Survey of Temporal Information Retrieval and Related Applications

R. CAMPOS
Polytechnic Institute of Tomar, Tomar, Portugal
LIAAD – INESC TEC
Center of Mathematics, University of Beira Interior, Covilhã, Portugal
G. DIAS
HULTECH/GREYC, University of Caen Basse-Normandie, France
A. M. JORGE
LIAAD – INESC TEC
DCC-FCUP, University of Porto, Porto, Portugal
AND
A. JATOWT
Graduate School of Informatics, Kyoto University, Japan
Japan Science and Technology Agency, Japan

Temporal information retrieval has been a topic of great interest in recent years. Its purpose is to improve the effectiveness of information retrieval methods by exploiting temporal information in documents and queries. In this article, we present a survey of the existing literature on temporal information retrieval. In addition to giving an overview of the field, we categorize the relevant research, describe the main contributions, and compare different approaches. We organize existing research to provide a coherent view, discuss several open issues, and point out some possible future research directions in this area. Despite significant advances, the area lacks a systematic arrangement of prior efforts and an overview of state-of-the-art approaches. Moreover, an effective end-to-end temporal retrieval system that exploits temporal information to improve the quality of the presented results remain undeveloped.

Categories and Subject Descriptors: H 3.3 **[Information Storage and Retrieval]**: Information Search and Retrieval – *Clustering*, *Query Formulation*, *Retrieval Models*, *Search Process*.

General Terms: Algorithms, Theory

Additional Key Words and Phrases: Temporal Information Retrieval, Temporal Queries, Query Classification, Time-Based Clustering, Temporal Search Engines

ACM File Format:
CAMPOS, R.., DIAS, G., JORGE, A.M., AND JATOWT, A. 2014. Temporal Information Retrieval. *ACM Comput. Surv.* 6, 3, Article 9 (April 2014), 45 pages. DOI = 10.1145/1290002.1290003 http://doi.acm.org/10.1145/1290002.1290003

This research was funded by Project NORTE-07-0124-FEDER-000059 which is financed by the North Portugal Regional Operational Programme (ON.2 – O Novo Norte), under the National Strategic Reference Framework (NSRF), through the European Regional Development Fund (ERDF), and by national funds, through the Portuguese funding agency (FCT). It was also financed by the ERDF through the COMPETE Programme (operational programme for competitiveness), by National Funds through the FCT within project « FCOMP-01-0124-FEDER-037281» and by the Center of Mathematics, University of Beira Interior, within project «PEst-OE/MAT/UI0212/2014»

Author addresses: R. Campos, ICT Department, Polytechnic Institute of Tomar, Tomar, Portugal, LIAAD – INESC TEC, Portugal E-mail: ricardo.campos@inescporto.pt; G. Dias, HULTECH/GREYC, University of Caen Basse-Normandie, France; E-mail: gael.dias@unicaen.fr; A.M. Jorge, LIAAD – INESC TEC, DCC – Faculty of Sciences, University of Porto, Portugal; E-mail: amjorge@fc.up.pt; A. Jatowt, Graduate School of Informatics, Kyoto University, Japan; E-mail: adam@dl.kuis.kyoto-u.ac.jp

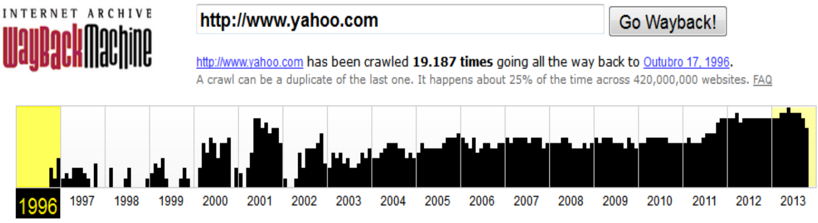
© 2001 ACM 1530-0226/07/0900-ART9 $5.00 DOI 10.1145/1290002.1290003 http://doi.acm.org/10.1145/1290002.1290003

1. INTRODUCTION

Information retrieval studies search for mechanisms to provide users with the most relevant documents from an existing collection. The user information needs are expressed by a query, typically in a short textual form. In recent years the issue of time has been gaining importance within search contexts, leading to a new research area known as temporal information retrieval (*T-IR*) that comprises a number of different challenges. In general, T-IR aims to satisfy search needs by combining the traditional notion of *document relevance* with *temporal relevance*. For example, users may require documents that describe the past (e.g., queries about historical figures), documents containing the most recent, up-to-date information (e.g., queries about weather forecasts or currency rates), or even future-related information (e.g., queries about planned events in a certain area). Information science researchers [Metzger 2007] tend to consider timeliness or currency as one of the five key aspects that determine a document's quality; the others are relevance, accuracy, objectivity, and coverage. The value of information and its quality are intrinsically time-dependent.

The huge volume of the web, however, makes T-IR a difficult task. First, since the web is constantly changing, maintaining up-to-date indexes is becoming more and more difficult. Second, a clear understanding of the temporal nature of queries is difficult due to query ambiguity, different temporal characteristics of queries or even unknown users' expectations towards the temporality of search results. Third, it is not easy to retrieve web documents so that their temporal dimension will meet the user temporal intent underlying the query. Nevertheless, researchers started to address the problem of retrieving web pages that are not only topically relevant but also created during (or that refer to) the most relevant time periods. They also approached the problem of determining various temporal dimensions of documents and queries. These contributions can greatly benefit the process of indexing documents, as well as the ranking of web search results or the clustering of documents.

The importance of considering temporal aspects in IR and the need for a continuous search for effective T-IR solutions becomes clear in light of the recent emergence of numerous *temporal initiatives and applications*. One of the first is the Internet Archive project [Kahle 1997] which is compiling a digital library of websites. Its objective is to store past versions of websites based on their periodical crawls. The archive has been used by computer scientists, information scientists and historians as a way to preserve, provide access, search, extract and visualize the different past versions of a web page. The information collected reportedly grows at a rate of 100 terabytes each month reaching an impressive number of over 350 billion archived web pages. Fig. 1 gives an overview of the increasing number of crawls for the example URL, www.yahoo.com. Each year in the timeline is divided into twelve black bars representing volumes of monthly crawls.



http://www.yahoo.com

http://www.yahoo.com has been crawled **19.187 times** going all the way back to Outubro 17, 1996.

A crawl can be a duplicate of the last one. It happens about 25% of the time across 420,000,000 websites.

Fig. 1: Results of Internet Archive for Yahoo! website, extracted from https://archive.org/

Further evidence is the recent development of research projects that address the archiving, the analysis, and the access to the temporal web. Examples include ARCOMEM[1], LAWA[2], LiWA[3] and LivingKnowledge[4]. There is also much research on using temporal information for exploration and search purposes. For instance MIT has developed SIMILE Timeline Visualization[5], a web widget prototype for visualizing temporal data.

Within the context of knowledge bases, YAGO2[6] provides a search interface to seek temporal and spatial knowledge facts. Fig. 2 shows an example of the SVG-based browser interface for the query "*David_Beckham*". We can observe that this famous athlete was born on 1975-05-02 by looking at the relation <wasBornOnDate>.

[1] http://www.arcomem.eu/ [March 27, 2014]
[2] http://www.lawa-project.eu/ [March 27, 2014]
[3] http://www.liwa-project.eu/ [March 27, 2014]
[4] http://livingknowledge.europarchive.org/ [March 27, 2014]
[5] http://www.simile-widgets.org/timeline/ [March 27, 2014]
[6] http://www.mpi-inf.mpg.de/yago-naga/yago/demo.html [March 27, 2014]

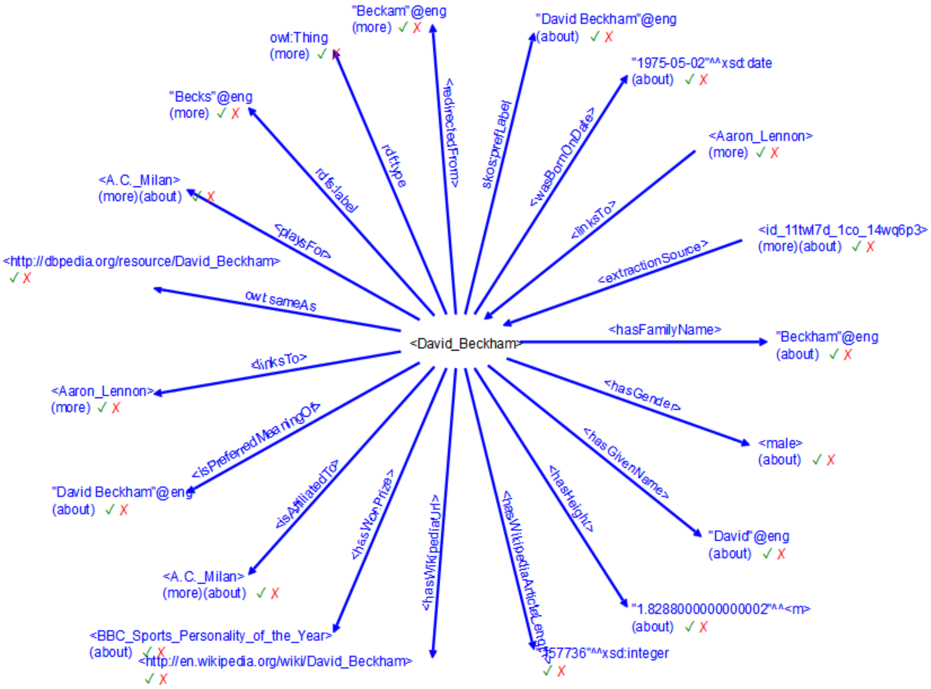
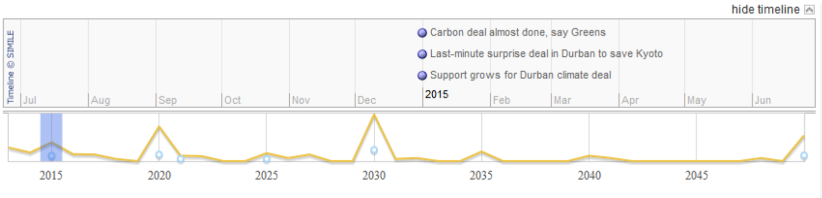


Fig. 2: YAGO2 interface for query "*David_Beckham*", extracted from https://gate.d5.mpi-inf.mpg.de/webyagospotlx/SvgBrowser.

Recorded Future[7] and Yahoo!'s Time Explorer[8] [Matthews *et al.* 2010] application (Fig. 3) are other examples of tools concerning the retrieval of future-related information.



Carbon deal almost done, say Greens

Last-minute surprise deal in Durban to save Kyoto

Support grows for Durban climate deal

Fig. 3: Time Explorer: predictions about climate change, extracted from http://fbmya01.barcelonamedia.org:8080/future/examples.jsp

More recently, Google NGram Viewer[9] [Michel *et al.* 2011] (Fig. 4) was released as a visualization tool that shows the rise and fall of particular keywords across a temporally arranged collection constructed from over five million books. All of these large-scale projects clearly suggest T-IR's high interest to the public and that it constitutes a promising new research area.

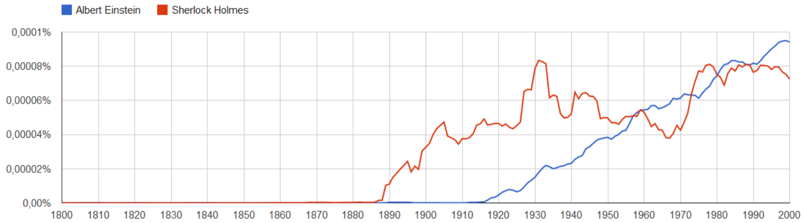


Fig. 4: Google Book Ngram viewer for the queries "*Albert Einstein*" and "*Sherlock Holmes*", extracted from https://books.google.com/ngrams

In addition, note the creation of annotation standard corpora like the TimeBank [Pustejovsky *et al.* 2006], annotation schemas such as TimeML[10] [Pustejovsky *et al.* 2003], and the development of temporal taggers [Strötgen and Gertz 2010a]. Another indication of T-IR's importance is the realization of an increasing number of contests and workshops that focus on the temporal aspects of information. For the former, different competitions have been proposed, such as the Message Understanding Conference (*MUC*) with specific tracks on the identification of temporal expressions (MUC6 and MUC7), the Automated Content Extraction (*ACE*) evaluation program, organized by the National Institute of Standards and Technology (*NIST*), the Time Expression Recognition and Normalization (*TERN*) that has been recently associated with the Text Analysis Conferences (*TAC*), and TempEval within the SemEval competition. WWW Temporal Web Analytics workshop (*TWAW* 2011, 2012, 2013 and 2014) and the SIGIR Time-Aware Information Access workshops (*TAIA* 2012, 2013 and 2014) are examples of seminars dedicated to temporal information search and processing. A detailed description of existing evaluation challenges, annotation schemes, and datasets can be found in [Mazur 2012, Costa 2013].

[7] https://www.recordedfuture.com/ [March 27, 2014]

[8] http://fbmya01.barcelonamedia.org:8080/future/ [March 27, 2014]

[9] http://books.google.com/ngrams [March 27, 2014]

[10] http://www.timeml.org/site/index.html [March 27, 2014]

Despite clear improvement of search and retrieval applications in the temporal context, one can still find examples when the returned search results do not satisfy the user information needs due to problems of a temporal nature. For example, the search results might contain obsolete data, even though a user was actually searching for fresh information (e.g., queries about weather forecasts, stock prices, traffic conditions). They might also contain past-related information even though the user's search intent was directed to the future (e.g., queries about a company's future plans that return results about previous events or predictions that have already become invalid or obsolete).

Based on all these factors, an upsurge of applications is expected in the near future, mostly concerning temporal information exploration, new forms of presenting the search results, and applications concerning temporally-focused retrieval in micro-blogs (e.g., Blog, Twitter and Facebook posts).

In particular, various research studies have already been proposed in different sub-areas of T-IR, such as user query understanding [Jones and Diaz 2007, Metzler *et al.* 2009], temporal web snippet generation [Alonso *et al.* 2009b, Alonso, Gertz and Baeza-Yates 2011], the temporal ranking of documents [Li and Croft 2003, Berberich *et al.* 2005, Dong *et al.* 2010a-b, Elsas and Dumais 2010, Dai *et al.* 2011, Kanhabua and Nørvåg 2012], temporal clustering [Alonso and Gertz 2006, Campos *et al.* 2012b], future-related information retrieval [Baeza-Yates 2005, Jatowt *et al.* 2009, Radinski and Horvitz 2013], and temporal image retrieval [Dias *et al.* 2012].

To the best of our knowledge, this is the first comprehensive review of the state-of-the-art in T-IR. Our main objective is to provide an overview of the research carried out in T-IR and to present a number of open issues as well as proposing promising directions that in our opinion offer much research potential, even though remain generally unexplored by the scientific community. Although already prior work described the field and indicated future promising directions [Alonso, Strötgen, Baeza-Yates and Gertz 2011], we provide a larger overview of related works and comparatively present them in various contexts. A recent tutorial[11] [Radinski *et al.* 2013] at the WSDM 2013 conference as well as a Wikipedia article entitled "Temporal information retrieval" [12] can also be considered attempts to overview and systematize the field.

Like most recent efforts, we focus on T-IR within the context of the web. Since this is a relatively new area, no comprehensive overview exists that positions the existing research in the field. Given this, we introduce a set of models to serve as a framework to enable comparisons between different approaches in the web context.

This paper is intended for academics and practitioners interested in Information Retrieval (*IR*) who do not have a detailed knowledge of T-IR and lack an organized overview of this research area. It also may provide value to scientific professionals interested in a broad snapshot of this field. Since we also indicate several directions and open challenges for future research, we hope it could become a source of new ideas for researchers who are already working on related topics or for those who are considering their own research studies.

This survey is organized as follows. Section 2 gives a high-level overview of T-IR and formalizes the definitions of time, events, and timelines. We introduce the notion of

[11] http://research.microsoft.com/en-us/um/people/sdumais/WSDM2013-Tutorial_Final.pdf [March 27, 2014]

[12] http://en.wikipedia.org/wiki/Temporal_information_retrieval [March 27, 2014]

temporal expressions and discuss the extraction of temporal information from texts, further listing available temporal taggers. Section 3 presents different approaches that can be used to extract time features within web collections. More specifically, we distinguish among metadata, content, and usage approaches and emphasize their main characteristics, challenges, and available data sources. Section 4 explores the key advances in temporal web information retrieval. We extensively detail in a classical IR fashion, the different approaches used in the execution of any process of a web T-IR, i.e., crawling, indexing, query processing, and ranking. In addition, we present works in the field of temporal clustering, temporal text classification, temporal search engines, as well as recent research conducted in the scope of future information retrieval. Section 5 highlights possible future trends and unexplored areas of temporal IR, which although already proposed, still lack further developments. Finally, Section 6 concludes this survey with some final remarks.

2. MODELS OF TEMPORAL ANNOTATIONS OF DOCUMENTS

In the following sections, we introduce different temporal dimensions. Section 2.1 provides a simplified definition of the concepts related to the notion of time. Section 2.2 describes the underlying relation between time and events. Section 2.3 introduces timelines as a means of graphically representing the effects of time's passage. Section 2.4 describes different types of temporal expressions occurring in texts. Finally, Section 2.5 outlines the methodologies behind the extraction of temporal information that are usually used as preprocessing stages in T-IR systems.

2.1. Notion of Time

One of the first works to present a formal model for temporal references was presented by [Bruce 1972]. He defined time as an ordered pair, (time, ≤), where time is a set whose elements are called time points and ≤ is a relation that partially orders time. Another formal approach is the work of [Allen 1983], which introduces the notion of time intervals rather than fixed time points and describes a set of thirteen possible temporal relationships between two time intervals.

In a less formal way, time can be seen as an inherent construct in human life since our thinking is often defined as chronologically arranged events stretching from past, to the present, and to the future. Each instance of time is a point-in-time value, where a single day is often considered an atomic time unit. Atomic units can be grouped into larger units from the finest granularity to the coarsest significant granularity: day (*D*), week (*w*), month (*M*), semester (*s*), quarter (*q*), year (*Y*), decade (*de*), and century (*c*). Note that a day can also include other time points, such as hours, minutes, seconds, fractions of a second, and so forth.

Time values can be physically represented in a calendar, which is a timekeeping system that organizes time into several different granularities. The most widely used calendar in the world today is the Gregorian (also called the Christian calendar). In some countries, this calendar is substituted for or complemented with local ones (e.g., Jewish, Hindu, Chinese, and Islamic calendars, to name a few). Following the ISO-8601:2004[13] standard, a date in the Gregorian calendar is usually represented in the form of *YYYY-MM-DD*, where [*YYYY*] indicates a four-digit year, [*MM*] indicates a two-digit month, and [*DD*] indicates a two-digit day of that month. Although less common, the date representation can also include the number of the week. In this case, the month is

[13] http://www.iso.org/iso/date_and_time_format [March 27, 2014]

replaced by the corresponding week, which results in format *YYYY-Www-DD*, where *ww* represents the week's number from *W01* to *W52*. Specialized calendars also exist, such as fiscal, sports, business, or academic ones.

When addressing the time issue within the realm of database research areas, two types of time are usually distinguished: *valid* and *transaction* [Snodgrass and Ahn 1985]. Valid time is related to the period of time during which events occur in real life, i.e., the time of the fact itself, and transaction time refers to the specific time when the fact was stored in a database. In the web context, *focus time* is the time mentioned or implicitly referred to in the content of web pages, and we regard it as a counterpart of valid time. Naturally, since a web page can refer to different points in time, its focus time is better represented by a set of time intervals delimited by the document's oldest and newest temporal references rather than as a single point in time. Transaction time, on the other hand, is treated as a parallel of a *document timestamp*, i.e., the point in time when the web page was either created (creation time - *ct*), modified (last-modified date - *lmd*), or published (publication time - *pt*). For instance, we may have an interval bounded by the [*initial focus time*, *final focus time*] of the document, but also by the [*ct*, *lmd*] and [*pt*, *lmd*] time references. However, in cases where neither the focus time nor the timestamp can be determined, we can also consider the birth time as the first crawling date and the end time as the most recent crawling time.

We also consider *reading time* and *document age* as additional types of time. In web search scenarios, a document's reading time is assumed to be the same as the time a search query was issued since users often access search results immediately after performing a search. On the other hand, a document's age is the difference between the reading time and the timestamp.

Fig. 5 shows a visual example of different types of time. We start by analyzing document *Doc1,* which was created in 2011. Knowing that the current reading time is 2013, this document is considered two years old. The temporal references in its content define its focus time as equal to 2006. Naturally, since the document content can be re-edited, the document's focus time can change with time. In this sense, a document that previously included references to past time can later refer to the present, to future time, or to other past dates.

In the same figure, document *Doc2* represents a document whose focus time is defined by a time interval. Created in 2012, it currently contains future-related information, since its focus time is defined by a time interval [2014, 2015].

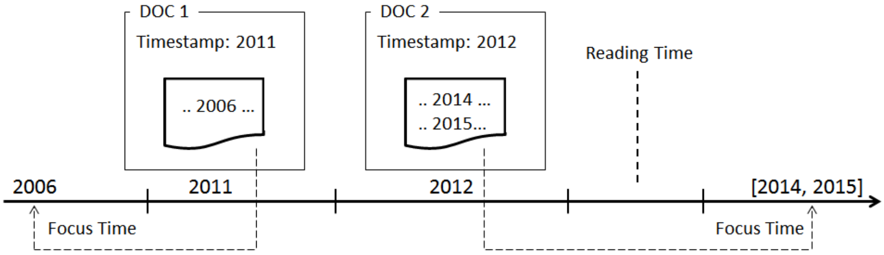


Fig. 5: Different types of time in documents.

In the next section we describe the underlying association between time and events.

2.2. Time and Events

Time is inherently associated with events [Setzer and Gaizauskas 2000]. A simple example is the phrase "**An airplane** coming from Brazil **fell** into the Atlantic Ocean **on Monday**", which uses the temporal expression *on Monday* as a point in time that defines the fall of an airplane. We define an event in a general way as a change ("occurrence") that happens at a given time in a given place and that could be thus mapped into a bi-dimensional spatio-temporal view. Events are usually considered a change or a disruption of a normal course that is important to society or to a group of people; thus they are worth reporting and publicizing.

2.2.1. Topic Detection and Tracking

One initial effort to automatically determine and track events was introduced by [Allan *et al.* 1998] through the Topic Detection and Tracking (*TDT*) initiative. The goal of the TDT project, which was one of the first important research initiatives related to news management, is the exploration of techniques that identify the occurrence of new events to follow their track over time. [Swan and Allan 1999, Swan and Jensen 2000] proposed a classical hypothesis test to discover time-dependent features that identify the important topics in text documents. [Makkonen and Ahonen-Myka 2003] suggested an alternative solution by comparing one document with another through a temporal similarity measure. [Kumaran and Allan 2004], on the other hand, detected new stories by measuring the degree overlap of one story with those that occurred in the past. [Shaparenko *et al.* 2005] correlated topic events with texts used in a document collection to provide an overview of how topics evolve over time. The underlying assumption is that as events change, the text used in documents will change as well. Changes in a text are detected by a K-Means clustering algorithm, where each cluster represents an important topic. The popularity of the topic over time is given by the number of documents that fall into each cluster. More recently, [Vandenbussche and Teissèdre 2011] introduced an experimental end-user prototype as a first step for query-event retrieval by offering users the possibility of querying a specific music dataset (enriched with web semantic data) for events occurring in a given time period at a specific location.

2.3. Timelines

A sequence of events is usually represented in a timeline. A timeline, also known as a chronology, is a graphic representation listing important events within a particular time span. Timelines are particularly useful to give a topic an historical context and to provide a comprehensive temporal understanding of it. An example of a timeline is what a user would construct to represent the history of Haitian earthquakes (Fig. 6).

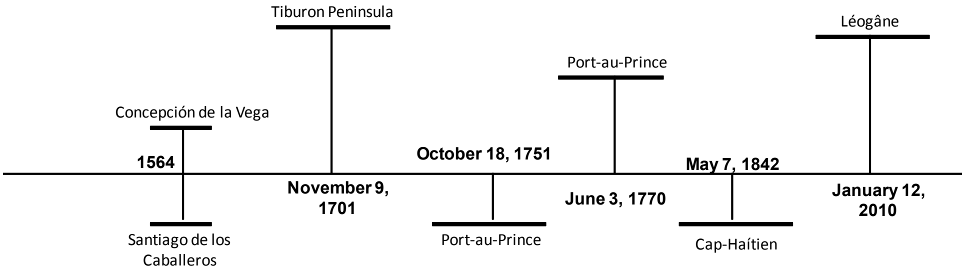


Fig. 6: Timeline for "*Haiti earthquakes*".

Depending on their purpose, timelines of different granularities can be constructed, either more fined-grained (e.g., quarters, semesters, months, weeks, and days) or more coarse-grained (decades and centuries). In our example, we use year, month, and day granularities to represent query "*Haiti earthquakes*". In what follows we describe different types of temporal expressions.

2.4. Temporal Expressions

Temporal expressions are a rich form of natural language that can be defined as a sequence of tokens with temporal meaning. The greatest difficulty in developing an automatic system for detecting temporal expressions is the large diversity of ways in which time can be expressed [Mazur 2012]. Following the work of [Schilder and Habel 2005] we distinguish among the following:

- Explicit Temporal Expressions
- Implicit Temporal Expressions
- Relative Temporal Expressions

Explicit temporal expressions, which were first referenced in 1993 [Setzer and Gaizauskas 2000] during MUC-5 [Advanced Research Projects Agency 1993], denote a precise moment in time and can be anchored on timelines without further knowledge. Based on the granularity level, we may have "*2009*" for the year's granularity, "*December 2009*" for the month's granularity, and "*25.12.2009*" for the day's granularity.

Implicit expressions are often associated with events carrying an implicit temporal nature. They are often difficult to position in time due to the lack of a clear temporal purpose or an unambiguously associated time point. For example, such expressions as "*Christmas Day*" embody a temporal nature that is not explicitly specified. Therefore, as observed by [Alonso *et al.* 2009b], these expressions require that at least a year appears somewhere close in the text to establish accurate temporal values.

Relative temporal expressions, which were referenced for the first time in 1998 during MUC-7 [Setzer and Gaizauskas 2000], depend on the document publication date or another date near in the context. For instance, the expressions *today*, *last Thursday*, or *45 minutes after* are all relative to the document timestamp or to the absolute dates occurring nearby in the text. As such, finding the document timestamp or related explicit temporal expressions is important, so that the expression can be mapped directly on the timeline as an explicit expression. An example is the normalization process of the expression *today*, based on the document creation time (e.g., "*2012.12.19*"). Even though such information is usually available in news documents, it is particularly difficult to locate within web documents, as we discuss in Section 3.1.1. Besides, access to the document timestamp or even to contextual clues, although important, might not be enough in the case of more ambiguous phrases. An example is the expression "*on Thursday*", which, as observed by [Alonso, Strötgen, Baeza-Yates, and Gertz 2011], can either refer to the previous or to the next Thursday.

2.5. Temporal Information Extraction

The identification of temporal information is a non-trivial task that requires the pre-processing stage of a document that usually involves four steps. The first one is *Tokenization,* which divides the text into words or phrases. The second is the *Sentence Extraction* process that identifies the set of all sentences in texts. The third is the part-of-speech tagging (*POS*) process where tokens are assigned a part-of-speech. Finally, the fourth step, Named-entity Recognition (*NER*), identifies the proper nouns in documents.

Interestingly, temporal expressions have also been part of the NER process. However, since 2004, after the introduction of the TERN task as part of the ACE program, Temporal Information Extraction (*T-IE*) has become an independent task. As such, once text processing is underway, the T-IE process can start. It consists of three main tasks. The first is the *Extraction* or *Recognition* of temporal expressions. The second is *Normalization* to unify the different ways in which temporal expressions can be expressed. Finally, the last task called *Temporal Annotation* expresses temporal expressions in a standard format. The result is a set of texts where temporal expressions are usually annotated with TimeML [Pustejovsky *et al.* 2003], which is a temporal formal specification XML language. Fig. 7 shows the entire process. Note that not all the pre-processing steps are always necessary to perform temporal information extraction.

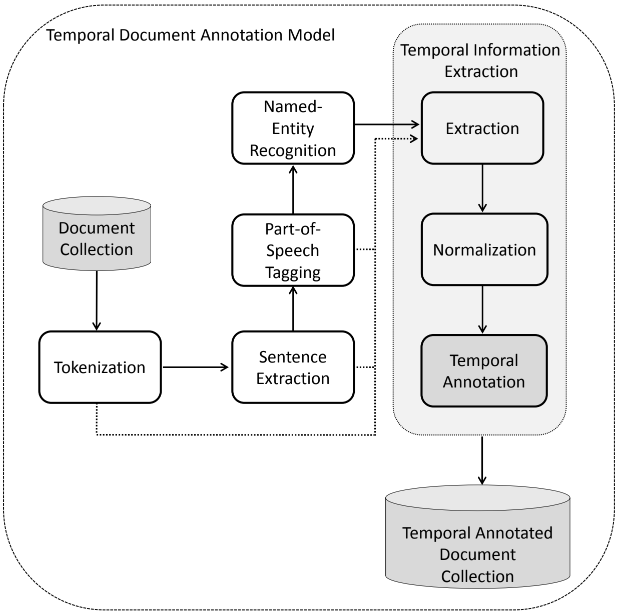


Fig. 7: Temporal document annotation model.

The overall T-IE process is usually conducted by temporal taggers, which follow rule-based approaches that are based on regular expressions or local grammar-based techniques and usually involve hard work by experts. In the last few years, temporal taggers have become an important research area. However, the fact that they rely on language-specific solutions makes them difficult to build. Hence, most available temporal taggers are useful for only one language (typically English) and for one domain (usually, the news domain). Other challenges involve determining the document creation time, delimiting, classifying, and normalizing temporal expressions, recognizing events or determining their temporal order [Costa 2013]. Moreover, the lack of an extensive collection of texts annotated with temporal information covering different languages forms an additional problem. Additionally, the multitude of different forms in which human language allows temporal information to be conveyed [Mazur 2012] as well as language intricacy and ambiguity complicate the task of tagging temporal information in texts more than simply finding the part-of-speech functions of words.

The following temporal taggers have been proposed: *TempEx* [Mani and Wilson 2000], *GUTime*[14], *Annie*[15], *HeidelTime*[16] [Strötgen and Gertz 2010a], and *SuTime*[17] [Chang and Manning 2012]. When evaluating temporal taggers, the task of extracting and normalizing temporal expressions can be measured individually. The former aims to correctly identify temporal expressions. The latter aims to normalize the temporal expression to a standard format and its complexity depend on the type of temporal expression extracted (see Section 2.4 for the different types of temporal expressions). For both tasks, precision (*P*), recall (*R*) and F_β-measure (F_β) are computed according to the formulas below. *TP* (True Positives) is the number of expressions correctly identified by the system as temporal expressions. *FP* (False Positives) is the number of expressions wrongly identified by the system as temporal expressions, and *FN* (False Negatives) is the number of expressions wrongly identified as non-temporal expressions.

$$P = \frac{TP}{TP+FP} \qquad R = \frac{TP}{TP+FN} \qquad F_\beta = \frac{(\beta^2+1)\times P\times R}{\beta^2 P+R}$$

TempEx was the first temporal expression tagger to be developed. It is a rule-based model that extracts temporal information, particularly explicit (e.g., *December 24, 2009*) and relative temporal expressions (e.g., *Monday*), and labels them with TIMEX2 tags. First, the document is tokenized into words and sentences, and part-of-speech tagged. Each sentence is then passed on to a module that identifies time expressions and thereafter to a discourse processing module which solves context-dependent time expressions such as indexicals. Another temporal tagger is GUTime which extends the capabilities of TempEx [Mani and Wilson 2000] by adding TIMEX3 tags. GUTime has been evaluated on the TERN 2004 training corpus and achieved F_1-measure scores of 0.85 and 0.82 for temporal expression recognition and normalization, respectively. *Annie* was also developed in 2002 as part of the GATE[18] distribution [Cunningham *et al.* 2002]. More recently, SuTime and HeidelTime have been developed based on a rule-based system to extract and normalize temporal expressions. SuTime is optimized for English texts and HeidelTime is a multi-lingual temporal tagger (English, German, and Dutch) that is adapted not only to the news domain but also to narrative documents. Both have been evaluated in the TempEval-2 challenge and achieved competitive results. For the extraction process, SuTime achieved the best performance with a score of 0.92 in terms of F_1-measure, while HeidelTime obtained 0.86. In contrast, HeidelTime achieved the best performance for the normalization process with an F_1-measure of 0.85 as opposed SuTime which achieved a score of 0.82. A detailed description of the existing approaches can be found in [Strötgen and Gertz 2012, Mazur 2012].

While temporal taggers play an important role in temporal information processing, some works simply use straightforward regular expressions to look for temporal instances. Indeed, for certain applications there may be no need to use temporal taggers since they may require very specific information, such as year mentions in texts.

[14] http://www.timeml.org/site/tarsqi/modules/gutime/download.html [March 27, 2014]
[15] http://www.aktors.org/technologies/annie/ [March 27, 2014]
[16] http://dbs.ifi.uni-heidelberg.de/index.php?id=form-downloads [March 27, 2014]
[17] http://nlp.stanford.edu:8080/sutime/process [March 27, 2014]
[18] http://gate.ac.uk/download/index.html [March 27, 2014]

There has also been research on extracting relational facts (e.g., *Barack Obama is the president of the USA*) from text corpora, which has led to the emergence of large knowledge bases, such as DBpedia[19] [Auer *et al.* 2007], YAGO[20] [Suchanek *et al.* 2007] as well as commercial ones, such as Freebase[21] [Bollacker *et al.* 2007] and Wolfram Alpha[22]. These knowledge bases provide countless factual relations among entities, such as people or locations. However, they often ignore the temporal dimension (e.g., *Barack Obama has been the president of the USA since 2008*) and mostly focus on identifying the most salient facts. One consequence is the inconsistency of information, since some facts might only have been true for a particular time (e.g., the relation [Bill Clinton, president of the USA] is only valid from 1993 to 2001).

Leveraging such temporal information enables researchers to create knowledgeable retrieval mechanisms that support entity-level temporal queries instead of keyword-based ones. Such a new paradigm will likely improve the effectiveness of the results and the user experience by answering such queries as "*Who got 2nd place at Ballon d'Or in 2010?*" or "*Which player made the most assists to Cristiano Ronaldo at the Real Madrid FC during the 2012/2013 season?*"

Despite the importance of time for information retrieval, research on time-sensitive fact extraction has only recently been addressed. A few recent works [Wang *et al.* 2010, Hoffart *et al.* 2011, Wang *et al.* 2011, Wang *et al.* 2012, Talukdar *et al.* 2012, and Kuzey and Weikum 2012] have explored such temporal information for the automatic development of temporal knowledge bases. One of the first works addressing this problem was developed by [Wang *et al.* 2010] with the T-YAGO knowledge base, which is an extension of YAGO [Suchanek *et al.* 2007]. T-YAGO uses regular expressions to extract temporal facts from semi-structured data attached to Wikipedia articles, such as infoboxes. Although an interesting first approach, its restriction to the football domain and the fact that it does not support the extraction of information from free text corpora limits its scope. Another extension of the YAGO knowledge base is the YAGO2 system [Hoffart *et al.* 2011], which focuses on temporal and spatial knowledge by gathering information from Wikipedia[23] infobox attributes, WordNet[24] and GeoNames.[25] However, like its predecessor it fails to extract information from free text corpora, which is a limitation.

[Wang *et al.* 2011] proposed PRAVDA to automatically harvest temporal facts from textual web sources, especially from news articles and biography texts. It uses a pattern-based approach to extract the temporal candidates of facts. Then label propagation, a semi-supervised learning algorithm, computes the confidence scores of the candidate facts. As above, [Wang *et al.* 2012] employs a methodology that combines label propagation and an integer linear program that incorporates temporal constraints among correlated events to determine noisy facts (e.g., Cristiano Ronaldo cannot play football for Manchester United and Real Madrid at the same time). In contrast, CoTS [Talukdar *et al.* 2012] temporally scopes relational facts based on change detection in a time series of

[19] http://dbpedia.org [March 27, 2014]

[20] http://www.mpi-inf.mpg.de/yago-naga/yago/ [March 27, 2014]

[21] http://www.freebase.com [March 27, 2014]

[22] http://www.wolframalpha.com [March 27, 2014]

[23] http://www.wikipedia.org/ [March 27, 2014]

[24] http://wordnet.princeton.edu/ [March 27, 2014]

[25] http://www.geonames.org [March 27, 2014]

the number of facts from the Google Books Ngram [Michel *et al.* 2011] and Gigaword [Graff *et al.* 2003] datasets. As done above, an integer linear program incorporating temporal constraints temporally scopes the correlated facts and guarantees their temporal consistency.

Another recent approach is the work of [Kuzey and Weikum 2012], which is an extension of T-YAGO and YAGO2. As well as harvesting temporal facts and events from Wikipedia, especially from the infoboxes of articles devoted to named events (e.g., historical events, conferences, etc), it also extracts temporal facts from free text, namely, from the full content of Wikipedia articles. This allows the construction of an enhanced knowledge base when compared to T-YAGO and YAGO2.

3. EXTRACTING TEMPORAL INFORMATION FROM WEB RESOURCES

With the advent of the web, the world's largest collection of data, a huge amount of temporal data has become available. This information can be found within a number of different web sources, from web query logs to collections of web pages or social networks such as Twitter. In this section, we describe approaches for extracting temporal features from web resources and consider three different approaches which are usually related to the type of underlying collection: web documents for *metadata* and *content* techniques, and web query logs for *usage* methodologies.

3.1. Metadata

The metadata approach extracts time information from a document's metadata. This includes the document's creation time, its publication time, and the last-modified date. But it may also embody the extraction of additional temporal information from the document structure, especially information extracted from the URL of the document or from the anchor text itself[26]. This information is usually available from news collections. One of the best known news sources is the New York Times Annotated Corpus [Sandhaus 2008], which spans 20 years of newspapers between 1987 and 2007, including 1.8 million articles (more than 1.5 million manually annotated) and 650,000 article summaries.

While metadata information may be quite useful to solve relative temporal expressions found in a document's content (e.g., "*today*") and to normalize them with a concrete date (e.g., "*2012/12/31*"), it may often be inadequate since the timestamp of a document (creation, publication, or modification time) may differ significantly from its focus time, i.e., its content. A simple example is a document published in "*2009*" whose content concerns the year "*2011*".

In addition, metadata information is particularly difficult to obtain from less structured collections, such as web pages, as opposed to news articles. One reason for this, as observed by [Nunes *et al.* 2007], is due to the fact that web servers typically do not provide other temporal information than the crawling date. An alternative solution is to extract metadata information from the document content, for instance, searching for temporal expressions preceded by the phrase "*last-modified*". This procedure demands a rule definition for each different case or language, which may be quite unfeasible for real-world applications.

[26] Note that Metadata simply refers to the structured information embedded in the web source, excluding any reference to the document's content. This is the typical definition used in the T-IR field and should not be compared to the terminology used in digital libraries (e.g., Dublin Core).

3.2. Content

The content approach focuses on the analysis and extraction of temporal features within web contents, usually to determine a document's focus time. This includes looking for information within web pages, within posts in web micro-blog collections or different past page versions stored in web archives. When seeking a content-based collection, a good starting point is to consider the Clueweb09[27] dataset, which consists of 1 billion web pages in 10 different languages or its newer version Clueweb12[28]. Further available data sources can be found on Datamob[29] and Kevin Chai's[30] websites. Many collections are listed as open room for research on a number of different dimensions, including the temporal one.

Unlike metadata approaches, the content approach implies an increased level of difficulty since it usually involves linguistic analysis of texts, as discussed in Section 2.5. Since the web is heterogeneous, multi-lingual, multi-cultural, and highly multi-domain, ambiguity is common. An illustrative example is the expression "*New Year*" which refers to a different point in time in the USA or in China. Other problems are related to multi-lingual time formats (e.g., "*December 31, 2012*" is translated to "*31 de Dezembro de 2012*" in Portuguese). In this case, one should build a time tagger for each language. Moreover, similar to the application of part-of-speech taggers, problems might surface when applying temporal taggers to micro-blog collections, such as web snippets or tweets. Indeed, their application may eventually result in poor outcomes, mostly due to a lack of background, which is caused by the small number of characters allowed for such sources (e.g., 140-tweet posts) and the specific language used to write these texts (e.g., "*tomorrow*" may be transcribed by "*Tomoz*"[31]).

[Jatowt and Yeung 2011, Campos *et al.* 2011b, Dias *et al.* 2011] recently conducted studies to understand the amount of temporal evidence embedded in web documents. [Jatowt and Yeung 2011] studied the typical granularity and the range of temporal expressions in a collection of online news articles and concluded that news articles are more likely to contain daily temporal expressions referring to the present, the immediate past, or the immediate future. This is shown in Fig. 8, where most of the detected temporal expressions occur in a two-year long time window with regard to the document timestamp and the near past and the near future tend to be referred to by fine granularity expressions, such as days.

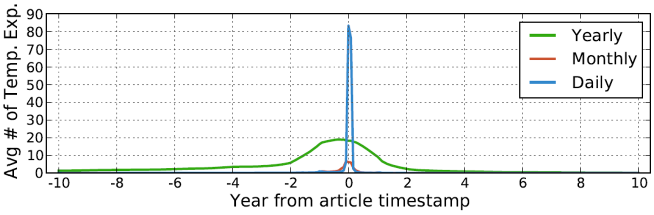


Fig. 8: Reference time of temporal expressions in news articles based on their granularity in relation to article timestamp.

[27] http://lemurproject.org/clueweb09/ [March 27, 2014]

[28] http://lemurproject.org/clueweb12/ [March 27, 2014]

[29] http://datamob.org/datasets [March 27, 2014]

[30] http://kevinchai.net/datasets [March 27, 2014]

[31] http://en.wikipedia.org/wiki/SMS_language [March 27, 2014]

[Campos *et al.* 2011b] studied temporal evidence within a collection of 62,842 web snippets. They observed that roughly 10% contain explicit temporal information and that, similar to the work of [Jatowt and Yeung 2011], most of the temporal expressions found within such collections are from the near past or the near future. [Dias *et al.* 2011] analyzed a set of 508 web snippets looking for future dates. Their research results show that 82.48% of the future dates are related to the near future (i.e., a few months after the query time) and only 17.52% are related to a further future (i.e., at least one year after the query time).

3.3. Usage

Finally, the usage approach considers the extraction of temporal information within queries in a twofold perspective: *query timestamp* and *query focus time*.

The query timestamp, which is the date when the query was issued, is mostly used to understand changes in query popularity and changing search intent. This information is usually obtained from web query logs, which are the flat sets of files that record server activity over time.

The query focus time is the content time of the query, i.e., the time to which the user's query refers. Two types of queries are considered: (*1*) *explicit* temporal queries and (*2*) *implicit* temporal queries. Explicit temporal queries indicate a certain time period and contain a concrete date (e.g., "*Sapporo Olympics 1972*") or an easily resolved temporal expression. Such queries represent about 1.5% of all queries [Nunes *et al.* 2008]. Another investigation [Campos *et al.* 2011b] reduced this value to 1.21% by arguing that some of these queries actually contain false positive temporal expressions (e.g., "*form 1412*" or "office 1997" in the context of the software product). Even though retrieving relevant documents related to explicit temporal queries appears to be a straightforward process, it can be problematic when the query contains such false positive temporal expressions. Indeed, retrieving the most relevant documents after the formulation of this type of query is challenging. One possible solution is to make a temporal index that contains time intervals associated with each crawled document. These time intervals can be used to adjust the score of the document with regard to the query's explicit temporal intent. As such, the documents of Blaise Pascal delimited by [1623, 1662] would not be retrieved for query "*Blaise Pascal 1450*" since the query time span falls outside the document boundaries. Although such a solution may be a step toward achieving a fully integrated temporal information retrieval solution, it still does not solve the problem of false positive temporal query expressions whose time period fits within the document's time span. In that case, a more elaborate approach that determines the time interval or a set of time intervals of a document is needed. We discuss this further in Section 4.6.

Implicit temporal queries point to a certain time period that does not contain an explicit date (e.g., "*Sapporo Olympics*" or "*Battle of Stalingrad*"). Following the work of [Jones and Diaz 2007], such queries may be divided into three categories: (*1*) atemporal queries, which are those not sensitive to time or that remain constant over time (e.g., "*rabbit*"); (*2*) temporal unambiguous queries, which are characterized by pointing to a concrete time period (e.g., "*first moon landing*"), and (*3*) temporal ambiguous queries, which indicate either periodical events (occurring on a recurring basis, e.g., "*SIGIR*") or aperiodic events (occasional peaks of popularity lacking periodicity, e.g., "*Haiti earthquakes*").

Several research efforts have studied the profile of implicit temporal queries. [Jones and Diaz 2007] asked a few annotators to classify 51 TREC *ad hoc* queries using TREC topic descriptions and concluded that 54% of the queries belong to the atemporal class and 46% belong to the temporally ambiguous one. They also developed an automatic classification of queries on the basis of the time profiles of news document collections and obtained approximately the same values. [Metzler *et al.* 2009] in turn, relied on web query logs to infer the implicit temporal values of queries based on similar explicit temporal queries, concluding that only 7% have an implicit temporal nature. A more recent study [Kulkarni *et al.* 2011] explored query intent changes over 10-week time spans and found that 10% of the queries never spiked (atemporal queries), 47% spiked once (temporal unambiguous queries), and 43% spiked multiple times (temporal ambiguous queries). In contrast to the above studies, [Campos *et al.* 2011c] addressed the profile of implicit temporal queries based on a collection of web snippets and found that 75% are atemporal and 25% have an implicit temporal nature.

Although relying on web query logs may be a valid solution to infer the temporal value of implicit temporal queries, access to real-world query logs outside big industrial labs is particularly difficult and a huge impediment to information retrieval research. As pointed out by [Callan and Moffat 2012], this is mostly due to legal concerns about privacy issues. In their report on the use of proprietary data, they point to the AOL incident[32] as one of the reasons for companies to embrace caution about providing query log data. In 2006, AOL Research released a file containing 21,011,240 queries for over 650,000 users collected over a 3-month period, including anonymized user ids, the time at which the query was submitted for search, the rank of the item on which the user clicked, and its corresponding URL. Even though the file was only intended for research purposes, it had to be removed from the Internet after a journalist from the New York Times Journal identified an individual, solely based on the available information.

Despite this incident, other search query collections have been publicly provided over the last few years. Microsoft, for example, released three large-scale datasets. The first two[33], MSLR-WEB30k and MSLR-WEB10K, were intended for ranking purposes and consist of 30,000 and 10,000 queries respectively. QRU-1[34] [Li *et al.* 2012], on the other hand, promotes query representation and understanding and can be used in a variety of tasks, such as query rewriting, query suggestion, query segmentation, and query expansion. A further possibility is to access Google Trends[35] or the New York Times' most popular search queries[36]. Since none of these collections includes a wide range of explicit temporal queries, the process of inferring the temporal nature of queries implicitly formulated is hampered. It also opens a wide field for debate about user search intents. Indeed, as stated by [Campos *et al.* 2011b], the simple fact that a query is year-qualified does not necessarily mean that it has a temporal intent (e.g., "*Microsoft office 2007*" or "*HP 1430*") or that the associated year is correlated with the query (e.g., "*football World Cup 2012*" – there was no World Cup in 2012). A further challenge is that, contrary to the extraction of information within metadata or web contents, which simply requires a set of web search results, extracting temporal information from web query logs implies that some versions of the query have already been issued, thus

[32] http://en.wikipedia.org/wiki/AOL_search_data_leak [March 27, 2014]

[33] http://research.microsoft.com/en-us/projects/mslr/ [March 27, 2014]

[34] http://research.microsoft.com/en-us/downloads/d6e8c8f2-721f-4222-81fa-4251b6c33752/default.aspx [March 27, 2014]

[35] http://www.google.com/trends/hottrends [March 27, 2014]

[36] http://www.nytimes.com/most-popular-searched [March 27, 2014]

contributing to query-dependency. To tackle all of these problems, several alternative directions have been proposed. A more detailed discussion on this topic is provided in Section 4.3.

Next we introduce research studies in a number of different T-IR areas and pinpoint some of the crucial shortcomings of each.

4. TEMPORAL WEB INFORMATION RETRIEVAL

Information retrieval studies the process of searching for relevant information. Typically, this involves looking for information within texts, structured databases, or the web. Documents are then processed regarding their similarity with the query and displayed as ranked documents, clusters, or similar structures. The general IR framework consists of the following four main steps: (*1*) document processing, (*2*) indexing, (*3*) query processing, and (*4*) ranking documents.

In this section, we cover the basics of building a web IR system geared toward the temporal dimension and refer to some relevant applications of T-IR. Since we already treated the document temporal processing stage in Sections 2 and 3, we will not describe it here. The remainder of this section is structured as follows. Section 4.1 introduces related studies on crawling and web archives. Section 4.2 offers an overview of temporal indexing, and Section 4.3 presents research devoted to query processing. Section 4.4 shows the recent improvements achieved in the field of temporal ranking. Recent advances in temporal clustering, temporal text classification, temporal search engines, and future-information retrieval are described in Sections 4.5, 4.6, 4.7, and 4.8 respectively.

4.1. Web Crawling and Web Archiving

The first step of any IR process is to crawl the web by fetching the content of pages. This is done by a software component that is often called a web crawler or a web spider. For each web page, the web crawler captures a snapshot of it at a specific time. Although performing this task is relatively easy, the huge number of web documents raises challenges. The evolution of web content has been widely studied over the years [Baeza-Yates *et al.* 2002], [Cho and Garcia-Molina 2003], [Fetterly *et al.* 2003], [Ntoulas *et al.* 2004], [Bordino *et al.* 2008], [Adar *et al.* 2009], [Elsas and Dumais 2010] and [Kulkarni *et al.* 2011]. Overall, the results show that the web is constantly changing mainly due to the creation of more and more new pages and, to a lesser extent, the modification of the content of existing ones.

In such a dynamic environment, web archives gain increased importance to preserve documents and prevent information loss. They contain information about how the web has evolved over time and can greatly benefit researchers who are recreating a particular historical period of the web. One of the first initiatives in this direction was proposed in 1996 through the Internet Archive project [Kahle 1997] and has saved more than 357 billion web pages. Access to archival content is enabled by the WayBack Machine, which allows particular versions of a given web page to be found. Unfortunately, the system still does not allow a free text query search. Another interesting project, started in 2004, is the Internet Memory Foundation[37] (formerly European Archive) which provides a large-scale, open memory of the Internet. Several other research projects, like ARCOMEM, LAWA, LiWA, and LivingKnowledge (introduced in Section 1), have been conducted with

[37] http://internetmemory.org/en/ [March 27, 2014]

European funding. Many countries have also launched a number of national web archiving initiatives[38] ([Gomes *et al.* 2011]). [Masanès 2006] offered a comprehensive overview of the methods, tools, standards, and the difficulties inherent in the development of a web archive system. Another useful summary is the work of [Gomes *et al.* 2013] who described some lessons learned while developing the Portuguese Web archive. They focused on web data acquisition, ranking search results, and user interface design.

A wealth of research also exists on usage scenarios and applications of content stored in web archives. For example, [McCown and Nelson, 2008] proposed using them to recover lost information on the web and [Van de Sompel *et al.* 2009] introduced a framework called *Memento*[39] in which archived resources can seamlessly be reached by their original URI.

Since the web is constantly changing, it is also becoming more difficult for search engines to maintain up-to-date indexes, which threatens the effectiveness of the search process and the usefulness of search results. This is particularly evident for recency-sensitive queries, for which the relevant set of documents changes frequently. This problem, known as the freshness of search results, is related to the notion of the credibility of presented information. For example, [Yamamoto *et al.* 2007] demonstrated that issuing a query following a real-world change (e.g., nomination of a new president or a change in such numerical values as a country's population or the number of EU states) may still result in outdated information being retrieved due to the self-correcting latency of the web.

One possibility is to give crawlers the ability to detect document age, so that their schedule becomes more precise. However, detecting the freshness of a web page can be quite difficult since temporal metadata are neither necessarily provided nor trustworthy. Indeed, as discussed in Section 3.1, it is generally quite difficult to determine (with a high degree of certainty) trustworthy metadata (i.e., document creation time, document publication time, or last-modified date) based on information extracted during the crawling process. This gives rise to a new challenge called temporal text classification, whose main goal is to determine the time of undated documents. A more detailed analysis of this task can be found in Section 4.6.

4.2. Indexing

Before indexing, documents must first be converted into a standard format. The component responsible for this task is the document processing module that requires the execution of four steps: tokenization, stopping, stemming, and information extraction, which were already described in Section 2.5.

The core of the indexing process is the inversion module which transforms document-term pairs into term-document ones. As described in the survey conducted by [Zobel and Moffat 2006], this is usually done on top of an inverted index structure. Several models can be adopted for this purpose. In Fig. 9, we depict an example of an inverted index structure, where the dictionary terms are associated with a posting that contains two numbers: document *d* and term position *p*.

[38] http://en.wikipedia.org/wiki/List_of_Web_Archiving_Initiatives [March 27, 2014]

[39] http://mementoweb.org/ [March 27, 2014]

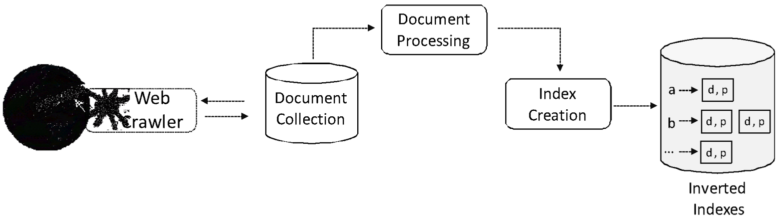


Fig. 9: Indexing process.

Although extensively diffused in text search systems, a completely comprehensive search feature is often missing in most time search mechanisms. One reason is that this process is treated as an inherent part of document processing, when in fact it should be an independent task, supported by a two-layered index framework, where both document and temporal features are considered. This problem was first addressed by [Berberich *et al.* 2007*a,* Berberich *et al.* 2007*b*] who proposed a solution for a time-travel text search by extending the inverted file index to make it operational for temporal searches. Temporal information is explicitly incorporated in the posting list as part of the position item. The postings are thus of the form (d, p, [t_b, t_e]) where d is the document, p is the positional information indicating where a term appears in the document, and [t_b, t_e] is a time interval, where t_b is the birth time of the document and t_e is its end time. This research was later extended by [Anand *et al.* 2012] to allow the incremental addition of new document versions without rebuilding the index structure.

In addition to the above studies, other proposals have been implemented. [Song and JaJa 2008], for example, proposed a novel indexing structure based on the concept of multi-version B-trees and a duplicate detection algorithm to avoid storing duplicate web content examined between two consecutive crawls. They proposed a key consisting of an URL and a time interval [t_b, t_e] during which the corresponding web content has not changed. [Jin *et al.* 2008] also proposed a temporal search engine to answer temporal ranged queries. Their proposal is supported on a hybrid temporal text index for web pages, where time and text keywords are grouped into one uniform index structure based on a MAP21 index [Nascimento and Dunham 1999]. [Arikan *et al.* 2009] suggested the creation of two types of indexes: one to store text documents as direct and inverted indexes (discarding all documents that do not contain any temporal expressions related to a temporal query) and a second to store the temporal data extracted from the content of the documents by regular expressions (Section 2.5). [Pasca 2008] proposed a temporal question answering system where both the dates as well as the text are stored in a fact repository. Last, [Matthews *et al.* 2010] suggested the creation of two indices, one for each document in the collection and one for each sentence. For the sentence level index, a content date is computed based on the document's focus time. If this information is not available, then the document publication time is considered.

4.3. Query Processing

The process of searching for information is inherently temporal. Even though some user information needs may be explicitly expressed, most are implicit by nature. In this section, we provide an overview of the relevant literature regarding the estimation of temporal intent behind user search queries since different studies have been proposed to solve this problem. Following the work of [Cheng *et al.* 2013] we categorize past research into two classes: (*1*) works that target recency-sensitive queries, for which users expect documents to be both topically relevant and up-to-date (fresh) and (*2*) works that target time-sensitive queries, where the results are preferably from a specific time period.

4.3.1. Recency-Sensitive Queries

The importance of information timeliness was recently studied by [Joho *et al.* 2013]. In their survey, nearly half of the 110 respondents stated that the information for which they search is related to the present (information on the same day); yet a significant fraction also searches for future-related or past-related information. 60% of the subjects confirmed that the freshness of search results was important in their recent search activities on the web, usually in response to the formulation of spiky queries (e.g., "*Halloween*" on October 31), fast changing phenomena queries (e.g., "*weather Miami*" or "*dollar/yen rate*") or news-related ones (e.g., "*Nelson Mandela"* at his death). With respect to a document, the property of recency is inversely proportional to the time that has passed since its creation. Detecting whether a query requires special treatment in terms of recency analysis is difficult, given that user search intents are usually underspecified. This may actually constitute a problem because, as reported by [Efron and Golovchinsky 2011], applying temporal approaches to non recency-sensitive queries might harm the quality of the search results. Motivated by this situation, several works have been proposed. Most pioneering approaches have tackled this problem based on newswire collections using query volume and the number of published documents as an indicator of the query's recency sensitiveness. [Diaz 2009], for example, proposed a solution that detects the news intent of a query by studying its dynamics and its click-through rates and modeling the click probability by logistic regression. They tackled a specific problem, called a news aggregated search, which refers to the integration of fresh content extracted from news article collections into "regular" web search results. [König *et al.* 2009], on the other hand, proposed a supervised learning method to estimate the click-through rate for news search results. [Dong *et al.* 2010a] used an automatic classifier to detect whether an incoming query is a breaking-news one. Even though all of these research studies perform well in the specific context of news, a more general solution that addresses this problem by resorting to any type of documents, such as "regular" web pages or micro-blog collections remain undeveloped. Zhang *et al*. 2010], for example, developed a machine learning method that combines multiple features into a classifier to determine queries occurring at regular time intervals, such as public events, lottery drawings, public holidays, tv programs, and so on. Features are derived from web query logs based on time series analysis, where the query frequency is measured at constant intervals. Similarly, [Styskin *et al.* 2011] trained a regression model classifier based on 30 features to predict the need for retrieving the recent contents for a given query. [Shokouhi 2011] detected seasonal queries using time series analysis.

A further problem is related to queries that despite non-spiky behavior may still benefit from retrieving more recent documents (e.g., "*fashionable haircuts*" and "*phone call prices*"). However, in this case, volume-based techniques cannot by applied, since the number of published documents or issued queries does not clearly reflect their temporality. To tackle this problem [Cheng *et al.* 2013] estimated query timeliness using the term distribution change of a query's relevant documents over time. They focused on recency-sensitive queries that are not driven by news events and for which there are no major spikes in query or document volumes over time.

4.3.2. Time-Sensitive Queries

Other works focus more on *time-sensitive* queries, where the results are preferably from a specific time period. Time-sensitive queries include those whose set of search intents (or the main intent) changes over time, such that the relevant answers may also vary over time. They can also be interpreted as seasonal queries [Shokouhi 2011], i.e., cyclic queries related to seasonal events. For example, the query "*Halloween*" issued on October 29 most likely indicates that the user is looking for a costume or a party, but the same query in December probably indicates that the user simply wants to learn about Halloween. Thus the Wikipedia page is relevant in this case.

A clear understanding of the query temporal nature offers search engines the chance to decide whether to return more historical or more recent information to provide the most relevant results. Several alternative directions for identifying relevant query time periods have been previously explored in the literature. The methods proposed to solve this problem can be broadly classified into three different classes: (*1*) metadata, (*2*) usage, and (*3*) content approaches.

Following a metadata approach, [Jones and Diaz 2007] used a language model solution and a collection of web news documents to model the period of time that is relevant to a query. More specifically, they estimate distribution $P(t|q)$, where t is the day relevant to query q. They adopt a relevance modeling solution that considers, not only the probability of the document's relevance, given by $P(q|d)$, but also the temporal information about the document, given by $P(t|d)$, where t is the day relevant to that document (*0* if t day is not equal to the document timestamp and *1* otherwise). [Kanhabua and Nørvåg 2010] proposed three different methods to determine the time of implicit temporal queries: (*1*) dating queries using only query keywords, (*2*) dating queries using the retrieved top-k documents, and (*3*) dating queries using the timestamp of the retrieved top-*k* documents. [Dakka *et al.* 2012] proposed a solution which takes into account the publication times of documents to identify the important time intervals that are likely to be of interest to an implicit temporal query. Time is incorporated into language models to assign an estimated relevance value to each time period. They also built a framework that divides each document d into content component c_d and temporal component t_d, where $P(c_d,t_d|q)$ represents the probability that c_d is topically relevant to query q in time period t_d.

Unfortunately, all of these approaches rely on the creation date of documents as correct temporal signals, which are far from credible in most cases. Moreover, such information is not available in many documents. Finally, as observed by [Kanhabua and Nørvåg 2010], the fact that all these studies are built on top of temporal language models involves drawbacks concerning document collection. In particular, the used documents must be timestamped and cover the time period of the queries.

An alternative solution to using metadata was proposed by [Vlachos *et al.* 2004] who developed a method to discover important time periods using the query logs of a commercial search engine. Likewise, [Metzler *et al.* 2009] suggested mining query logs to identify implicit temporal information needs. They proposed a weighted measure that considers the number of times query q is pre- and post-qualified with given year y. A query is then implicitly year-qualified if it is qualified by at least two different years (e.g., "*Miss Universe 1990*" and "*Miss Universe 1991*"). A relevance value is then given for each year found in a document. Based on this, they proposed a time-dependent ranking model that explicitly adjusts the score of a document in favor of those matching the users' implicit temporal intents. The referred study addresses an interesting solution because it introduces the notion of correlation between a query and a year. However, the approach lacks query coverage since it depends on the analysis of query logs. [Shokouhi and Radinsky 2012] proposed a time-sensitive approach for query auto-completion by applying time series analysis. Their results show that predicting the popularity of queries by time series analysis and periodicity estimation is more reliable than straightforwardly using information on past query popularity derived from web query logs.

While the above models rely on spikes in the distribution of relevant documents or queries, none extracted temporal information from web contents. To the best of our knowledge only one study [Campos *et al.* 2012a] has taken up this research so far. They proposed a language-independent temporal similarity measure called *GTE,* which, based on corpus statistics, associates relevant date(s) to a query while filtering out non-relevant ones.

4.4. Temporal Ranking

Estimating the relevance of a document can greatly benefit from the introduction of temporal aspects into ranking models. Based on such observation, researchers have started to address the problem of returning documents that are not only topically relevant but that are also from the most important time periods. Under this assumption several works have been proposed that can be broadly divided into two classes: (*1*) those favoring more recent documents (recency-sensitive ranking) and (*2*) those that target documents from different time periods (time-sensitive ranking). Both classes are closely related to the categories defined in the query processing section since the method used to identify queries with a need of recency or time-sensitive treatment is intrinsically related to the methods used in the ranking stage.

4.4.1. Recency-Sensitive Ranking

One of the first works that implemented recency-sensitive ranking was proposed by [Li and Croft 2003]. In their study, they introduced the notion of time-based language models as an extension of work proposed by [Ponte and Croft 1998] to favor documents created in recent time. Instead of assuming uniform prior probabilities in the retrieval model, they assign document priors using an exponential decay distribution over the creation dates of documents. Thus, documents with a more recent creation date are assigned a higher probability. [Berberich *et al.* 2005] introduced two approaches based on link analysis, T-Rank Light and T-Rank, taking into account both the freshness (i.e., the timestamps of the most recent updates) and the activity (i.e., update rates) of pages as well as links to retrieve recent documents. Similarly, [Cho *et al.* 2005] relied on web link structure, especially its evolution, to propose a new ranking metric to solve the problem pointed out by [Baeza-Yates *et al.* 2002], who demonstrated the temporal bias of PageRank. In its traditional form, the PageRank algorithm fails to promote newly created relevant web pages because acquiring links usually requires considerable time. [Li *et al.* 2008] tried to improve PageRank by assigning a non-fixed dumping factor governed by a function that depends on the time that elapsed since the last update of the pages.

Next [Zhang *et al.* 2009] described a re-ranking score adjustment to give a ranking boost to fresh documents. The overall process assumes implicit temporal queries as input and relies on the extraction of temporal features from documents, especially from their titles, URLs, and anchor texts. Documents with more recent dates occurring in these fields are thus ranked higher. [Dai and Davison 2010] estimated the web page authority by determining the variation of page and in-link freshness over time and incorporated this information into a temporal ranking probabilistic model called T-Fresh.

Within the context of learning to rank, [Dong *et al.* 2010a] proposed a retrieval system to answer recency-sensitive breaking news queries. Such queries are first classified with regard to their recency sensitivity before being sent to the ranker. Document freshness is taken into account by combining multiple temporal features that represent document recency and the recency demotion of <query, urls> pairs in the training stage. The training data are then used to learn a ranking function. The models proposed incorporate regular ranking trained data to solve the problem of insufficient recency information.

The methods put forward by [Dong *et al.* 2010b] and [Inagaki *et al.* 2010] use user click feedback features to identify how document relevance varies over time. More concretely, [Dong *et al.* 2010b] incorporated fresh URLs extracted from Twitter into a general web search system. Using the features and the labeled <*query*,*url*> training data pairs, a machine learning ranking algorithm can predict the appropriate ranking of the search results for unseen queries. [Inagaki *et al.* 2010] also proposed a set of novel temporal click features and query reformulation chains to improve the machine learning recency-sensitive ranking by favoring URLs that have been of recent interest for the user's recency-sensitive query.

In contrast to [Dong *et al.* 2010a], who selected one particular ranker per query type, [Dai *et al.* 2011] proposed a framework where each query is run against a set of rankers. Consequently, weights vary based on the temporal profile of a query, thus minimizing the risk of poor performance when queries are misclassified in terms of recency intent. Another recent research is the work of [Styskin *et al.* 2011] who relied on a recency-sensitive query classifier to apply result diversification by combining ordinary search results with fresh documents. [Efron and Golovchinsky 2011], in turn, proposed an extension of the query likelihood model that considered not only when a document was published but also the relationship between the publication time and the query. They also proposed temporally informed smoothing, so that older documents that are further from the target time associated with the query are smoothed more aggressively.

Within the context of micro-blogging, [Efron 2012] proposed survival analysis [40] to incorporate recency information into document ranking by following a query-dependent approach. While their results remain preliminary, their research opens up debate for future research directions.

More recently, [Cheng *et al.* 2013] presented a language ranking model that incorporates the timeliness factor to retrieve fresh recent results for non-spike timely queries. The proposed model can be used for different query types and does not depend on the distribution of documents over time since the timeliness factor of a query is determined using the term's distribution change of a query's relevant documents over time.

4.4.2. Time-Sensitive Ranking

Rather than concentrating on the retrieval of fresh documents, other studies propose more general time-sensitive frameworks where the results are adjusted upon longer time periods. [Perkiö *et al.* 2005] automatically detected topical trends and their importance over time within a news corpus using a simple variant of TF.IDF. These trends are then used as the basis for temporally adaptive rankings; the ranking of the results for query *q* at time *t* should promote documents whose most prominent topics are the same as the most active topics within the whole corpus at time *t*. [Jin *et al.* 2008] proposed a new ranking algorithm to sort results by applying a linear interpolation of three factors: text similarity, temporal information, and page importance. Text similarity represents the

[40] A branch of statistics applied in many fields, also called reliability analysis, which studies the amount of time until one or more events happen.

ranking scores of text relevance and depends on the frequency of query keywords and their corresponding locations in web pages. Temporal similarity is the ranking score of temporal relevance based on the set of intersection conditions between the temporal query and the temporal expressions found in the web page. Page importance represents the ranking score of the importance of the web page based on the PageRank algorithm [Brin and Page 1998].

[Metzler *et al.* 2009] considered a web query log dataset and a set of document fields (e.g., title and anchor text) to estimate both the times of the query and the document. Based on this, they proposed a time-dependent ranking model to explicitly adjust the score of a document in favor of those matching the user's intent expressed by an implicit temporal query.

[Arikan *et al.* 2009] were the first to propose an approach that integrates temporal expressions extracted from the document content in a language modeling framework. Similarly, [Berberich *et al.* 2010] proposed a temporal retrieval model which integrates temporal expressions into query-likelihood language modeling. However unlike in [Arikan *et al.* 2009], uncertainty in temporal expressions is explicitly considered, both in the query and in the document, so that temporal expressions can be linked to the same time-point even if they are not exactly equal (e.g., "*1998*" and "*XX*"). [Elsas and Dumais 2010] developed a language model based ranking algorithm that incorporates the dynamics of document content changes using term frequency distribution over time. For example, although it may be advantageous for a recency-sensitive query to have a high weight set on recent terms, for navigational queries, it may be better to focus on content that is stable and present within many past versions. [Kanhabua and Nørvåg 2010] used a query's determined time to improve the re-ranking of the web page results. The idea behind this research is that documents with creation dates that closely match the query's time are more relevant in the temporal dimension and thus should be ranked higher. To achieve this goal, they proposed a mixture model that linearly combined both textual and temporal similarity.

[Aji *et al.* 2010] introduced a new term weighting model that uses the revision history analysis (RHA) of a document to redefine a term's importance, assuming that a term should be as relevant as the number of times it occurs in the different versions of a document gets higher. A decay factor is included so that the terms in older versions of the document get a higher value. RHA is then incorporated into BM25 and statistical language models, so that documents get ranked based on the importance of the terms in the past.

More recently, [Kanhabua and Nørvåg 2012] proposed a new approach by applying a time-senstive ranking model based on learning-to-rank techniques for explicit temporal queries. To learn the ranking model, they applied two classes of features: temporal and entity-based. For the temporal-based ones, both the document focus time and the timestamp are combined. Entity-based features, on the other hand, are used for inferring semantic similarity (such named entities as person, location, or organization). An unseen <document, query> pair is then ranked by the weighted sum of the feature scores. The results show that SVM^{MAP} learning-to-rank model outperforms the proposed method of [Berberich *et al.* 2010]. [Chang *et al.* 2012] re-ranked the search results based on user intents at different times of day using the user temporal click information obtained from query logs.

Other works have explored the temporal dimension in specific types of temporal ranking. For example, [Pasca 2008] proposed a temporal question answering system which defines regular expressions to detect dates that meet certain requirements. The dates found in the document content should provide direct answers for the user's query (e.g., "*When was the Taj Mahal built?*"). Documents are processed offline, and their content is stored in a fact repository. Whenever a new query is issued, the system matches the query through a Boolean search against the text stored in the repository, scores each match individually upon text heuristics, and aggregates the texts associated to the same date by combining the scores of the matching function. Dates with the highest score are then retrieved. In the particular context of temporal clustering (Section 4.5), [Alonso *et al.* 2009a] proposed a measure to rank documents within a cluster based on the number of times a query occurs in a sentence with explicit, implicit, and relative temporal expressions. [Kanhabua *et al.* 2011] also proposed a ranking model for future predictions using a learning-to-rank algorithm trained over a set of labeled query/prediction pairs. Many features are used to measure the similarity between a news article query (which is automatically generated) and the prediction. The query/prediction pair is ranked by the weighted sum of such feature scores as term similarity, entity-based similarity, topic similarity, and temporal similarity. Finally [Strötgen and Gertz 2013] presented a novel ranking approach that takes into account the proximity of text, temporal, and geographic query terms in documents to answer queries with a temporal and spatial information need, i.e., queries of the form "when and where did something happen?"

4.5. Temporal Clustering

In this section, we focus on the temporal clustering of web page results, a relatively new subfield of T-IR. To the best of our knowledge, only three studies have been proposed. The first [Alonso and Gertz 2006] represents a document as a vector of temporal attributes extracted from its metadata (e.g., creation time) and from its content (by applying an Annie temporal tagger, see Section 2.5). Documents are then clustered using a complete-link hierarchical clustering algorithm which results in a set of hierarchical clusters with two possible views: topical and temporal. [Alonso *et al.* 2009a] introduced *TCluster*, an overlapping clustering algorithm, where each document is associated with a temporal document consisting of a list of 3-tuples, <*E*,*C*,*P*>, where *E* is the list of all temporal expressions detected (explicit, implicit, and relative) within a document, *C* is the respective normalized time units (day, week, month, and year), and *P* is their positions. Clusters are formed by a set of documents sharing a year such that the more frequently the query occurs close to the set of temporal expressions, the more relevant the document is in the cluster.

Unfortunately, none of these works filtered out temporal patterns, which may lead to the selection of noisy information. A possible solution to this problem was first introduced by [Campos *et al.* 2012b, Campos *et al.* 2014] who identified the relevant temporal expressions extracted from web snippets by clustering in which documents were grouped into the same cluster if they share a common relevant year. The underlying methodology is based on GTE (Section 4.3.2): a temporal similarity measure that identifies the top relevant dates within a document while filtering out the irrelevant ones. The obtained results show that the introduction of GTE improves the quality of generated clusters by retrieving a higher number of relevant dates than previous approaches which consider all the temporal patterns found as relevant dates. However, since this proposal simply clusters documents on the basis of common dates, documents may not be topically related.

Temporal clustering has also been the subject of study in a number of diverse temporal applications. For example, [Shaparenko *et al.* 2005] tracked events over time by clustering. [Jatowt *et al.* 2009] introduced a clustering approach to summarize future-related information using text content, the content dates, and the timestamps of future predictions. Also, [Jatowt and Yeung 2011] proposed a clustering algorithm to detect future events based on the information extracted from a reference text corpus. Each instance extracted from the corpus is generated by first picking a topical cluster with a particular probability from which the terms and the temporal expression of the instance from it are generated. Each final cluster corresponds to a forecasted event with a certain expected period of its occurrence in the future.

4.6. Temporal Text Classification

In this section, we examine the process of determining a document's time dimension, which may be useful for improving several T-IR tasks, including crawling, indexing, user query understanding, and the ranking of web search results. As observed by [Kanhabua and Nørvåg 2008], this process can be divided into two categories: (*1*) determining a document's timestamp, i.e., the time when the document was created, published or last-modified and (*2*) determining its focus time, i.e., the time to which its contents refer. As previously stated in Section 3.1, these two times may differ significantly.

Determining a document's timestamp was first studied by [Toyoda and Kitsuregawa 2006] and [Nunes *et al.* 2007]. Both used neighboring pages to estimate the document creation date or the last-modified date, assuming that temporal information can be extracted from the web structure. In particular, [Toyoda and Kitsuregawa 2006] proposed a measure to estimate the document creation date based on the scores of linking web pages. [Nunes *et al.* 2007] employed link structure analysis by considering three types of features derived from the web, incoming links, outgoing links, and HTML src attributes (e.g., <img src="URL"), to date web documents for which the last-modified date is not available. Such an approach is limited however by two main constraints: (*1*) the need to have a huge set of external documents and (*2*) the requirement that the last-modified dates exist in a set of external sources, which is not guaranteed. Moreover, as stated in Section 3.1, it remains unclear whether this information is reliable. [Jong *et al.* 2005] and [Kanhabua and Nørvåg 2008] approached this problem by determining the time of non-timestamped documents, namely, news articles, resorting to temporal language models, which (see Section 4.3) suffer from some drawbacks. [Jatowt *et al.* 2007] proposed using web archives to compare the content of the current version of a web page with its past versions in order to estimate approximate age of different content elements on the page. Finally, [Garcia Fernandez *et al.* 2011] proposed a system to automatically determine the publication date of French historical documents based on unsupervised and supervised algorithms. For the former, chronological methods supported by clues in the text (a person's name, newly created or old words, or spelling reforms) and external resources (Wikipedia, Google books unigram, or a French dictionary) are used to determine the document publication date. For the latter, classification methods, such as Support Vector Machines (*SVM*) compute temporal similarities between the document and a training reference corpus (Google books Ngrams).

While some methods determine a document's timestamp, few tackled the problem of determining the document focus time, especially in the case of the lack or the scarcity of temporal expressions within the document content. Clearly, a more generic solution is needed. Such an approach was first addressed by [Jatowt *et al.* 2011]. In their study, the focus time of page p is the set of time periods resulting from the occurrence of events covered by the content of p. Events are detected by applying a clustering algorithm to the related news articles. The resulting clusters are then compared to the content of the web pages. The event's occurrence time is estimated as the average timestamp of the news articles belonging to the underlying cluster. More recent work on document focus time estimation [Jatowt *et al.* 2013a] takes a statistical approach based on data derived from large news collections. First, the association between a term and any year is calculated using the sets of sentences that contain explicit temporal expressions. For example, their study could determine that the term "*hitler*" has the strongest association with the time period between 1939 and 1945, or the term "*atomic*" has a strong association with 1945 while having weaker associations with other years. Next, the terms strongly associated with only a few years (e.g., names of events or entities specific to only one year or to short time periods) are found using temporal entropy and temporal kurtosis measures. Such terms are weighted to reflect their discriminative characteristics to estimate the document focus time. Finally, the focus time of a target document is calculated as the weighted average association of its terms with years.

Parallel to this, [Kawai *et al.* 2010] presented an approach to filter out noisy year expressions *y* from web snippets *s* that are temporally irrelevant to query *q* by applying machine learning techniques trained over a set of labeled <*s*,*q*,*y*> triplets, where each triplet is represented by a set of text features. Although the incorporation of a date filtering process is novel, their proposal does not determine the degree of relevance for each temporal pattern. An improved solution to this problem was done by [Strötgen *et al.* 2012], who proposed the first approach to identify the most relevant temporal expressions in text documents. Each temporal expression, which is extracted by applying the HeidelTime tagger (Section 2.5), is represented by a set of document and corpus-based features. The relevance of the temporal expressions is combined into a single relevance function based on a set of pre-defined heuristics.

Similarly, the works of [Alonso *et al.* 2009a] and [Campos *et al.* 2012b] introduced the notion of temporal clusters, which can also be used to associate each document with a given time span. Future work, however, must focus more on the identification of the most relevant expressions within longer web documents. These pose a few more challenges, mostly due to the possible sets of diverse topics that they may contain. One possible solution is to segment the text into different pieces based on the different topics discussed. Each part of the text can then be assigned a different time period.

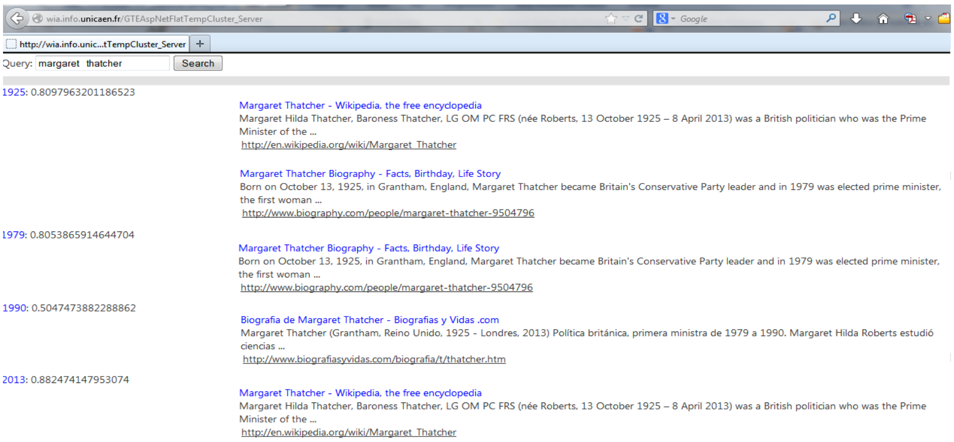
4.7. Temporal Search Engines

With the growth of research in temporal information retrieval, search engines have started to exploit time to improve their search processes. The first initiative, as pointed out by [Manica *et al.* 2012], which is still used today, pushed the most recent web pages to the top list of the results by freshness metrics that take into account the document timestamp (publication time or last-modified date). This approach is usually used by traditional search engines such as Google, Yahoo, and Bing in news domains or with spike phenomena, which tend to lose interest as time goes by. However, this method is of little relevance to users interested in more historical information. For instance, a user may type a query "*Football World Cup Brazil*" and be more interested in the competition held in Brazil in 1950 than in 2014. Traditional search engines however, will likely retrieve more recent web pages about the competition that will occur in 2014 instead of information about the 1950 tournament. Indeed, it is necessary to wade through more than 60 web pages in the Google engine search results to find the first reference to the 1950 event.

A more elaborate mechanism gives users the possibility to specify a point-in-time or a temporal interval of their interest to filter out results outside that time period. For such approaches, the time attribute is again the publication time or the last-modified date of the web page. Over the years, commercial search engines have adopted this solution. For example, Google has a time feature that allowed users to filter their search results. Yahoo, on the other hand, has been experimenting with basic temporal refinement in their web search engine to filter results by the publication time of the document (past day, past week, or past month). Although this solution may be very effective to filter in more detail recently published documents, it may prove inefficient when the user is seeking timely information about a given topic. For example, a user searching for information about "*Blaise Pascal*" will hardly obtain relevant data about the date of his death or about his well-known works when applying such a filter since recent information about him will tend to be scarce. Moreover, the fact that the user has to specify a given time period naturally represents a shortcoming in terms of the user experience. Other problems, as discussed in Section 3.1, are related to the gap that exists between the document timestamp and the time to which the document contents refer as well as the inherent difficulty of extracting timestamp information from such unstructured documents as web pages.

Obviously, users will greatly benefit if a search engine system can explore the temporal information within web pages. However, most popular search engines do not explicitly consider the use of the temporal information extracted from web pages, and the construction of an effective end-to-end temporal system remain proposed. This is particularly evident for implicit temporal queries (e.g., "*Haiti earthquake*", "*BP oil spill*" or "*Madagascar*") for which one would benefit if a comprehensive temporal contextualization of the topic is given.

To tackle this problem several research works have been proposed, leading to the emergence of a number of search engine temporal applications. Next we present a list of academic prototypes and focus on research works that offer a complete framework fully dedicated to T-IR, including indexing, query processing, and the ranking of web search results to answer user queries with temporal information needs. To the best of our knowledge, there are only a limited number of prototypes in this field. [Alonso and Gertz 2006] described a prototype that provides users with an alternative presentation of the results by a hit list of documents clustered by temporal attributes. [Alonso *et al.* 2007a] presented an exploratory search interface that uses timelines to explain and explore search results. [Berberich *et al.* 2007*b*] implemented FluxCapacitor a time-travel text search prototype which extends the inverted file index structure to deal with successive versions of the same document (e.g., searches on web archive collections). [Jin *et al.* 2008] introduced TISE, a temporal search engine that supports content time retrieval for Chinese web pages. [Vicente-Diez and Martinez 2009] proposed a temporal expression recognition and normalization system for Spanish contents that has been integrated into a web search engine prototype. Another work [Alonso *et al.* 2009a] outlined a prototype implementation as a web interface where users can explore results by clusters returned in response to a query. [Matthews *et al.* 2010] proposed Time Explorer, a timeline search tool that enables analysis within a news archive collection about how news topics change over time. [Kawai *et al.* 2010] proposed an on-demand search engine called ChronoSeeker, which allows users to find past/future events. Finally, [Campos *et al.* 2012b] presented GTE-Cluster, an online temporal search interface, which consistently allows searching for topics in a temporal perspective by clustering relevant temporal web search results. GTE-Cluster results can be graphically explored by a demo search engine interface made publicly available for research purposes. In Fig. 10 we show an example of the GTE-Cluster interface [Campos *et al.* 2014] for the query “*Margaret Thatcher*”. Examples of different timelines were already presented in the Introduction section.


wia.info.unicaen.fr/GTEAspNetFlatTempCluster_Server
Google
http://wia.info.unic...tTempCluster_Server
Search

Query: margaret thatcher

1925: 0.8097963201186523

Margaret Thatcher - Wikipedia, the free encyclopedia
Margaret Hilda Thatcher, Baroness Thatcher, LG OM PC FRS (née Roberts, 13 October 1925 – 8 April 2013) was a British politician who was the Prime Minister of the ...
http://en.wikipedia.org/wiki/Margaret_Thatcher

Margaret Thatcher Biography - Facts, Birthday, Life Story
Born on October 13, 1925, in Grantham, England, Margaret Thatcher became Britain's Conservative Party leader and in 1979 was elected prime minister, the first woman ...
http://www.biography.com/people/margaret-thatcher-9504796

1979: 0.8053865914644704

Margaret Thatcher Biography - Facts, Birthday, Life Story
Born on October 13, 1925, in Grantham, England, Margaret Thatcher became Britain's Conservative Party leader and in 1979 was elected prime minister, the first woman ...
http://www.biography.com/people/margaret-thatcher-9504796

1990: 0.5047473882288862

Biografia de Margaret Thatcher - Biografias y Vidas .com
Margaret Thatcher (Grantham, Reino Unido, 1925 - Londres, 2013) Política británica, primera ministra de 1979 a 1990. Margaret Hilda Roberts estudió ciencias ...
http://www.biografiasyvidas.com/biografia/t/thatcher.htm

2013: 0.882474147953074

Margaret Thatcher - Wikipedia, the free encyclopedia
Margaret Hilda Thatcher, Baroness Thatcher, LG OM PC FRS (née Roberts, 13 October 1925 – 8 April 2013) was a British politician who was the Prime Minister of the ...
http://en.wikipedia.org/wiki/Margaret_Thatcher

Fig. 10: GTE-Cluster interface for query “*Margaret Thatcher*”, extracted from http://wia.info.unicaen.fr/GTEAspNetFlatTempCluster_Server.

In Table I, we summarize the contributions made by different research works. We categorize each one by considering the temporal expression taggers used to extract temporal information, the extraction methodology, the base collection, the type of interface, and whether an explicit temporal query is required. Details of the overview of these approaches have already been given throughout the text.

Table I: Summary of research in temporal search engines.

Name	Extraction Methodology	Temporal Tagger	Temporal Queries	Evaluation Dataset	Interface
[Alonso and Gertz 2006]	Metadata Content	Annie	Implicit	N/A	Clustering
[Alonso *et al.* 2007a]	Metadata Content	N/A	Implicit	DBLP	Timeline
[Jin *et al.* 2008]	Metadata Content	TempEx	Explicit	N/A	List
[Berberich *et al.* 2007*b*]	Metadata	N/A	Explicit	English Wikipedia	List
[Vicente-Díez and Martínez 2009]	Content	N/A	Explicit	Newswire Articles	List
[Alonso *et al.* 2009a]	Metadata Content	GUTime	Implicit	DMOZ TimeBank	Clustering
[Kawai *et al.* 2010]	Content	Regular Expressions	Implicit	Web Snippets	Timeline
[Matthews *et al.* 2010]	Metadata Content	TARSQI	Implicit	N/A	Timeline
[Campos *et al.* 2012b, Campos *et al.* 2014]	Content	Regular Expressions	Implicit	Web Snippets	Clustering

There has also been work on the search and the retrieval of geographic and temporal information. [Strötgen and Gertz 2010b], for example, presented a prototype system called TimeTrails for the extraction, querying, storage, and exploration of the spatio-temporal information stored in text documents. YAGO2 [Hoffart *et al.* 2011] provided a search interface to seek temporal and spatial knowledge facts in knowledge bases. The concept of terminology evolution has also recently received attention from researchers. In particular, [Holzmann *et al.* 2012] developed *fokas,* which is a search engine that offers user query suggestion terms of the original query based on a named entity evolution procedure.

4.8. Future-related Information Retrieval

The task of supporting searches for future-related information is critical since a significant number of users are looking for content about future events [Joho *et al.* 2013]. Indeed, future information retrieval is a promising T-IR trend that offers many advantages, especially for supporting decision makers. To name a few, imagine a person who wishes to buy a Toyota car and needs to know whether the company is planning to release a new model. In another example, a prospective house buyer might like information about key urban changes scheduled for housing areas.

The study of the retrieval and the processing of future-related information from text collections have only recently begun. [Baeza-Yates 2005] was the first to suggest a future search engine and *future-related information retrieval*. He proposed to extract future temporal expressions from news articles and represented documents using tuples of time segments and the confidence probabilities of future events. [Jatowt *et al.* 2009] proposed two methods to summarize future-related information in web pages and news articles. The first extracts future-related information about any entity by issuing queries containing future dates and the entity name to search engines and clusters the returned results. The second method focuses on the periodicity analysis of recurring events in news article collections to forecast future occurrences. [Jatowt and Yeung 2011] extended the above concepts by taking into account the uncertainty of a piece of future-related information. In particular, the proposed clustering approach not only considered the textual but also the temporal similarities between sentences referring to future events.

From the information retrieval viewpoint, [Matthews *et al.* 2010] proposed Time Explorer, a search engine that lets users search in the future and analyze the future evolution of topics. [Kawai *et al.* 2010] analyzed effective ways to automatically categorize future-related information in documents using supervised learning, and [Kanhabua *et al.* 2011] proposed a learned ranking model for news predictions that considers the weighted sum of a number of feature scores.

Some research [Jatowt *et al.* 2010, Jatowt and Yeung 2011, Dias *et al.* 2011, Campos *et al.* 2011a, Jatowt *et al.* 2013b] has also been conducted to understand the characteristics of the future-related information in news articles and on the web. For instance, [Jatowt *et al.* 2010] analyzed future-related information on the web by showing the distribution of hit counts obtained from web search engines for queries containing future dates as well as by listing terms that appear frequently with different future years. This work compares the amount and the typical topics of information related to the near or distant future and finds that significant amount of near-term future-related information refers to the events scheduled to happen until the end of current calendar year. Their study was later extended by the cross-lingual comparison and the sentiment analysis of future-related information on the web as well as topical comparison with the future-related content in news articles [Jatowt et al, 2013b]. [Jatowt and Yeung 2011] studied the time range to which future references refer on average in news articles and the granularity of these temporal expressions as a function of the temporal distance from the article creation date (Fig. 8). [Dias *et al.* 2011] and [Campos *et al.* 2011a] discussed whether web snippets can be used to understand the future temporal nature of text queries and described the results of applying classification and clustering algorithms to group informative, schedules, and rumors.

Finally, the methods advocated by [Weerkamp and de Rijke 2012] and [Radinski and Horvitz 2013] tackled the problem of predicting future activities. [Weerkamp and de Rijke 2012] for example, explored the use of Twitter to predict upcoming events that users may perform in the near future, based on tweet messages referring to a future time (e.g., "*Excited for bodypump class tonight!*"). The extraction of time references from twitter messages, however, can be a particularly difficult task mostly due to its informal communication style nature and short message length (e.g., "2nite" instead of "tonight"). [Radinski and Horvitz 2013] instead, aims to predict future events by mining 22 years of news stories from the NYT archive corpus toward the goal of identifying significant increases in the likelihood of disease outbreaks, deaths and violence.

5. PROMISING RESEARCH DIRECTIONS

Quite a few challenges remain to be explored in T-IR. In the following we describe some future trends in information retrieval including some references to studies that have been developed so far.

Credibility As the world continues to change, time-sensitive information can rapidly become invalid. Particularly, future-related information is inherently uncertain in contrast to past-related information. One problem is the validity of future-related information, which arises from the gap between the timestamp (e.g., creation time) and its reading time. For example, imagine a sentence about Toyota planning to establish a new plant in Thailand. Suppose that, actually, this prediction soon afterwards became outdated (e.g., the company decided to cancel the previous plan). An unsuspecting reader might be easily misled when reading it. Users would benefit from automatic warnings when encountering future-related information that has become invalid. To filter out "already happened future-related information" and to eliminate invalid, obsolete predictions one can compare such information with the reports of occurred events and with newer, related predictions [Kanazawa *et al.* 2011].

Other solutions that measure the trustworthiness of temporal information could be based on the document type and derivation, for example, putting more emphasis on news articles from major and reputable newspapers and less on articles published in less credible blogs or documents of unknown authorship. In general, credibility estimation can be improved by considering the source of the article, its linguistic style, citation count, etc. In addition, paying attention to the timestamp of predictions is critical because newer information is more reliable than old information.

Such modal expressions as "*might*", "*could*" or "*is likely to*" are often used when news articles mention future events to indicate different levels of the certainty of events or different levels of confidence put into the predictions by the document's author. Naturally, weighting instances by the modal expressions found near temporal expressions might improve accuracy. Moreover, often events are not totally independent from one another. It is not uncommon to see sentences in the form of "A will occur if B and C happen". In other words, the probability of one event may be dependent on the probabilities of other related events.

The uncertainty of a future event influences the precision of temporal expressions in news articles. When an event is very likely to happen, usually the date and the time mentioned in news articles are more exact (finer granularity). Considering the distance to the event's occurrence date and the actual granularity of the temporal expressions used to describe it might provide additional evidence.

Memory Studies and Computational History The temporal information in text collections can also provide a wealth of information for historical and memory studies. *Collective memory* [Halbwachs 1992] can be analyzed in a similar way as the analysis of collected predictions discussed above. With current text mining techniques, it has now become possible to measure society's attention and focus when it comes to remembering past events and topics. One way to do this is by extracting the context of temporal expressions that refer to the past, whether recent or distant, from large-scale collections that reflect the current concerns and interests of society, such as book-based ngram datasets (Google Books Ngram), blog datasets, or web page collections [Au Yeung and Jatowt 2011]. Typical historical studies are conducted on the old documents stored in archived collections, which are often digitized and subjected to OCR; on the other hand, generating collective images of the past on the basis of current sources could serve as additional, complementary information. This line of research can be extended in many ways, for example, by capturing more implicit remembrances of the past (e.g., historical events or person names) rather than explicit ones in the form of dates. We can study sentiment levels associated with the past, with certain events, or historical entities or compare the collective images of the past in different document genres (e.g., blogs, books, news) as well as conduct cross-country comparisons. In related work, [Kanhabua 2013] studied the problem of collective forgetting, especially the notion of forgetting employed as a means for making archival decisions on what should and should not be preserved.

Related is the emerging field of *computational history* [Michel *et al.* 2011, Hoffmann 2013, Au Yeung and Jatowt 2011] that uses digital historical texts or other artifacts to provide new types of knowledge or information interpretation either for general purposes or for supporting historians. New computer science techniques can be proposed to verify and validate historical assumptions. Some examples are exploratory interfaces over long-term document collections for supporting the work of history and social scientists [Odijk *et al.* 2012, Reinanda *et al.* 2013, Michel *et al.* 2011, Matthews *et al.* 2010] or data mining approaches for large-scale data analysis [Cook *et al.* 2012, Au Yeung and Jatowt 2011, Huet *et al.* 2013]. Several interdisciplinary events have also started to appear, such as *Digital Humanities Conference*[41], *Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities*[42] or the *Workshop on Histoinformatics*[43].

Temporal Text Similarity With the increasing number of digitalized archives, new applications have arisen over the past few years. However, most fail to exploit the inherent temporal issue of historical collections. An interesting way to look at the present is to compare it to the past. In particular, when long periods of time are concerned, it may be important to understand the relation between old concepts and recent ones, for example, helping users choose appropriate queries when searching for collections of documents written in the distant past (the vocabulary mismatch problem in web archive searches). For example, at the beginning of the 19th century, coaches were the cars of the

[41] http://adho.org/conference [March 27, 2014]

[42] http://sighum.science.ru.nl/latech2013/ [March 27, 2014]

[43] http://www.histoinformatics.org [March 27, 2014]

20th century. Currently, this word is usually associated to trainers, whether in athletic contexts or in business, showing that text similarity cannot rely on words but rather on concepts. This task can be defined as proposing models that enable researchers to discover text semantic similarity over time to understand how a given event is intimately related to its evolution. This research direction opens many interesting challenges at the frontier of natural language processing and information retrieval. Already some computational advances have been made. For example, [Berberich *et al.* 2009] developed a technique to reformulate user queries that rests on a novel measure of across-time semantic similarity, contributing to minimize the problem of terminology evolution when searching through web archives. [Radinsky *et al.* 2011] proposed a new semantic relatedness model, Temporal Semantic Analysis (TSA), which constructs a time series for each word of the NYT collection on the assumption that two words are highly related if their time series are related as well. In another work, [Tahmasebi *et al.* 2012] introduced NEER, an unsupervised method for the named entity evolution recognition independent of external knowledge sources. [Odijk *et al.* 2012] demonstrated the environment for visualizing term evolution for understanding how the meaning of words changes over time.

Time-focused Visual Search Interfaces A relatively large research focus has been put on using temporal information for exploration and search purposes, as previously stated in Section 1. One of the first efforts in this field was proposed by [Cousins and Kahn 1991] through Time Line Browser, which provides a basis for the development of further models. That work was followed by [Karam 1994] and [Plaisant *et al.* 1996] with LifeLines, a general visualization environment for visualizing the summaries of personal histories in the health and legal justice fields. SIMILE as shown in Fig. 11 is an example of an end-user visualization and navigation tool for temporal document collections. This widget is relatively easy to use and works with XML data.

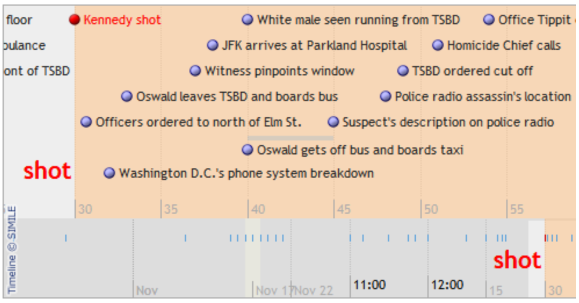

floor
oulance
ont of TSBD
shot
Kennedy shot
White male seen running from TSBD
Office Tippit
JFK arrives at Parkland Hospital
Homicide Chief calls
Witness pinpoints window
TSBD ordered cut off
Oswald leaves TSBD and boards bus
Police radio assassin's location
Officers ordered to north of Elm St.
Suspect's description on police radio
Oswald gets off bus and boards taxi
Washington D.C.'s phone system breakdown
30
35
40
45
50
55
Timeline © SIMILE
Nov
Nov 17
Nov 22
11:00
12:00
15
30
shot

Fig. 11: Timeline of John F. Kennedy assassination, extracted from http://simile-widgets.org/timeline/

Further research on this topic should focus on answering which is the best way to display such information. Listing documents, timelines, temporal or/and topic clusters? Using term clouds with encoded temporal information, for example, by adding tiny time series plots under each term [Lee *et al.* 2010]. This issue remains unanswered as the temporal visualization of documents is still underexploited by internet users.

Temporal IR and Social Network Service Also important is the emergence of micro-blog collections, like tweets or Facebook posts that usually include temporal information.

This new type of data poses, however, some new problems, mostly due to its short message length.

Within the overall context of Twitter, various types of research efforts have been recently conducted on temporal issues ranging from real-time event detection [Sakaki *et al.* 2010, Whiting and Alonso 2012, Osborne *et al.* 2012] to tweet classification [Takemura and Tajima 2012] or tweet-based timeline generation [Alonso and Shiells, 2013]. For example, [Sakaki *et al.* 2010] monitored tweets to detect earthquakes in real-time. [Whiting and Alonso 2012] identified events based on Twitter hashtags to construct timelines. [Osborne *et al.* 2012], on the other hand, used Wikipedia logs to improve event detection from Twitter streams. They also discovered that Wikipedia tends to lag about two hours behind Twitter in terms of page tweets and page views related to the same real-world events. Another recent approach is the work of [Takemura and Tajima 2012], who categorized tweets into different classes based on their information value and decay over time. In this way, the highly dynamic characteristic of tweets can be better assessed to promote tweets with high informational value at given time points. For instance, the "it is raining outside" message has little informational value on the next day, but the "Hawaii is beautiful" message will retain its value for a long time. [Alonso and Shiells, 2013] demonstrated the concept of using Twitter to automatically summarize such events as World Cup matches through timeline generation by finding important or interesting time periods within these events.

Popular social networks like Facebook or Twitter also include temporal evidence in their timelines and message posts. An important aspect of this is the availability of trustworthy temporal data that allows posts to be arranged from the latest to the oldest. Another important issue concerns the privacy control of users over their posts. Some want to make theirs more private while others prefer to make theirs more visible [Bauer *et al.* 2013, Ayalon and Toch 2013]. In the future, more effective use of social graph data (such as those enabled by Facebook social graph API access) is expected to improve search effectiveness, including the temporal aspects of retrieved information [Ugander *et al.* 2011, Bakshy *et al.* 2012]. A further interesting aspect concerns social-based search and social-based recommendation which may be seen as an essential part of the leading search engines in the coming years. For example, the temporal aspects of Amazon's suggestion "*other users like you*" or likewise systems should be considered because users interests and trends change continuously, thus evidencing an obvious dynamic behaviour.

T-IR Standardized Tasks Research in information retrieval and extraction is often fostered through the availability of standardized, open test collections and the proposals of task challenges, thanks to which research communities can compare diverse approaches to the same problems. Likewise, in the area of temporal IR and temporal information extraction, several research challenges have been proposed. TREC Temporal Summarization[44] (TempSum) task is composed of two subtasks: Sequential Update Summarization and Value Tracking. The former requires finding timely, sentence-level, reliable, relevant, and non-redundant updates about a given developing event. The latter subtask tracks the values of event-related attributes with high importance to the event. Examples include the number of fatalities or the financial impact of an event. Both subtasks have clear temporal characteristics because the discovered updates have to be timely and relevant.

[44] http://www.trec-ts.org [March 27, 2014]

TREC Knowledge Base Acceleration[45] (KBA) is a task proposal for filtering large streams of text to find documents that can support the updating of knowledge bases like Wikipedia, Facebook, LinkedIn, etc. Its subtask called Streaming Slot Filling seeks techniques to track the attributes and relations of a selected entity over time. Similar to the Temporal Summarization Task, the recency of information is critical. Both TempSum and KBA are constrained to the information on a given past event or entity and focus on a particular type of information, such as event attributes or event relations. On the other hand, the Temporal Information Access[46] (Temporalia) task hosted by NTCiR[47] asks participants to categorize queries into predefined sets of classes, such as temporal queries, non-temporal queries, past- or future-related queries, or recency-sensitive queries. Temporalia also introduces a second subtask of time-sensitive ranking of news articles for different sets of temporal queries. Finally, the GeoTime[48] challenge answers mixed geo-temporal information and needs to be represented by such questions as "When and where did George Kennan die?" or "When and where were the last three Winter Olympics held?" The temporal component of the answers was in the form of a date or a period/interval type variable.

Temporal Aspects of Web Snippets Other recent research is also related to constructing effective query-based summaries of results. For example, [Alonso *et al.* 2009b, Alonso, Gertz and Baeza-Yates 2011] introduced the notion of temporal web snippets, where the usual text is partly replaced by a number of relevant temporal expressions. [Svore *et al.* 2012], on the other hand, include recent temporal content in web snippet texts. Their results suggest that for trending queries, displaying new temporal content can be quite useful for users.

Temporal Web Image Retrieval Another important topic is temporal image retrieval, which is defined as a process that retrieves sets of web images based on temporal intent behind text queries. Like in document retrieval, image retrieval queries may sometimes have an explicit or implicit temporal intent. A temporal query can be used to obtain images of past or future objects (e.g., car).

We divide the problem of returning images, which satisfy temporal text queries, into the following subtasks: (1) detecting and recognizing the temporal intent of a user query, (2) finding relevant images, and (3) returning relevant images that satisfy the temporal intent in the query. Step (1) resembles the task of temporal intent detection within queries for searching textual documents, although it may need to be adapted for image searches. Step (2) has been well studied so far, and many successful methods have been proposed. Of particular interest is step (3) that filters out images obtained from step (2) that are not representative of the required time period. When assuming the existence of annotated image collections with timestamp metadata, this step essentially contains any of the temporal ranking methods used for text documents described in Section 4.4. However, in unstructured collections such as the web, many images do not contain explicit metadata that can be easily retrieved. Thus methods must be proposed that automatically estimate

[45] http://trec-kba.org/trec-kba-2013.shtml [March 27, 2014]

[46] https://sites.google.com/site/ntcirtemporalia/ [March 27, 2014]

[47] http://research.nii.ac.jp/ntcir/index-en.html [March 27, 2014]

[48] http://metadata.berkeley.edu/NTCIR-GeoTime/ [March 27, 2014]

the timestamps of images to be scaled to large-size collections. Solving this step would also help satisfy users who are interested in the evolution of entities by generating timeline-like overviews that contain representative images for significant years.

[Dias *et al.* 2012] proposed two approaches for solving this problem. In the first they used ephemeral clustering (post-retrieval clustering) to cluster web search results on the fly as they are returned by a web search engine (i.e., text or images). In the specific context of Temporal Web Image Retrieval (*T-WIR*), web image results are retrieved by temporal query expansion (e.g., the query "*Olympic Games*" is expanded to "*Olympic Games 2012*", "*Olympic Games 2008*" and so on). In the second method, they estimate the approximate age of images by SVMs trained over a collection of temporally annotated old images. A similar method was also investigated by [Palermo *et al.* 2012]. They extended their research by comparing the results obtained by their classification task and those of a user survey where untrained humans classified the same set of photos by decade. Last, an approach for satisfying time-sensitive queries was also recently proposed for image retrieval [Kim and Xing 2013]. More specifically, the authors extracted temporal patterns from Flickr datasets (e.g., the time when the photo was taken) to rank images when an explicit temporal query is issued.

Temporal Query Similarity Can two queries be considered similar based on the temporal features shared by the documents they return? The central idea here is to infer if two queries are semantically related based mainly on their temporal information. This issue can be illustrated by such queries as "*war*" and "*peace*" that are related over time, although they usually appear in different documents. A possible application in this scope is query expansion.

Time Period Query Expansion Predicting a query's temporal intent is a critical step to decide appropriate ranking. Thus it is of high importance to develop temporal predictive models that identify queries that may benefit from personalized time-sensitive results (see Section 4.3). However, none of the research studies has proposed time period query expansion, which is mostly due to the fact that systems continue to adopt a simplistic approach that reduces temporal expressions to a single point in time rather than to a time span. Within this context, detecting periods for entities is certainly an interesting challenge that may receive attention in the next few years. For example, the query "*Obama*" might suggest a set of period queries "*Obama 1961 - 2003*", "*Obama Illinois senator 1997 - 2004*" or "*Obama president 2008 - 2012*".

Temporal Diversity Another challenge is developing an approach that provides users with diversified results depending on query intents. Within this scope, we should consider different dimensions, such as topicality, spatiality, and of course, temporality. Gathering all these dimensions into a single model seems a promising research area both for web search and the visualization of web search results. A recent work [Berberich and Bedathur 2013] explores the concept of temporal diversification and proposes an approach in which search results are composed of documents that were published at diverse times of interest to the query.

6. CONCLUSION

Time is obviously one key dimension of our lives, and timeliness is one fundamental feature of information quality. In recent years, time has been gaining increased importance in information retrieval and in a large number of its sub-areas. However, despite the fact that documents are full of temporal expressions and many have strong temporal characteristics, current IR systems still do not sufficiently exploit this information. As an example, when a user's information quest includes temporal aspects, traditional IR systems may fail because they continue to treat temporal expressions as normal text terms.

Consider the following elucidative example. To the query "*FIFA World Cup*", a traditional IR system barely returns a document concerning "*FIFA World Cup in 1994*" but it has no difficulty retrieving more current results, such as 2014 or even 2018. Another example is the query "*FIFA World Cup Germany*" which mostly returns results related to 2006 as opposed to 1974, due to the typically high importance of recency, thus downplaying the subject's historical perspective. What these two examples show is that neglecting the temporal dimension is a key search signal that some content has been omitted. On the other hand, since it may prevent returning relevant documents more or less uniformly distributed over time, IR systems poorly obtain the historical or up-to-date perspectives of some subjects.

In this survey we overviewed the important advances in a new IR sub-field. We first outlined the crucial concepts related to the notion of time, calendar systems, handling temporal expressions in texts as well as the different types of sources of temporal information on the web. We then surveyed existing research that deals with the temporal aspects of both search queries and documents and the diverse ways of generating temporally enhanced search results. Finally, we provided a list of promising research directions.

Despite the growing importance of the area, this recent research trend is still without immediate or at least visible effects for average users since most of the researches developed so far have a rather specific scope. Thus, a number of significant advances must be made before search engines can entirely understand the temporality of a query and correctly reflect it in their returned results. We particularly emphasize the detection of the implicit intents inherent in temporal queries, the development of retrieval models that include temporal features extracted from web documents, and the presentation of the results based on the query type. A further problem is related to the difficulty of evaluating research proposals, since in many cases the community still lacks a gold standard to which most of the approaches can be compared.

REFERENCES

Advanced Research Projects Agency. Software and Intelligent Systems Technology Office. 1993. In Proceedings of the 5th Conference on Message Understanding. MUC-5. Baltimore, Maryland, USA. August 25-27: Morgan Kaufmann Publishers.

Adar, E., Teevan, J., Dumais, S. T., and Elsas, J. L. 2009. The Web Changes Everything: Understanding the Dynamics of Web Content. In Proceedings of the WSDM'09. Barcelona, Spain. February 9-12: ACM Press. 282-291.

Aji, A., Wang, Y., Agichtein, E., and Gabrilovich, E. 2010. Using the Past to Score the Present: Extending Term Weighting Models through Revision History Analysis. In Proceedings of the CIKM'10. Toronto, Canada. October 26-30: ACM Press. 629-638

Allan, J., Carbonell, J., Doddington, G., and Yamron, J. 1998. Topic Detection and Tracking Pilot Study Final Report. In Proceedings of the DARPA Broadcast News Transcription and Understanding Workshop. Lansdowne, Virginia, USA. February. 194-218.

Allen, F. J. 1983. Maintaining Knowledge about Temporal Intervals. CACM: Communications of the ACM, 26(11): ACM Press. 832-843.

Alonso, O. and Gertz, M. 2006. Clustering of Search Results using Temporal Attributes. In Proceedings of the SIGIR'06. Seattle, USA. August 6-11: ACM Press. 597-598.

Alonso, O., Baeza-Yates, R., and Gertz, M. 2007a. Exploratory Search Using Timelines. In Proceedings of the Workshop on Exploratory Search and Computer Human Interaction associated to CHI'07: SIGCHI. San Jose, California, USA. April 29: ACM Press.

Alonso, O., Gertz, M., and Baeza-Yates, R. 2007b. On the Value of Temporal Information in Temporal Information Retrieval. In SIGIR Forum, 41(2). 35-41.

Alonso, O., Gertz, M., and Baeza-Yates, R. 2009a. Clustering and Exploring Search Results using Timeline Constructions. In Proceedings of the CIKM'09. China. November 2-6: ACM Press.

Alonso, O., Baeza-Yates, R., and Gertz, M. 2009b. Effectiveness of Temporal Snippets. In Proceedings of the WSRSP'09 Workshop associated to WWW'09. Madrid. Spain: ACM Press.

Alonso, O., Gertz, M., and Baeza-Yates, R. 2011. Enhancing Document Snippets Using Temporal Information. In Proceedings of the ISSPIR'11. Lecture Notes in Computer Science, 7024/2011: Springer-Verlag. 26-31.

Alonso, O., Strötgen, J., Baeza-Yates, R., and Gertz, M. 2011. Temporal Information Retrieval: Challenges and Opportunities. In Proceedings of the TWAW'11 Workshop associated to WWW'11. Hyderabad, India. March 28.

Alonso, O. and Shiell, K. 2013. Timelines as Summaries of Popular Scheduled Events. In Proceedings of the TempWeb'13 Workshop associated to WWW'13. Rio de Janeiro, Brazil. May 13. 1037-1044.

Anand, A., Bedathur, S., Berberich, K., and Schenkel, R. 2012. Index Maintenance for Time-Travel Text Search. In Proceedings of the SIGIR'12. Portland, USA. August 12-16: ACM Press. 235-243.

Arikan, I., Bedathur, S., and Berberich, K. 2009. Time Will Tell: Leveraging Temporal Expressions in Information Retrieval. In Proceedings of the WSDM'09. Barcelona, Spain. February 09-12: ACM Press.

Au Yeung, C-m. and Jatowt, A. 2011. Studying How the Past is Remembered: Towards Computational History through Large Scale Text Mining. In Proceedings of the CIKM'11. Glasgow, Scotland, UK. October 24-28: ACM Press. 1231-1240.

Auer, S., Bizer, C., Kobilarov, G., Lehmann, J., Cyganiak, R., and Ives, Z. 2007. DBpedia: a Nucleus for a Web of Open Data. In Proceedings of the ISWC'07. Busan, Korea. November 11-15: Springer-Verlag. 722-735.

Ayalon, O. and Toch, E. 2013. Retrospective Privacy: Managing Longitudinal Privacy in Online Social Networks. In Proceedings of the SOUPS'13. Newcastle, UK. July 24-26: ACM Press. 1-13.

Baeza-Yates, R., Saint-Jean, F., and Castillo, C. 2002. Web Structure, Dynamics and Page Quality. In Proceedings of the SPIRE'02. Lecture Notes in Computer Science, 2476/2002: Springer-Verlag. 117-130.

Baeza-Yates, R. 2005. Searching the Future. In Proceedings of the Mathematical/Formal Methods in Information Retrieval Workshop associated to SIGIR'05. Salvador, Brazil. August 15-19: ACM Press.

Bakshy, E., Rosenn, I., Marlow, C., and Adamic, L. 2012. The Role of Social Networks in Information Diffusion. In Proceedings of the WWW'12. Lyon, France. April 16-20: ACM Press. 519-528.

Bauer, L., Cranor, L. F., Komanduri, S., Saranga, M., Michelle, L., Reiter, M. K., Sleeper, M., and Ur, B. 2013. The Post Anachronism: The Temporal Dimension of Facebook Privacy. In Proceedings of the WPES'13. Berlin, Germany. November 4: ACM Press.

Berberich, K., Vazirgiannis, M., and Weikum, G. 2005. Time-Aware Authority Ranking. Internet Mathematics, 2(3). 301-332.

Berberich, K., Bedathur, S., Neumann, T., and Weikum, G. 2007a. A Time Machine for Text Search. In Proceedings of the SIGIR'07. Amsterdam, Netherlands. July 23-27: ACM Press. 519-526.

Berberich, K., Bedathur, S., Neumann, T., and Weikum, G. 2007b. FluxCapacitor: Efficient Time-Travel Text Search. In Proceedings of the VLDB'07. Vienna, Austria. September 23-28. 1414-1417.

Berberich, K., Bedathur, S., Sozio, M., and Weikum, G. 2009. Bridging the Terminology Gap in Web Archive Search. In Proceedings of the WebDB'09 associated to SIGMOD'09. Rhode Island, USA. June 28.

Berberich, K. and Bedathur, S. 2010. A Language Modeling Approach for Temporal Information Needs. In Proceedings of the ECIR'10. Lecture Notes in Computer Science, 5993/2010: Springer-Verlag. 13-25.

Berberich, K. and Bedathur, S. 2013. Temporal Diversification of Search Results. In Proceedings of the TAIA'13 Workshop associated to SIGIR'13. Dublin, Ireland. August 1.

Bollacker, K., Tufts, P., Pierce, T., and Cook, R. 2007. A Platform for Scalable, Collaborative, Structured Information Integration. In Proceedings of the IIWeb'07 Workshop associated to AAAI'07. Vancouver, Canada. July 23: AAAI Press. 22-27.

Bordino, I., Boldi, P., Donato, D., Santini, M., and Vigna, S. 2008. Temporal Evolution of the UK Web. In Proceedings of the 1st International Workshop on AND'08 associated to ICDM'08. Pisa, Italy. December 19: IEEE Computer Society Press. 909-918.

Brin, S. and Page, L., 1998. The Anatomy of a Large-Scale Hypertextual Web Search Engine. In Proceedings of the WWW'98. Computer Networks: The International Journal of Computer and Telecommunications Networking, 30(1-7). 107-117.

Bruce, B. C. 1972. A Model for Temporal References and its Application in a Question Answering Program. Artificial Intelligence, 3: Elsevier. 1-25.

Callan, J. and Moffat, A. 2012. Panel on use of Proprietary Data. In ACM SIGIR Forum 46(2), 10-18.

Campos, R., Dias, G., and Jorge, A. M. 2011a. An Exploratory Study on the impact of Temporal Features on the Classification and Clustering of Future-Related Web Documents. In Proceedings of the EPIA'11. Lecture Notes in Artificial Intelligence - Progress in Artificial Intelligence, 7026/2011: Springer-Verlag. 581-596.

Campos, R., Dias, G., and Jorge, A. M. 2011b. What is the Temporal Value of Web Snippets? In Proceedings of the TWAW'11 Workshop associated to WWW'11. Hyderabad, India. March 28.

Campos, R., Jorge, A., and Dias, G. 2011c. Using Web Snippets and Query-logs to Measure Implicit Temporal Intents in Queries. In Proceedings of the QRU'11 Workshop associated to SIGIR'11. Beijing, China. July 28. 13-16.

Campos, R., Dias, G., Jorge, A. M., and Nunes, C. 2012a. GTE: A Distributional Second-Order Co-Occurrence Approach to Improve the Identification of Top Relevant Dates. In Proceedings of the CIKM'12. Maui, Hawaii. October 29-November 02: ACM Press. 2035–2039.

Campos, R., Jorge, A. M., Dias, G., and Nunes, C. 2012b. Disambiguating Implicit Temporal Queries by Clustering Top Relevant Dates in Web Snippets. In Proceedings of the IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technology. Macau, China. December 4-7: IEEE Computer Society Press, 1-8.

Campos, R., Dias, G., Jorge, A. M., and Nunes, C. 2014. GTE-Cluster: A Temporal Search Interface for Implicit Temporal Queries. In Proceedings of the ECIR'14. Lecture Notes in Computer Science - Advances in Information Retrieval, 8416/2014: Springer-Verlag. 775-779.

Chang, P-T., Huang, Y-C., Yang, C.-L., Lin, S-D., and Cheng, P-J. 2012. Learning-Based Time-Sensitive Re-Ranking for Web Search. In Proceedings of the SIGIR'12. Portland, USA. August 12-16: ACM Press. 1101-1102.

Chang, A. and Manning, C. 2012. SUTIME: A Library for Recognizing and Normalizing Time Expressions. In Proceedings of the LREC'12. Istanbul, Turkey. May 23-25.

Cheng, S., Arvanitis, A., and Hristidis, V. 2013. How Fresh Do You Want Your Search Results? In Proceedings of the CIKM'13. San Francisco, USA. October 27-November 01: ACM Press, 1271-1280

Cho, J. and Garcia-Molina, H. 2003. Estimating Frequency of Change. ACM Transactions on Internet Technology, 3(3). 256-290.

Cho, J., Roy, S., and Adams, R. 2005. Page Quality: In Search of an Unbiased Web Ranking. In Proceedings of the SIGMOD'05. Baltimore, USA. June 13-16: ACM Press. 551-562.

Costa, F. 2013. Processing Temporal Information in Unstructured Documents. PhD thesis, Universidade de Lisboa. May 31. 1-281.

Cook, J., Das Sarma, A., Fabrikant, A., and Tomkins, A. 2012. Your Two Weeks of Fame and your Grandmother's. In Proceedings of the WWW'12. Lyon, France. April 16-20: ACM Press. 919-928.

Cousins, S. and Kahn, M. 1991. The Visual Display of Temporal Information. Artificial Intelligence in Medicine, 3(6). 341-357.

Cunningham, H., Maynard, D., Bontcheva, K., and Tablan, V. 2002. GATE: A Framework and Graphical Development Environment For Robust NLP Tools And Applications. In ACL'02. Philadelphia, USA. July 6-12: Association for Computational Linguistics. 168-175.

Dai, N. and Davison, B. 2010. Freshness Matters: In Flowers, Food, and Web Authority. In Proceedings of the SIGIR'10. Geneva, Switzerland. July 19-23: ACM Press. 114-121.

Dai, N., Shokouhi, M., and Davison, B. D. 2011. Learning to Rank for Freshness and Relevance. In Proceedings of the SIGIR'11. Beijing, China. July 24-28: ACM Press. 95-104.

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

Dias, G., Campos, R., and Jorge, A. 2011. Future Retrieval: What Does the Future Talk About? In Proceedings of the ENIR'11 Workshop associated to SIGIR'11. Beijing, China. July 28.

Dias, G., Moreno, J. G., Jatowt, A., and Campos, R. 2012. Temporal Web Image Retrieval. In Proceedings of the SPIRE'12. Lecture Notes in Computer Science, 7608/2012: Springer-Verlag. 199-204.

Diaz, F. 2009. Integration of News Content into Web Results. In Proceedings of the WSDM'09. Barcelona, February 9-12: ACM Press. 182-191.

Dong, A., Chang, Y., Zheng, Z., Mishne, G., Bai, J., Zhang, R., Buchner, K., Liao, C., and Diaz, F. 2010a. Towards Recency Ranking in Web Search. In Proceedings of the WSDM'10. New York, USA. Feb. 3-6: ACM Press. 11-20.

Dong, A., Zhang, R., Kolari, P., Jing, B., Diaz, F., Chang, Y., Zheng, Z., and Zha, H. 2010b. Time is of the Essence: Improving Recency Ranking Using Twitter Data. In Proceedings of the WWW'10. Raleigh, USA. April 26-30: ACM Press. 331-340.

Efron, M. and Golovchinsky, G. 2011. Estimation Methods for Ranking Recent Information. In Proceedings of the SIGIR'11. Beijing, China. July 24-28: ACM Press. 495-504.

Efron, M. 2012. Query-Specific Recency Ranking: Survival Analysis for Improved Microblog Retrieval. In Proceedings of the TAIA'12 Workshop associated to SIGIR'12. Portland, USA. August 16.

Elsas, J. L. and Dumais, S. T. 2010. Leveraging Temporal Dynamics of Document Content in Relevance Ranking. In Proceedings of the WSDM'10. New York, USA. February 03-06: ACM Press. 1-10.

Fetterly, D., Manasse, M., Najork, M., and Wiener, J. 2003. A Large-Scale Study of the Evolution of Web Pages. In Proceedings of the WWW'03. Budapest, Hungary. May 20-24: ACM Press. 669-678.

Garcia-Fernandez, A., Ligozat, A-L., Dinarelli, M., and Bernhard, D. 2011. When was it Written? Automatically Determining Publication Dates. In Proceedings of the SPIRE'11. Lecture Notes in Computer Science, 7024/2011: Springer-Verlag. 221-236.

Gomes, D., Miranda, J., and Costa, M. 2011. A Survey on Web Archiving Initiatives. In Proceedings of the TPDL'11. Berlin, Germany. September 25-29: Springer-Verlag. 408-420.

Gomes, D., Costa, M., Cruz, D., Miranda, J., and Fontes, S. 2013. Creating a Billion-scale Searchable Web Archive. In Proceedings of the TempWeb'13 Workshop associated to WWW'13. Rio de Janeiro, Brazil. May 13. 1059-1066.

Graff, D., Kong, J., Chen, K., and Maeda, K. 2003. English Gigaword. Linguistic Data Consortium, Philadelphia. USA.

Halbwachs, M. On Collective Memory. The University of Chicago Press, 1992.

Hoffart, J., Suchanek, F.M., Berberich, K., and Weikum, G. 2013. YAGO2: A spatially and temporally enhanced knowledge base from Wikipedia. Artificial Intelligence, 194. 28-61.

Hoffmann, L., 2013. Looking back at big data. CACM: Communications of the ACM, 56(4): ACM Press. 21-23.

Holzmann, H., Gossen, G., and Tahmasebi, N. 2012. Formerly Known As - A Search Engine Incorporating Named Entity Evolution. In Proceedings of the Coling'12. Mumbai, India. December 8-15: ACL. 215-222.

Huet, T., Biega, J., and Suchanek, F. M. 2013. Mining History with Le Monde. In Proceedings of the AKBC'13 Workshop associated to CIKM'13. San Francisco, USA. October 27-28: ACM Press. 49-53.

Inagaki, Y., Sadagopan, N., Dupret, G., Dong, A., Liao, C., Chang, Y., and Zheng, Z. 2010. Session Based Click Features for Recency Ranking. In Proceedings of the AAAI'10. Atlanta, USA. July 11-15: AAAI Press. 1334-1339.

Jatowt, A., Kawai, Y., and Tanaka, K. 2007. Detecting Age of Page Content. In Proceedings of the WIDM'07 Workshop associated to CIKM'07. Portugal. November 9: ACM Press. 137-144.

Jatowt, A., Kawai, H., Kanazawa, K., Tanaka, K., and Kunieda, K. 2009. Supporting Analysis of Future-Related Information in News Archives and the Web. In Proceedings of the JCDL'09. Austin, USA. June 15-19: ACM Press. 115-124.

Jatowt, A., Kawai, H., Kanazawa, K., Tanaka, K., and Kunieda, K. 2010. Analyzing Collective View of Future, Time-referenced Events on the Web. In Proceedings of the WWW'10. Raleigh, USA. April 26-30: ACM Press. 1123-1124.

Jatowt, A., Kawai, Y., and Tanaka, K. 2011. Calculating Content Recency based on Timestamped and Non-Timestamped Sources for Supporting Page Quality Estimation. In Proceedings of the SAC'11. Taiwan. March 21-25: ACM Press.

Jatowt, A. and Yeung, C. M. 2011. Extracting Collective Expectations about the Future from Large Text Collections. In Proceedings of the CIKM'11. Glasgow, Scotland, UK. October 24-28: ACM Press. 1259-1264.

Jatowt, A., Yeung, C. M., and Tanaka, K. Estimating Document Focus Time. 2013a. In Proceedings of the CIKM'13, San Francisco, USA. October 27-November 01: ACM Press. 2273-2278.

Jatowt, A., Kawai, H., Kanazawa, K., Tanaka, K., Kunieda, K., and Yamada, K. 2013b. Multilingual, Longitudinal Analysis of Future-related Information on the Web. In Proceedings of the ICCC'13. Kyoto, Japan. September 16-18: IEEE Computer Society Press. 27-32.

Jin, P., Lian, J., Zhao, X., and Wan, S. 2008. TISE: A Temporal Search Engine for Web Contents. In Proceedings of the ISIITA'08. China. December 21-22: IEEE Computer Society Press. 220-224.

Jones, R. and Diaz, F. 2007. Temporal Profiles of Queries. ACM Transactions on Information Systems, 25(3). Article No.: 14.

Joho, H., Jatowt, A., and Blanco, R. 2013. A Survey of Temporal Web Search Experience. In Proceedings of the TempWeb'13 Workshop associated to WWW'13. Rio de Janeiro, Brazil. May 13. 1101-1108.

Jong, F., Rode, H., and Hiemstra, D. 2005. Temporal Language Models for the Disclosure of Historical Text. In Proceedings of the AHC'05. Amsterdam, Netherlands. September 14-17. 161-168.

Kahle, B. 1997. Preserving the Internet. Scientific American Magazine, 276(3). 72-73.

Kanazawa, K., Jatowt, A., and Tanaka, K. 2011. Improving Retrieval of Future-Related Information in Text Collections. In Proceedings of the WI-IAT'11. Lyon, France. August 22-27: IEEE Computer Society. 278-283.

Kanhabua, N. and Nørvåg, K. 2008. Improving Temporal Language Models for Determining Time of Non-timestamped Documents. In Proceedings of the ECDL'08. Lecture Notes in Computer Science - Research and Advanced Technology for Digital Libraries, 5173/2008: Springer-Verlag. 358-370.

Kanhabua, N. and Nørvåg, K. 2010. Determining Time of Queries for Re-Ranking Search Results. In Proceedings of ECDL'10. Glasgow, Scotland, UK. September 6-10. 261-272

Kanhabua, N., Blanco, R., and Matthews, M. 2011. Ranking Related News Predictions. In Proceedings of the SIGIR'11 Beijing, China. July 24-28: ACM Press. 755-764.

Kanhabua, N. and Nørvåg, K. 2012. Learning to Rank Search Results for Time-Sensitive Queries. In Proceedings of the CIKM'12. Maui, Hawaii. October 29 - November 02: ACM Press. 2463–2466.

Kanhabua, N., Niederee, C., and Siberski, W. 2013. Towards Concise Preservation by Managed Forgetting: Research Issues and Case Study. In Proceedings of the iPRES'13. Lisbon, Portugal. September 2-6.

Karam, M. 1994. Visualization Using Timelines. In Proceedings of the ISSTA'94 associated to SIGSOFT. Seattle, Washington, USA. August 17-19: ACM Press. 125-137.

Kawai, H., Jatowt, A., Tanaka, K., Kunieda, K., and Yamada, K. 2010. ChronoSeeker: Search Engine for Future and Past Events. In Proceedings of the UIMC'10. Suwon, Republic of Korea. January 14-15: ACM Press. 166-175.

Kim G. and Xing E. P. Time-Sensitive Web Image Ranking and Retrieval via Dynamic Multi-Task Regression. 2013. In Proceedings of the WSDM'13. Rome, Italy. February 04-08: ACM Press. 163-172.

König, A. C., Gamon, M., and Wu, Q. 2009. Click-Through Prediction for News Queries. In Proceedings of the SIGIR'09. Boston, USA. July 19-23: ACM Press. 347-354.

Kulkarni, A., Teevan, J., Svore, K. M., and Dumais, S. T. 2011. Understanding Temporal Query Dynamics. In Proceedings of the WSDM'11. Hong Kong, China. February 9-12: ACM Press. 167-176.

Kumaran, G. and Allan, J. 2004. Text Classification and Named Entities for New Event Detection. In Proceedings of the SIGIR'04. Sheffield, UK. July 25-29: ACM Press, 297-304.

Kuzei, E. and Weikum, G. 2012. Extraction of Temporal Facts and Events from Wikipedia. In Proceedings of the TempWeb'12 Workshop associated to WWW'12. Lyon, France. April 17: ACM Press. 25-32.

Lee, B., Riche, N. H., Karlson, A. K., and Carpendale, S. SparkClouds: Visualizing Trends in Tag Clouds 2010. In IEEE Transactions on Visualization and Computer Graphics, 16(6). 1182-1189.

Li, X. and Croft, W. B. 2003. Time-Based Language Models. In Proceedings of the CIKM'03. New Orleans, Louisiana, USA. November 2-8: ACM Press. 469-475.

Li, H., Xu, G., Croft, W. B., Bendersky, M., Wang, Z., and Viegas, E. 2012. A Public Dataset for Promoting Query Representation and Understanding Research. In WSCD'12. Seattle, USA. February 12.

Li, X, Liu, B., and Yu., P. 2008. Time Sensitive Ranking with Application to Publication Search. In ICDM'08. Pisa, Italy. December 15-19: IEEE Computer Society Press. 893-898.

Makkonen, J. and Ahonen-myka, H. 2003. Utilizing Temporal Information in Topic Detection and Tracking. In ECDL'03. Lecture Notes in Computer Science, 2769/2004: Springer-Verlag. 393-404.

Manica, E., Dorneles C.F., and Galante, R. 2012. Handling Temporal Information in Web Search Engines. SIGMOD Record, 41(3): ACM Press. 15-23.

Mani, I. and Wilson, G. 2000. Robust Temporal Processing of News. In ACL'00. Hong Kong, China. October 1-8: Association for Computational Linguistics. 69-76.

Masanès, J. 2006. Web Archiving: Springer-Verlag. 1-234.

Matthews, M., Tolchinsky, P., Blanco, R., Atserias, J., Mika, P., and Zaragoza, H. 2010. Searching through time in the New York Times. In Proceedings of the HCIR'10 Workshop. New Brunswick, USA. August 22. 41-44.

Mazur, P. 2012. Broad-Coverage Rule-Based Processing of Temporal Expressions. PhD thesis, Australia Macquarie University. March. 1-245.

McCown, F. and Nelson, M.L. 2008. Recovering a Website's Server Components from the Web Infrastructure. In Proceedings of JCDL'08. Pittsburgh, USA. June 16-20: ACM Press. 124-133.

Metzger, M. J. 2007. Making sense of credibility on the Web: Models for evaluating online information and recommendations for future research. Journal of the American Society for Information Science and Technology, 58(13). 2078-2091.

Metzler, D., Jones, R., Peng, F., and Zhang, R. 2009. Improving Search Relevance for Implicitly Temporal Queries. In Proceedings of the SIGIR’09. Boston, USA. July 19-23: ACM Press. 700-701.

Michel, J. B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., Pickett, J. P., Hoiberg, D., Clancy, D., Norvig, P., Orwant, J., Pinker, S., Nowak, M. A., and Aiden, E. L. 2011. Quantitative Analysis of Culture Using Millions of Digitized Books. Science, 331(6014). 176-182.

Nascimento, M. A. and Dunham, M. H. 1999. Indexing Valid Time Via B+ -trees – The MAP21 Approach. IEEE Transactions on Knowledge and Engineering, 11(6), 929-947.

Ntoulas, A., Cho, J., and Olston, C. 2004. What's New on the Web?: the Evolution of the Web from a Search Engine Perspective. In Proceedings of the WWW’04. New York, USA: ACM Press. 1-12.

Nunes, S., Ribeiro, C., and David, G. 2007. Using Neighbors to Date Web Documents. In Proceedings of the WIDM’07 Workshop associated to CIKM'07. Lisboa, Portugal. November 9: ACM Press. 129-136.

Nunes, S., Ribeiro, C., and David, G. 2008. Use of Temporal Expressions in Web Search. In Proceedings of the ECIR’08. Lecture Notes in Computer Science, 4956/2008: Springer-Verlag. 580-584.

Odijk, D., Santucci, G., de Rijke, M., Angelini, M., and Granato, G. 2012. Exploring Word Meaning through Time. In Proceedings of the TAIA’12 Workshop associated to SIGIR’12. Portland, USA. August 16.

Osborne, M., Petrovic, S., McCreadie, R., Macdonald, C., and Ounis, I. 2012. Bieber no more: First Story Detection using Twitter and Wikipedia. In Proceedings of the TAIA’12 Workshop associated to SIGIR’12. Portland, USA. August 16.

Palermo, F., Hays, J., and Efros, A. A. 2012. Dating Historical Color Images. In Proceedings of the ECCV’12. Lecture Notes in Computer Science, 7577/2012: Springer-Verlag. 499-512.

Pasca, M. 2008. Towards Temporal Web Search. In Proceedings of the ACM Symposium on Applied Computing. Fortaleza, Ceara, Brazil. March 16-20: ACM Press. 1117-1121.

Perkiö, J., Buntine, W., and Tirri, H. 2005. A Temporally Adaptative Content-Based Relevance Ranking Algorithm. In Proceedings of the SIGIR’05. Salvador, Brazil. August 15-16: ACM Press. 647-648.

Plaisant, C., Milash, B., Rose, A., Widoff, S., and Shneiderman, B. 1996. LifeLines: Visualizing Personal Histories. In Proceedings of the SIGCHI’96. Vancouver, Canada. April 13-18: ACM Press. 221-227.

Ponte, J.M., and Croft, W. B. 1998. A Language Modeling Approach to Information Retrieval. In Proceedings of the SIGIR’98. Melbourne, Australia. August 24-28: ACM Press. 275-281.

Pustejovsky, J., Castaño, J., Ingria, R., Sauri, R., Gaizauskas, R. J., Setzer, A., and Katz, G. 2003. TimeML: Robust Specification of Event and Temporal Expression in Text. In Proceedings of the IWCS’03. Tilburg, Netherlands. January 15-17. 28-34.

Pustejovsky, J., Verhagen, M., Sauri, R., Littman, J., Gaizauskas, R. J., Katz, G., Mani, I., Knippen, R., and Setzer, A. 2006. TimeBank 1.2. Linguistic Data Consortium, Philadelphia. USA.

Radinsky, K., Agichtein, E., Gabrilovich, E., and Markovitch, S. 2011. A Word at a Time: Computing Word Relatedness using Temporal Semantic Analysis. In Proceedings of the WWW’11. Hyderabad, India. March 28-April 1: ACM Press. 337-346.

Radinsky, K. and Horvitz, E. 2013. Mining the Web to Predict Future Events. In Proceedings of the WSDM’13. Rome, Italy. February 4-8: ACM Press. 255-264.

Radinsky, K., Diaz, F., Dumais, S., Shokouhi, M., Dong, A., and Chang, Y. 2013. Temporal Web Dynamics and its Application to Information Retrieval. In Proceedings of the WSDM’13. Rome, Italy. February 4-8: ACM Press. 781-782.

Reinanda, R., Odijk, D., and de Rijke, M. 2012. Exploring Entity Associations Over Time. In Proceedings of the TAIA’12 Workshop associated to SIGIR’12. Portland, USA. August 16.

Sakaki, T., Okazaki, M., and Matsuo, Y. 2010. Earthquake Shakes Twitter Users: Real-time Event Detection by Social Sensors. In Proceedings of the WWW’10. Raleigh, USA. April 26-30: ACM Press. 851-860.

Sandhaus, E. 2008. The New York Times Annotated Corpus. Linguistic Data Consortium. Philadelphia, USA.

Schilder, F. and Habel, C. 2005. From Temporal Expressions to Temporal Information: Semantic Tagging of News Messages. In Mani, I., Pustejovsky, J., and Gaizauskas, R. (eds), The Language of Time: A Reader. Oxford University Press. 533-544.

Setzer, A. and Gaizauskas, R. J. 2000. Annotating Events and Temporal Information in Newswire Texts. In Proceedings of the LREC'00. Athens, Greece. May 31-June 2: ELDA.

Shaparenko, B., Caruana, R., Gehrke, J., and Joachims, T. 2005. Identifying Temporal Patterns and Key Players in Document Collections. In Proceedings of the TDM'05 Workshop associated to ICDM'05. Houston, USA. November 27-30: IEEE Computer Society Press. 165-174.

Shokouhi, M. 2011. Detecting Seasonal Queries by Time-Series Analysis. In Proceedings of the SIGIR'11 Beijing, China. July 24-28: ACM Press. 1171-1172.

Shokouhi, M., and Radinsky, K. 2012. Time-Sensitive Query Auto-Completion. In Proceedings of the SIGIR'12. Portland, USA. August 12 - 16: ACM Press. 601-610.

Snodgrass, R. and Ahn, I. 1985. A Taxonomy of Time Databases. In Proceedings of the SIGMOD'85. Austin, Texas, USA. May 28 - 31: ACM Press. 236-246.

Song, S. and JaJa, J. 2008. Archiving Temporal Web Information: Organization of Web Contents for Fast Access and Compact Storage. Technical Report UMIACS-TR-2008-08, University of Maryland Institute for Advanced Computer Studies, Maryland, MD, USA.

Strötgen, J. and Gertz, M. 2010a. HeidelTime: High Quality Rule-based Extraction and Normalization of Temporal Expressions. In Proceedings of the IWSE'10 associated to ACL'10. Uppsala, Sweden. July 11-16. 321-324.

Strötgen, J. and Gertz, M. 2010b. TimeTrails: a System for Exploring Spatio-Temporal Information in Documents. The Proceedings of the VLDB Endowment, 3 (1-2): VLDB Endowment. 1569-1572.

Strötgen, J., and Gertz, M. 2012. Multilingual and cross-domain temporal tagging. Language Resources and Evaluation. 1-30.

Strötgen, J. and Gertz, M. Proximity^2-aware Ranking for Textual, Temporal, and Geographic Queries. In Proceedings of CIKM'13. San Francisco, USA. October 27-November 01: ACM Press, 739-744.

Strötgen, J., Alonso, O., and Gertz, M. 2012. Identification of Top Relevant Temporal Expressions in Documents. In Proceedings of the TempWeb'12 Workshop associated to WWW'12. Lyon, France. April 17: ACM Press. 33-40.

Suchanek, F. M., Kasneci, G., and Weikum, G. 2007. Yago: a Core of Semantic Knowledge. In Proceedings of the WWW'07. Banff, Alberta, Canada. May 8-12: ACM Press. 697-706.

Svore, K. M., Teevan, J., Dumais, S. T., and Kulkarni, A. 2012. Creating Temporally Dynamic Web Search Snippets. In Proceedings of the SIGIR'12. Portland, USA. August 12-16: ACM Press. 1045-1046.

Swan, R. and Allan, J. 1999. Extracting Significant Time-Varying Features from Text. In Proceedings of the CIKM'99. Kansas City, USA. November 2-6: ACM Press. 38-45.

Swan, R. and Jensen, D. 2000. TimeMines: Constructing Timelines with Statistical Models of Word Usage. In Proceedings of the TM'00 Workshop associated to KDD'00. Boston, USA. August 20-23: ACM Press. 73-80.

Styskin, A., Romanenko, F., Vorobyev, F., and Serdyukov, P. 2011. Recency Ranking by Diversification of Result Set. In Proceedings of the CIKM'11. Glasgow, Scotland, UK. October 24-28: ACM Press. 1949-1952.

Tahmasebi, N., Gossen, G., Kanhabua, N., Holzmann, H., and Risse, T. 2012. NEER: An Unsupervised Method for Named Entity Evolution Recognition. In Proceedings of the Coling'12. Mumbai, India. December 8-15: ACL. 2553-2568.

Takemura, H. and Tajima, K. 2012. Tweet Classification based on their Lifetime Duration. In Proceedings of the CIKM'12. Maui, Hawaii. October 29-November 02: ACM Press. 2367–2370.

Talukdar, P. P., Wijaya, D., and Mitchell, T. 2012. Coupled Temporal Scoping of Relational Facts. In Proceedings of the WSDM'12. Seattle, USA. February 8-12: ACM Press. 73-82.

Toyoda, M. and Kitsuregawa, M. 2006. What's Really New on the Web? Identifying New Pages from a Series of Unstable Web Snapshots. In Proceedings of the WWW'06. Edinburgh, Scotland, UK. May 23-26: ACM Press. 233-241.

Ugander, J., Karrer, B., Backstrom, L., and Marlow, C. 2011. The Anatomy of the Facebook Social Graph. Computing Research Repository Journal. CoRR abs/1111.4503.

Van de Sompel, H., Nelson, M.L., Sanderson, R., Balakireva, L. L., Ainsworth, S., and Shankar, H. 2009. Memento: Time Travel for the Web. CoRR abs/0911.1112.

Vandenbussche, P-Y. and Teissèdre, C. 2011. Events Retrieval Using Enhanced Semantic Web Knowledge. In Proceedings of the DeRiVE'11 Workshop associated to ISWC2011. Bonn, Germany. October 23.

Vicente-Diez, M. T., and Martinez, P. 2009. Temporal Semantics Extraction for Improving Web Search. In Proceedings of the DEXA'09. Linz, Austria. August 31-September 4: IEEE. 69-73.

Vlachos, M., Meek, C., Vagena, Z., and Gunopulos, D. 2004. Identifying Similarities, Periodicities and Bursts for Online Search Queries. In Proceedings of the ICMD'04. Paris, France. June 13-18: ACM Press. 131-142.

Wang, Y., Zhu, M., Qu, L., Spaniol, M., and Weikum, G. 2010. Timely YAGO: Harvesting, Querying, and Visualizing Temporal Knowledge from Wikipedia. In Proceedings of the EDBT'10. Lausanne, Switzerland. March 22-26: ACM Press. 697-700.

Wang, Y., Yang, B., Qu, L., Spaniol, M., and Weikum, G. 2011. Harvesting Facts from Textual Web Sources by Constrained Label Propagation. In Proceedings of the CIKM'11. Glasgow, Scotland, UK. October 24-28: ACM Press. 837-846.

Wang, Y., Maximilian, D., Spaniol, M., and Weikum, G. 2012. Coupling Label Propagation and Constraints for Temporal Fact Extraction. In Proceedings of the ACL'12. Jeju Island, Korea. July 8-14: Association for Computational Linguistics. 233-237.

Weerkamp, W. and de Rijke, M. 2012. Activity Prediction: A Twitter-based Exploration. In Proceedings of the TAIA'12 Workshop associated to SIGIR'12. Portland, USA. August 16.

Whiting, S. and Alonso, O. 2012. Hashtags as Milestones in Time. In Proceedings of the TAIA'12 Workshop associated to SIGIR'12. Portland, USA. August 16.

Yamamoto, Y., Tezuka, T., Jatowt, A., and Tanaka, K. 2007. Honto? Search: Estimating Trustworthiness of Web Information by Search Results Aggregation and Temporal Analysis. In Proceedings of the joint 9th Asia-Pacific Web and 8th international conference on Web-age information management conference on Advances in data and Web management. Huang Shan, China. June 16-18. 253-264.

Zhang, R., Chang, Y., Zheng, Z., Metzler, D., and Nie, J-Y. 2009. Search Result Re-ranking by Feedback Control Adjustment for Time-sensitive Query. In Proceedings of the NAACL-HLT'09. Boulder, USA. May 31-June 5: Association for Computational Linguistics. 165-168.

Zhang, R., Konda, Y., Dong, A., Kolari, P., Chang, Y., and Zheng, Z. 2010. Learning Recurrent Event Queries for Web Search. In Proceedings of the EMNLP'10. Massachusetts, USA. October 9-11: Association for Computational Linguistics. 1129-1139.

Zobel, J. and Moffat, A. 2006. Inverted Files for Text Search Engines. ACM Computing Surveys, 38(2): ACM Press. 1-56.