Information Retrieval with Verbose Queries

Manish Gupta
Microsoft
gmanish@microsoft.com

Michael Bendersky
Google
bemike@google.com

the essence of knowledge
Boston — Delft

Foundations and Trends® in Information Retrieval

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
United States
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is

M. Gupta and M. Bendersky. *Information Retrieval with Verbose Queries.* Foundations and Trends® in Information Retrieval, vol. 9, no. 3-4, pp. 209–354, 2015.

This Foundations and Trends® issue was typeset in LaTeX using a class file designed by Neal Parikh. Printed on acid-free paper.

ISBN: 978-1-68083-045-3

Full text available at: http://dx.doi.org/10.1561/1500000050

Foundations and Trends® in Information Retrieval

Volume 9, Issue 3-4, 2015

Editorial Board

Editorial Scope

Topics

Foundations and Trends® in Information Retrieval publishes survey and tutorial articles in the following topics:

- Applications of IR
- Architectures for IR
- Collaborative filtering and recommender systems
- Cross-lingual and multilingual IR
- Distributed IR and federated search
- Evaluation issues and test collections for IR
- Formal models and language models for IR
- IR on mobile platforms
- Indexing and retrieval of structured documents
- Information categorization and clustering
- Information extraction
- Information filtering and routing
- Metasearch, rank aggregation, and data fusion
- Natural language processing for IR
- Performance issues for IR systems, including algorithms, data structures, optimization techniques, and scalability
- Question answering
- Summarization of single documents, multiple documents, and corpora
- Text mining
- Topic detection and tracking
- Usability, interactivity, and visualization issues in IR
- User modelling and user studies for IR
- Web search

Information for Librarians

Foundations and Trends® in Information Retrieval, 2015, Volume 9, 5 issues. ISSN paper version 1554-0669. ISSN online version 1554-0677. Also available as a combined paper and online subscription.

Foundations and Trends® in Information Retrieval
Vol. 9, No. 3-4 (2015) 209–354

DOI: 10.1561/1500000050

Information Retrieval with Verbose Queries

Manish Gupta
Microsoft
gmanish@microsoft.com

Michael Bendersky
Google
bemike@google.com

Contents

Abstract

Recently, the focus of many novel search applications has shifted from short keyword queries to verbose natural language queries. Examples include question answering systems and dialogue systems, voice search on mobile devices and entity search engines like Facebook's Graph Search or Google's Knowledge Graph. However the performance of textbook information retrieval techniques for such verbose queries is not as good as that for their shorter counterparts. Thus, effective handling of verbose queries has become a critical factor for adoption of information retrieval techniques in this new breed of search applications.

Over the past decade, the information retrieval community has deeply explored the problem of transforming natural language verbose queries using operations like reduction, weighting, expansion, reformulation and segmentation into more effective structural representations. However, thus far, there was not a coherent and organized survey on this topic. In this survey, we aim to put together various research pieces of the puzzle, provide a comprehensive and structured overview of various proposed methods, and also list various application scenarios where effective verbose query processing can make a significant difference.

M. Gupta and M. Bendersky. *Information Retrieval with Verbose Queries.* Foundations and Trends® in Information Retrieval, vol. 9, no. 3-4, pp. 209–354, 2015.

DOI: 10.1561/1500000050.

Preface

Information retrieval with verbose natural language queries has gained a lot of interest in recent years both from the research community and the industry. Search with verbose queries is one of the key challenges for many of the current most advanced search platforms, including question answering systems (Watson or Wolfram Alpha), mobile personal assistants (Siri, Cortana and Google Now), and entity-based search engines (Facebook Graph Search or Knowledge Graph). Therefore, we believe that this survey is very timely and should be interesting to readers from both academia as well as industry.

Scope of the Survey

We cover an exhaustive list of techniques to handle verbose queries. Intuitively verbose queries are long. Also empirical observations show that often times long queries are verbose in nature. We use the terms "verbose" queries and "long" queries interchangeably in this survey.

In order to stay focused, following is a list of related topics that we do not cover as part of this survey.

- Automatic Speech Recognition (ASR)
- Processing null queries other than verbose queries

- Methods (e.g., [Yang et al., 2009] and [Tsagkias et al., 2011]) and applications (e.g., [Yih et al., 2006]) which consider documents as queries
- Query processing tasks for short queries which do not need any non-trivial modification to be applicable to long queries
- Community-based question-answering systems

Development of the Survey

Many tutorials and surveys dedicated to general query handling or query log analysis have been conducted by researchers in information retrieval and web mining. However, all of them focus on short queries; none of these have explicitly focused on long verbose queries. This survey is based on a full-day tutorial offered by the authors at the 38^{th} International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2015). The slides for the tutorial can be obtained from `http://research.microsoft.com/pubs/241895/gupta15_verbose.pptx`.

This survey is entirely based on previously published research and publicly available datasets, rather than the internal practices of the respective employers of the authors. As such, it should prove useful for both practitioners and academic researchers interested in reproducing the reported results.

Audience

Researchers in the field of information retrieval will benefit the most, as this survey will give them an exhaustive overview of the research in the direction of handling verbose web queries. We believe that the survey will give the newcomers a complete picture of the current work, introduce important research topics in this field, and inspire them to learn more. Practitioners and people from the industry will clearly benefit from the discussions both from the methods perspective, as well as from the point of view of applications where such mechanisms are starting to be applied.

After reading the survey, the audience will be able to appreciate and understand the following.

- What are the interesting properties of complex natural language verbose queries
- Challenges in effective information retrieval with verbose queries
- State-of-the-art techniques for verbose query transformations that yield better expected search performance
- State-of-the-art ranking methods for verbose queries, including supervised learning-to-rank methods
- What user/industry segments can be affected by better retrieval with verbose queries and what are the possible applications

Writing Style

We have tried to make the survey as self-contained as possible. However, for some sections, we have deliberately adopted a reference paper writing style, to enable a holistic overview of the research field. In such cases, we discuss those pieces of work from a more general and abstract standpoint, and advise the readers to go through the referenced papers for details. We provide a basic introduction to preliminary information retrieval concepts, graphical models and dependency parsing in the Appendices.

1

Introduction

Web search has matured significantly in the past two decades. Beyond the ten blue links, search engines display a large amount of heterogeneous information including direct factual answers, task panes, image answers, news answers, video answers, social results, related searches, etc. Broadly, queries to a search engine can be divided into two parts: head and tail. Head queries are the highly popular queries while the tail queries occur with a low frequency in the query log. Although the head queries are handled very elegantly by the popular search engines, there is a large room for improvement when handling the tail queries, a part of which return no results.

1.1 Null Queries

Null queries are queries for which the search engine returns zero results. This could be because of the following reasons.

- Query verbosity
- Mismatch between the searcher and the publisher vocabulary

- Unavailability of relevant documents (temporally, or general rarity)
- Inability of the naïve users to formulate appropriate queries

In this survey, we focus on the verbosity aspect of such "null" or difficult to handle queries. We use the terms "verbose" queries and "long" queries interchangeably. This work focuses on verbose queries as well as on long queries which may or may not be verbose.

1.2 Verbose Queries are Frequent

As shown in Figure 2.1, the percentage of the total query traffic follows a power law distribution with respect to the query length [Arampatzis and Kamps, 2008, Bailey et al., 2010], i.e., for a query Q,

$$p(|Q|) = C|Q|^{-s},\ for\ |Q| \geq k_0 \tag{1.1}$$

where $|Q|$ is the query length in words, C is a normalizing constant, s is the slope, k_0 is the lower bound from which the power law holds.

We consider queries with five or more words as verbose or long queries. In 2006, Yahoo! claimed that 17% of the queries contained five or more words.[1]. Figure 2.1 shows that ∼15% queries contain five or more words.

Popular usage of speech-based personal assistants like Cortana, Siri, and Google Now attract an even higher percentage of verbose queries. Crestani and Du [2006] and Yi and Maghoul [2011] analyzed the properties of written versus spoken queries which were manually generated by participants to satisfy TREC topic information needs. They found that while written queries had an average length of 9.54 and 7.48 words with and without stop words respectively, spoken queries had an average length of 23.07 and 14.33 words respectively. Voice queries were considerably longer than the typed mobile queries.

While most of the verbose queries are explicitly asked by the users, some of them are implicit. Users ask verbose queries explicitly in a large

[1] `http://www.zdnet.com/blog/micro-markets/yahoo-searches-more-sophisticated-and-specific/27`

number of scenarios. Advanced users searching for an exhaustive list of relevant documents in medical literature or patent documents often use verbose comprehensive queries. Naïve users like children or the elderly are not trained to ask short queries to search engines and hence end up using full sentence queries. Community-based question answering platforms also attract long queries. Sometimes users end up using long queries implicitly. Long queries could be an outcome of cut-and-paste behavior. For example, a user just found some text on some topic (say a news headline) and fires it as a query to find related news articles. Similarly, to find a relevant image for a paragraph in a textbook, one may fire the entire paragraph as a query to the search engine. We discuss both the implicit and explicit examples of verbose queries in more details in §9.

1.3 Search Engine Performance for Verbose Queries

Past research in information retrieval found that long queries increase the retrieval performance. However, for web search queries, many researchers have observed that search engines perform poorly on verbose queries. The reasons for poor performance are as follows.

- High degree of query specificity. To satisfy their specific (or narrow) needs, users put additional non-redundant information in verbose queries. But since there are not many web-pages to satisfy such highly specific information needs, it is difficult for search engines to surface the right results.

- Term redundancy or extraneous terms (lot of noise). Often times, verbose queries contain a lot of noise, such as extraneous terms that users believe are important to conveying their information needs, but in fact are confusing to automatic systems.

- Rarity of verbose queries. Most search engines optimize for highly popular (or head) queries. Since verbose queries are rare, search engine algorithms are not tweaked to always perform well for them.

- Lack of sufficient natural language parsing. Longer queries can be answered more effectively if the semantics can be understood using natural language understanding techniques. However, search engines currently do not perform such deep parsing because (a) they are optimized for short queries for which deep natural language parsing is not required, and (b) such deep parsing has performance implications.

- Difficulty in distinguishing between the key and complementary concepts. A verbose query can have multiple concepts. The performance can be improved if the results that contain key concepts are shown at the top. However, identifying key concepts from a verbose query is challenging.

Hence, a large number of efforts have been made to understand such long queries in a more effective manner.

1.4 Datasets

Most of the papers in this area have used the TREC datasets for evaluating their approaches. ROBUST04, W10g, GOV2, ClueWeb-09-Cat-B, TREC123, and CERC are the most popular TREC[2] datasets. ROBUST04 is a Newswire collection, while W10g, GOV2 and ClueWeb-09-Cat-B are web collections. TREC123 is a collection of documents from TREC disks 1 and 2. CERC is the CSIRO Enterprise Research Collection (CERC), a crawl of *.csiro.au (public) web sites conducted in March 2007 and used in the 2007 edition of the TREC Enterprise track. Table 1.1 gives a summary of the dataset statistics. Each of these datasets contain relevance judgments for multiple topics (or queries). The judgments are for multiple documents and are binary or graded (e.g., non-relevant, relevant, highly relevant). TREC topics illustrate the difference between a keyword query and a description query. A TREC topic consists of several parts, each of which corresponds to a certain aspect of the topic. In the example at Figure 1.1, we consider the title (denoted ⟨title⟩) as a keyword query on the topic, and the de-

[2] `http://trec.nist.gov`

Collection	Content	#Docs	Topics
Robust04	Newswire	528155	250
W10g	Web	1692096	100
GOV2	Web	25205179	150
ClueWeb-09-Cat-B	Web	50220423	150
TREC123	TREC disks 1 and 2	742611	150
CERC	Enterprise Documents from *.csiro.au	370715	50

Table 1.1: Statistics for TREC Datasets

scription of the topic (denoted ⟨desc⟩) as a natural language description of the information request. In general, the description field is intended to model what a searcher might first say to someone who will actually help them with their search. The verbose description is therefore often used as the verbose query. Another popular similar dataset is the NTCIR-4/5 English-English ad-hoc IR tasks dataset with an average length of 14 query words for description queries.

Some of the recent papers have also used real web query logs [Balasubramanian et al., 2010, Parikh et al., 2013, Yang et al., 2014]. A few researchers have also used document paragraphs or passages as verbose queries [Agrawal et al., 2011, Lee and Croft, 2012, Gupta, 2015].

1.5 Metrics

A variety of standard information retrieval metrics have been used to evaluate the methods for verbose query processing. Most of the researchers that use TREC datasets evaluate their methods using Mean Average Precision (MAP), Mean Reciprocal Rank (MRR), and Precision@K measures against the relevance judgments. Researchers using query logs also use Normalized Discounted Cumulative Gain (NDCG) with respect to the original long query as a metric. We provide a short description of these metrics in §A.6.

```
<num> Number 829
<title> Spanish Civil War support
<desc> Provide information on all kinds of material
international support provided to either side in the
Spanish Civil War.
```

Figure 1.1: An Example of ⟨title⟩ and ⟨desc⟩ Parts of a TREC Topic

1.6 Organization of the Survey

In this survey we present an organized summary of efforts towards improved information retrieval for verbose queries. We begin with a study of the specific properties of verbose queries (§2) which makes them especially challenging in information retrieval applications. Next, we discuss six main ways of handling long queries – query reduction to a single sub-query, query reduction to multiple sub-queries, query weighting, query expansion, query reformulation, and query segmentation in §3 to §8. Table 1.2 shows examples of each of the techniques.

Long verbose queries can be reduced to a single sub-query which could be, for example, the most important noun phrase in the query (§3). Or the long query could be processed to extract multiple short queries (§4). Rather than reducing queries by dropping terms from long queries, each term could be assigned a weight proportional to its importance (§5). Another way to handle long queries is to add concept words to the original query to make the intent clearer (§6). If the words used in the long queries are very specific, they could be completely reformulated to a new query which could potentially match a larger number of documents (§7). Finally, a verbose query can contain multiple pieces of the user information need. Such a query could be segmented and then each such segment can be reduced, weighted, expanded or reformulated to get desired results (§8). For each of these techniques, we group together related methods and present comparisons of these methods. We put together various domains in which verbose queries are frequent, and also discuss how various verbose query processing techniques have been used to handle them (§9). We conclude this survey with a brief

Technique	Original Query	Modified Query
Query Reduction to a Single Sub-query (§3)	ideas for breakfast menu for a morning staff meeting	breakfast meeting menu ideas
Query Reduction to a Multiple sub-queries (§4)	identify any efforts proposed or undertaken by world governments to seek reduction of iraqs foreign debt	reductions iraqs foreign debt, iraqs foreign debt
Query Weighting (§5)	civil war battle reenactments	civil:0.0889, war:0.2795, battle:0.1310, reenactments:0.5006
Query Expansion (§6)	staining a new deck	staining a new deck Shopping/Home and Garden/Home Improvement
Query Reformulation (§7)	how far is it from Boston to Seattle	distance from Boston to Seattle
Query Segmentation (§8)	new ac adapter and battery charger for hp pavilion notebook	new, ac adapter, and, battery charger, for, hp pavilion notebook

Table 1.2: Examples of Various Techniques for Handling Verbose Queries

Notation	Meaning
$Q = \{q_1, q_2, \ldots, q_n\}$	Original verbose query
P^Q	Power set of Q
P	A sub-query of Q
C	Collection
$\|C\|$	Number of words in C
N	Number of documents in C
$m(P, M)$	Target measure of effectiveness of ranking function M for query P
$tf(q_i)$	Term frequency of q_i in C.
$tf_d(q_i)$	Term frequency of q_i in document or document collection d.
$df(q_i)$	Document frequency of q_i in C.
$T_M(Q)$	Top M relevant documents for query Q.

Table 1.3: Table of Notations

overview of future research directions (§10). Table 1.3 presents a list of frequent notations that we use in this survey.

1.7 Summary

Query verbosity is one of the main reasons for zero results returned by search engines. Verbose queries occur in multiple domains and are increasing with increase in usage of speech-based personal assistants. Currently, search engines perform poorly for such long verbose queries. Hence, a large number of efforts have been made to understand such long queries in a more effective manner. In this survey we present an organized summary of efforts towards improved information retrieval for verbose queries.

Suggested Further Reading: [Arampatzis and Kamps, 2008]: Query length analysis and distribution fitting for multiple datasets; [Crestani and Du, 2006]: Comparison between written and spoken queries in terms of length, duration, part-of-speech, aptitude to describe relevant documents, and retrieval effectiveness; http://trec.nist.gov/: Details of the various TREC datasets.

References

Elena Agapie, Gene Golovchinsky, and Pernilla Qvarfordt. Leading People to Longer Queries. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems (SIGCHI)*, pages 3019–3022, New York, NY, USA, 2013. ACM.

Rakesh Agrawal, Sreenivas Gollapudi, Anitha Kannan, and Krishnaram Kenthapadi. Enriching Textbooks with Images. In *Proc. of the* 20^{th} *ACM Intl. Conf. on Information and Knowledge Management (CIKM)*, pages 1847–1856, New York, NY, USA, 2011. ACM.

James Allan, Jamie Callan, W Bruce Croft, Lisa Ballesteros, John Broglio, Jinxi Xu, and Hongming Shu. INQUERY at TREC-5. In *Proc. of the* 5^{th} *Text REtrieval Conference (TREC)*, pages 119–132. NIST, 1996.

Giambattista Amati. *Probability Models for Information Retrieval based on Divergence From Randomness.* PhD thesis, University of Glasgow, 2003.

Giambattista Amati, Claudio Carpineto, and Giovanni Romano. Query Difficulty, Robustness, and Selective Application of Query Expansion. In *Proc. of the* 25^{th} *European Conf. on Information Retrieval (ECIR)*, pages 127–137, Berlin, Heidelberg, 2004. Springer-Verlag.

Avi Arampatzis and Jaap Kamps. A Study of Query Length. In *Proc. of the* 31^{st} *Annual Intl. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR)*, pages 811–812, New York, NY, USA, 2008. ACM.

Judith L Bader and Mary Frances Theofanos. Searching for Cancer Information on the Internet: Analyzing Natural Language Search Queries. *Journal of Medical Internet Research*, 5(4):e31, 2003.

Peter Bailey, Ryen W White, Han Liu, and Giridhar Kumaran. Mining Historic Query Trails to Label Long and Rare Search Engine Queries. *ACM Transactions on the Web (TWEB)*, 4(4):15, 2010.

Niranjan Balasubramanian, Giridhar Kumaran, and Vitor R Carvalho. Exploring Reductions for Long Web Queries. In *Proc. of the* 33^{rd} *Intl. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR)*, pages 571–578, New York, NY, USA, 2010. ACM.

Krisztian Balog, Wouter Weerkamp, and Maarten de Rijke. A Few Examples go a Long Way: Constructing Query Models from Elaborate Query Formulations. In *Proc. of the* 31^{st} *Annual Intl. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR)*, pages 371–378. ACM, 2008.

Michael Bendersky. *Information Retrieval with Query Hypegraphs.* Ir, University of Massachusetts Amherst, July, 2012.

Michael Bendersky and W Bruce Croft. Discovering Key Concepts in Verbose Queries. In *Proc. of the* 31^{st} *Annual Intl. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR)*, pages 491–498, New York, NY, USA, 2008. ACM.

Michael Bendersky and W Bruce Croft. Analysis of Long Queries in a Large Scale Search Log. In *Proc. of the 2009 Workshop on Web Search Click Data (WSCD)*, pages 8–14, New York, NY, USA, 2009. ACM.

Michael Bendersky and W Bruce Croft. Modeling Higher-Order Term Dependencies in Information Retrieval using Query Hypergraphs. In *Proc. of the* 35^{th} *Intl. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR)*, pages 941–950, New York, NY, USA, 2012. ACM.

Michael Bendersky, Donald Metzler, and W Bruce Croft. Learning Concept Importance using a Weighted Dependence Model. In *Proc. of the* 3^{rd} *ACM Intl. Conf. on Web Search and Data Mining (WSDM)*, pages 31–40, New York, NY, USA, 2010. ACM.

Michael Bendersky, W Bruce Croft, and David A Smith. Joint Annotation of Search Queries. In *Proc. of the* 49^{th} *Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (HLT)*, pages 102–111, Stroudsburg, PA, USA, 2011a. Association for Computational Linguistics.

Michael Bendersky, Donald Metzler, and W Bruce Croft. Parameterized Concept Weighting in Verbose Queries. In *Proc. of the* 34^{th} *Intl. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR)*, pages 605–614, New York, NY, USA, 2011b. ACM.

Michael Bendersky, Donald Metzler, and W Bruce Croft. Effective Query Formulation with Multiple Information Sources. In *Proc. of the* 5^{th} *ACM Intl. Conf. on Web Search and Data Mining (WSDM)*, pages 443–452, New York, NY, USA, 2012. ACM.

Paul N. Bennett, Ryen W. White, Wei Chu, Susan T. Dumais, Peter Bailey, Fedor Borisyuk, and Xiaoyuan Cui. Modeling the Impact of Short- and Long-term Behavior on Search Personalization. In *Proc. of the* 35^{th} *Intl. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR)*, pages 185–194, New York, NY, USA, 2012.

Shane Bergsma and Qin Iris Wang. Learning Noun Phrase Query Segmentation. In *Proc. of the 2007 Joint Conf. on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL)*, volume 7, pages 819–826, Prague, Czech Republic, 2007. Association for Computational Linguistics.

Francesco Bonchi, Raffaele Perego, Fabrizio Silvestri, Hossein Vahabi, and Rossano Venturini. Recommendations for the Long Tail by Term-Query Graph. In *Proc. of the* 20^{th} *Intl. Conf. Companion on World Wide Web (WWW)*, pages 15–16, New York, NY, USA, 2011. ACM.

Thorsten Brants and Alex Franz. Web 1T 5-gram Version 1. `https://catalog.ldc.upenn.edu/LDC2006T13`, 2006.

Andrei Z Broder, David Carmel, Michael Herscovici, Aya Soffer, and Jason Zien. Efficient query evaluation using a two-level retrieval process. In *Proceedings of the* 12^{th} *International Conference on Information and Knowledge Management*, pages 426–434. ACM, 2003.

Zhe Cao, Tao Qin, Tie-Yan Liu, Ming-Feng Tsai, and Hang Li. Learning to Rank: From Pairwise Approach to Listwise Approach. In *Proc. of the* 24^{th} *Intl. Conf. on Machine Learning (ICML)*, pages 129–136, New York, NY, USA, 2007. ACM.

Yan Chen and Yan-Qing Zhang. A Query Substitution-Search Result Refinement Approach for Long Query Web Searches. In *Proc. of the 2009 IEEE/WIC/ACM Intl. Joint Conf. on Web Intelligence and Intelligent Agent Technology-Volume 01 (WI-IAT)*, pages 245–251, Washington, DC, USA, 2009. IEEE Computer Society.

Sungbin Choi, Jinwook Choi, Sooyoung Yoo, Heechun Kim, and Youngho Lee. Semantic concept-enriched dependence model for medical information retrieval. *Journal of Biomedical Informatics*, 47(0):18 – 27, 2014.

Kenneth Church and William Gale. Inverse Document Frequency (idf): A Measure of Deviations from Poisson. In *Natural Language Processing using Very Large Corpora*, pages 283–295. Springer, 1999.

Fabio Crestani and Heather Du. Written versus Spoken Queries: A Qualitative and Quantitative Comparative Analysis. *Journal of the American Society for Information Science and Technology (JASIST)*, 57(7):881–890, 2006.

Ronan Cummins, Mounia Lalmas, Colm O'Riordan, and Joemon M Jose. Navigating the User Query Space. In *Proc. of the* 18^{th} *Intl. Symposium on String Processing and Information Retrieval (SPIRE)*, pages 380–385, Berlin, Heidelberg, 2011. Springer-Verlag.

Ronan Cummins, Jiaul H. Paik, and Yuanhua Lv. A Pólya Urn Document Language Model for Improved Information Retrieval. *ACM Transactions on Information Systems (TOIS)*, 33(4):21:1–21:34, May 2015.

Van Dang and Bruce W Croft. Query Reformulation using Anchor Text. In *Proc. of the* 3^{rd} *ACM Intl. Conf. on Web Search and Data Mining (WSDM)*, pages 41–50, New York, NY, USA, 2010. ACM.

Van Dang, Michael Bendersky, and W. Bruce Croft. Two-Stage Learning to Rank for Information Retrieval. In *Proc. of the* 35^{th} *European Conf. on Advances in Information Retrieval (ECIR)*, pages 423–434, Berlin, Heidelberg, 2013. Springer-Verlag.

Sudip Datta and Vasudeva Varma. Tossing Coins to Trim Long Queries. In *Proc. of the* 34^{th} *Intl. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR)*, pages 1255–1256, New York, NY, USA, 2011. ACM.

David A. Ferrucci, Eric W. Brown, Jennifer Chu-Carroll, James Fan, David Gondek, Aditya Kalyanpur, Adam Lally, J. William Murdock, Eric Nyberg, John M. Prager, Nico Schlaefer, and Christopher A. Welty. Building Watson: An Overview of the DeepQA Project. *AI Magazine*, 31:59–79, 2010.

Kristofer Franzen and Jussi Karlgren. Verbosity and Interface Design. Technical Report T2000:04, Swedish Institute of Computer Science, 2000.

Tatiana Gossen, Thomas Low, and Andreas Nürnberger. What are the Real Differences of Children's and Adults' Web Search. In *Proc. of the* 34^{th} *Annual Intl. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR)*, pages 1115–1116. ACM, 2011.

Jiafeng Guo, Gu Xu, Hang Li, and Xueqi Cheng. A Unified and Discriminative Model for Query Refinement. In *Proc. of the* 31^{st} *Annual Intl. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR)*, pages 379–386, New York, NY, USA, 2008. ACM.

Manish Gupta. CricketLinking: Linking Event Mentions from Cricket Match Reports to Ball Entities in Commentaries. In *Proc. of the* 38^{th} *Annual Intl. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR)*, New York, NY, USA, 2015. ACM.

Matthias Hagen, Martin Potthast, Benno Stein, and Christof Braeutigam. The Power of Naïve Query Segmentation. In *Proc. of the* 33^{rd} *Annual Intl. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR)*, pages 797–798, New York, NY, USA, 2010. ACM.

Samuel Huston and W Bruce Croft. Evaluating Verbose Query Processing Techniques. In *Proc. of the* 33^{rd} *Intl. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR)*, pages 291–298, New York, NY, USA, 2010. ACM.

Mohit Iyyer, Jordan Boyd-Graber, Leonardo Claudino, Richard Socher, and Hal Daumé III. A Neural Network for Factoid Question Answering over Paragraphs. In *Proc. of the 2014 Intl. Conf. on Empirical Methods in Natural Language Processing (EMNLP)*, 2014.

Jing Jiang and Chengxiang Zhai. An Empirical Study of Tokenization Strategies for Biomedical Information Retrieval. *Information Retrieval*, 10(4-5): 341–363, 2007.

Rosie Jones and Daniel C Fain. Query Word Deletion Prediction. In *Proc. of the* 26^{th} *Annual Intl. ACM SIGIR Conf. on Research and Development in Informaion Retrieval (SIGIR)*, pages 435–436, New York, NY, USA, 2003. ACM.

Rosie Jones, Benjamin Rey, Omid Madani, and Wiley Greiner. Generating Query Substitutions. In *Proc. of the 15th Intl. Conf. on World Wide Web (WWW)*, pages 387–396, New York, NY, USA, 2006. ACM.

Ron Kohavi and George H John. Wrappers for feature subset selection. *Artificial intelligence*, 97(1):273–324, 1997.

Giridhar Kumaran and James Allan. A Case For Shorter Queries, and Helping Users Create Them. In *Proc. of the Human Language Technologies: The Annual Conference of the North American Chapter of the Association for Computational Linguistics (HLT-NAACL)*, pages 220–227, Stroudsburg, PA, USA, 2007. The Association for Computational Linguistics.

Giridhar Kumaran and James Allan. Effective and Efficient User Interaction for Long Queries. In *Proc. of the* 31^{st} *Annual Intl. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR)*, pages 11–18, New York, NY, USA, 2008. ACM.

Giridhar Kumaran and Vitor R Carvalho. Reducing Long Queries using Query Quality Predictors. In *Proc. of the* 32^{nd} *Intl. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR)*, pages 564–571, New York, NY, USA, 2009. ACM.

Tessa Lau and Eric Horvitz. Patterns of Search: Analyzing and Modeling Web Query Refinement. In *Proc. of the* 7^{th} *Intl. Conf. on User Modeling (UM)*, pages 119–128, Secaucus, NJ, USA, 1999. Springer-Verlag New York, Inc.

Victor Lavrenko and W Bruce Croft. Relevance Models in Information Retrieval. In *Language Modeling for Information Retrieval*, pages 11–56. Springer, Netherlands, 2003.

Matthew Lease. An Improved Markov Random Field Model for Supporting Verbose Queries. In *Proc. of the* 32^{nd} *Intl. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR)*, pages 476–483, New York, NY, USA, 2009. ACM.

Matthew Lease, James Allan, and W Bruce Croft. Regression Rank: Learning to Meet the Opportunity of Descriptive Queries. In *Proc. of the* 31^{th} *European Conf. on IR Research on Advances in Information Retrieval (ECIR)*, pages 90–101, Berlin, Heidelberg, 2009. Springer-Verlag.

Chia-Jung Lee and W Bruce Croft. Generating Queries from User-Selected Text. In *Proc. of the* 4^{th} *Symposium on Information Interaction in Context (IIiX)*, pages 100–109, New York, NY, USA, 2012. ACM.

Chia-Jung Lee, Ruey-Cheng Chen, Shao-Hang Kao, and Pu-Jen Cheng. A Term Dependency-based Approach for Query Terms Ranking. In *Proc. of the* 18^{th} *ACM Conf. on Information and Knowledge Management (CIKM)*, pages 1267–1276, New York, NY, USA, 2009a. ACM.

Chia-Jung Lee, Yi-Chun Lin, Ruey-Cheng Chen, and Pu-Jen Cheng. Selecting Effective Terms for Query Formulation. In *Proc. of the* 5^{th} *Asia Information Retrieval Symposium on Information Retrieval Technology (AIRS)*, pages 168–180, Berlin, Heidelberg, 2009b. Springer-Verlag.

Jin Ha Lee. Analysis of User Needs and Information Features in Natural Language Queries seeking Music Information. *Journal of the American Society for Information Science and Technology (JASIST)*, 61(5):1025–1045, 2010.

Chee Wee Leong and Silviu Cucerzan. Supporting Factual Statements with Evidence from the Web. In *Proc. of the* 21^{st} *ACM Intl. Conf. on Information and Knowledge Management (CIKM)*, pages 1153–1162, New York, NY, USA, 2012. ACM.

Yunyao Li, Huahai Yang, and HV Jagadish. Constructing a Generic Natural Language Interface for an XML Database. In *Proc. of the 2006 Conf. on Advances in Database Technology (EDBT)*, pages 737–754, Berlin, Heidelberg, 2006. Springer-Verlag.

Christina Lioma and Iadh Ounis. A Syntactically-based Query Reformulation Technique for Information Retrieval. *Information processing & management (IPM)*, 44(1):143–162, 2008.

K Tamsin Maxwell and W Bruce Croft. Compact Query Term Selection using Topically Related Text. In *Proc. of the* 36^{th} *Intl. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR)*, pages 583–592, New York, NY, USA, 2013. ACM.

Donald Metzler and W. Bruce Croft. A Markov Random Field Model for Term Dependencies. In *Proc. of the* 28^{th} *Annual Intl. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR)*, pages 472–479, New York, NY, USA, 2005. ACM.

Nikita Mishra, Rishiraj Saha Roy, Niloy Ganguly, Srivatsan Laxman, and Monojit Choudhury. Unsupervised Query Segmentation using Only Query Logs. In *Proc. of the* 20^{th} *Intl. Conf. on World Wide Web (WWW)*, pages 91–92, New York, NY, USA, 2011. ACM.

Liqiang Nie, Shuicheng Yan, Meng Wang, Richang Hong, and Tat-Seng Chua. Harvesting visual concepts for image search with complex queries. In *Proceedings of the 20th ACM International Conference on Multimedia*, MM '12, pages 59–68, New York, NY, USA, 2012. ACM. ISBN 978-1-4503-1089-5. . URL `http://doi.acm.org/10.1145/2393347.2393363`.

Daan Odijk, Edgar Meij, Isaac Sijaranamual, and Maarten de Rijke. Dynamic Query Modeling for Related Content Finding. In *Proc. of the* 38^{th} *Annual Intl. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR)*. ACM, 2015.

Jiaul H. Paik and Douglas W. Oard. A Fixed-Point Method for Weighting Terms in Verbose Informational Queries. In *Proc. of the* 23^{rd} *ACM Conf. on Information and Knowledge Management (CIKM)*, pages 131–140, New York, NY, USA, 2014. ACM.

Nish Parikh, Prasad Sriram, and Mohammad Al Hasan. On Segmentation of E-Commerce Queries. In *Proc. of the* 22^{nd} *ACM Intl. Conf. on Information and Knowledge Management (CIKM)*, pages 1137–1146, New York, NY, USA, 2013. ACM.

Jae Hyun Park and W Bruce Croft. Query Term Ranking based on Dependency Parsing of Verbose Queries. In *Proc. of the* 33^{rd} *Intl. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR)*, pages 829–830, New York, NY, USA, 2010. ACM.

Jae Hyun Park, W Bruce Croft, and David A Smith. A Quasi-Synchronous Dependence Model for Information Retrieval. In *Proc. of the* 20^{th} *ACM Intl. Conf. on Information and Knowledge Management (CIKM)*, pages 17–26, New York, NY, USA, 2011. ACM.

Fuchun Peng, Nawaaz Ahmed, Xin Li, and Yumao Lu. Context Sensitive Stemming for Web Search. In *Proc. of the* 30^{th} *Annual Intl. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR)*, pages 639–646, New York, NY, USA, 2007. ACM.

Nina Phan, Peter Bailey, and Ross Wilkinson. Understanding the Relationship of Information Need Specificity to Search Query Length. In *Proc. of the* 30^{th} *Annual Intl. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR)*, pages 709–710, New York, NY, USA, 2007. ACM.

Jay M Ponte and W Bruce Croft. A Language Modeling Approach to Information Retrieval. In *Proc. of the* 21^{st} *Annual Intl. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR)*, pages 275–281, New York, NY, USA, 1998. ACM.

Knut Magne Risvik, Tomasz Mikolajewski, and Peter Boros. Query Segmentation for Web Search. In *Proc. of the* 12^{th} *Intl. Conf. on World Wide Web (WWW)*, New York, NY, USA, 2003. ACM.

Daniel Sheldon, Milad Shokouhi, Martin Szummer, and Nick Craswell. LambdaMerge: Merging the Results of Query Reformulations. In *Proc. of the* 4^{th} *ACM Intl. Conf. on Web Search and Data Mining (WSDM)*, pages 795–804, New York, NY, USA, 2011. ACM.

Gyanit Singh, Nish Parikh, and Neel Sundaresan. Rewriting Null E-Commerce Queries to Recommend Products. In *Proc. of the* 21^{st} *Intl. Conf. Companion on World Wide Web (WWW)*, pages 73–82, New York, NY, USA, 2012. ACM.

Mark D Smucker and James Allan. Lightening the Load of Document Smoothing for Better Language Modeling Retrieval. In *Proc. of the* 29^{th} *Annual Intl. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR)*, pages 699–700. ACM, 2006.

Richard Socher, Cliff C. Lin, Andrew Y. Ng, and Christopher D. Manning. Parsing Natural Scenes and Natural Language with Recursive Neural Networks. In *Proc. of the* 26^{th} *Intl. Conf. on Machine Learning (ICML)*, 2011.

Charles Sutton and Andrew McCallum. Introduction to Conditional Random Fields for Relational Learning. In Lise Getoor and Ben Taskar, editors, *Introduction to Statistical Relational Learning*. MIT Press, 2006.

Krysta M. Svore, Pallika H. Kanani, and Nazan Khan. How Good is a Span of Terms?: Exploiting Proximity to Improve Web Retrieval. In *Proc. of the* 33^{rd} *Annual Intl. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR)*, pages 154–161, New York, NY, USA, 2010. ACM.

Bin Tan and Fuchun Peng. Unsupervised Query Segmentation using Generative Language Models and Wikipedia. In *Proc. of the* 17^{th} *Intl. Conf. on World Wide Web (WWW)*, pages 347–356, New York, NY, USA, 2008. ACM.

Sergio D. Torres, Djoerd Hiemstra, and Pavel Serdyukov. An Analysis of Queries Intended to Search Information for Children. In *Proc. of the* 3^{rd} *Symposium on Information Interaction in Context (IIiX)*, pages 235–244, New York, NY, USA, 2010. ACM.

Manos Tsagkias, Maarten De Rijke, and Wouter Weerkamp. Hypergeometric Language Models for Republished Article Finding. In *Proc. of the* 34^{th} *Intl. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR)*, pages 485–494. ACM, 2011.

Xuanhui Wang and ChengXiang Zhai. Mining Term Association Patterns from Search Logs for Effective Query Reformulation. In *Proc. of the* 17^{th} *ACM Conf. on Information and Knowledge Management (CIKM)*, pages 479–488, New York, NY, USA, 2008. ACM.

Xiaobing Xue and W Bruce Croft. Modeling Subset Distributions for Verbose Queries. In *Proc. of the* 34^{th} *Intl. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR)*, pages 1133–1134, New York, NY, USA, 2011. ACM.

Xiaobing Xue and W Bruce Croft. Generating Reformulation Trees for Complex Queries. In *Proc. of the* 35^{th} *Intl. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR)*, pages 525–534, New York, NY, USA, 2012. ACM.

Xiaobing Xue, Jiwoon Jeon, and W. Bruce Croft. Retrieval Models for Question and Answer Archives. In *Proc. of the* 31^{st} *Annual Intl. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR)*, pages 475–482, New York, NY, USA, 2008. ACM.

Xiaobing Xue, Samuel Huston, and W Bruce Croft. Improving Verbose Queries using Subset Distribution. In *Proc. of the* 19^{th} *ACM Intl. Conf. on Information and Knowledge Management (CIKM)*, pages 1059–1068, New York, NY, USA, 2010. ACM.

Xiaobing Xue, Yu Tao, Daxin Jiang, and Hang Li. Automatically Mining Question Reformulation Patterns from Search Log Data. In *Proc. of the* 50^{th} *Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 187–192, Stroudsburg, PA, USA, 2012. Association for Computational Linguistics.

Bishan Yang, Nish Parikh, Gyanit Singh, and Neel Sundaresan. A Study of Query Term Deletion Using Large-Scale E-commerce Search Logs. In *Proc. of the* 36^{th} *European Conf. on Information Retrieval (ECIR)*, pages 235–246, Berlin, Heidelberg, 2014. Springer-Verlag.

Yin Yang, Nilesh Bansal, Wisam Dakka, Panagiotis Ipeirotis, Nick Koudas, and Dimitris Papadias. Query by Document. In *Proc. of the* 2^{nd} *ACM Intl. Conf. on Web Search and Data Mining (WSDM)*, pages 34–43, New York, NY, USA, 2009. ACM.

Jeonghe Yi and Farzin Maghoul. Mobile Search Pattern Evolution: The Trend and the Impact of Voice Queries. In *Proc. of the* 20^{th} *Intl. Conf. Companion on World Wide Web (WWW)*, pages 165–166, New York, NY, USA, 2011. ACM.

Wen-tau Yih, Joshua Goodman, and Vitor R Carvalho. Finding Advertising Keywords on Web Pages. In *Proc. of the* 15^{th} *Intl. Conf. on World Wide Web (WWW)*, pages 213–222, New York, NY, USA, 2006. ACM.

Chengxiang Zhai and John Lafferty. A Study of Smoothing Methods for Language Models applied to Ad hoc Information Retrieval. In *Proc. of the* 24^{th} *Annual Intl. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR)*, pages 334–342, New York, NY, USA, 2001. ACM.

Chengxiang Zhai and John Lafferty. A Study of Smoothing Methods for Language Models applied to Information Retrieval. *ACM Transactions on Information Systems (TOIS)*, 22(2):179–214, 2004.

Le Zhao and Jamie Callan. Term Necessity Prediction. In *Proc. of the* 19^{th} *ACM Intl. Conf. on Information and Knowledge Management (CIKM)*, pages 259–268, New York, NY, USA, 2010. ACM.

Index