_d$ queries where $q \in Q_U$ holds. Queries $q_i \in N(q_e, P_d)$ are processed to serve as tags (dimensions of the characteristic vector describing explicit intentional queries) for the corresponding intentional query q_e. After stop words have been removed, the remaining tokens are combined into a set of words and form a tag set $T(q_e)=\{t_1, t_2 \dots t_m\}$ of the explicit intentional query q_e. In addition to parameter P_d, we introduce the parameter P_i that denotes the intersection size between explicit intentional queries and neighboring queries. This parameter can be considered as a quality filter. Tokens of one query are only admitted to the tag set $T(q_e)$ if the query shares at least P_i tokens with q_e. Let q_e be "lose weight fast", q_u be "weight loss supplements" and $P_i = 1$: q_e and q_u share one common term ("weight"). Consequently, the tokens of q_u are considered tags for q_e, i.e. $T(q_e)$ = {"weight", "loss", "supplement"}. We suspect this parameter to be related to the quality of the tags admitted to the tag set and consequently related to the quality of the entire model. This yields a characteristic vector of tags for each explicit intentional query based on session-neighborhood.

Figure 1 shows a bipartite graph that was partly generated from the query log excerpt in Table 2 with a parameter setting $P_d = 3$ and $P_i = 1$. The graph illustrates relations between explicit intentional queries and meaningful terms in the session neighborhood, representing characteristic term vectors for explicit intentional queries. The example also shows that the neighborhood-based approach is agnostic to misspellings. The bipartite graph is useful in at least two ways: Bottom-up, it can help to produce intentional query suggestions based on co-occurrence (e.g. "upplements" → "lose weight fast"). Top-down, the graph can help to transform explicit intentional queries into implicit ones (which is not further pursued in this paper). Note that $q_{u,3}$ and $q_{u,4}$ both represent explicit intentional queries and are therefore neglected in the graph generation process.

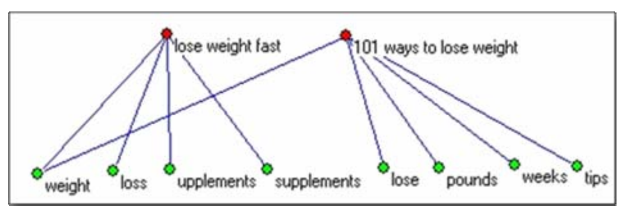

Figure 1: Bipartite graph partly generated from search query log excerpt in Table 2 with parameter setting P_d=3 and P_i=1.

Similarity between an input query ("upplements") and a number of explicit intentional queries ("lose weight fast") can now be calculated with traditional similarity metrics. Again, we experimented with different similarity measures and opted for the Jaccard similarity measure due to insignificant differences between the measures. In neighbourhood-based intentional query suggestion, we calculate Jaccard similarity in the following way:

$$S_G(q_A, q_B) = \frac{|T(q_A) \cap T(q_B)|}{|T(q_A) \cup T(q_B)|}$$

where $T(q_A)$ and $T(q_B)$ are the respective token sets representing two queries.

3.4 Query Suggestion based on User Intent

When input queries are processed by our algorithm, both similarity measures are calculated. In our approach, a linear combination determines the overall similarity between an input query and every explicit intentional query in our dataset yielding a ranked list of potential user intentions. The parameter α defines the impact of each measure:

$$S(q_A, q_B) = \alpha * S_T(q_A, q_B) + (1 - \alpha) * S_G(q_A, q_B)$$

In this work we do not intend to identify an optimized parameter set to generate the model. We rather chose a simple parameter set for the purpose of seeking answers to the exploratory questions of this paper. Future work might explore the utility of parameter variations in greater depth.

The parametric algorithm for Intentional Query Suggestion can be described by the function IQS → f (P_d, P_i, α). We used following parameter setting: $P_d = 3$, $P_i = 1$ and $\alpha = 0.5$ in our experiments. An evaluation of the selected model is provided in Section 3.5.

3.5 Evaluation

We conducted a user study to learn more about the quality of intentions that were suggested by our algorithm. Annotators were asked to categorize the 10 top-ranked suggested explicit intentional queries for 30 queries into one of the following two relevance classes.

Relevance Classes:

(1) Potential User Intention: the suggested query represents a plausible intention behind a short query.

Initial Query	*Intentional Query Suggestions*
"anime"	"draw anime", "draw manga"
"playground mat"	"buy playground equipment", "build a swing set"

or the suggested query represents an unlikely yet still related user intention as illustrated by following examples:

Initial Query	*Intentional Query Suggestions*
"Boston herald"	"getting around Boston", "sightseeing in Boston"
"ginseng coffee"	"moving coffee stains", "fix my keyboard"

(2) Clear Misinterpretation: the suggested query has no relation with the initial query. Suggestions that do not conform to our definition (see Section 3) are assigned this category as well.

Initial Query	*Intentional Query Suggestions*
"Boston herald"	"care for Boston fern", "flying to Nantucket"
"playground mat"	"raise money for our playground", "weave a basket fifth grade project"

30 queries of length 1 or 2 were randomly drawn from the MSN search query log. The prospective queries were filtered with regard to (i) reasonableness, i.e. discarding queries such as "wiseco" or "drinkingmate" and to (ii) non American raters, i.e. discarding queries such as "target" or "espn".

In order to evaluate intentional query suggestions that are provided by our algorithm, we calculated the percentage of correct suggestions, i.e. query suggestions that were assigned to relevance class 1. Achieved precision values are illustrated in Table 3.

Table 3: Precision values of our algorithm as rated by three human annotators (X, Y and Z).

	X	Y	Z
Precision	0.61	0.73	0.8

The average precision amounts to 0.71, i.e. in seven out of ten cases the algorithm returns a potential user intention.

In addition, we calculated the inter-rater agreement κ [8] between all pairs of human subjects X, Y, and Z. Cohen's κ measures the average pair-wise agreement corrected for chance agreement when classifying N items into C mutually exclusive categories. Cohen's κ formula reads:

$$\kappa = \frac{P(O) - P(C)}{1 - P(C)}$$

where P(O) is the proportion of times that a hypothesis agrees with a standard (or another rater), and P(C) is the proportion of times that a hypothesis and a standard would be expected to agree by chance. The κ value is constrained to the interval [-1,1]. A κ-value of 1 indicates total agreement, 0 indicates agreement by chance and -1 indicates total disagreement. Table 4 shows the achieved κ-values in our human subject study.

Table 4: Kappa values amongst three annotators (X, Y and Z) for the two relevance classes.

	X-Y	X-Z	Y-Z
Cohen's Kappa (κ)	0.6416	0.5125	0.6703

The κ-values (see Table 4) range from 0.51 to 0.67 (0.61 on average) containing two values above 0.6 indicating some level of agreement.

4. PRELIMINARY RESULTS

In this section we discuss two potential implications of Intentional Query Suggestion for web search: First, diversity of search results has recently gained importance in web search [9]. For example in informational queries, web search results should not provide monolithic search result sets but rather cover as many different aspects (topics) as possible. We are interested in exploring the influence of explicit intentional queries on the diversity of search result sets. If result sets of explicit intentional queries would be more diverse, Intentional Query Suggestion could help to better focus and guide searchers' intent in exploratory searches.

Second, click through rates have been frequently used as a proxy for measuring relevance in large document collections (cf. [10]). We are interested in studying whether explicit intentional queries would yield other/better click-through rates than implicit intentional queries. If explicit intentional queries would yield higher click-through rates, making user intent more explicit would represent an interesting new mechanism to improve search engine performance.

4.1 Influence on Diversity of Search Results

We examine the diversity within search results by calculating the intersection size between different URL result sets produced by different/same query suggestion mechanisms. Two experiments were conducted, seeking answers to the following questions:

(i) *Intersection between different Query Suggestion Mechanisms:* How many URLs (top level domains only) intersect between URL result sets retrieved by 1) the original queries, 2) the corresponding Yahoo! expanded queries and 3) the corresponding intentional query suggestions?

(ii) *Intersection within same Query Suggestion Mechanisms:* How many URLs (top level domains only) intersect between result sets that were retrieved by different query suggestions (produced by the same query suggestion mechanism) regarding one original query?

400 queries of length 1 or 2 were randomly drawn from the MSR search query log. Following constraints were made: original queries (i) should yield at least 10 suggestions by our algorithm, (ii) should not contain misspellings and (iii) must not be 'adult' phrases. For each selected query, the top 10 suggestions were produced by using the Yahoo! API and by the Intentional Query Suggestion algorithm. We processed the top 50 result URLs for each suggestion, totalling 500 URLs per selected query. Searches were conducted by applying the Yahoo! BOSS API[3]. In order to compare the original query results with both expanded results sets, 500 resulting URLs are retrieved for every original query. For each query, we calculated how many URLs are shared on average between the URL result sets taking into account only unique URLs as well as only top level domains of the resulting set. Again, we used Jaccard as a metric for intersection/similarity. The averaged results over all candidate queries are shown in Table 5.

Table 5: Average intersection sizes for URL sets of original queries and their corresponding suggestions.

Compared URL result sets	Avg.Inter-section
Original Queries vs. Yahoo! Suggestions	0.1911
Original Queries vs. Intentional Suggestions	0.0467
Yahoo! Suggestions vs. Intentional Suggestions	0.0511

The results in Table 5 imply that original query results share more URLs with results from Yahoo! expanded queries than with results yielded by queries that reflect potential user intent. This suggests that if queries are expanded by user intent more diverse result sets can be achieved. In addition, we calculated the inner intersection size of the result sets, i.e. the overlap between different result sets produced by the same suggestion mechanism. The results were again averaged over all queries and are shown in Table 6.

The results in Table 6 suggest that queries expanded by Yahoo! yield more overlapping URLs than queries expanded by user intent. These results suggest that queries that express a specific intention lead to more diverse results than queries that attempt to approximate the expected document content to retrieve.

Table 6: Average intersection sizes for URL sets expanded by Yahoo! Suggestions and Intentional Query Suggestion.

Compared URL result sets	Average Intersection
Yahoo! Suggestions	0.103
Intentional Query Suggestion	0.026

Considering the presented results, we can speculate that search processes could be made more focused if the searchers' intention is explicitly included in the search process. It appears that intentional query suggestions diversify search results and cover a wider range of topics than Yahoo!'s suggestions.

4.2 Influence on Click-Through

To study the influence of explicit intentional queries on click through, we analyzed the number of click-through events for different token lengths. We obtained the click-through numbers for different token lengths in the MSR query dataset and created the following token length bins: one token queries, two token queries, three to four token queries, five token queries, six to ten token queries and queries consisting of more than ten tokens (excluding explicit intentional queries). Five token queries were of particular interest, since the average length of queries in our Explicit Intentional Query Dataset amounts to 5.33 tokens. For each category, a random sample of 5,000 queries was drawn from the MSN search query log and all corresponding click-through events were registered and counted. Table 7 shows the number of click through events for each bin and also for the set of explicit intentional queries.

Table 7: Click-through distribution for different query lengths and explicit intentional queries

	Implicit Intentional Queries						*Explicit Intent. Queries*
Query Length	*1*	*2*	*3-4*	*5*	*6-10*	*>10*	*5.33*
#click-through	855,649	358,327	64,313	5,559	2,728	960	7,236

It can be observed that explicit intentional queries appear to have a ~ 30% higher number of click through events (#click-through = 7,236) than implicit intentional queries of comparable length (length 5, #click-through = 5,559). The higher click-through numbers of explicit intentional queries suggest that such queries retrieve more relevant results, which appears to be an interesting finding and preliminary evidence for the potential utility of intentional query suggestions.

5. RELATED WORK

Two areas of research are particularly relevant to our work: Studies of search intent in query logs and query suggestion.

Studies of search intent in query logs: Peter Norvig discussed[4] search intent as one of the outstanding problems in the future of search. One interpretation of understanding the users' needs is to

[3] http://developer.yahoo.com/search/boss/

[4] Interview in the Technology Review (Monday, July 16, 2007)

understand the intentions behind search queries. Intentional query suggestions could be regarded as a first step in this direction by helping users to make their search intent more explicit. In previous years, several different definitions of user intent emerged [6], [10], [12],[25]. Broder [6] for example introduced a high level taxonomy of search intent by categorizing search queries into three categories: navigational, informational and transactional. This has stimulated a series of follow up research on automatic query categorization by [18], [13], [15], [12] and [23]. Evolutions of Broder's taxonomy include collapsing categories, adding categories [5] and/or focusing on subsets only [18]. In contrast to Broder, we do not incorporate high-level categories of search intent but rather focus on instances of user intentions (informational vs. "things to consider when buying a car").

He et al. [12] used syntactic structures, i.e. verb-object pairs, to classify queries into Broder's categories. In a similar way, Strohmaier et al. [25] employed part-of-speech trigrams as features to extract instances of user intentions in search query logs. In this paper, user intent is understood as a certain type of verb phrases that explicitly state the user's goal. Downey et al [10] view the information seeking process differently: Actions that follow a search query are proposed as characterizations of the searcher's information goal. The last URL visited in a search session serves as a proxy for the user intent. While their approach is useful to study user behavior during search sessions, it can not easily be used in an interactive way - to enable users to make their search intent more explicit.

In addition to studies of user intent, research on *query suggestion* is related to our work as well. Query expansion [27], query substitution [14], query recommendation [4] and query refinement [17] are different concepts that share a similar objective: transforming an initial query into a 'better' query that is capable of satisfying the searcher's information need by retrieving more relevant documents. We deviate from these traditional approaches that focus on query vs. expected documents by focusing on queries and potential user intentions. Xu et al. [27] for example employed local and global documents in query expansion by applying the measure of global analysis to the selection of query terms in local feedback. Query suggestion is closely related to query substitution as well where the original query is extended by new search terms to narrow the search scope. Jones et al. [14] investigated a query substitution mechanism that does not exhibit query drift which represents a common drawback of query expansion techniques. The authors make use of search query sessions to infer relations between queries.

Baeza-Yates et al. [4] proposed an approach that suggests related queries based on query log data and clustering. Former queries were transformed into a new term-vector representation by taking into account the content of the clicked URLs. Another approach reported in [17] employed anchor texts for the purpose of query refinement. It is based on the observation that queries and anchor texts are highly similar. Query transformation techniques have already spread to other areas such as question answering [1].

Work on query suggestion has recently been done by [20], [22]. Both papers apply their algorithms on bipartite graphs (user - query and/or query - URL) that were generated from search query logs. In a similar way, our work generates a bipartite graph from a search query log. However, our approach focuses on explicit intentional queries and their implicit intentional query neighborhood, thereby focusing on explicit user intent rather than the generation of syntactic or semantic query suggestions.

6. CONCLUSIONS

While there is a significant body of research on understanding user intent during search ([6], [23], [13], [5], [18], [10], [7]), to the best of our knowledge, the application of user intent to query suggestion is a novel idea which has not been studied yet. In this paper, we introduce and define the concept of *Intentional Query Suggestion* and present a prototypical algorithm as first evidence for the feasibility of this idea. In a number of experiments, we.could highlight interesting differences to traditional query suggestion mechanisms: 1) *Differences in the diversity of search results*. Our results suggest that intentional query expansions can be used to diversify result sets. One implication of this finding is that search engine vendors might be able to make search processes more focused if the searchers' intention is explicitly included in the search process. *2) Different click-through distributions for explicit intentional queries.* Our experiments showed a higher click-through ratio for explicit intentional queries compared to implicit intentional queries of similar length. The higher click-through numbers suggest that such queries retrieve more relevant results. This interesting finding might inspire novel ways to approach query suggestion in the future.

Our results could be relevant for a number of currently open research problems. 1) *Query disambiguation*: Similar to Allan [2], where the problem of query disambiguation was approached by posing questions, Intentional Query Suggestion could provide a mechanism to identify the original user goal during search. 2) *Search intent*: A better understanding of the user's intent could give search engine vendors a better picture of users' needs. In the long run, approximating user intent could help making search more focused and prevent topic drift. 3) *Search session*: Along with a better understanding of users' search intent, new, more useful definitions of search sessions might be necessary. New definitions could differ from existing definitions by, for example, putting emphasis on a set of coherent, goal-related queries rather than time-based notions, where multi-tasking behavior of users is hard to capture. 4) *Evaluation*: Kinney et al. [16] point out the difficulty of finding expert annotators when it comes to annotating web search results for evaluation purposes. In order to alleviate the annotation task, the authors proposed statements that described the user intent behind a query. Intentional query suggestion might serve as a link between plain queries and the intent statements by offering a list of empirically-grounded, plausible user intentions.

7. ACKNOWLEDGMENTS

We would like to thank Microsoft Research for providing the search query log and Peter Prettenhofer for his support in extracting the Explicit Intentional Query Dataset. This work is funded by the FWF Austrian Science Fund Grant P20269 *TransAgere*. The Know-Center is funded within the Austrian COMET Program under the auspices of the Austrian Ministry of Transport, Innovation and Technology, the Austrian Ministry of Economics and Labor and by the State of Styria. COMET is managed by the Austrian Research Promotion Agency FFG.

8. REFERENCES

[1] Agichtein E., Lawrence S. and Gravano L. Learning search engine specific query transformations for question answering. In 'WWW '01: Proceedings of the 10th

International Conference on World Wide Web', ACM, New York, NY, USA, pp. 169--178, 2001.

[2] Allan J. and Raghavan H. Using part-of-speech patterns to reduce query ambiguity. In 'Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval', ACM Press New York, NY, USA, pp. 307--314, 2002.

[3] Baeza-Yates R. and Ribeiro-Neto B. *Modern Information Retrieval*, AddisonWesley, 1999.

[4] Baeza-Yates R., Hurtado C.A. and Mendoza M. Query recommendation using query logs in search engines. In Lindner W., Mesiti M., Türker C., Tzitzikas Y. and Vakali A., 'EDBT Workshops', Springer, pp. 588--596, 2004.

[5] Baeza-Yates R., Calderón-Benavides L. and González-Caro C. The intention behind web queries. In *String Processing and Information Retrieval*, pp. 98--109, 2006.

[6] Broder A. A taxonomy of web search. In *ACM SIGIR Forum* 36(2), pp. 3--10, 2002.

[7] Chang Y., He, K., Yu S. and Lu, W. Identifying user goals from web search results. In 'WI '06: Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence', IEEE Computer Society, Washington, DC, USA, pp. 1038--1041, 2006.

[8] Cohen, J. A coefficient of agreement for nominal scales. In *Educational and Psychological Measurement* 20(1), 37, 1960.

[9] Crabtree, D. W., Andreae, P. and Gao, X. Exploiting underrepresented query aspects for automatic query expansion. In 'KDD '07: Proceedings of the 13th ACM SIGKDD international conference on Knowledge Discovery and Data Mining', ACM, New York, NY, USA, pp. 191--200, 2007.

[10] Downey, D., Liebling, D. and Dumais, S. Understanding the relationship between searchers, queries and information goals. 'CIKM '08: Proceedings of the 17th ACM Conference on Information and Knowledge Management', ACM, New York, NY, USA, 2008.

[11] Ferber, R. *Information Retrieval*, Dpunkt.Verlag, ISBN 978-3898642132, 2003.

[12] He, K., Chang, Y. and Lu, W. Improving identification of latent user goals through search-result snippet classification. In 'WI '07: Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence', IEEE Computer Society, Washington, DC, USA, pp. 683--686, 2007.

[13] Jansen, B. J., Booth, D. L. and Spink, A. Determining the informational, navigational, and transactional intent of web queries. In *Inf. Process. Manage.* 44(3), pp. 1251--1266, 2008.

[14] Jones, R., Rey, B., Madani, O. and Greiner, W. Generating query substitutions. In 'WWW '06: Proceedings of the 15th International Conference on World Wide Web', ACM, New York, NY, USA, pp. 387--396, 2006.

[15] Kang, I. and Kim, G. Query type classification for web document retrieval. In 'SIGIR '03: Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval', ACM, New York, NY, USA, pp. 64--71, 2003.

[16] Kinney, K. A., Huffman, S. B. and Zhai, J. How evaluator domain expertise affects search result relevance judgments. In 'CIKM '08: Proceedings of the 17th ACM Conference on Information and Knowledge Management', ACM, New York, NY, USA, pp. 591--598, 2008.

[17] Kraft, R. and Zien, J. Mining anchor text for query refinement. In 'WWW '04: Proceedings of the 13th International Conference on World Wide Web', ACM, New York, NY, USA, pp. 666--674, 2004.

[18] Lee, U., Liu, Z. and Cho, J. Automatic identification of user goals in web search. In 'WWW '05: Proceedings of the 14th International Conference on World Wide Web', ACM, New York, NY, USA, pp. 391--400, 2005.

[19] Liu, H. and Singh, P. 'ConceptNet — A practical commonsense reasoning tool-kit'. In *BT Technology Journal* 22(4), pp. 211--226, 2004.

[20] Ma, H., Yang, H., King, I. and Lyu, M. R. Learning latent semantic relations from clickthrough data for query suggestion. In 'CIKM '08: Proceedings of the 17th ACM Conference on Information and Knowledge Management', ACM, New York, NY, USA, pp. 709--718, 2008.

[21] Manning, C. D., Raghavan, P. and Schütze, H. *Introduction to Information Retrieval*, Cambridge University Press, 2008.

[22] Mei, Q., Zhou, D. and Church, K. Query suggestion using hitting time. In 'CIKM '08: Proceedings of the 17th ACM Conference on Information and Knowledge Management', ACM, New York, NY, USA, pp. 469--478, 2008.

[23] Rose, D. E. and Levinson, D. Understanding user goals in web search. In 'WWW '04: Proceedings of the 13th International Conference on World Wide Web', ACM, New York, NY, USA, pp. 13--19, 2004.

[24] Strohmaier, M., Prettenhofer, P. and Lux, M. Different degrees of explicitness in intentional artifacts - studying user goals in a large search query log. In 'CSKGOI'08: Proceedings of the Workshop on Commonsense Knowledge and Goal Oriented Interfaces, in conjunction with IUI'08', Canary Islands, Spain, 2008.

[25] Strohmaier, M., Prettenhofer, P. and Kröll, M. Acquiring explicit user goals from search query logs. In 'International Workshop on Agents and Data Mining Interaction ADMI' 08, in conjunction with WI '08', 2008.

[26] Strzalkowski, T. and Carballo, J. Natural Language Information Retrieval: TREC-5 Report. In 'Text REtrieval Conference', pp. 164--173, 1998.

[27] Xu, J. and Croft, W. B. Query expansion using local and global document analysis. In 'SIGIR '96: Proceedings of the 19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval', ACM, New York, NY, USA, pp. 4--1, 1996