Temporal Information Retrieval

Nattiya Kanhabua
L3S Research Center
Hanover, Germany
kanhabua@L3S.de

Roi Blanco
Yahoo Labs
London, United Kingdom
roi@yahoo-inc.com

Kjetil Nørvåg
NTNU
Trondheim, Norway
Kjetil.Norvag@idi.ntnu.no

now
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

N. Kanhabua, R. Blanco, and K. Nørvåg . *Temporal Information Retrieval.* Foundations and Trends® in Information Retrieval, vol. 9, no. 2, pp. 91–208, 2015.

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

ISBN: 978-1-68083-033-0

Foundations and Trends® in Information Retrieval
Volume 9, Issue 2, 2015

Editorial Board

Full text available at: http://dx.doi.org/10.1561/1500000043

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. 2 (2015) 91–208

DOI: 10.1561/1500000043

now
the essence of knowledge

Temporal Information Retrieval

Nattiya Kanhabua
L3S Research Center
Hanover, Germany
kanhabua@L3S.de

Roi Blanco
Yahoo Labs
London, United Kingdom
roi@yahoo-inc.com

Kjetil Nørvåg
NTNU
Trondheim, Norway
Kjetil.Norvag@idi.ntnu.no

Contents

Abstract

Temporal dynamics and how they impact upon various components of information retrieval (IR) systems have received a large share of attention in the last decade. In particular, the study of relevance in information retrieval can now be framed within the so-called *temporal IR approaches*, which explain how user behavior, document content and scale vary with time, and how we can use them in our favor in order to improve retrieval effectiveness. This survey provides a comprehensive overview of temporal IR approaches, centered on the following questions: *what* are temporal dynamics, *why* do they occur, and *when* and *how* to leverage temporal information throughout the search cycle and architecture. We first explain the general and wide aspects associated to temporal dynamics by focusing on the web domain, from content and structural changes to variations of user behavior and interactions. Next, we pinpoint several research issues and the impact of such temporal characteristics on search, essentially regarding processing dynamic content, temporal query analysis and time-aware ranking. We also address particular aspects of temporal information extraction (for instance, how to timestamp documents and generate temporal profiles of text). To this end, we present existing temporal search engines and applications in related research areas, e.g., exploration, summarization, and clustering of search results, as well as future event retrieval and prediction, where the time dimension also plays an important role.

N. Kanhabua, R. Blanco, and K. Nørvåg . *Temporal Information Retrieval.* Foundations and Trends® in Information Retrieval, vol. 9, no. 2, pp. 91–208, 2015. DOI: 10.1561/1500000043.

1

Introduction

During the last decade, information retrieval has been successful in providing everybody with easy access to the vast amount of information available on the Web. As illustrated in Figure 1.1, creating, handling, and sharing information on the Web has seen the unprecedented growth and change in recent years. Cornerstones for such development are new technical devices and corresponding changes in our everyday behaviors. Digital photos and videos create large data volumes and numerous artifacts. Participative content generation and sharing in Web 2.0 solutions and social interaction via networks and platforms have gained wide acceptance, ranging from media-specific sharing (e.g., Flickr) over text and video distribution channels (e.g., Twitter and Youtube) up to web-based documentation and sharing of nearly complete life histories as encouraged by Facebook.

While the current way of accessing the Web comprise a good baseline, an optimal access to the evolving Web requires new models and algorithms for retrieval, exploration, and analytics which go far beyond what is needed to access the current state of the Web. This includes taking into account the time dimension, structured semantic information available on the Web, as well as social media and network information.

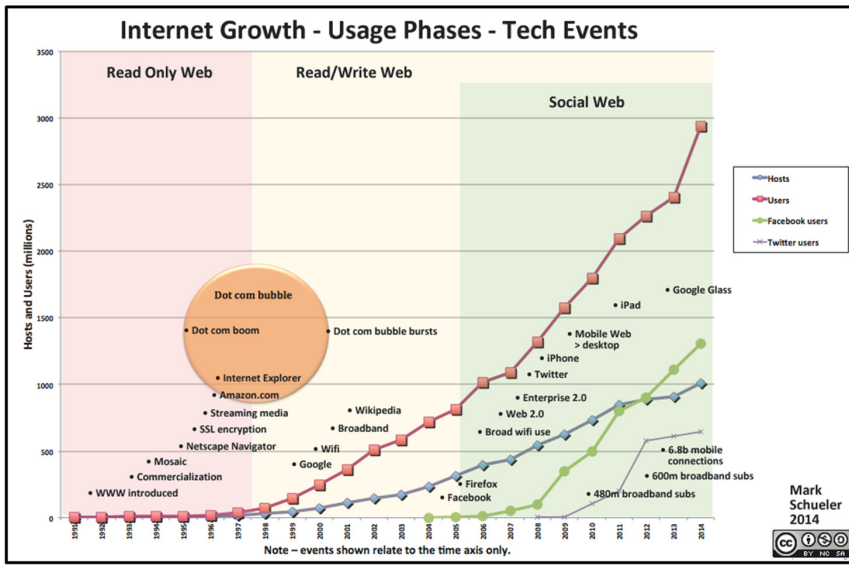

Figure 1.1: Internet Growth/Usage Phases/Tech Events (created by Mark Schueler, used with permission).

1.1 Temporal Dynamics

It is noteworthy that the time dimension has strong influence in many domains, e.g., Topic Detection and Tracking (TDT) (Allan et al., 1998; He et al., 2007), and Emerging Trend Detection (ETD) (Berry, 2003). However, in this context, we focus on the impact of time on the Web and we explain the evolution of the Web and its impact on web search and data mining before going into the details of temporal information retrieval. We will then discuss the scope and aim of this survey, and present the organization of the rest of the survey.

The Web has evolved in many aspects including its size, content, structure, and how it is accessed by people through web search engines. Such evolution has been previously discussed in (Ke et al., 2006; Risvik and Michelsen, 2002).

In this work, we aim at providing a comprehensive survey and answers to the following questions: *what* are temporal web dynamics, *why* do they occur, *when* and *how* to leverage the time dimension through-

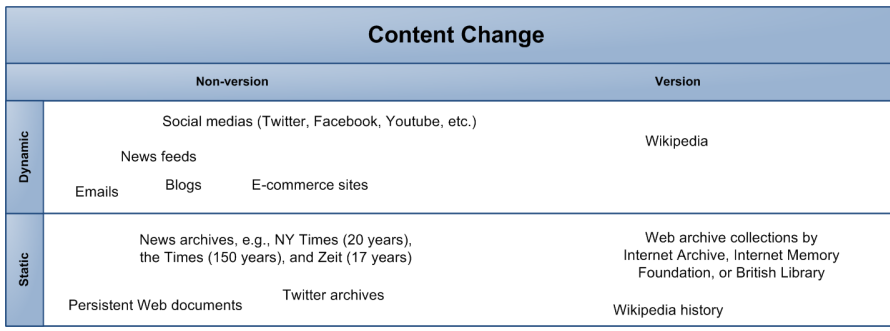

Figure 1.2: Categorization of documents by the degrees of content change, i.e., static or dynamic.

out the search cycle and architecture. For this purpose, we begin by explaining the general and wide aspects associated to temporal web dynamics, namely, the evolution of the Web categorized by its changes of 1) content and structure, and 2) user querying behavior.

1.1.1 Content and Structure Changes

The content of the Web, changes constantly over time, e.g., web documents are added, modified or deleted continuously. National and international initiatives have recognized this need and started to collect and preserve parts of the Web (Gomes et al., 2011; Costa et al., 2013). The most prominent one is the Internet Archive, which has collected more than 456 billion web pages (as of April 15, 2015) since 1996. Two important European initiatives include 1) the Internet Memory Foundation providing a set of smaller crawls for specific topics, domains and projects and 2) the British Library that aims at preserving national web content.

As illustrated in Figure 1.2, we categorize document collections, such as, personal homepages, corporate websites, Wikipedia articles and blogs, with respect to the various degrees of change, and whether the document creators or web sites keep different versions of each hosted document, i.e., versioning vs. non-versioning. On one hand, web archives are created by periodically visiting and crawling publicly available web pages. A web archive contains documents with multiple versions since

the new version of a document will be added into the archive repository when re-crawling. On the other hand, a web archive can have just one or the latest version for each document due to non-versioning policies, e.g., news archives, or real-time web data, such as, Twitter messages.

In parallel to content changing, the link structure of the Web also evolves (Dai and Davison, 2010a). The changes of content and structure affect basic processes like crawling and indexing, but also the computation of graph-based authority measures used for document ranking or spam detection.

1.1.2 Changes in User Behavior

Temporal web dynamics are related to user querying behavior in at least two ways. First, search traffic for particular queries varies over time and might present certain temporal patterns, such as, spikes, periodicity (e.g., weekly or monthly), seasonality and trends. Examples of sporadic or spiky queries are *breaking news* (e.g., iran, japan, earthquake), *celebrities* (e.g., beyonce, lady gaga), and *short-span events* (e.g., marathon, lollapalooza). Periodic or seasonal queries are, for instance, *annual events* (e.g., earth day, march madness, april fools' day, pgatour) and *television series* (e.g., american idol, crystal bowersox, dancing with the stars). Queries representing trends consist of *anticipated events* (e.g., iphone 7, mlb, miss usa), *past recent events* (e.g., easter ideas, final four), and *current events* (e.g., tax extension, presidential candidates).

Second, many queries are *time-sensitive queries*, which contain underlying temporal information needs that do not exhibit a temporal pattern in search streams. In other words, a time-sensitive query can be inferred to a particular time period, for example, an initial query Brazil FIFA World Cup might be later reformulated as 2014 FIFA World Cup. We categorize such queries with underlying temporal information needs into two types: 1) an explicit temporal query having temporal criteria explicitly provided by users (Berberich et al., 2007; Nørvåg, 2004), and 2) an implicit temporal query with no temporal criteria provided (Campos et al., 2012a; Kanhabua and Nørvåg, 2010a). An example of explicit temporal query is U.S. Presidential election 2016, whereas an implicit temporal query, e.g., Brazil FIFA World Cup, is likely to refer to the

most recent World Cup event in 2014 or the historical event in 1950. Note that, the temporal intent of the latter type can be determined using temporal information extraction techniques. Several studies of real-world user query logs have shown that temporal queries comprises a significant fraction of web search queries. For example, Zhang et al. (2010) showed that 13.8% of queries contain explicit time (Nunes et al. (2008) reported 1.5%) and 17.1% of queries have a temporal intent implicitly provided (7% reported by Metzler et al. (2009)).

Understanding temporal search intent is a challenging task that is the first step for applying an appropriate time-aware ranking method. In addition to the change in information needs, user interactions in the social Web are highly dynamic over time, e.g., comments, likes, interests as well as users' profiles. This affects how user interests/profiles should be modeled by taking into account such dynamics.

1.2 Scope and Aim of this Survey

This survey gives a comprehensive overview of the most important aspects of temporal information retrieval. It describes techniques involved in the complete pipeline of processing, from obtaining web documents, document processing and indexing, information extraction, and querying. It also gives an overview of application areas showing that its use extends well beyond simply searching web archives.

In addition to giving an extensive overview, we also intend that this survey should be self-contained enough to be used as lectures/teaching material for researchers that want to get acquainted with the research area. The survey can be read and understood by anybody with basic information retrieval knowledge, but should also be of use for more advanced researchers wanting to understand in more detail this field of research. As such it extends previous overviews of challenges and opportunities in temporal information retrieval (Alonso et al., 2011b), and the survey by Campos et al. (2014b).

The remainder of this survey is organized as follows: Section 2 describes research problems for the pre-processing step of temporal document collections, i.e., dynamic crawling and temporal indexing of web

documents. Section 3 presents current approaches to identifying and extracting of temporal information useful for leveraging in temporal information retrieval. Section 4 describes approaches to determining the temporal intents of queries, the effect of terminology changes over time, as well as query performance prediction for temporal queries. Section 5 describes a comparison of different time-aware ranking methods. Section 6 presents applications in information retrieval and related research areas where the time dimension also plays an important role, e.g., temporal analytics and exploration, temporal summarization, temporal clustering of search results, and future event retrieval and prediction. Section 7 concludes the survey and discusses possible research topics beyond what have been addressed in the survey. Finally, in Appendix A, we present existing research resources and recent evaluation workshops organized in the field of temporal information retrieval.

References

Adar, E., J. Teevan, S. T. Dumais, and J. L. Elsas (2009), 'The Web Changes Everything: Understanding the Dynamics of Web Content'. In: *Proceedings of the Second ACM International Conference on Web Search and Data Mining*. pp. 282–291.

Aji, A., Y. Wang, E. Agichtein, and E. Gabrilovich (2010), 'Using the Past to Score the Present: Extending Term Weighting Models Through Revision History Analysis'. In: *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*. pp. 629–638.

Alici, S., I. S. Altingovde, R. Ozcan, B. B. Cambazoglu, and O. Ulusoy (2011), 'Timestamp-based Result Cache Invalidation for Web Search Engines'. In: *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*. pp. 973–982.

Alici, S., I. S. Altingovde, R. Ozcan, B. B. Cambazoglu, and O. Ulusoy (2012), 'Adaptive Time-to-live Strategies for Query Result Caching in Web Search Engines'. In: *Proceedings of the 34th European Conference on Advances in Information Retrieval*. pp. 401–412.

Allan, J., R. Papka, and V. Lavrenko (1998), 'On-line New Event Detection and Tracking'. In: *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. pp. 37–45.

Alonso, O., R. Baeza-Yates, and M. Gertz (2009a), 'Effectiveness of Temporal Snippets'. In: *World Wide Web Conference Series*.

Alonso, O. and M. Gertz (2006), 'Clustering of Search Results Using Temporal Attributes'. In: *SIGIR '06: Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval*. pp. 597–598.

Alonso, O., M. Gertz, and R. Baeza-Yates (2011a), 'Enhancing Document Snippets Using Temporal Information'. In: *Proceedings of the 18th International Conference on String Processing and Information Retrieval*. pp. 26–31.

Alonso, O., M. Gertz, and R. A. Baeza-Yates (2009b), 'Clustering and Exploring Search Results Using Timeline Constructions'. In: *Proceedings of the 18th ACM conference on Information and knowledge management*. pp. 97–106.

Alonso, O., J. Strötgen, R. A. Baeza-Yates, and M. Gertz (2011b), 'Temporal Information Retrieval: Challenges and Opportunities'. In: *Proceedings of the 1st International Temporal Web Analytics Workshop (TWAW 2011)*.

Amodeo, G., G. Amati, and G. Gambosi (2011a), 'On Relevance, Time and Query Expansion'. In: *Proceedings of the 20th ACM International Conference on Information and Knowledge Management*. pp. 1973–1976.

Amodeo, G., R. Blanco, and U. Brefeld (2011b), 'Hybrid Models for Future Event Prediction'. In: *Proceedings of the 20th ACM International Conference on Information and Knowledge Management*. pp. 1981–1984.

Anand, A., S. Bedathur, K. Berberich, and R. Schenkel (2010), 'Efficient Temporal Keyword Search over Versioned Text'. In: *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*. pp. 699–708.

Anand, A., S. Bedathur, K. Berberich, and R. Schenkel (2011), 'Temporal Index Sharding for Space-time Efficiency in Archive Search'. In: *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*. pp. 545–554.

Anick, P. G. and R. A. Flynn (1992), 'Versioning a Full-text Information Retrieval System'. In: *Proceedings of the 15th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. pp. 98–111.

Arun, R., V. Suresh, C. E. Veni Madhavan, and M. N. Narasimha Murthy (2010), 'On Finding the Natural Number of Topics with Latent Dirichlet Allocation: Some Observations'. In: *Proceedings of the 14th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining - Volume Part I*. pp. 391–402.

Au Yeung, C.-m. and A. Jatowt (2011), 'Studying How the Past is Remembered: Towards Computational History Through Large Scale Text Mining'. In: *Proceedings of the 20th ACM International Conference on Information and Knowledge Management*. pp. 1231–1240.

Baeza-Yates, R. A. (2005), 'Searching the Future'. In: *Proceedings of SIGIR workshop on mathematical/formal methods in information retrieval MF/IR*.

Baeza-Yates, R. A. and B. A. Ribeiro-Neto (2011), *Modern Information Retrieval - the concepts and technology behind search, Second edition*. Pearson Education Ltd., Harlow, England.

Bar-Yossef, Z. and N. Kraus (2011), 'Context-sensitive Query Autocompletion'. In: *Proceedings of the 20th International Conference on World Wide Web*. pp. 107–116.

Barbosa, L., A. C. Salgado, F. de Carvalho, J. Robin, and J. Freire (2005), 'Looking at Both the Present and the Past to Efficiently Update Replicas of Web Content'. In: *Proceedings of the 7th Annual ACM International Workshop on Web Information and Data Management*. pp. 75–80.

Beitzel, S. M., E. C. Jensen, A. Chowdhury, O. Frieder, and D. Grossman (2007), 'Temporal Analysis of a Very Large Topically Categorized Web Query Log'. *J. Am. Soc. Inf. Sci. Technol.* **58**(2), 166–178.

Beitzel, S. M., E. C. Jensen, A. Chowdhury, D. Grossman, and O. Frieder (2004), 'Hourly Analysis of a Very Large Topically Categorized Web Query Log'. In: *Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. pp. 321–328.

Berberich, K., S. Bedathur, T. Neumann, and G. Weikum (2007), 'A Time Machine for Text Search'. In: *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*. pp. 519–526.

Berberich, K., S. J. Bedathur, O. Alonso, and G. Weikum (2010), 'A Language Modeling Approach for Temporal Information Needs'. In: *Proceedings of the 32nd European Conference on IR Research on Advances in Information Retrieval*. pp. 13–25.

Berberich, K., S. J. Bedathur, M. Sozio, and G. Weikum (2009), 'Bridging the Terminology Gap in Web Archive Search'. In: *Proceedings of the 12th International Workshop on the Web and Databases (WebDB)*.

Berberich, K., M. Vazirgiannis, and G. Weikum (2005), 'Time-Aware Authority Ranking'. *Internet Mathematics* **2**(3).

Berry, M. (2003), *Survey of Text Mining: Clustering, Classification, and Retrieval*. Springer.

Beyer, H. and K. Holtzblatt (1998), *Contextual Design: Defining Customer-centered Systems*. Morgan Kaufmann Publishers Inc.

Bian, J., X. Li, F. Li, Z. Zheng, and H. Zha (2010), 'Ranking Specialization for Web Search: A Divide-and-conquer Approach by Using Topical RankSVM'. In: *Proceedings of the 19th International Conference on World Wide Web*. pp. 131–140.

Blanco, R., E. Bortnikov, F. Junqueira, R. Lempel, L. Telloli, and H. Zaragoza (2010), 'Caching Search Engine Results over Incremental Indices'. In: *Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. pp. 82–89.

Blei, D. M., A. Y. Ng, and M. I. Jordan (2003), 'Latent dirichlet allocation'. *J. Mach. Learn. Res.* **3**, 993–1022.

Bøhn, C. and K. Nørvåg (2010), 'Extracting Named Entities and Synonyms from Wikipedia'. In: *Proceedings of the 24th IEEE International Conference on Advanced Information Networking and Applications AINA'2010*. pp. 1300–1307.

Boldi, P., B. Codenotti, M. Santini, and S. Vigna (2004), 'UbiCrawler: A Scalable Fully Distributed Web Crawler'. *Softw. Pract. Exper.* **34**(8), 711–726.

Box, G. E. P. and G. Jenkins (1990), *Time Series Analysis, Forecasting and Control*. Holden-Day, Incorporated.

Broder, A. Z., N. Eiron, M. Fontoura, M. Herscovici, R. Lempel, J. McPherson, R. Qi, and E. J. Shekita (2006), 'Indexing Shared Content in Information Retrieval Systems'. In: *Proceedings of the 10th International Conference on Extending Database Technology*. pp. 313–330.

Cambazoglu, B. B., F. P. Junqueira, V. Plachouras, S. Banachowski, B. Cui, S. Lim, and B. Bridge (2010), 'A Refreshing Perspective of Search Engine Caching'. In: *Proceedings of the 19th International Conference on World Wide Web*. pp. 181–190.

Campos, R., G. Dias, A. Jorge, and C. Nunes (2012a), 'GTE: a Distributional Second-Order Co-Occurrence Approach to Amprove the Identification of Top Relevant Dates in Web Snippets'. In: *Proceedings of the 21st ACM International Conference on Information and Knowledge Management*. pp. 2035–2039.

Campos, R., G. Dias, A. Jorge, and C. Nunes (2014a), 'GTECluster: A Temporal Search Interface for Implicit Temporal Queries'. In: *Proceedings of the 36th European Conference on IR Research*. pp. 775–779.

Campos, R., G. Dias, A. M. Jorge, and A. Jatowt (2014b), 'Survey of Temporal Information Retrieval and Related Applications'. *ACM Comput. Surv.* **47**(2), 15:1–15:41.

Campos, R., G. Dias, A. M. Jorge, and C. Nunes (2014c), 'GTE-Rank: Searching for Implicit Temporal Query Results'. In: *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management*. pp. 2081–2083.

Campos, R., A. M. Jorge, G. Dias, and C. Nunes (2012b), 'Disambiguating Implicit Temporal Queries by Clustering Top Relevant Dates in Web Snippets'. In: *Proceedings of the The 2012 IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technology - Volume 01*. pp. 1–8.

Cao, J., T. Xia, J. Li, Y. Zhang, and S. Tang (2009), 'A Density-Based Method for Adaptive LDA Model Selection'. *Neurocomput.* **72**(7-9), 1775–1781.

Carmel, D. and E. Yom-Tov (2010), *Estimating the Query Difficulty for Information Retrieval*. Morgan & Claypool Publishers.

Carterette, B. and P. Chandar (2009), 'Probabilistic Models of Ranking Novel Documents for Faceted Topic Retrieval'. In: *Proceedings of the 18th ACM Conference on Information and Knowledge Management*. pp. 1287–1296.

Ceroni, A., V. Solachidis, C. Niederée, O. Papadopoulou, N. Kanhabua, and V. Mezaris (2015), 'To Keep or not to Keep: An Expectation–oriented Photo Selection Method for Personal Photo Collections'. In: *Proceedings of International Conference on Multimedia Retrieval*.

Chambers, N. (2012), 'Labeling Documents with Timestamps: Learning from Their Time Expressions'. In: *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Long Papers - Volume 1*. pp. 98–106.

Chang, A. X. and C. Manning (2012), 'SUTime: A library for recognizing and normalizing time expressions'. In: *Proceedings of the Eighth International Conference on Language Resources and Evaluation*. pp. 3735–3740.

Chang, F., J. Dean, S. Ghemawat, W. C. Hsieh, D. A. Wallach, M. Burrows, T. Chandra, A. Fikes, and R. E. Gruber (2006), 'Bigtable: A Distributed Storage System for Structured Data'. In: *Proceedings of the 7th USENIX Symposium on Operating Systems Design and Implementation - Volume 7*. pp. 15–15.

Cheng, S., A. Arvanitis, and V. Hristidis (2013), 'How Fresh Do You Want Your Search Results?'. In: *Proceedings of the 22Nd ACM International Conference on Conference on Information and Knowledge Management*. pp. 1271–1280.

Cho, J. and H. Garcia-Molina (2000), 'The Evolution of the Web and Implications for an Incremental Crawler'. In: *Proceedings of the 26th International Conference on Very Large Data Bases*. pp. 200–209.

Cho, J. and H. Garcia-Molina (2003a), 'Effective Page Refresh Policies for Web Crawlers'. *ACM Trans. Database Syst.* **28**(4), 390–426.

Cho, J. and H. Garcia-Molina (2003b), 'Estimating Frequency of Change'. *ACM Trans. Internet Technol.* **3**(3), 256–290.

Ciglan, M. and K. Nørvåg (2010), 'WikiPop: Personalized Event Detection System Based on Wikipedia Page View Statistics'. In: *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*. pp. 1931–1932.

Cleveland, R. B., W. S. Cleveland, J. E. McRae, and I. Terpenning (1990), 'STL: A Seasonal-Trend Decomposition Procedure Based on Loess'. *Journal of Official Statistics* **6**, 3–73.

Cormack, G. V. and M. R. Grossman (2014), 'Evaluation of Machine-learning Protocols for Technology-assisted Review in Electronic Discovery'. In: *Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval*. pp. 153–162.

Corso, G. M. D., A. Gullí, and F. Romani (2005), 'Ranking a stream of news'. In: *Proceedings of the 14th international conference on World Wide Web*. pp. 97–106.

Costa, M., F. Couto, and M. Silva (2014), 'Learning Temporal-dependent Ranking Models'. In: *Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval*. pp. 757–766.

Costa, M., D. Gomes, F. Couto, and M. Silva (2013), 'A Survey of Web Archive Search Architectures'. In: *Proceedings of the 22nd International Conference on World Wide Web (Companion)*. pp. 1045–1050.

Croft, W. B., D. Metzler, and T. Strohman (2009), *Search Engines: Information Retrieval in Practice*. Addison-Wesley Publishing Company, 1st edition.

Cronen-Townsend, S., Y. Zhou, and W. B. Croft (2002), 'Predicting Query Performance'. In: *Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval*. pp. 299–306.

Dai, N. and B. D. Davison (2010a), 'Freshness Matters: in Flowers, Food, and Web Authority'. In: *Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval*. pp. 114–121.

Dai, N. and B. D. Davison (2010b), 'Mining Anchor Text Trends for Retrieval'. In: *Proceedings of the 32nd European Conference on Advances in Information Retrieval*. pp. 127–139.

Dai, N., M. Shokouhi, and B. D. Davison (2011), 'Learning to Rank for Freshness and Relevance'. In: *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*. pp. 95–104.

Dakka, W., L. Gravano, and P. G. Ipeirotis (2012), 'Answering General Time-Sensitive Queries'. *IEEE Transactions on Knowledge and Data Engineering* **24**(2), 220–235.

de Boer, V., M. van Someren, and B. J. Wielinga (2010), 'Extracting Historical Time Periods from the Web'. *Journal of the American Society for Information Science and Technology* **61**, 1888–1908.

de Jong, F., H. Rode, and D. Hiemstra (2005), 'Temporal Language Models for the Disclosure of Historical Text'. In: *Proceedings of the 16th International Conference of the Association for History and Computing*. pp. 161–168.

Demartini, G., M. M. S. Missen, R. Blanco, and H. Zaragoza (2010), 'Entity Summarization of News Articles'. In: *Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. pp. 795–796.

Deng, H., I. King, and M. R. Lyu (2009), 'Entropy-biased Models for Query Representation on the Click Graph'. In: *Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval*. pp. 339–346.

Diaz-Aviles, E., L. Drumond, L. Schmidt-Thieme, and W. Nejdl (2012), 'Real-time Top-N Recommendation in Social Streams'. In: *Proceedings of the sixth ACM conference on Recommender systems*. pp. 59–66.

Dong, A., Y. Chang, Z. Zheng, G. Mishne, J. Bai, R. Zhang, K. Buchner, C. Liao, and F. Diaz (2010a), 'Towards Recency Ranking in Web Search'. In: *Proceedings of the third ACM international conference on Web search and data mining*. pp. 11–20.

Dong, A., R. Zhang, P. Kolari, J. Bai, F. Diaz, Y. Chang, Z. Zheng, and H. Zha (2010b), 'Time is the Essence: improving Recency Ranking using Twitter Data'. In: *Proceedings of the 19th international conference on World wide web*. pp. 331–340.

Dou, Z., R. Song, J.-Y. Nie, and J.-R. Wen (2009), 'Using Anchor Texts with Their Hyperlink Structure for Web Search'. In: *Proceedings of the 32Nd International ACM SIGIR Conference on Research and Development in Information Retrieval*. pp. 227–234.

Efron, M. (2013), 'Query Representation for Cross-temporal Information Retrieval'. In: *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval*. pp. 383–392.

Efron, M., J. Lin, J. He, and A. de Vries (2014), 'Temporal Feedback for Tweet Search with Non-parametric Density Estimation'. In: *Proceedings of the 37th International ACM SIGIR Conference on Research and Development in Information Retrieval*. pp. 33–42.

Elsas, J. L. and S. T. Dumais (2010), 'Leveraging Temporal Dynamics of Document Content in Relevance Ranking'. In: *Proceedings of the third ACM international conference on Web search and data mining*. pp. 1–10.

Erkan, G. and D. R. Radev (2004), 'LexRank: Graph-based Lexical Centrality As Salience in Text Summarization'. *J. Artif. Int. Res.* **22**(1), 457–479.

Ernst-Gerlach, A. and N. Fuhr (2006), 'Generating Search Term Variants for Text Collections with Historic Spellings'. In: *Proceedings of the 28th European Conference on IR Research on Advances in Information Retrieval*. pp. 49–60.

Ernst-Gerlach, A. and N. Fuhr (2007), 'Retrieval in text collections with historic spelling using linguistic and spelling variants'. In: *Proceedings of the 7th ACM/IEEE-CS joint conference on Digital libraries*. pp. 333–341.

Fagni, T., R. Perego, F. Silvestri, and S. Orlando (2006), 'Boosting the Performance of Web Search Engines: Caching and Prefetching Query Results by Exploiting Historical Usage Data'. *ACM Trans. Inf. Syst.* **24**(1), 51–78.

Ferron, M. and P. Massa (2012), 'Psychological Processes Underlying Wikipedia Representations of Natural and Manmade Disasters'. In: *Proceedings of WikiSym '12*.

Fetterly, D., M. Manasse, M. Najork, and J. Wiener (2003), 'A Large-scale Study of the Evolution of Web Pages'. In: *Proceedings of the 12th International Conference on World Wide Web*. pp. 669–678.

Fontoura, M., R. Lempel, R. Qi, and J. Y. Zien (2007), 'Inverted Index Support for Numeric Search'. *Internet Mathematics* **3**(2), 153–185.

Garcia-Fernandez, A., A.-L. Ligozat, M. Dinarelli, and D. Bernhard (2011), 'When was it Written? Automatically Determining Publication Dates'. In: *Proceedings of the 18th international conference on String processing and information retrieval*. pp. 221–236.

Ge, T., B. Chang, S. Li, and Z. Sui (2013), 'Event-Based Time Label Propagation for Automatic Dating of News Articles'. In: *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*. pp. 1–11.

Georgescu, M., N. Kanhabua, D. Krause, W. Nejdl, and S. Siersdorfer (2013), 'Extracting Event-Related Information from Article Updates in Wikipedia'. In: *Proceedings of the 35th European conference on Advances in Information Retrieval*. pp. 254–266.

Goldberg, Y. and J. Orwant (2013), 'A dataset of syntactic-ngrams over time from a very large corpus of english books'. In: *Proceedings of Second Joint Conference on Lexical and Computational Semantics (* SEM)*, Vol. 1. pp. 241–247.

Gomes, D., J. a. Miranda, and M. Costa (2011), 'A Survey on Web Archiving Initiatives'. In: *Proceedings of the 15th International Conference on Theory and Practice of Digital Libraries: Research and Advanced Technology for Digital Libraries*. pp. 408–420.

Griffiths, T. L. and M. Steyvers (2004), 'Finding scientific topics'. *Proceedings of the National Academy of Sciences* **101**(Suppl. 1), 5228–5235.

Hamilton, J. D. (1994), *Time series analysis*, Vol. 2. Princeton university press Princeton.

Hauff, C. and L. Azzopardi (2005), 'Age Dependent Document Priors in Link Structure Analysis'. In: *Proceedings of the 27th European conference on Advances in Information Retrieval*. pp. 552–554.

Hauff, C., L. Azzopardi, and D. Hiemstra (2009), 'The Combination and Evaluation of Query Performance Prediction Methods'. In: *Proceedings of the 31st European Conference on IR Research on Advances in Information Retrieval*, Vol. 5478 of *ECIR '09*. pp. 301–312.

Hauff, C., L. Azzopardi, D. Hiemstra, and F. de Jong (2010), 'Query Performance Prediction: Evaluation Contrasted with Effectiveness'. In: *Proceedings of the 32nd European Conference on IR Research on Advances in Information Retrieval*. pp. 204–216.

Hauff, C., D. Hiemstra, and F. de Jong (2008), 'A Survey of Pre-Retrieval Query Performance Predictors'. In: *Proceedings of the 17th ACM conference on Information and knowledge management*. pp. 1419–1420.

He, J. and T. Suel (2011), 'Faster Temporal Range Queries over Versioned Text'. In: *Proceeding of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*. pp. 565–574.

He, J., H. Yan, and T. Suel (2009), 'Compact Full-text Indexing of Versioned Document Collections'. In: *Proceedings of the 18th ACM Conference on Information and Knowledge Management*. pp. 415–424.

He, Q., K. Chang, and E.-P. Lim (2007), 'Analyzing Feature Trajectories for Event Detection'. In: *Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. pp. 207–214.

Herscovici, M., R. Lempel, and S. Yogev (2007), 'Efficient Indexing of Versioned Document Sequences'. In: *Proceeding of the 29th European Conference on IR Research on Advances in Information Retrieval*. pp. 76–87.

Hiemstra, D. (2002), 'Term-specific Smoothing for the Language Modeling Approach to Information Retrieval: The Importance of a Query Term'. In: *Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. pp. 35–41.

Hoffart, J., F. M. Suchanek, K. Berberich, and G. Weikum (2013), 'YAGO2: A Spatially and Temporally Enhanced Knowledge Base from Wikipedia'. *Artif. Intell.* **194**, 28–61.

Holt, C. C. (2004), 'Forecasting Seasonals and Trends by Exponentially Weighted Moving Averages'. *International Journal of Forecasting* **20**(1), 5–10.

Holzmann, H., N. Tahmasebi, and T. Risse (2013), 'BlogNEER: Applying Named Entity Evolution Recognition on the Blogosphere?'. In: *Proceedings of the 3rd International Workshop on Semantic Digital Archives*. pp. 28–39.

Jatowt, A., É. Antoine, Y. Kawai, and T. Akiyama (2015), 'Mapping Temporal Horizons: Analysis of Collective Future and Past related Attention in Twitter'. In: *Proceedings of the 24th International Conference on World Wide Web*. pp. 484–494.

Jatowt, A. and C.-m. Au Yeung (2011), 'Extracting Collective Expectations About the Future from Large Text Collections'. In: *Proceedings of the 20th ACM International Conference on Information and Knowledge Management*. pp. 1259–1264.

Jatowt, A., C.-M. Au Yeung, and K. Tanaka (2013), 'Estimating Document Focus Time'. In: *Proceedings of the 22Nd ACM International Conference on Conference on Information & Knowledge Management*. pp. 2273–2278.

Jatowt, A., Y. Kawai, and K. Tanaka (2005), 'Temporal Ranking of Search Engine Results'. In: *Proceedings of the 6th International Conference on Web Information Systems Engineering*. pp. 43–52.

Jatowt, A., Y. Kawai, and K. Tanaka (2007), 'Detecting Age of Page Content'. In: *Proceedings of the 9th Annual ACM International Workshop on Web Information and Data Management*. pp. 137–144.

Jatowt, A. and K. Tanaka (2012), 'Large Scale Analysis of Changes in English Vocabulary over Recent Time'. In: *Proceedings of the 21st ACM International Conference on Information and Knowledge Management*. pp. 2523–2526.

Joho, H., A. Jatowt, and R. Blanco (2014a), 'NTCIR Temporalia: a Test Collection for Temporal Information Access Research'. In: *Proceedings of the 23rd International Conference on World Wide Web (Companion)*. pp. 845–850.

Joho, H., A. Jatowt, R. Blanco, H. Naka, and S. Yamamoto (2014b), 'Overview of NTCIR-11 Temporal Information Access (Temporalia) Task'. In: *Proceedings of the NTCIR-11 Conference, Tokyo, Japan*.

Joho, H., A. Jatowt, and B. Roi (2013), 'A Survey of Temporal Web Search Experience'. In: *Proceedings of the 22nd International Conference on World Wide Web (Companion)*. pp. 1101–1108.

Jones, R. and F. Diaz (2007), 'Temporal Profiles of Queries'. *ACM Trans. Inf. Syst.* **25**.

Kalczynski, P. J. and A. Chou (2005), 'Temporal Document Retrieval Model for Business News Archives'. *Inf. Process. Manage.* **41**, 635–650.

Kaluarachchi, A. C., A. S. Varde, S. Bedathur, G. Weikum, J. Peng, and A. Feldman (2010a), 'Incorporating terminology evolution for query translation in text retrieval with association rules'. In: *Proceedings of the 19th ACM international conference on Information and knowledge management*. pp. 1789–1792.

Kaluarachchi, A. C., A. S. Varde, J. Peng, and A. Feldman (2010b), 'Intelligent Time-Aware Query Translation for Text Sources'. In: *Proceedings of the 24th AAAI Conference on Artificial Intelligence*.

Kanhabua, N., R. Blanco, and M. Matthews (2011), 'Ranking Related News Predictions'. In: *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval*. pp. 755–764.

Kanhabua, N. and W. Nejdl (2013), 'Understanding the Diversity of Tweets in the Time of Outbreaks'. In: *Proceedings of the 22nd International Conference on World Wide Web (Companion Volume)*. pp. 1335–1342.

Kanhabua, N. and W. Nejdl (2014), 'On the Value of Temporal Anchor Texts in Wikipedia'. In: *SIGIR 2014 Workshop on Temporal, Social and Spatially-aware Information Access (TAIA'2014)*.

Kanhabua, N., T. N. Nguyen, and W. Nejdl (2015), 'Learning to Detect Event-Related Queries for Web Search'. In: *Proceedings of the 24th International Conference on World Wide Web Companion - Companion Volume*. pp. 1339–1344.

Kanhabua, N., T. N. Nguyen, and C. Niederée (2014), 'What Triggers Human Remembering of Events?: A Large-scale Analysis of Catalysts for Collective Memory in Wikipedia'. In: *Proceedings of the 14th ACM/IEEE-CS Joint Conference on Digital Libraries*. pp. 341–350.

Kanhabua, N. and K. Nørvåg (2008), 'Improving Temporal Language Models for Determining Time of Non-timestamped Documents'. In: *Proceedings of the 12th European conference on Research and Advanced Technology for Digital Libraries*. pp. 358–370.

Kanhabua, N. and K. Nørvåg (2010a), 'Determining Time of Queries for Re-ranking Search Results'. In: *Proceedings of the 14th European conference on Research and advanced technology for digital libraries*. pp. 261–272.

Kanhabua, N. and K. Nørvåg (2010b), 'Exploiting Time-based Synonyms in Searching Document Archives'. In: *Proceedings of the 10th ACM/IEEE Joint Conference on Digital Libraries*. pp. 79–88.

Kanhabua, N. and K. Nørvåg (2010c), 'QUEST: Query Expansion using Synonyms over Time'. In: *Proceedings of the 2010 European conference on Machine learning and knowledge discovery in databases: Part III*. pp. 595–598.

Kanhabua, N. and K. Nørvåg (2011), 'Time-based Query Performance Predictors'. In: *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval*. pp. 1181–1182.

Kanhabua, N. and K. Nørvåg (2012), 'Learning to Rank Search Results for Time-sensitive Queries'. In: *Proceedings of the 21st ACM International Conference on Information and Knowledge Management*. pp. 2463–2466.

Kanhabua, N., S. Romano, and A. Stewart (2012), 'Identifying Relevant Temporal Expressions for Real-World Events'. In: *SIGIR 2012 Workshop on Time-aware Information Access (TAIA'2012)*.

Ke, Y., L. Deng, W. Ng, and D.-L. Lee (2006), 'Web Dynamics and Their Ramifications for the Development of Web Search Engines'. *Computer Networks* **50**(10), 1430–1447.

Keikha, M., S. Gerani, and F. Crestani (2011a), 'TEMPER: A Temporal Relevance Feedback Method'. In: *Proceedings of the 33rd European Conference on IR Research on Advances in Information Retrieval*. pp. 436–447.

Keikha, M., S. Gerani, and F. Crestani (2011b), 'Time-based Relevance Models'. In: *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval*. pp. 1087–1088.

Kleinberg, J. (2003), 'Bursty and Hierarchical Structure in Streams'. *Data Min. Knowl. Discov.* **7**, 373–397.

Koen, D. B. and W. Bender (2000), 'Time Frames: Temporal Augmentation of the News'. *IBM Systems Journal* **39**(3&4), 597–616.

Kotsakos, D., T. Lappas, D. Kotzias, D. Gunopulos, N. Kanhabua, and K. Nørvåg (2014), 'A Burstiness-aware Approach for Document Dating'. In: *Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval*. pp. 1003–1006.

Kraaij, W. (2005), 'Variations on Language Modeling for Information Retrieval'. *SIGIR Forum* **39**(1), 61.

Kulkarni, A., J. Teevan, K. M. Svore, and S. T. Dumais (2011), 'Understanding Temporal Query Dynamics'. In: *Proceedings of the Forth International Conference on Web Search and Web Data Mining*. pp. 167–176.

Kullback, S. and R. A. Leibler (1951), 'On Information and Sufficiency'. *Ann. Math. Statist.* **22**(1), 79–86.

Kumar, A., M. Lease, and J. Baldridge (2011), 'Supervised Language Modeling for Temporal Resolution of Texts'. In: *Proceedings of the 20th ACM international conference on Information and knowledge management*. pp. 2069–2072.

Lampos, V. and N. Cristianini (2012), 'Nowcasting Events from the Social Web with Statistical Learning'. *ACM Trans. Intell. Syst. Technol.* **3**(4), 72:1–72:22.

Lavrenko, V. and W. B. Croft (2001), 'Relevance based language models'. In: *Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval*. pp. 120–127.

Lester, N., A. Moffat, and J. Zobel (2008), 'Efficient Online Index Construction for Text Databases'. *ACM Trans. Database Syst.* **33**(3), 19:1–19:33.

Li, C. and A. Sun (2014), 'Fine-grained Location Extraction from Tweets with Temporal Awareness'. In: *Proceedings of the 37th International ACM SIGIR Conference on Research and Development in Information Retrieval*. pp. 43–52.

Li, C., Y. Wang, P. Resnick, and Q. Mei (2014), 'ReQ-ReC: High Recall Retrieval with Query Pooling and Interactive Classification'. In: *Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval*. pp. 163–172.

Li, X. and W. B. Croft (2003), 'Time-based Language mMdels'. In: *Proceedings of the 12th international conference on Information and knowledge management*. pp. 469–475.

Llidó, D., R. B. Llavori, and M. J. A. Cabo (2001), 'Extracting Temporal References to Assign Document Event-Time Periods'. In: *Proceedings of the 12th International Conference on Database and Expert Systems Applications*. pp. 62–71.

Lumezanu, C., N. Feamster, and H. Klein (2012), '#bias: Measuring the Tweeting Behavior of Propagandists'. In: *Proceedings of the Sixth International Conference on Weblogs and Social Media*.

Mani, I. and G. Wilson (2000), 'Robust Temporal Processing of News'. In: *Proceedings of the 38th Annual Meeting on Association for Computational Linguistics*. pp. 69–76.

Manning, C. D., P. Raghavan, and H. Schütze (2008), *Introduction to Information Retrieval*. Cambridge University Press.

Matthews, M., P. Tolchinsky, R. Blanco, J. Atserias, P. Mika, and H. Zaragoza (2010), 'Searching Through Time in the New York Times'. In: *HCIR Workshop on Bridging Human-Computer Interaction and Information Retrieval*.

Mazeika, A., T. Tylenda, and G. Weikum (2011), 'Entity Timelines: Visual Analytics and Named Entity Evolution'. In: *Proceedings of the 20th ACM International Conference on Information and Knowledge Management*. pp. 2585–2588.

McCreadie, R., C. Macdonald, and I. Ounis (2014), 'Incremental Update Summarization: Adaptive Sentence Selection Based on Prevalence and Novelty'. In: *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management*. pp. 301–310.

Meeter, M., J. M. J. Murre, and S. M. J. Janssen (2005), 'Remembering the News: Modeling Retention Data from a Study with 14,000 Participants'. *Memory & Cognition* **33**(5), 793–810.

Metzler, D., R. Jones, F. Peng, and R. Zhang (2009), 'Improving Search Relevance for Implicitly Temporal Queries'. In: *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*. pp. 700–701.

Mishra, N., R. W. White, S. Ieong, and E. Horvitz (2014), 'Time-critical Search'. In: *Proceedings of the 37th International ACM SIGIR Conference on Research and Development in Information Retrieval*. pp. 747–756.

Nguyen, T. N. and N. Kanhabua (2014), 'Leveraging Dynamic Query Subtopics for Time-aware Search Result Diversification'. In: *Proceedings of the 36th European Conference on Advances in Information Retrieval*. pp. 222–234.

Niederée, C., N. Kanhabua, F. Gallo, and R. H. Logie (2015), 'Forgetful Digital Memory: Towards Brain-Inspired Long-Term Data and Information Management'. *(To appear) SIGMOD Record*.

Nørvåg, K. (2004), 'Supporting Temporal Text-Containment Queries in Temporal Document Databases'. *Journal of Data & Knowledge Engineering* **49**(1), 105–125.

Nørvåg, K. and A. O. Nybø (2005), 'Improving Space-Efficiency in Temporal Text-Indexing'. In: *Proceedings of 10th International Conference on Database Systems for Advanced Applications*. pp. 791–802.

Nørvåg, K. and A. O. Nybø (2006), 'DyST: Dynamic and Scalable Temporal Text Indexing'. In: *Proceedings of the 13th International Symposium on Temporal Representation and Reasoning*. pp. 204–211.

Ntoulas, A., J. Cho, and C. Olston (2004), 'What's New on the Web?: The Evolution of the Web from a Search Engine Perspective'. In: *Proceedings of the 13th International Conference on World Wide Web*. pp. 1–12.

Nunes, S., C. Ribeiro, and G. David (2007), 'Using Neighbors to Date Web Documents'. In: *Proceedings of the 9th annual ACM international workshop on Web information and data management*. pp. 129–136.

Nunes, S., C. Ribeiro, and G. David (2008), 'Use of Temporal Expressions in Web Search'. In: *Proceedings of the 30th European Conference on IR Research on Advances in Information Retrieval*. pp. 580–584.

Odijk, D., G. Santucci, M. d. Rijke, M. Angelini, and G. Granato (2012), 'Time-Aware Exploratory Search: Exploring Word Meaning through Time'. In: *SIGIR 2012 Workshop on Time-aware Information Access*.

Olston, C. and M. Najork (2010), 'Web Crawling'. *Found. Trends Inf. Retr.* **4**(3), 175–246.

Parikh, N. and N. Sundaresan (2008), 'Scalable and Near Real-time Burst Detection from eCommerce Queries'. In: *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. pp. 972–980.

Peetz, M.-H. and M. de Rijke (2013), 'Cognitive Temporal Document Priors'. In: *Proceedings of the 35th European conference on Advances in Information Retrieval*. pp. 318–330.

Peetz, M.-H., E. Meij, and M. Rijke (2014), 'Using Temporal Bursts for Query Modeling'. *Information Retrieval* **17**(1), 74–108.

Pentzold, C. (2009), 'Fixing the Floating Gap: The Online Encyclopedia Wikipedia as a Global Memory Place'. *Memory Studies* **2**(2), 255–272.

Perkiö, J., W. Buntine, and H. Tirri (2005), 'A temporally adaptive content-based relevance ranking algorithm'. In: *Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval*. pp. 647–648.

Pickens, J., M. Cooper, and G. Golovchinsky (2010), 'Reverted Indexing for Feedback and Expansion'. In: *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*. pp. 1049–1058.

Ponte, J. M. and W. B. Croft (1998), 'A Language Modeling Approach to Information Retrieval'. In: *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. pp. 275–281.

Radinsky, K., E. Agichtein, E. Gabrilovich, and S. Markovitch (2011), 'A Word at a Time: Computing Word Relatedness Using Temporal Semantic Analysis'. In: *Proceedings of the 20th International Conference on World Wide Web*. pp. 337–346.

Radinsky, K. and E. Horvitz (2013), 'Mining the Web to Predict Future Events'. In: *Proceedings of the Sixth ACM International Conference on Web Search and Data Mining*. pp. 255–264.

Radinsky, K., K. Svore, S. Dumais, J. Teevan, A. Bocharov, and E. Horvitz (2012), 'Modeling and Predicting Behavioral Dynamics on the Web'. In: *Proceedings of the 21st international conference on World Wide Web*. pp. 599–608.

Rayson, P., D. Archer, and N. Smith (2005), 'VARD versus WORD: A Comparison of the UCREL Variant Detector and Modern Spellcheckers on English Historical Corpora'. In: *Corpus Linguistics 2005*.

Richardson, M. (2008), 'Learning About the World Through Long-term Query Logs'. *ACM Trans. Web* **2**(4), 21:1–21:27.

Risvik, K. M. and R. Michelsen (2002), 'Search engines and Web dynamics'. *Computer Networks* **39**(3), 289–302.

Rybak, J., K. Balog, and K. Nørvåg (2014), 'Temporal Expertise Profiling'. In: *Proceedings of the 36th European Conference on IR Research*. pp. 540–546.

Sakaki, T., M. Okazaki, and Y. Matsuo (2010), 'Earthquake Shakes Twitter Users: Real-time Event Detection by Social Sensors'. In: *Proceedings of the 19th international conference on World wide web*. pp. 851–860.

SalahEldeen, H. M. and M. L. Nelson (2013), 'Carbon dating the web: estimating the age of web resources'. In: *Proceedings of the 22nd international conference on World Wide Web (Companion)*. pp. 1075–1082.

Sang, E. T. K. and J. Bos (2012), 'Predicting the 2011 Dutch Senate Election Results with Twitter'. In: *Proceedings of the Workshop on Semantic Analysis in Social Media*. pp. 53–60.

Schilder, F. and C. Habel (2003), 'Temporal Information Extraction for Temporal Question Answering'. In: *New Directions in Question Answering, Papers from 2003 AAAI Spring Symposium*. pp. 35–44.

Shan, D., W. X. Zhao, R. Chen, B. Shu, Z. Wang, J. Yao, H. Yan, and X. Li (2012), 'EventSearch: A System for Event Discovery and Retrieval on Multitype Historical Data'. In: *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. pp. 1564–1567.

Shaparenko, B., R. Caruana, J. Gehrke, and T. Joachims (2005), 'Identifying Temporal Patterns and Key Players in Document Collections'. In: *Proceedings of IEEE ICDM Workshop on Temporal Data Mining: Algorithms, Theory and Applications*. pp. 165–174.

Shokouhi, M. (2011), 'Detecting Seasonal Queries by Time-series Analysis'. In: *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*. pp. 1171–1172.

Shokouhi, M. and K. Radinsky (2012), 'Time-sensitive Query Autocompletion'. In: *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval*. pp. 601–610.

Sipos, R., A. Swaminathan, P. Shivaswamy, and T. Joachims (2012), 'Temporal Corpus Summarization Using Submodular Word Coverage'. In: *Proceedings of the 21st ACM International Conference on Information and Knowledge Management*. pp. 754–763.

Song, W., Y. Zhang, H. Gao, T. Liu, and S. Li (2011), 'HITSCIR System in NTCIR-9 Subtopic Mining Task'.

Spina, D., J. Gonzalo, and E. Amigó (2014), 'Learning Similarity Functions for Topic Detection in Online Reputation Monitoring'. In: *Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval*. pp. 527–536.

Stefanidis, K., E. Ntoutsi, M. Petropoulos, K. Nørvåg, and H.-P. Kriegel (2013), 'A Framework for Modeling, Computing and Presenting Time-Aware Recommendations'. *T. Large-Scale Data- and Knowledge-Centered Systems* **10**, 146–172.

Strötgen, J., O. Alonso, and M. Gertz (2012), 'Identification of Top Relevant Temporal Expressions in Documents'. In: *Proceedings of the 2nd Temporal Web Analytics Workshop*. pp. 33–40.

Strötgen, J. and M. Gertz (2010), 'HeidelTime: High Quality Rule-based Extraction and Normalization of Temporal Expressions'. In: *Proceedings of the 5th International Workshop on Semantic Evaluation*. pp. 321–324.

Strötgen, J. and M. Gertz (2012), 'Event-centric Search and Exploration in Document Collections'. In: *Proceedings of the 12th ACM/IEEE-CS Joint Conference on Digital Libraries*. pp. 223–232.

Styskin, A., F. Romanenko, F. Vorobyev, and P. Serdyukov (2011), 'Recency Ranking by Diversification of Result Set'. In: *Proceedings of the 20th ACM International Conference on Information and Knowledge Management*. pp. 1949–1952.

Svore, K. M., J. Teevan, S. T. Dumais, and A. Kulkarni (2012), 'Creating Temporally Dynamic Web Search Snippets'. In: *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval*. pp. 1045–1046.

Tahmasebi, N., G. Gossen, N. Kanhabua, H. Holzmann, and T. Risse (2012), 'NEER: An Unsupervised Method for Named Entity Evolution Recognition'. In: *Proceedings the 24th International Conference on Computational Linguistics*. pp. 2553–2568, ACL.

Tahmasebi, N., T. Iofciu, T. Risse, C. Niedereé, and W. Siberski (2008), 'Terminology Evolution in Web Archiving: Open Issues'. In: *Proceedings of the 8th IWAW*.

Tilahun, G., A. Feuerverger, and M. Gervers (2012), 'Dating Medieval English Charters'. *The Annals of Applied Statistics* **6**(4), 1615–1640.

Tran, G. B., T. Tran, N.-K. Tran, M. Alrifai, and N. Kanhabua (2013), 'Leveraging Learning To Rank in an Optimization Framework for Timeline Summarization'. In: *SIGIR 2013 Workshop on Time-aware Information Access (TAIA'2013)*.

Tran, N. K., A. Ceroni, N. Kanhabua, and C. Niederée (2015), 'Back to the Past: Supporting Interpretations of Forgotten Stories by Time-aware Re-Contextualization'. In: *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*. pp. 339–348.

UzZaman, N., R. Blanco, and M. Matthews (2012), 'TwitterPaul: Extracting and Aggregating Twitter Predictions'. *CoRR* **abs/1211.6496**.

Verhagen, M., I. Mani, R. Sauri, J. Littman, R. Knippen, S. B. Jang, A. Rumshisky, J. Phillips, and J. Pustejovsky (2005), 'Automating Temporal Annotation with TARSQI'. In: *Proceedings of ACL'2005*.

Vlachos, M., C. Meek, Z. Vagena, and D. Gunopulos (2004), 'Identifying Similarities, Periodicities and Bursts for Online Search Queries'. In: *Proceedings of the 2004 ACM SIGMOD International Conference on Management of Data*. pp. 131–142.

White, R. W., P. N. Bennett, and S. T. Dumais (2010), 'Predicting Short-term Interests Using Activity-based Search Context'. In: *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*. pp. 1009–1018.

Whiting, S. and J. M. Jose (2014), 'Recent and Robust Query Auto-completion'. In: *Proceedings of the 23rd International Conference on World Wide Web*. pp. 971–982.

Whiting, S., K. Zhou, J. Jose, and M. Lalmas (2013), 'Temporal Variance of Intents in Multi-faceted Event-driven Information Needs'. In: *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval*. pp. 989–992.

Yan, R., X. Wan, J. Otterbacher, L. Kong, X. Li, and Y. Zhang (2011), 'Evolutionary Timeline Summarization: A Balanced Optimization Framework via Iterative Substitution'. In: *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*. pp. 745–754.

Yu, P. S., X. Li, and B. Liu (2004), 'On the Temporal Dimension of Search'. In: *Proceedings of the 13th international World Wide Web conference on Alternate track papers & posters*. pp. 448–449.

Zhai, C. and J. D. Lafferty (2004), 'A study of smoothing methods for language models applied to information retrieval'. *ACM Trans. Inf. Syst.* **22**(2), 179–214.

Zhang, J. and T. Suel (2007), 'Efficient Search in Large Textual Collections with Redundancy'. In: *Proceedings of the 16th International Conference on World Wide Web*. pp. 411–420.

Zhang, R., Y. Konda, A. Dong, P. Kolari, Y. Chang, and Z. Zheng (2010), 'Learning Recurrent Event Queries for Web Search'. In: *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*. pp. 1129–1139.

Zhao, X. W., Y. Guo, R. Yan, Y. He, and X. Li (2013), 'Timeline Generation with Social Attention'. In: *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval*. pp. 1061–1064.

Zhou, K., S. Whiting, J. M. Jose, and M. Lalmas (2013), 'The Impact of Temporal Intent Variability on Diversity Evaluation'. In: *Proceedings of the 35th European Conference on Advances in Information Retrieval*. pp. 820–823.