Artificial Intelligence Review (2023) 56:9115–9140
https://doi.org/10.1007/s10462-023-10400-y

Time expression recognition and normalization: a survey

Xiaoshi Zhong[1] · Erik Cambria[2]

Published online: 24 January 2023

Abstract

Time information plays an important role in the areas of data mining, information retrieval, and natural language processing. Among the linguistic tasks related to time expressions, time expression recognition and normalization (TERN) is fundamental for other downstream tasks. Researchers from these areas have devoted considerable effort in the last two decades to define the problem of time expression analysis, design the standards for time expression annotation, build annotated corpora for time expressions, and develop methods to identify time expressions from free text. While there are some surveys concerned with the development of time information extraction, retrieval, and reasoning, to the best of our knowledge, there is no survey focusing on the TERN development. We fill in this blank. In this survey, we review previous researches, aiming to draw an overview of the development of time expression analysis and discuss the role that time expressions play in different areas. We focus on the task of recognizing and normalizing time expressions from free text and investigate three kinds of methods that researchers develop for TERN, namely rule-based methods, traditional machine-learning methods, and deep-learning methods. We will also discuss some factors about TERN development, including TIMEX type factor, language factor, and domain and textual factors. After that, we list some useful datasets and softwares for both tasks of TER and TEN as well as TERN and finally outline some potential directions of future research. We hope that this survey can help those researchers who are interested in TERN quickly gain a comprehensive understanding of the development of TERN and its potential research directions.

Keywords Information extraction · Time expressions · Rule-based methods · Machine-learning methods · Deep-learning methods

✉ Erik Cambria
cambria@ntu.edu.sg

Xiaoshi Zhong
xszhong@bit.edu.cn

[1] School of Computer Science and Technology, Beijing Institute of Technology, Beijing, China

[2] School of Computer Science and Engineering, Nanyang Technological University, Singapore, Singapore

1 Introduction

1.1 Importance of time expression recognition and normalization

Time information plays an important role in the areas of data mining, information retrieval, and natural language processing (Smith 1978; Enc 1986; Hinrichs 1987; Grishman and Sundheim 1996; Chinchor 1998a, b; Mani 2003; Wong et al. 2005; Lim et al. 2019; Leeuwenberg and Moens 2019). It is involved in many linguistic tasks, including time expression recognition and normalization (Verhagen et al. 2007, 2010; UzZaman et al. 2013; Bethard et al. 2015, 2016, 2017; Zhong et al. 2017; Zhong and Cambria 2018), timeline construction (Do et al. 2012; Li and Cardie 2014; Minard et al. 2015), temporal relation extraction (Mani et al. 2006; Chambers et al. 2007; Verhagen et al. 2007), temporal information retrieval (Alonso et al. 2011; Campos et al. 2014; Lin et al. 2014; Strotgen 2015; Rahoman and Ichise 2018; Khan et al. 2018), temporal event extraction and ordering (Verhagen et al. 2007, 2010; UzZaman et al. 2013; Cheng and Miyao 2018; Lee et al. 2018; Naik et al. 2019; Niu et al. 2020; Liu et al. 2021; Breitfeller et al. 2021; Su et al. 2021), temporal reasoning (Leeuwenberg and Moens 2019; Qin et al. 2021), temporal graph mining (Guillou et al. 2020), temporal query expansion Shokouhi and Radinsky (2012), and temporal question answering (Jia et al. 2018, 2021). Among these linguistic tasks, time expression recognition and normalization (TERN) is a crucial and fundamental task that significantly affects the downstream tasks. In this survey, we focus on reviewing the development of this fundamental task: TERN.

The study of time expressions (also known as "temporal expressions" or "timex" in brief) could be dated back to 1978 when Carlota S. Smith analyzes the syntactic structures of time expressions and their semantic interpretation in English (Smith 1978). After that there are some works concerning different aspects of time expressions, such as referential analysis (Enc 1986) and compositional semantics (Hinrichs 1987). Extensive studies of time expressions start from the sixth and seventh Message Understanding Conference (MUC-6 and MUC-7), in which Grishman and Sundheim (1996) and Chinchor (1998a, 1998b) formally define the task of identifying time expressions from free text, together with the tasks of identifying the entity names and number expressions as well as other information extraction tasks. After MUC-6 and MUC-7, researchers from diverse fields (e.g., data mining, information retrieval, natural language processing, and bioinformatics) have devoted tremendous effort to specify annotation standards for time expressions (Mani et al. 2001; Ferro et al. 2001; Ferro 2001; Pustejovsky et al. 2003a; Ferro et al. 2005; Pustejovsky et al. 2010; Styler et al. 2014; Laparra et al. 2018a; Kamila et al. 2018; Sakaguchi et al. 2018; Ocal et al. 2022a), develop annotated corpora for time expressions (Pustejovsky et al. 2003b; Mazur and Dale 2010; Styler et al. 2014; Hasanuzzaman et al. 2014; MacAvaney et al. 2018; Kamila et al. 2018; Sakaguchi et al. 2018; Grabar and Hamon 2019; Aumiller et al. 2022; Ocal et al. 2022b), and organize shared tasks to address the problems of recognizing and normalizing time expressions from free text (Chinchor 1998a, b; Negri and Marseglia 2004; Zhao et al. 2010; Verhagen et al. 2007, 2010; UzZaman et al. 2013; Bethard et al. 2015, 2016, 2017; Kuzey et al. 2016; Laparra et al. 2018b).

We present here a survey for previous researches about TERN, including early years' syntactic and semantic analysis of time expressions and recent 26 years' (from 1996 to 2022) development of TERN. There are some surveys that concern the development of time information extraction, retrieval, and reasoning (Mani 2003; Wong et al. 2005; Alonso et al. 2011; Campos et al. 2014; Leeuwenberg and Moens 2019; Lim et al. 2019). For example, Mani (2003) broadly review the development of temporal information extraction up to 2003. Wong et al. (2005) review the development of temporal information extraction and time aspect in artificial intelligence and its applications up to 2005. Alonso et al. (2011) highlight the challenges and opportunities in temporal information retrieval. Campos et al. (2014) review recent years' research and its related applications in temporal information retrieval. Lim et al. (2019) broadly review the development of extracting methods for temporal points (i.e., time expressions), events, and temporal relations. Leeuwenberg and Moens (2019) focus on surveying the development of temporal reasoning for various temporal information extraction from unstructured text. However, these surveys either investigate the role that time expressions play in other linguistic tasks with an implicit assumption that time expressions have been precisely recognized and normalized or investigate a variety of linguistic tasks that are related to time information extraction, in which TERN occupies only a small part of the content, with a lack of focus on TERN. To the best of our knowledge, there is no survey that focuses on recent advances on the TERN task, especially no survey that reviews the challenges and implicit problems on TERN. This survey fills this blank. We review the development of TERN from early years to currently, with more attention on recent advances, highlight the challenges and bottleneck problems that limit a more advance on this task, and provide some insights on potential directions of future research for TERN. We hope that this survey can help those researchers who are interested in TERN quickly gain a comprehensive understanding of the development of TERN and its potential research directions. We also hope that TERN could be significantly improved in next few years so as to improve the accuracy of downstream tasks.

1.2 Definition of time expression recognition and normalization

The time expressions we are concerned with in this survey are those expressions that explicitly express time information, such as "2022", "early years", and "26 years". The "explicitly" means that we do not consider those expressions with implicit time information, such as "The time Arnold reached Quebec City" and "the moment when I got my paper accepted" because these expressions are extremely descriptive. The most widely used scheme to categorize time expressions is the TimeML scheme with markups <TIMEX3> and </TIMEX3> to classifies time expressions into DATE, TIME, DURATION, and SET (Pustejovsky et al. 2003a). DATE includes complete or partial date expressions (e.g., "2022" and "September 1") while TIME includes complete or partial expressions of time of the day (e.g., "13:25:43" and "this morning"); DURATION includes complete or partial expressions of intervening time between two end-points of a time interval (e.g., "30 days" and "26 years") while SET includes complete or partial expressions of periodic temporal sets that represent times occurring frequently (e.g., "monthly" and "every Sunday").

The TERN task aims to recognize those time expressions from unstructured text and then normalize the recognized time expressions to a standard format. For example, the following is a piece of text taken from the TimeBank corpus with the document create time (DCT) being March 22, 2013 (Pustejovsky et al. 2003b):

> In that group, 177 out of every 100,000 were hospitalized with flu-related illness in the past several months. That's more than 2 1/2 times higher than any other recent season. This flu season started in early December, a month earlier than usual, and peaked by the end of year. Since then, flu reports have been dropping off throughout the country.

The TERN task aims to recognize the time expressions “the past several months”, “recent season”, “early December”, “a month”, and “the end of the year” occurring in the text, and normalize them to the standard format, as shown below. The recognition and normalization results are presented here with the TimeML format.

<?xml version=“1.0” ?>
<TimeML xmlns:xsi=“http://www.w3.org/2001/XMLSchema-instance”xsi:noNamespaceSchemaLocation=“http://timeml.org/timeMLdocs/TimeML_1.2.1.xsd”>
<DOCID>AP_20130322</DOCID> <DCT>
<TIMEX3 functionInDocument=“CREATION_TIME” temporalFunction=“false” tid=“t0” type=“DATE” value=“2013-03-22”>March 22, 2013</TIMEX3>
</DCT>
<TEXT>
In that group, 177 out of every 100,000 were hospitalized with flu-related illness in <TIMEX3 tid=“t1” type=“DURATION” value=“PXM”>the past several months</TIMEX3>. That’s more than 2 1/2 times higher than any other <TIMEX3 tid=“t2” type=“SET” value=“XXXX-WI”>recent season </TIMEX3>. This flu season started in <TIMEX3 tid=“t3” type=“DATE” value=“2012-12”>early December</TIMEX3>, <TIMEX3 tid=“t4” type=“DURATION” value=“P1M”>a month</TIMEX3> earlier than usual, and peaked by <TIMEX3 tid=“t5” type=“DATE” value=“2012”>the end of year </TIMEX3>. Since then, flu reports have been dropping off throughout the country.
</TEXT>
</TimeML>

The TERN task can be viewed as containing two sub-tasks, namely time expression recognition (TER) and time expression normalization (TEN). TER aims to recognize time expressions from unstructured text while TEN aims to normalize the recognized time expressions to standard format.

1.3 Challenges of time expression recognition and normalization

There are several challenges in resolving the TERN problem, or the problems of TER and TEN. We discuss those main challenges and point out which ones have been solved or partially solved and which ones are still unsolved. The challenges in solving the TER problem are listed below.

(1) Collecting keywords that constitute time expressions (almost solved)
(2) Various written forms of time keywords and time expressions (partially solved)
(3) Countless of digit-based time expressions (partially solved)
(4) Inconsistent annotations, especially the boundary of time expressions (unsolved)

The challenges in solving the TEN problem are listed as follows.

(1) Collecting keywords and their normalization forms of time expressions (partially solved)
(2) Ambiguous meaning of time elements that constitute time expressions (unsolved)
(3) Inconsistent annotations, especially the normalizing form of time expressions (unsolved)

(4) Inconsistent reference dates within the same document (unsolved, not even be noticed)

1.4 Structure of this survey

The survey is organized as follows. Section 1 overall describes the importance, definition, and challenges of TERN. Section 2 briefly depicts early years' studies on syntactic structures and semantic interpretation of time expressions, which is fundamental for practicing the recognition and normalization of time expressions. Section 3 covers the development from 1996 to 2022, in the perspective of different factors. Section 5.1 summarizes the development of time expression recognition, in which we investigate three kinds of methods; after that Sect. 5.2 look at the development of time expression normalization, in which rule-based methods play a dominate role. Section 4 depicts the evaluations that are designed for the task TERN and Sect. 7 draw the conclusion and highlight the potential direction in future research.

2 Interpretation of time expressions

This section introduces the study by Smith (1978) on syntactic structures and semantic interpretations of time expressions, which is fundamental for the practical development of time expression recognition and normalization.

2.1 Three notions of time: speech time, reference time, and event time

Smith (1978) starts with Hans Reichenbach's view that temporal specification involves three notions of time: speech time, reference time, and event time (Reichenbach 1947). Speech time (ST) is the time at which a given sentence is uttered, that is, the moment of utterance. Reference time (RT) is the time indicated by a sentence, which need not be the same as speech time. Event time (ET) refers to the moment at which the relevant event or state occurs, which need not to be the same as reference time. In Example (1), Event time and reference time are the same, and they are prior to speech time.

Example 1 Marilyn won the prize last week.

Example 2 Marilyn had already won the prize last week.

In Example (2), all the three times are different: speech time is the moment of utterance, reference time is the last week, while event time is an unspecified time prior to last week.

To understand the temporal specification of a sentence, we need to know the values of the three kinds of times, and their relations to each other. Two times may be simultaneous, or one may precede the other: reference time may be (but need not to be) simultaneous with speech time, and event time may be (but need not to be) simultaneous with reference time.

2.2 Three relational values: past, present, and future

Time expressions fall into three classes according to whether they indicate anteriority, simultaneity, or posteriority. The three classes correspond three relational values: past,

Table 1 Three relational values of time expressions

	Anteriority	Simultaneity	Posteriority
Tense type	Past tense	Present tense	Future tense
Adverbials	Yesterday, - ago, last -	Now, at this moment	Tomorrow, next -, in -
Prepositions	Before, etc	At, on, etc	After, etc
Reference-time type	Past reference time	Present reference time	Future reference time

Fig. 1 Phrase structure rules for time adverbials in time expressions

Time adverbial	→	(Temporal) + (Habitual)
Temporal	→	PP
PP	→	P{NP|S}(PP)
Habitual	→	(Frequency) + (Temporal)
Frequency	→	Number + Unit
Unit	→	Indef Det + Noun

present, and future. For example, “right now” is simultaneous with speech time, indicating the present; “yesterday” precedes speech time, indicating the past; and “tomorrow” follows speech time, indicating the future. Table 1 briefly shows the three relational values of time expressions and their corresponding tenses, adverbials, and prepositions.

2.3 Syntactic rules for temporal elements

The syntactic rules relevant to time expressions are few, mainly including phrase structure rules for time adverbials and the auxiliary, with the constraints on co-occurrence. The phrase structure rules, illustrated in Fig. 1, will generate the appropriate structures for time adverbials that are used in time expressions.

3 Observations: 1996 to 2022

The computational research aiming at automatically recognizing time expressions from free text forms a numerous and heterogeneous pool of strategies, methods, and representations. This section reviews the development of TERN in terms of three factors: time expression types, languages, and domain and textual types.

3.1 Timex type factor

3.1.1 TIMEX

In the expression “Time Expression”, the word “Time” aims to restrict the task to those expressions involved in time information. The MUC-6 and MUC-7 (Grishman and Sundheim 1996; Chinchor 1998a, b) concern only two types of time expressions: DATE and TIME. DATE includes the complete or partial date expressions while TIME includes the complete or partial expressions of the time of day. The time expressions

Table 2 Seven attributes defined in the TIMEX2 annotation scheme

Attribute	Description
VAL	Indicate the value of a time expression
MOD	Modifier in a time expression and is used in conjunction with other attributes
SET	Indicate sets of times that recur regularly or irregularly
PERIODICITY	Indicate sets of regularly recurring times
GRANULARITY	Represent the unit of time denoted by each set member and applies to sets of both regularly and irregularly recurring times
NON_SPECIFIC	Indicate the expression is non specific or concrete
COMMENT	Intend to record annotator comments in reference annotations

in text are annotated by the pair of markups <TIMEX> and </TIMEX>. Following examples show the annotation format for the two types of time expressions.

- <TIMEX TYPE="DATE">February 1975</TIMEX>
- <TIMEX TYPE="DATE">third quarter of 1991</TIMEX>
- <TIMEX TYPE="TIME">5 p.m. EST</TIMEX>
- <TIMEX TYPE="TIME">twelve o'clock noon</TIMEX>

3.1.2 TIMEX2

In 2001, Lisa Ferro et al. from MITRE Corporation extend the work of Chinchor (1998a) to consider the meaning of time expressions by replacing the TYPE (i.e., DATE vs. TIME) categorization attribute with a set of attributes to represent the actual time indicated by the expressions (Mani et al. 2001; Ferro et al. 2001; Ferro 2001; Ferro et al. 2005). They use the pair of markups <TIMEX2> and </TIMEX2> with seven attributes (i.e., VAL, MOD, SET, COMMENT, PERIODICITY, GRANULARITY, and NON_SPECIFIC; they are described in Table 2 and some examples are given in the following example time expressions) to annotate time expressions in text, in which VAL is the most important attribute whose value follows the ISO 8601 time standard. Following examples show the TIMEX2 format.

- <TIMEX2 VAL="1999-08-03">two weeks</TIMEX2> from <TIMEX2 VAL="1999-07-20">next Tuesday</TIMEX2>
- I tutored an English student <TIMEX2 VAL="1998-WXX-4" SET="YES" GRANULARITY="G1D">some Thursdays</TIMEX2> in <TIMEX2 VAL="1999">1998</TIMEX2>.
- <TIMEX2 VAL="PT30M" MOD="LESS_THAN">almost half an hour</TIMEX2>
- <TIMEX2 VAL="1998-SU">Last summer</TIMEX2>, I went to the beach on <TIMEX2 VAL="1998-WXX-6" SET="YES" GRANULARITY="G1D"> numerous Saturdays</TIMEX2>.

3.1.3 TIMEX3

In 2003, Pustejovsky et al. (2003a) introduce TimeML, a specification language that extends TIMEX and TIMEX2 to include both of time expressions and events and uses the pair of markups <TIMEX3> and </TIMEX3> to annotate the time expressions.[1] TIMEX3 includes four types of time expressions, namely DATE, TIME, SET, and DURATION, with value attribute (and some other attributes). Following shows some examples of the TIMEX3 annotation format.

- Shot down the plane on <TIMEX3 tid="t1" type="DATE" value="1994-04-06" temporalFunction="false" functionInDocument="NONE">April 6, 1994 </TIMEX3>.
- A Brooklyn woman was killed <TIMEX3 tid="t2" type="TIME" value="1998-02-12TEV" temporalFunction="true" functionInDocument="NONE" anchorTimeID="t0">Thursday evening</TIMEX3> when...
- The official Iraqi News Agency gives the <TIMEX3 tid="t3" type="SET" value="XXXX-XX-XX">daily</TIMEX3> tally of inspections.
- In <TIMEX3 tid="t4" type="DURATION" value="PT1H" mod="LESS_THAN" temporalFunction="false" functionInDocument="NONE">less than one hour </TIMEX3>

Both TIMEX2 and TIMEX3 are developed as the annotation guidelines for creating normalized representations of time expressions (and their connection to events) in free text. Although they are complex and require much effort to master them well, they has been widely accepted as standards and extensively applied to the research related to time expressions in many languages. Up to now, we find in the literature that time expressions receive attentions in at least 23 languages.

3.2 Language factor

The majority of research in TERN is devoted to the study of English (Chinchor 1998a; Pustejovsky et al. 2003a, b; Negri and Marseglia 2004; Ferro et al. 2005; Kolomiyets and Moens 2009; Pustejovsky et al. 2010; Styler et al. 2014; Mazur and Dale 2010; Verhagen et al. 2007, 2010; UzZaman et al. 2013; Bethard et al. 2015, 2016, 2017; Laparra et al. 2018a, b). Besides English, a proportion of research concerns the language independence and multilingualism problems. In the TempEval-2 shared task (Verhagen et al. 2010), six languages are of interest to investigate, and they include Chinese, English, Italian, French, Korean, and Spanish.

Outside the shared tasks, Chinese is well studied and presented in English and Chinese literature (Wu et al. 2005a, b; He et al. 2007, 2008; Wu et al. 2010; Li et al. 2014; Yin and Jin 2017; Pan et al. 2020). Similarly, French, Italian, and Korean are strongly represented and boosted in series of works (Vazov 2001; Baldwin 2002; Jang et al. 2004; Lavelli et al. 2005; Im et al. 2009; Caselli et al. 2011; Bittar et al. 2011; Moriceau and Tannier 2014;

[1] Besides TIMEX3, TimeML also contains other three data structures: LINK, EVENT, and SIGNAL. This survey focuses on TIMEX3; for other three data structures, please refer to (Pustejovsky et al. 2003a) for details.

Manfredi et al. 2014; Nzali et al. 2015; Jeong et al. 2016; Lim and Choi 2017, 2019). Many other languages receive attention as well: Arabic (Boudaa et al. 2018), Basque (Altuna et al. 2017), Brazilian (de Azevedo et al. 2018a, b), Catalan (Taule et al. 2008), Croatian (Skukan et al. 2014), Dutch (van de Camp and Christiansen 2012), German (Strötgen and Gertz 2011), Hindi (Kamila et al. 2018), Persian (Mansouri et al. 2018), Portuguese (Costa and Branco 2012; de Azevedo et al. 2018a, b; Tissot et al. 2019), Romanian (Forascu and Tufis 2012), Russian Funkner and Kovalchuk (2020), Spanish (Saquete et al. 2002; Taule et al. 2008; Sauri et al. 2010; Strötgen et al. 2013; Najafabadipour et al. 2019; de-la Cuadra et al. 2019; Navas-Loro and Rodríguez-Doncel 2020), Swedish (Berglund 2004), Turkish (Emirali and Karslıgil 2022), Uyghur (Murat et al. 2018), Ukrainian (Grabar and Hamon 2018, 2019), and Vietnamese (Strötgen et al. 2014). Some works concern the TERN problem in multilingual text (Wilson et al. 2001; Saquete et al. 2004; Negri et al. 2006; Taule et al. 2008; Llorens et al. 2010; Strotgen and Gertz 2013, 2015; Strötgen et al. 2013, 2014; Starý et al. 2020; Lange et al. 2020, 2022; Cao et al. 2022).

Table 3 summarizes the languages that researchers conduct researchers on TERN. Currently, there are only about 21 languages that researchers construct useful datasets for TERN. The number 23 is not low, indicating that there exist enough useful datasets and algorithms for researchers who are interested in TERN in different languages especially low-resource languages. But the number 23 is still not high, in comparison with the whole languages used in the world, indicating that TERN and its downstream tasks need researchers from more language background to contribute to this area.

3.3 Domain and textual type factor

The very first studies are mainly concerned with the TERN problem in formal text like news articles (Chinchor 1998a; Setzer and Gaizauskas 2000; Pustejovsky et al. 2003b; Boguraev et al. 2007; Parker et al. 2011; Fu and Dhonnchadha 2020). Later on, studies are gradually concerned with TERN in other domains and textual types. Mazur and Dale (2010) collect English articles from Wikipedia about famous wars and annotate the time expressions for domain-specific time expression analysis; the corpus is called WikiWars. Similarly, Strötgen and Gertz (2011) develop the WikiWarsDE, which includes Wikipedia articles in German in the war domain. Strötgen and Gertz (2012) analyze time expressions in the text from colloquial short message service (SMS) and scientific biomedical documents while Degaetano-Ortlieb and Strötgen (2017) analyze the time expressions in the scientific literature and their diachronic variation over a time span of about 350 years. Tabassum et al. (2016) analyze the time expressions in the tweets which are informal text. Zhong et al. (2017) and Zhong and Cambria (2018) analyze the time expressions across formal and informal text and comprehensive and specific domain text. Kim et al. (2020) analyze time expressions in novels while Navas-Loro et al. (2019) and Loro (2021) analyzes time expressions in legal documents. Zarcone et al. (2020) and Alam et al. (2021) are concerned with time expressions in the voice assistant domain. Nzali et al. (2015), Strotgen (2015), Zhong and Cambria (2018), Olex et al. (2019), and Alam et al. (2021) analyze time expressions in cross-domain or cross-dataset scenarios.

A series of researches have devoted tremendous effort on time expressions in clinical text (including the clinical reports) and public health (Jindal and Roth 2013; Lin et al. 2013; Roberts et al. 2013; Sohn et al. 2013; Miller et al. 2015; Sun et al. 2015; Bethard et al. 2015, 2016, 2017; Lee et al. 2018; MacAvaney et al. 2018; Laparra et al. 2018b; Hao et al. 2018; Viani et al. 2018; de-la Cuadra et al. 2019; Najafabadipour et al. 2019;

Table 3 Languages on which researchers conduct researches about TERN

Language	Literature
English	Grishman and Sundheim (1996), Chinchor (1998a), Pustejovsky et al. (2003a, 2003b), Negri and Marseglia (2004), Ferro et al. (2005), Kolomiyets and Moens (2009), Pustejovsky et al. (2010), Styler et al. (2014), Mazur and Dale (2010), Verhagen et al. (2007, 2010), UzZaman et al. (2013), Bethard et al. (2015, 2016, 2017), Laparra et al. (2018a, 2018b)
Chinese	Wu et al. (2005a, 2005b), He et al. (2007, 2008), Wu et al. (2010), Verhagen et al. (2010), Li et al. (2014), Yin and Jin (2017), Pan et al. (2020)
French	Vazov (2001), Baldwin (2002), Bittar et al. (2011), Verhagen et al. (2010), Moriceau and Tannier (2014), Nzali et al. (2015)
Italian	Lavelli et al. (2005), Caselli et al. (2011), Verhagen et al. (2010), Manfredi et al. (2014)
Korean	Jang et al. (2004), Im et al. (2009), Verhagen et al. (2010), Jeong et al. (2016), Lim and Choi (2017, 2019)
Arabic	Boudaa et al. (2018)
Basque	Altuna et al. (2017)
Brazilian	de Azevedo et al. (2018a, b)
Catalan	Taule et al. (2008)
Croatian	Skukan et al. (2014)
Dutch	van de Camp and Christiansen (2012)
German	Strötgen and Gertz (2011), Almasian et al. (2022)
Hindi	Kamila et al. (2018)
Persian	Mansouri et al. (2018)
Portuguese	Costa and Branco (2012), de Azevedo et al. (2018a, b), Tissot et al. (2019)
Romanian	Forascu and Tufis (2012)
Russain	Funkner and Kovalchuk (2020)
Spanish	Saquete et al. (2002), Taule et al. (2008), Sauri et al. (2010), Verhagen et al. (2010), Strötgen et al. (2013), Najafabadipour et al. (2019), de-la Cuadra et al. (2019), Navas-Loro and Rodríguez-Doncel (2020)
Swedish	Berglund (2004)
Turkish	Emirali and Karslıgil (2022)
Uyghur	Murat et al. (2018)
Ukrainian	Grabar and Hamon (2018, 2019)
Vietnamese	Strötgen et al. (2014)
Multilingual	Wilson et al. (2001), Saquete et al. (2004), Negri et al. (2006), Taule et al. (2008), Llorens et al. (2010), Strotgen and Gertz (2013, 2015), Strötgen et al. (2013, 2014), Starý et al. (2020), Lange et al. (2020, 2022), Cao et al. (2022)

Zhong et al. 2019; Dupuis et al. 2020; Fu et al. 2020; Niu et al. 2020; Pan et al. 2020; Viani et al. 2020; Shim et al. 2021).

Table 4 summarizes the domains and textual types about which researchers have conducted investigation on TERN. It shows that researchers are mainly concerned with TERN in news articles and clinical records and public health. To the best of our knowledge, there is no research by now about TERN in sports, biographies, and entertainments, which are three important domains, because people are very much interested in different kinds of sports and entertainment activities and famous people in history and nowadays. Researchers who are interested in TERN and have some domain knowledge could consider to contribute to TERN in the three domains.

Table 4 Domain and textual types on which researchers conduct researches about TERN

Domain/text type	Literature
News articles	Grishman and Sundheim (1996), Chinchor (1998a), Pustejovsky et al. (2003a), Setzer and Gaizauskas (2000), Pustejovsky et al. (2003b), Negri and Marseglia (2004), Boguraev et al. (2007), Pustejovsky et al. (2010), Verhagen et al. (2007, 2010), UzZaman et al. (2013), Parker et al. (2011), Fu and Dhonnchadha (2020)
Wars	Mazur and Dale (2010), Strötgen and Gertz (2011), Grabar and Hamon (2019)
Colloquial SMS	Strötgen and Gertz (2012)
Scientific literature	Strötgen and Gertz (2012)
Tweets	Tabassum et al. (2016), Zhong et al. (2017), Zhong and Cambria (2018)
Novels	Navas-Loro et al. (2019)
Legal documents	Navas-Loro et al. (2019), Loro (2021)
Voice assistants	Zarcone et al. (2020), Alam et al. (2021)
Cross-domains	Zhong and Cambria (2018), Olex et al. (2019), Alam et al. (2021)
Clinical records and public health	Jindal and Roth (2013), Lin et al. (2013), Roberts et al. (2013), Sohn et al. (2013), Miller et al. (2015), Sun et al. (2015), Bethard et al. (2015, 2016, 2017), Lee et al. (2018), MacAvaney et al. (2018), Laparra et al. (2018b), Hao et al. (2018), Viani et al. (2018), de-la Cuadra et al. (2019), Najafabadipour et al. (2019), Zhong et al. (2019), Dupuis et al. (2020), Fu et al. (2020), Niu et al. (2020), Pan et al. (2020), Viani et al. (2020), Shim et al. (2021)

4 Evaluation of TERN

The evaluation shared tasks of TERN are essential to the progress in TERN. Many techniques are proposed to rank systems based on their capability to annotate a text like an expert linguist. In this section, we take a look at six main measures that are used for MUC, ACE, TempEval, and Clinical TempEval competitions. Before that, let us summarize the task from the point of view of evaluation.

4.1 MUC evaluation

In the MUC-7 event (Chinchor 1998a), a system is scored on two axes: its ability to find the correct type (TYPE) and its ability to find exact text (TEXT). A correct TYPE is credited if an entity is assigned the correct type, regardless of boundaries as long as there is an overlap. A correct TEXT is credited if entity boundaries are correct, regardless of the type. For both TYPE and TEXT, three measures are kept: the number of correct answers, the number of actual system guesses, and the number of possible entities in the solution.

The final MUC score is the micro-averaged f-measure (MAF), which is the harmonic mean of precision and recall calculated over all entity slots on both axes. A micro-averaged measure is performed on all entity types without distinction (errors and successes for all entity types are summed together). The harmonic mean of two numbers is never higher than the geometrical mean. It also tends toward the least number, minimizing the impact of large outliers and maximizing the impact of small ones. The F-measure therefore tends to privilege balanced systems.

In MUC, precision is calculated as COR/ACT and the recall is COR/POS. For the previous example, COR = 4 (2 TYPE + 2 TEXT), ACT = 10 (5 TYPE + 5 TEXT), and POS = 10 (5 TYPE + 5 TEXT). The precision is therefore 40%, the recall is 40%, and the MAF is 40%. However, these metrics are complicated and later on researchers tend not to use them in reporting results about TERN.

4.2 ACE evaluation

ACE has a complex evaluation procedure. It includes mechanisms for dealing various evaluation issues (e.g., partial match and wrong type). The ACE task definition is also more elaborated than previous tasks at the level of named entity "subtypes" and "class" as well as entity mentions, but these supplemental elements will be ignored.

Basically, each entity type has a parameterized weight and contributes up to a maximal proportion (MAXVAL) of the final score (e.g., if each person is worth 1 point and each organization is worth 0.5 point, then it takes two organizations to counterbalance one person in the final score).

4.3 TempEval evaluation

TempEval-2 (Verhagen et al. 2010), organized as the SemEval-2010 Task 13, aims to automatically identify all the time expressions, events, and temporal relations within a text, in six languages, namely Chinese, English, Italian, French, Korean, and Spanish. In this survey, we focus on time expressions. It uses three standard metrics to evaluate the TER performance: Precision, Recall, and F_1, and uses Accuracy to evaluate the performance of attribute classification and value normalization.

$$Precision = TP/(TP + FP) \tag{1}$$

$$Recall = TP/(TP + FN) \tag{2}$$

$$F_1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{3}$$

$$Accuracy = \frac{Number\ of\ correct\ answers}{Number\ of\ total\ answers} \tag{4}$$

where TP is the number of tokens that are part of an extent in both key and response, FP is the number of tokens that are part of an extent in the response but not in the key, and FN is the number of tokens that are part of an extent in the key but not in the response.

4.4 Clinical TempEval evaluation

Clinical TempEval Evaluations (Bethard et al. 2015, 2016, 2017; Laparra et al. 2018b) are a series of evaluations that consider time information extraction, including TERN, in the clinical domain. They are concerned with different types of time expressions in clinical documents but use the same evaluation metrics (i.e., precision, recall, F_1, and accuracy) as standard TempEval evaluations as mentioned above.

Table 5 Evaluations about TERN

Evaluation	Literature
MUC	Grishman and Sundheim (1996), Chinchor (1998a)
ACE	Mazur and Dale (2010), Strötgen and Gertz (2011), Grabar and Hamon (2019)
TempEval	Verhagen et al. (2007, 2010), UzZaman et al. (2013)
Clinical TempEval	Bethard et al. (2015, 2016, 2017), Laparra et al. (2018b)

Table 5 summarizes the evaluations and shared tasks that researchers organize to resolve the TERN problem alone with the literature. There are four series of shared tasks regarding TERN, namely MUC, ACE, TempEval, and Clinical TempEval, and the TempEval series receives most attention and have a significant impact on the TERN development, including setup some standards for TERN research.

5 Time expression recognition and normalization

Realizing the importance of time expression in text analysis and its subsequent applications, researchers from diverse areas organize TempEval evaluations and other shared tasks to automatically recognize time expressions from unstructured text and then normalize them into standard format. Although most approaches address the time expression recognition and normalization as an end-to-end task, we discuss the two sub-tasks separately, because the two sub-tasks lie at different levels of linguistic analysis. Specifically, the recognition task lies at the level of syntactic analysis while the normalization task lies at the level of semantic analysis. According to Noam Chomsky's "Syntactic Structures," pp. 93–94 (Chomsky 1957), syntax is not related to semantics and semantics does not affect syntax. Therefore, it is better to discuss the syntactic task and semantic task separately, and in this way, we can understand each of the two sub-tasks deeply.

5.1 Time expression recognition

The methods for TER are mainly categorized into three kinds: rule-based methods, traditional machine learning-based methods, and deep learning-based methods.

5.1.1 Rule-based methods

Rule-based time taggers like TempEx, GUTime, HeidelTime, and SUTime mainly predefine a set of time-related words and regular expression patterns (Mani and Wilson 2000; Verhagen et al. 2005; Strötgen and Gertz 2010; Chang and Manning 2012). HeidelTime hand-crafts rules with time resources like weekdays, seasons, and months, and leverages language clues like part-of-speech (POS) to identify time expression and then normalize them to the standard form in a pipeline Unstructured Information Management Architecture (UIMA[2]) (Strötgen and Gertz 2010). SUTime (Chang and Manning 2012) designs

[2] http://uima.apache.org.

deterministic rules using a cascade finite automata (Hobbs et al. 1997) on regular expressions over tokens (Chang and Manning 2014). It first identifies individual words, then expands them to chunks, and finally to the whole time expressions. Other rule-based taggers include the FSS-TimEx (Zavarella and Tanev 2013), which uses finite-state rule cascades to recognize time expressions, and SynTime (Zhong et al. 2017), which defines a token type system to describe the constituents of time expressions and designs a set of heuristic rules over the token types to recognize the time expressions. The rule-based time taggers achieve very good performance in the TempEval exercises, in which HeidelTime achieves the highest F_1 of 86% in TempEval-2 (Verhagen et al. 2010) and SUTime achieves the highest F_1 of 91.3% at relaxed match in the TempEval-3 (UzZaman et al. 2013). The type-based method SynTime achieves much better results on the TempEval-3, WikiWars, and Tweets datasets compared with both of rule-based and learning-based taggers (Zhong et al. 2017).

5.1.2 Machine learning-based methods

Machine learning based methods mainly extract features from the text and on the features apply the statistical models for recognizing time expressions. Example features include character features (e.g., first 3 characters and last 3 characters), word features (e.g., current word, previous word, and subsequent word), syntactic features (e.g., part-of-speech and noun phrase chunks), semantic features (e.g., lexical semantics and semantic role), and gazetteer features (e.g., match in a dictionary) (Llorens et al. 2010; Filannino et al. 2013; Bethard 2013). The statistical models include Markov logic network, logistic regression, support vector machines, maximum entropy, and conditional random fields (Llorens et al. 2010; UzZaman and Allen 2010; Jung and Stent 2013; Filannino et al. 2013; Bethard 2013). Those methods mainly leverage information from the labeled data under the framework of supervised learning. Some of them obtain good performance, and even achieve the highest F_1 of 82.71% in terms of strict match in the TempEval-3 competition (Bethard 2013).

Outside TempEval evaluations, Angeli et al. (2012) leverage compositional grammar and employ a EM-style approach to learn a latent parser for time expression recognition. In the method named UWTime, Lee et al. (2014) handcraft a combinatory categorial grammar (CCG) Steedman (1996) to define a set of lexicon with rules and use L1-regularization to learn linguistic context. The two methods explicitly use linguistic information. In Lee et al. (2014), especially, CCG could capture rich structure information of language, similar to the rule-based methods. Tabassum et al. (2016) focus on resolving the dates in tweets, and use distant supervision to recognize time expressions. Kuzey et al. (2016) formulate the TERN problem as an entity-like temponym resolution problem and develop an integer linear program to jointly infer temponym mappings to timeline and knowledge base so as to resolve time expressions. Zhong and Cambria (2018) define a constituent-based tagging scheme TOMN, which indicates the constituents of time expressions defined in SynTime (Zhong et al. 2017), to model time expressions (Zhong 2020); they also apply this idea to model both time expressions and named entities (Zhong et al. 2020, 2022; Zhong and Cambria 2021; Zhou et al. 2022). Ning et al. (2018) formulate TER as a text-chunking problem and use a basic machine learning method to recognize time expression chunks. Ding et al. (2019) propose to automatically generate abstracted patterns and then use the extended budgeted maximum coverage model to select appropriate patterns for recognizing time expressions.

5.1.3 Deep learning-based methods

With the popularity of applying neural networks and deep-learning methods to numerous tasks, many researchers use neural networks, word embeddings, and graph embeddings to model and recognize time expressions from unstructured text (Kim and Jeong 2016; Lin et al. 2017; Etcheverry and Wonsever 2017; Vashishth et al. 2018; Hossain et al. 2018; Chen et al. 2019; Kim et al. 2020; Lange et al. 2020; Patra et al. 2020; Almasian et al. 2021, 2022; Cao et al. 2022). Kim and Jeong (2016) employ element-type and Jordan-type recurrent neural networks (RNN) to extract both time expressions and events from textual data. Lin et al. (2017) propose to represent time expressions with single pseudo-tokens for convolutional neural networks (CNNs) and use these representation to extract temporal relations. Etcheverry and Wonsever (2017) leverage distributed representations and artificial neural networks to model time expressions and investigate the number of layers, sizes, and normalization techniques in recognizing Spanish time expressions from unstructured text. Vashishth et al. (2018) use a graph convolutional network (GCN) with jointly exploiting syntactic and temporal graph structures of document to infer the creating date of documents. Hossain et al. (2018) develop a system using long short-term memory (LSTM) recurrent neural network (RNN) along with word embedding to extract time expressions from the TempEval-2 textual data. Chen et al. (2019) try to model time expressions with pre-trained word representations and explore the need of contextualization and training resource requirements to recognize time expressions from free text. Kim et al. (2020) investigate multilingual methods with adversarial alignment to model time expressions in multiple languages in a common embedding space and recognize time expressions in the cross-lingual scenario. Patra et al. (2020) exploit a sequence-to-sequence encoder with contextual entity embeddings and negation constraints to resolute date-time entities in scheduling. Almasian et al. (2021, 2022) introduce two popular pre-trained transformer-based models BERT (Devlin et al. 2019) and RoBERTa (Liu et al. 2019) to model time expressions and general temporal information. Cao et al. (2022) instead leverage two state-of-the-art multilingual models mBERT (Devlin et al. 2019) and XLMR (Conneau et al. 2020) to model time expressions in multiple languages by transferring knowledge from multiple source languages to the low-resource target language.

Table 6 summarizes the three kinds of methods that are developed to resolve the TER task. It shows that in relatively early years, researchers tend to design rules and use traditional

Table 6 Three kinds of methods developed for time expression recognition

Method	Literature
Rule methods	TempEx (Mani and Wilson 2000), GUTime (Verhagen et al. 2005), HeidelTime (Strötgen and Gertz 2010; Strötgen et al. 2013), SUTime (Chang and Manning 2012, 2013), SynTime (Zhong et al. 2017)
Machine-learn methods	TIPS (Llorens et al. 2010), TRIPS-TRIO (UzZaman and Allen 2010), Angeli et al. (2012), ATT (Jung and Stent 2013), ManTime (Filannino et al. 2013), ClearTK (Bethard 2013), UWTime (Lee et al. 2014), TweeTIME (Tabassum et al. 2016), Temponym (Kuzey et al. 2016), TOMN (Zhong and Cambria 2018), CogCompTime (Ning et al. 2018), PTime (Ding et al. 2019)
Deep-learn methods	Kim and Jeong (2016), Lin et al. (2017), Etcheverry and Wonsever (2017), Vashishth et al. (2018), Chen et al. (2019), Kim et al. (2020), Lange et al. (2020), Patra et al. (2020), Almasian et al. (2021, 2022), XLTime (Cao et al. 2022)

machine-learning methods for TER while in recent years researchers prefer to apply deep learning-based methods to model time expressions.

5.2 Time expression normalization

Those methods that are developed for time expression normalization in the TempEval exercises are mainly based on rules (Mani and Wilson 2000; Verhagen et al. 2005; Strötgen and Gertz 2010; Llorens et al. 2010; UzZaman and Allen 2010; Chang and Manning 2013; Filannino et al. 2013; Bethard 2013). Because the rule systems have high similarity, Llorens et al. (2012) suggest to construct a large knowledge base as a public resource for the normalization task. Some researchers treat the normalization process as a learning task and use machine learning methods (Lee et al. 2014; Tabassum et al. 2016). Lee et al. (2014) use AdaGrad algorithm (Duchi et al. 2011) and Tabassum et al. (2016) use a log-linear algorithm to normalize time expressions. Sun et al. (2015) use traditional machine-learning models to normalize time expressions as a two multi-label classification task in clinical narratives; but their models achieve the accuracy of type classification at 57.2%. Laparra et al. (2018b) specifically organize a SemEval shared task focusing on parsing time normalization and use the SCATE scheme (Laparra et al. 2018a) to represent time expressions instead of the standard TimeML scheme, but only Olex et al. (2018) submit their results. Laparra et al. (2018a) propose a new annotation scheme SCATE and construct a new dataset with the same name SCATE for time expression normalization; they also use a character-level multi-output neural network to normalize time expressions under their SCATE scheme and corpus. Tissot et al. (2019) propose variant forms of trapezoidal and hexagonal fuzzy membership models to normalize time expressions in Portugueses and English. Ding et al. (2021) propose to automatically generate normalization rules from annotated data with common surface forms to normalize time expressions. Lange et al. (2022) use masked language models to normalize time expressions in multilingual scenarios and apply their models on low-resource languages. All these normalization researches assume that all the time expressions within a document have the same reference date and this reference date is generally set by the document creating date (DCT). However, there are many cases that within the same document, different time expressions may have different reference date. There are few researches that are concerned with this challenge except that Zhao et al. (2010) normalize time expressions with choosing dynamic reference time. This challenge is actually a bottleneck for TEN in long-text documents. We hope that researchers can pay some attention on this issue in future research.

Table 7 summarizes the three main kinds of methods that are proposed to normalize time expressions in textual data. It shows that rule-based methods are still dominant in the TEN task with good performance in comparison with other two kinds of methods. Some methods (i.e., CogCompTime) that use learning models for TER still design rules for TEN. There are also some deep-learning methods that are developed for TEN. But these traditional machine-learning methods and deep-learning methods need to tackle the issues of countless of digit-based time expressions, and the correspondence between time elements and their values in time expressions.

Table 7 Three kinds of methods developed for time expression normalization

Method	Literature
Rule methods	TempEx (Mani and Wilson 2000), GUTime (Verhagen et al. 2005), TIPS (Llorens et al. 2010), TRIPS-TRIO (UzZaman and Allen 2010), HeidelTime (Strötgen and Gertz 2010; Strötgen et al. 2013), SUTime (Chang and Manning 2012, 2013), ManTime (Filannino et al. 2013), CogCompTime (Ning et al. 2018), Chrono (Olex et al. 2018)
Machine-learn methods	Zhao et al. (2010), Angeli et al. (2012), ATT (Jung and Stent 2013), UWTime (Lee et al. 2014), TweeTIME (Tabassum et al. 2016), Temponym (Kuzey et al. 2016), Tissot et al. (2019), ARTime (Ding et al. 2021)
Deep-learn methods	SCATE (Laparra et al. 2018a), Lange et al. (2022)

Table 8 Softwares for time expression recognition and normalization

Method	Type	Address
HeidelTime (Strötgen et al. 2013)	Rule	https://github.com/HeidelTime/heideltime
SUTime (Chang and Manning 2012)	Rule	https://nlp.stanford.edu/software/sutime.shtml
ClearTK (Bethard 2013)	Machine learning	https://cleartk.github.io/cleartk/
UWTime (Lee et al. 2014)	Machine learning	https://lil.cs.washington.edu/uwtime/
SynTime (Zhong et al. 2017)	Rule	https://github.com/xszhong/syntime
TOMN (Zhong and Cambria 2018)	Machine learning	https://github.com/xszhong/tomn
CogCompTime (Ning et al. 2018)	Machine learning	https://github.com/qiangning/CogCompTime
Chrono (Olex et al. 2018)	Rule	https://github.com/AmyOlex/Chrono
ParsTime (Mansouri et al. 2018)	Rule	https://github.com/BehroozMansouri/ParsTime
PTime (Ding et al. 2019)	Machine learning	http://ws.nju.edu.cn/ptime
Kim et al. (2020)	Deep Learning	https://github.com/allenkim/what-time-is-it
ARTime (Ding et al. 2021)	Machine learning	https://github.com/nju-websoft/ARTime
DATEing (Aumiller et al. 2022)	Machine learning	https://github.com/satya77/Temporal_Tagger_Service
Almasian et al. (2021)	Deep Learning	https://github.com/satya77/Transformer_Temporal_Tagger
XLTime (Cao et al. 2022)	Deep Learning	https://github.com/YuweiCao-UIC/XLTime

6 TERN datasets and softwares

Available high-quality datasets and softwares about TERN are crucial for researchers to contribute to this area. We list here as many as possible resources that are developed for TERN. Table 8 lists some publicly available softwares for the TERN task while Table 9 reports some datasets that are usually used for TERN. A portal website for the TERN task is http://timexportal.wikidot.com/, where researchers can find some old-fashion but useful resources, such as literature, datasets, and systems.

Table 9 Annotated corpora for time expression recognition and normalization

Corpus	Brief description	Timex type	Address
MUC-7	Available at LDC under the catalogue number LDC2001T02	TIMEX	https://catalog.ldc.upenn.edu/LDC2001T02
TIDES	This corpus consists of two parts: (1) 95 Spanish dialogs and their English translations; (2) 193 documents of the TDT-2 corpus	TIMEX2	http://timexportal.wikidot.com/corpora-tides
ACE2004	Available at LDC under the catalogue number LDC2005T07	TIMEX2	https://catalog.ldc.upenn.edu/LDC2005T07
ACE2005	Available at LDC under the catalogue number LDC2006T06	TIMEX2	https://catalog.ldc.upenn.edu/LDC2006T06
ACE2007	Consist of selected domains in Arabic and Spanish	TIMEX2	https://catalog.ldc.upenn.edu/LDC2014T18
TimeBank	Available at LDC under the catalogue number LDC2006T08	TIMEX3	http://timexportal.wikidot.com/corpora-timebank12
THYME	Temporal annotation in the clinical domain	XML	https://github.com/stylerw/thymedata
WikiWars	A corpus of English Wikipedia articles about wars	TIMEX2	http://timexportal.wikidot.com/wikiwars
WikiWarsDE	A German version of WikiWars from German Wikipedia articles	TIMEX2	https://doi.org/10.11588/data/10026
ModeS TimeBank	A corpus of Modern Spanish (17th and 18th centuries) annotated under TimeML and SpatialML schemes	TIMEX3	https://catalog.ldc.upenn.edu/LDC2012T01
French TimeBank	Annotated with ISO-TimeML annotation standard	TIMEX3	https://gforge.inria.fr/projects/fr-timebank/
Tweets	A corpus about time expression collected from Twitter	TIMEX3	https://github.com/xszhong/syntime/tree/master/syntime/resources/tweets
PATE	A dataset of natural-language commands containing time expressions for voice assistants	TIMEX3	https://zenodo.org/record/3697930
RSDD-Time	Subset of diagnosed RSDD users with time information of the diagnosis made and whether the condition persists	TIMEX3	https://ir.cs.georgetown.edu/resources

7 Discussion and conclusion

The task of time expression recognition and normalization (TERN) has been thriving for more than 20 years. It aims at recognizing time expressions from free text and then normalizing the recognized time expressions to a machine readable format. In this survey, we have shown the diversity of languages, domains, textual genres and timex types covered in the literature. More than ten languages are studied. However, most of the works have concentrated on limited domains and textual genres such as news and web pages.

We have also provided an overview of the techniques employed to develop TERN systems, documenting the recent trend away from hand-created rules and traditional machine-learning approaches towards deep-learning methods. Handcrafted systems provide good performance at a relatively high system engineering cost. When supervised learning is used, a prerequisite is available to a large collection of annotated data. Such collections are available from the evaluation forums but remain rather rare and limited in domain and language coverage. Recent studies in the field have explored semi-supervised and unsupervised learning techniques that promise fast deployment for many timex types without the prerequisite of an annotated corpus. We have listed and categorized the features that are used in recognition and normalization algorithms. The use of an expressive and varied set of features turns out to be just as important as the choice of machine learning algorithms. Finally, we have also provided an overview of the evaluation methods that are used in the major forums of the TERN research community.

Acknowledgements This research is supported by the Agency for Science, Technology and Research (A*STAR) under its AME Programmatic Funding Scheme (Project #A18A2b0046).

References

Alam T, Zarcone A, Pado S (2021) New domain, major effort? How much data is necessary to adapt a temporal tagger to the voice assistant domain. In: International Conference on Computational Semantics, pp 144–154

Almasian S, Aumiller D, Gertz M (2021) Bert got a date: introducing transformers to temporal tagging. In: arXiv

Almasian S, Aumiller D, Gertz M (2022) Time for some german? pre-training a transformer-based temporal tagger for german. In: Proceedings of the Text2Story'22 Workshop, pp 83–90

Alonso O, Strotgen J, Baeza-Yates R, Gertz M (2011) Temporal information retrieval: challenges and opportunities. In: Proceedings of 1st International Temporal Web Analytics Workshop, pp 1–8

Altuna B, Aranzabe MJ, de Ilarraza AD (2017) Eusheideltime: time expression extraction and normalisation for basque. Proces Lenguaje Nat 59:15–22

Angeli G, Manning CD, Jurafsky D (2012) Parsing time: learning to interpret time expressions. In: Proceedings of 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp 446–455

Aumiller D, Almasian S, Pohl D, Gertz M (2022) Online dateing: a web interface for temporal annotations. In: Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp 3289–3294

Baldwin JA (2002) Learning temporal annotation of french news. Master's thesis, Graduate School of Arts and Sciences, Georgetown University

Berglund A (2004) Extracting temporal information and ordering events for swedish. Master's thesis

Bethard S (2013) Cleartk-timeml: a minimalist approach to tempeval 2013. In: Proceedings of the 7th International Workshop on Semantic Evaluation, pp 10–14

Bethard S, Derczynski L, Savova G, Pustejovsky J, Verhagen M (2015) Semeval-2015 task 6: clinical tempeval. In: Proceedings of the 9th International Workshop on Semantic Evaluation, pp 806–814

Bethard S, Savova G, Chen WT, Derczynski L, Pustejovsky J, Verhagen M (2016) Semeval-2016 task 12: clinical tempeval. In: Proceedings of SemEval-2016, pp 1052–1062

Bethard S, Savova G, Palmer M, Pustejovsky J (2017) Semeval-2017 task 12: clinical tempeval. In: Proceedings of the 11th International Workshop on Semantic Evaluation, pp 565–572

Bittar A, Amsili P, Denis P, Danlos L (2011) French timebank: an iso-timeml annotated reference corpus. In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics, pp 130–134

Boguraev B, Pustejovsky J, Ando R, Verhagen M (2007) Timebank evolution as a community resource for timeml parsing. Lang Resour Eval 41(1):91–115

Boudaa T, Marouani ME, Enneya N (2018) Arabic temporal expression tagging and normalization. In: Proceedings of International Conference on Big Data, Cloud and Applications, pp 546–557

Breitfeller L, Naik A, Rose C (2021) Stage: tool for automated extraction of semantic time cues to enrich neural temporal ordering models. In: arXiv

Campos R, Dias G, Jorge AM, Jatowt A (2014) Survey of temporal information retrieval and related applications. ACM Comput Surv 47(2):1–41

Cao Y, Groves W, Saha TK, Tetreault JR, Jaimes A, Peng H, Yu PS (2022) Xltime: a cross-lingual knowledge transfer framework for temporal expression extraction. In: NAACL, pp 1931–1942

Caselli T, Lenzi VB, Sprugnoli R, Pianta E, Prodanof I (2011) Annotating events, temporal expressions and relations in italian: the it-timeml experience for the ita-timebank. In: Proceedings of the 5th Linguistic Annotation Workshop, pp 143–151

Chambers N, Wang S, Jurafsky D (2007) Classifying temporal relations between events. In: Proceedings of the 45th Annual Meeting of the ACL on Interactive Poster and Demonstration Sessions, pp 173–176

Chang AX, Manning CD (2012) Sutime: a library for recognizing and normalizing time expressions. In: Proceedings of 8th International Conference on Language Resources and Evaluation, pp 3735–3740

Chang AX, Manning CD (2013) Sutime: evaluation in tempeval-3. In: Proceedings of the Second Joint Conference on Lexical and Computational Semantics (SEM), pp 78–82

Chang AX, Manning CD (2014) Tokensregex: defining cascaded regular expressions over tokens. Stanford University, Tech rep, Department of Computer Science

Chen S, Wang G, Karlsson B (2019) Exploring word representations on time expression recognition. Microsoft Research Asia, Tech rep

Cheng F, Miyao Y (2018) Inducing temporal relations from time anchor annotation. Proc NAACL-HLT 2018:1833–1843

Chinchor NA (1998a) Muc-7 named entity task definition. In: Proceedings of the 7th Message Understanding Conference, vol 29

Chinchor NA (1998b) Overview of muc-7/met-2. In: Proceedings of the 7th Message Understanding Conference

Chomsky N (1957) Syntactic structures. Mouton Publishers, Berlin

Conneau A, Khandelwal K, Goyal N, Chaudhary V, Wenzek G, Guzman F, Grave E, Ott M, Zettlemoyer L, Stoyanov V (2020) Unsupervised cross-lingual representation learning at scale. In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, pp 8440–8451

Costa F, Branco A (2012) Timebankpt: a timeml annotated corpus of portuguese. In: Proceedings of the 8th International Conference on Language Resources and Evaluation, pp 3727–3734

de Azevedo RF, Joao Pedro Santos Rodrigues MRdSR, Moro CMC (2018a) Temporal tagging of noisy clinical texts in brazilian. In: Proceedings of International Conference on Computational Processing of the Portuguese Language, pp 231–241

de Azevedo RF, Rodrigues JPS, da Silva Reis MR, Moro CMC, Paraiso EC (2018b) Temporal tagging of noisy clinical texts in brazilian portuguese. In: Proceedings of PROPOR, pp 231–241

de-la Cuadra AR, López-Cuadrado JL, González-Carrasco I, Ruiz-Mezcua B (2019) Sequence time expression recognition in the spanish clinical narrative. In: 2019 IEEE 32nd International Symposium on Computer-Based Medical Systems, pp 331–336

Degaetano-Ortlieb S, Strötgen J (2017) Diachronic variation of temporal expressions in scientific writing through the lens of relative entropy. In: Proceedings of International Conference of the German Society for Computational Linguistics and Language Technology, pp 259–275

Devlin J, Chang MW, Lee K, Kristina T (2019) Bert: pre-training of deep bidirectional transformers for language understanding. In: Proceedings of NAACL-HLT, pp 4171–4186

Ding W, Gao G, Shi L, Qu Y (2019) A pattern-based approach to recognizing time expressions. Proc AAAI Conf Artif Intell 33:6335–6342

Ding W, Chen J, Li J, Qu Y (2021) Automatic rule generation for time expression normalization. In: ACL, pp 3135–3144

Do QX, Lu W, Roth D (2012) Joint inference for event timeline construction. In: Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, pp 677–687

Duchi J, Hazan E, Singer Y (2011) Adaptive subgradient methods for online learning and stochastic optimization. J Mach Learn Res 12:2121–2159

Dupuis L, Bergou N, Tissot H, Velupillai S (2020) Relative and incomplete time expression anchoring for clinical text. In: Proceedings of the 3rd Clinical Natural Language Processing Workshop, pp 117–129

Emirali E, Karslıgil ME (2022) Using word embeddings in detection of temporal expressions in turkish texts. In: IEEE 2022 30th Signal Processing and Communications Applications Conference, pp 1–4

Enc M (1986) Towards a referential analysis of temporal expressions. Linguist Philos 9(4):405–426

Etcheverry M, Wonsever D (2017) Time expressions recognition with word vectors and neural networks. In: Proceedings of the 24th International Symposium on Temporal Representation and Reasoning, pp 1–12

Ferro L (2001) Tides instruction manual for the annotation of temporal expressions. MITRE, Tech rep

Ferro L, Mani I, Sundheim B, Wilson G (2001) Tides temporal annotation guidelines—version 1.0.2. MITRE, Tech rep

Ferro L, Gerber L, Mani I, Sundheim B, Wilson G (2005) Tides 2005 standard for the annotation of temporal expressions. MITRE, Tech rep

Filannino M, Brown G, Nenadic G (2013) Mantime: temporal expression identification and normalization in the tempeval-3 challenge. In: Proceedings of the 7th International Workshop on Semantic Evaluation, pp 53–57

Forascu C, Tufis D (2012) Romanian timebank: an annotated parallel corpus for temporal information. In: LREC, pp 3762–3766

Fu Y, Dhonnchadha EU (2020) A pattern-mining driven study on differences of newspapers in expressing temporal information. http://arXiv.org/2011.12265

Fu JT, Sholle E, Krichevsky S, Scandura J, Campion TR (2020) Extracting and classifying diagnosis dates from clinical notes: a case study. J Biomed Inform 110(103):569

Funkner AA, Kovalchuk SV (2020) Time expressions identification without human-labeled corpus for clinical text mining in russian. In: International Conference on Computational Science, pp 591–602

Grabar N, Hamon T (2018) Automatic detection of temporal information in Ukrainian general-language texts. CNRS University Lille and LIMSI University Paris-Saclay, Tech rep

Grabar N, Hamon T (2019) Wikiwars-ua: Ukrainian corpus annotated with temporal expressions. In: Proceedings of the 3d International Conference Computational Linguistics And Intelligent Systems, pp 22–31

Grishman R, Sundheim B (1996) Message understanding conference—6: a brief history. In: Proceedings of the 16th International Conference on Computational Linguistics

Guillou L, Vroe SBD, Hosseini MJ, Johnson M, Steedman M (2020) Incorporating temporal information in entailment graph mining. In: Proceedings of the Graph-based Methods for Natural Language Processing, pp 60–71

Hao T, Pan X, Gu Z, Qu Y, Weng H (2018) A pattern learning-based method for temporal expression extraction and normalization from multi-lingual heterogeneous clinical texts. BMC Med Inform Decis Mak 18:16–25

Hasanuzzaman M, Dias G, Ferrari S, Mathet Y (2014) Propagation strategies for building temporal ontologies. In: Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics, pp 6–11

He RF, Qin B, Liu T, Pan YQ, Li S (2007) Recognizing the extent of Chinese time expressions based on the dependency parsing and error-driven learning. J Chin Inf Process 21(5):36–40

He R, Qin B, Liu T, Pan Y, Li S (2008) A novel heuristic error-driven learning for recognizing Chinese time expression. J Chin Lang Comput 18(4):139–159

Hinrichs EW (1987) A compositional semantics of temporal expressions in English. In: Proceedings of the 25th Annual Meeting on Association for Computational Linguistics, pp 8–15

Hobbs JR, Appelt DE, Bear J, Israel D, Kameyama M, Stickel M, Tyson M (1997) Fastus: a cascaded finite-state transducer for extracting information from natrual-language text. In: Finite State Devices for Natural Language Processing, pp 383–406

Hossain T, Rahman MM, Islam SM (2018) Temporal information extraction from textual data using long short-term memory recurrent neural network. J Comput Technol Appl 9(2):1–6

Im S, You H, Jang H, Nam S, Shin H (2009) Ktimeml: specification of temporal and event expressions in Korean text. In: Proceedings of the 7th Workshop on Asian Language Resources, ACL-IJCNLP 2009, pp 115–122

Jang SB, Baldwin J, Mani I (2004) Automatic timex2 tagging of Korean news. ACM Trans Asian Lang Inf Process 3(1):51–65
Jeong YS, Joo WT, Do HW, Lim CG, Choi KS, Choi HJ (2016) Korean timeml and Korean timebank. In: Proceedings of LREC, pp 356–359
Jia Z, Abujabal A, Roy RS, Strotgen J, Weikum G (2018) Tempquestions: a benchmark for temporal question answering. In: Proceedings of the 2018 World Wide Web Conference Companion, pp 1057–1062
Jia Z, Pramanik S, Roy RS, Weikum G (2021) Complex temporal question answering on knowledge graphs. In: Proceedings of the 30th ACM international conference on information and knowledge management, pp 792–802
Jindal P, Roth D (2013) Extraction of events and temporal expressions from clinical narratives. J Biomed Inform 46:S13–S19
Jung H, Stent A (2013) Att1: temporal annotation using big windows and rich syntactic and semantic features. In: Proceedings of the 7th International Workshop on Semantic Evaluation, pp 20–24
Kamila S, Hasanuzzaman M, Ekbal A, Bhattacharyya P (2018) Tempo-hindiwordnet: a lexical knowledge-base for temporal information processing. ACM Trans Asian Low-Resour Lang Inf Process 18(2):1–22
Khan SUR, Islam MA, Aleem M, Iqbal MA (2018) Temporal specificity-based text classification for information retrieval. Turk J Electr Eng Comput Sci 26:2915–2926
Kim A, Pethe C, Skiena S (2020) What time is it? Temporal analysis of novels. In: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, pp 9076–9086
Kim ZM, Jeong YS (2016) Timex3 and event extraction using recurrent neural networks. In: Proceedings of the 2016 IEEE International Conference on Big Data and Smart Computing, pp 450–453
Kolomiyets O, Moens MF (2009) Meeting tempeval-2: shallow approach for temporal tagger. In: Proceedings of the NAACL HLT Workshop on Semantic Evaluation: Recent Achievements and Future Directions, pp 52–57
Kuzey E, Setty V, Strotgen J, Weikum G (2016) As time goes by: comprehensive tagging of textual phrases with temporal scopes. In: Proceedings of the 2016 World Wide Web Conference, pp 915–925
Lange L, Iurshina A, Adel H, Strötgen J (2020) Adversarial alignment of multilingual models for extracting temporal expressions from text. In: Proceedings of the 5th Workshop on Representation Learning for NLP, pp 103–109
Lange L, Strötgen J, Adel H, Klakow D (2022) Multilingual normalization of temporal expressions with masked language models. In: arXiv
Laparra E, Xu D, Bethard S (2018a) From characters to time intervals: new paradigms for evaluation and neural parsing of time normalizations. Trans Assoc Computat Linguist 6:343–356
Laparra E, Xu D, Bethard S, Elsayed AS, Palmer M (2018b) Semeval 2018 task 6: Parsing time normalizations. In: Proceedings of the 12th International Workshop on Semantic Evaluation, pp 88–96
Lavelli A, Magnini B, Negri M, Pianta E, Speranza M, Sprugnoli R (2005) Italian content annotation bank (i-cab): temporal expressions (v. 1.0). ITC-irst, Centro per la Ricerca Scientifica e Technologica Povo, Tech rep
Lee K, Artzi Y, Dodge J, Zettlemoyer L (2014) Context-dependent semantic parsing for time expressions. In: Proceedings of the 52th Annual Meeting of the Association for Computational Linguistics, pp 1437–1447
Lee HJ, Zhang Y, Jiang M, Xu J, Tao C, Xu H (2018) Identifying direct temporal relations between time and events from clinical notes. BMC Med Inform Decis Mak 18(2):49
Leeuwenberg A, Moens MF (2019) A survey on temporal reasoning for temporal information extraction from text. J Artif Intell Res 66:341–380
Li J, Cardie C (2014) Timeline generation: tracking individuals on twitter. In: Proceedings of the 23rd International Conference on World Wide Web, pp 643–652
Li H, Strötgen J, Zell J, Gertz M (2014) Chinese temporal tagging with heideltime. In: Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics, pp 133–137
Lim CG, Choi HJ (2017) Efficient temporal information extraction from korean documents. In: Proceedings of IEEE 18th International Conference on Mobile Data Management, p 366370
Lim CG, Choi HJ (2019) Korean time information analysis of hierarchical annotation rules from natural language text. In: 2019 IEEE International Conference on Big Data and Smart Computing (BigComp), pp 1–4
Lim CG, Jeong YS, Choi HJ (2019) Survey of temporal information extraction. J Inf Process Syst 15(4):931–956
Lin YK, Chen H, Brown RA (2013) Medtime: a temporal information extraction system for clinical narratives. J Biomed Inform 46:S20–S28

Lin S, Jin P, Zhao X, Yue L (2014) Exploiting temporal information in web search. Expert Syst Appl 41(2):331–341
Lin C, Miller T, Dligach D, Bethard S, Savova G (2017) Representations of time expressions for temporal relation extraction with convolutional neural networks. In: Proceedings of the 16th BioNLP Workshop, pp 322–327
Liu Y, Ott M, Goyal N, Du J, Joshi M, Chen D, Levy O, Lewis M, Zettlemoyer L, Stoyanov V (2019) Roberta: a robustly optimized bert pretraining approach. http://arxiv.org/1907.11692
Liu J, Xu J, Chen Y, Zhang Y (2021) Discourse-level event temporal ordering with uncertainty-guided graph completion. In: Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, pp 3871–3877
Llorens H, Saquete E, Navarro B (2010) Tipsem (english and spanish): evaluating crfs and semantic roles in tempeval-2. In: Proceedings of the 5th International Workshop on Semantic Evaluation, pp 284–291
Llorens H, Derczynski L, Gaizauskas R, Saquete E (2012) Timen: an open temporal expression normalisation resource. In: Proceedings of the 8th International Conference on Language Resources and Evaluation, pp 3044–3051
Loro MN (2021) Processing, identification and representation of temporal expressions and events in legal documents. PhD thesis, Escuela Tecnica Superior de Ingenieros Informatica
MacAvaney S, Desmet B, Cohan A, Soldaini L, Yates A, Zirikly A, Goharian N (2018) Rsdd-time: temporal annotation of self-reported mental health diagnoses. In: Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic, pp 168–173
Manfredi G, Strötgen J, Zell J, Gertz M (2014) Heideltime at eventi: tuning italian resources and addressing timeml's empty tags. In: Proceedings of the 4th International Workshop EVALITA, pp 39–43
Mani I (2003) Recent developments in temporal information. In: Proceedings of the International Conference on Recent Advances in Natural Language Processing, pp 45–60
Mani I, Wilson G (2000) Robust temporal processing of news. In: Proceedings of the 38th annual meeting on Association for Computational Linguistics, pp 69–76
Mani I, Wilson G, Ferro L, Sundheim B (2001) Guidelines for annotation temporal information. In: Proceedings of the first International Conference on Human Language Technology Research, pp 1–3
Mani I, Verhagen M, Wellner B, Lee CM, Pustejovsky J (2006) Machine learning of temporal relations. In: Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics, pp 753–760
Mansouri B, Zahedi MS, Campos R, Farhoodi M, Rahgozar M (2018) Parstime: Rule-based extraction and normalization of persian temporal expressions. In: European Conference on Information Retrieval, pp 715–721
Mazur P, Dale R (2010) Wikiwars: a new corpus for research on temporal expressions. In: Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing, pp 913–922
Miller T, Bethard S, Dligach D, Lin C, Savova G (2015) Extracting time expressions from clinical text. In: Proceedings of the 2015 Workshop on Biomedical Natural Language Processing, pp 81–91
Minard AL, Speranza M, Agirre E, Aldabe I, van Erp M, Magnini B, Rigau G, Urizar R (2015) Semeval-2015 task 4: timeline: Cross-document event ordering. In: 9th International Workshop on Semantic Evaluation (SemEval 2015), pp 778–786
Moriceau V, Tannier X (2014) French resources for extraction and normalizationg of temporal expressions with heideltime. In: Proceedings of the 9th International Conference on Language Resources and Evaluation, pp 3239–3243
Murat A, Yusup A, Iskandar Z, Yusup A, Abaydulla Y (2018) Applying lexical semantics to automatic extraction of temporal expressions in Uyghur. J Inf Process Syst 14(4):824
Naik A, Breitfeller L, Rose C (2019) Tddiscourse: a dataset for discourse-level temporal ordering of events. In: Proceedings of the SIGDial 2019 Conference, pp 239–249
Najafabadipour M, Zanin M, Rodríguez-González A, Gonzalo-Martín C, García BN, Calvo V, Bermudez JLC, Provencio M, Menasalvas E (2019) Recognition of time expressions in spanish electronic health records. In: IEEE International Symposium on Computer-Based Medical Systems, pp 69–74
Navas-Loro M, Rodríguez-Doncel V (2020) Annotador: a temporal tagger for Spanish. J Intell Fuzzy Syst 39(2):1979–1991
Navas-Loro M, Filtz E, Rodríguez-Doncel V, Polleres A, Kirrane S (2019) Tempcourt: evaluation of temporal taggers on a new corpus of court decisions. Knowl Eng Rev 34:1–25
Negri M, Marseglia L (2004) Recognition and normalization of time expressions: Itc-irst at tern 2004. ITC-irst, Tech rep
Negri M, Saquete E, Martinez-Barco P, Munoz R (2006) Evaluating knowledge-based approaches to the multilingual extension of a temporal expression normalizer. In: Proceedings of the Workshop on Annotating and Reasoning about Time and Events, pp 30–37

Ning Q, Zhou B, Feng Z, Peng H, Roth D (2018) Cogcomptime: a tool for understanding time in natural language. In: Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing: System Demonstrations, pp 72–77
Niu J, Ng V, Penn G, Rees EE (2020) Temporal histories of epidemic events (thee): a case study in temporal annotation for public health. In: LREC, pp 2223–2230
Nzali MDT, Tannier X, Neveol A (2015) Automatic extraction of time expressions across domains in french narratives. In: Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, pp 492–498
Ocal M, Perez A, Radas A, Finlayson MA (2022a) Holistic evaluation of automatic timeml annotators. In: Proceedings of the 13th Conference on Language Resources and Evaluation, pp 1444–1453
Ocal M, Radas A, Hummer J, Megerdoomian K (2022b) A comprehensive evaluation and correction of the timebank corpus. In: Proceedings of the 13th Conference on Language Resources and Evaluation, pp 2919–2927
Olex AL, Maffey LG, Morton N, McInnes BT (2018) Chrono at semeval-2018 task 6: a system for normalizing temporal expressions. In: Proceedings of the 12th International Workshop on Semantic Evaluation (SemEval-2018), pp 97–101
Olex AL, Maffey LG, McInnes BT (2019) Nlp whack-a-mole: challenges in cross-domain temporal expression extraction. Proc NAACL-HLT 2019:3682–3692
Pan X, Chen B, Weng H, Gong Y, Qu Y (2020) Temporal expression classification and normalization from Chinese narrative clinical texts: pattern learning approach. JMIR Med Inform 8(7):e17-652
Parker R, Graff D, Kong J, Chen K, Maeda K (2011) Engilish gigaword, fifth edn
Patra B, Fufa C, Bhattacharya P, Lee CC (2020) To schedule or not to schedule: extracting task specific temporal entities and associated negation constraints. In: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, pp 8445–8455
Pustejovsky J, Castano J, Ingria R, Sauri R, Gaizauskas R, Setzer A, Katz G, Radev D (2003a) Timeml: robust specification of event and temporal expressions in text. New Dir Quest Answering 3:28–34
Pustejovsky J, Hanks P, Sauri R, See A, Gaizauskas R, Setzer A, Sundheim B, Radev D, Day D, Ferro L, Lazo M (2003b) The timebank corpus. Corpus Linguist 2003:647–656
Pustejovsky J, Lee K, Bunt H, Romary L (2010) Iso-timeml: an international standard for semantic annotation. In: Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10), pp 394–397
Qin L, Gupta A, Upadhyay S, He L, Choi Y, Faruqui M (2021) Timedial: temporal commonsense reasoning in dialog. In: ACL/IJCNLP, pp 7066–7076
Rahoman MM, Ichise R (2018) A proposal of a temporal semantics aware linked data information retrieval framework. J Intell Inf Syst 50(3):573–95
Reichenbach H (1947) Elements of Symbolic Logic. University of California Berkeley Press
Roberts K, Rink B, Harabagiu SM (2013) A flexible framework for recognizing events, temporal expressions, and temporal relations in clinical text. J Am Med Inform Assoc 20(5):867–875
Sakaguchi T, Kawahara D, Kurohashi S (2018) Comprehensive annotation of various types of temporal information on the time axis. In: Proceedings of the 11th Conference on Language Resources and Evaluation
Saquete E, Martinez-Barco P, Munoz R (2002) Recognizing and tagging temporal expressions in spanish. In: In Proceedings of LREC Workshop on Annotation Standards for Temporal Information in Natural Language, pp 44–51
Saquete E, Martinez-Barco P, Munoz R (2004) Evaluation of the automatic multilinguality for time expression resolution. In: Proceedings of the 15th International Workshop on Database and Expert Systems Applications, pp 25–30
Sauri R, Saquete E, Pustejovsky J (2010) Annotating time expressions in Spanish. timeml annotation guidelines (version tempeval-2010). Barcelona Media—Innovation Center, Tech rep
Setzer A, Gaizauskas R (2000) Annotating events and temporal information in newswire texts. In: Proceedings of LREC, pp 1287–1294
Shim H, Lowet D, Luca S, Vanrumste B (2021) Synthetic data generation and multi-task learning for extracting temporal information from health-related narrative text. In: Proceedings of the 2021 EMNLP Workshop W-NUT: the Seventh Workshop on Noisy User-generated Text, pp 260–273
Shokouhi M, Radinsky K (2012) Time-sensitive query auto-completion. In: Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp 601–610
Skukan L, Glavas G, Snajder J (2014) Heideltime.hr: extracting and normalizing temporal expressions in Croatian. In: Proceedings of the 9th Slovenian Language Technologies Conference, pp 99–103
Smith CS (1978) The syntax and interpretation of temporal expressions in English. Linguist Philos 2(1):43–99

Sohn S, Wagholikar KB, Li D, Jonnalagadda SR, Tao C, Elayavilli RK, Liu H (2013) Comprehensive temporal information detection from clinical text: medical events, time, and tlink identification. J Am Med Inform Assoc 20(5):836–842

Starý M, Nevěřilová Z, Valčík J (2020) Multilingual recognition of temporal expressions. In: Proceedings of Recent Advances in Slavonic Natural Language Processing, pp 67–78

Steedman M (1996) Surface Structure and Interpretation. The MIT Press, Cambridge

Strotgen J (2015) Domain-sensitive temporal tagging for event-centric information retrieval. PhD thesis, Heidelberg University, Germany

Strötgen J, Gertz M (2010) Heideltime: high quality rule-based extraction and normalization of temporal expressions. In: Proceedings of the 5th International Workshop on Semantic Evaluation (SemEval'10), Association for Computational Linguistics, Stroudsburg, PA, pp 321–324

Strötgen J, Gertz M (2011) Wikiwarsde: A german corpus of narratives annotated with temporal expressions. In: Proceedings of German Society for Computational Linguistics and Language Technology, pp 129–134

Strötgen J, Gertz M (2012) Temporal tagging on different domains: challenges, strategies, and gold standards. In: Proceedings of 8th International Conference on Language Resources and Evaluation, pp 3746–3753

Strotgen J, Gertz M (2013) Multilingual and cross-domain temporal tagging. Lang Resour Eval 47(2):269–298

Strotgen J, Gertz M (2015) A baseline temporal tagger for all languages. In: Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, pp 541–547

Strötgen J, Zell J, Gertz M (2013) Heideltime: tuning English and developing Spanish resources. In: Proceedings of the Second Joint Conference on Lexical and Computational Semantics (SEM), pp 15–19

Strötgen J, Armiti A, Canh TV, Zell J, Gertz M (2014) Time for more languages: temporal tagging of Arabic, Italian, Spanish, and Vietnamese. ACM Trans Asian Lang Inf Process 13(1):1–21

Styler WF, Bethard S, Finan S, Palmer M, Pradhan S, de Groen PC, Erickson B, Miller T, Lin C, Savova G, Pustejovsky J (2014) Temporal annotation in the clinical domain. Trans Assoc Computat Linguist 2:143–154

Su BY, Hsu SL, Lai KY, jen Hsu JY (2021) Context-dependent semantic parsing for temporal relation extraction. In: arXiv

Sun W, Rumshisky A, Uzuner O (2015) Normalization of relative and incomplete temporal expressions in clinical narratives. J Am Med Inform Assoc 22(5):1001–1008

Tabassum J, Ritter A, Xu W (2016) Tweetime: a minimally supervised method for recognizing and normalizing time expressions in twitter. In: Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, pp 307–318

Taule M, Marti T, Recasens M (2008) Ancora: Multilevel annotated corpora for Catalan and Spanish. In: Proceedings of the 6th International Conference on Language Resources and Evaluation

Tissot H, Fabro MDD, Derczynski L, Roberts A (2019) Normalisation of imprecise temporal expressions extracted from text. Knowl Inf Syst 61:1361–1394

UzZaman N, Allen JF (2010) Trips and trios system for tempeval-2: extracting temporal information from text. In: Proceedings of the 5th International Workshop on Semantic Evaluation, pp 276–283

UzZaman N, Llorens H, Derczynski L, Verhagen M, Allen J, Pustejovsky J (2013) Semeval-2013 task 1: Tempeval-3: evaluating time expressions, events, and temporal relations. In: Proceedings of the 7th International Workshop on Semantic Evaluation, pp 1–9

van de Camp M, Christiansen H (2012) Resolving relative time expressions in dutch text with constraint handling rules. In: Proceedings of the 7th International Workshop on Constraint Solving and Language Processing, pp 74–85

Vashishth S, Dasgupta SS, Ray SN, Talukdar P (2018) Dating documents using graph convolution networks. In: Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, pp 1605–1615

Vazov N (2001) A system for extraction of temporal expressions from French texts based on syntactic and semantic constraints. In: Proceedings of the Workshop on Temporal and Spatial Information Processing, pp 14–21

Verhagen M, Mani I, Sauri R, Knippen R, Jang SB, Littman J, Rumshisky A, Phillips J, Pustejovsky J (2005) Automating temporal annotation with Tarqi. In: Proceedings of the ACL Interactive Poster and Demonstration Sessions., pp 81–84

Verhagen M, Gaizauskas R, Schilder F, Hepple M, Katz G, Pustejovsky J (2007) Semeval-2007 task 15: tempeval temporal relation identification. In: Proceedings of the 4th International Workshop on Semantic Evaluation, pp 75–80

Verhagen M, Sauri R, Caselli T, Pustejovsky J (2010) Semeval-2010 task 13: tempeval-2. In: Proceedings of the 5th International Workshop on Semantic Evaluation, pp 57–62

Viani N, Yin L, Kam J, Alawi A, Bittar AL, Dutta R, Patel R, Stewart RJ, Velupillai S (2018) Time expressions in mental health records for symptom onset extraction. In: Proceedings of the 9th International Workshop on Health Text Mining and Information Analysis

Viani N, Kam J, Yin L, Bittar A, Dutta R, Patel R, Stewart R, Velupillai S (2020) Temporal information extraction from mental health records to identify duration of untreated psychosis. J Biomed Seman 11(2):1–11

Wilson G, Mani I, Sundheim B, Ferro L (2001) A multilingual approach to annotating and extracting temporal information. In: Proceedings of the Workshop on Temporal and Spatial Information Processing, p 12

Wong KF, Xia Y, Li W, Yuan C (2005) An overview of temporal information extraction. Int J Comput Process Oriental Lang 18(2):137–152

Wu M, Li W, Chen Q, Lu Q (2005a) Normalizing Chinese temporal expressions with multi-label classification. In: Proceedings of the 2th International Conference on Natural Language Processing and Knowledge Engineering, pp 318–323

Wu M, Li W, Lu Q, Li B (2005b) Ctemp: a Chinese temporal parser for extracting and normalizing temporal information. In: Proceedings of Natural Language Processing-IJCNLP, pp 694–706

Wu T, Zhou Y, Huang X, Wu L (2010) Chinese time expression recognition based on automatically generated basic-time-unit rules. J Chin Inf Process 24(4):3–10

Yin B, Jin B (2017) A multi-label classification method on Chinese temporal expressions based on character embedding. In: Proceedings of the 4th International Conference on Information Science and Control Engineering, pp 51–54

Zarcone A, Alam T, Kolagar Z (2020) Pâté: a corpus of temporal expressions for the in-car voice assistant domain. In: LREC, pp 523–530

Zavarella V, Tanev H (2013) Fss-timex for tempeval-3: extracting temporal information from text. In: SemEval, pp 58–63

Zhao X, Jin P, Yue L (2010) Automatic temporal expression normalization with reference time dynamic-choosing. COLING 2010:1498–1506

Zhong X (2020) Time expression and named entity analysis and recognition. PhD thesis, Nanyang Technological University, Singapore

Zhong X, Cambria E (2018) Time expression recognition using a constituent-based tagging scheme. In: Proceedings of the World Wide Web Conference, pp 983–992

Zhong X, Cambria E (2021) Time expression and named entity recognition, socio-affective computing, vol 10. Springer, Berlin

Zhong X, Sun A, Cambria E (2017) Time expression analysis and recognition using syntactic token types and general heuristic rules. In: ACL, pp 420–429

Zhong X, Kaalia R, Rajapakse JC (2019) Go2vec: transforming go terms and proteins to vector representations via graph embeddings. BMC Genomics 20(9):1–10

Zhong X, Cambria E, Hussain A (2020) Extracting time expressions and named entities with constituent-based tagging schemes. Cogn Comput 12(4):844–862

Zhong X, Cambria E, Hussain A (2022) Does semantics aid syntax? An empirical study on named entity recognition and classification. Neural Comput Appl 34(11):8373–8384

Zhou R, Li X, He R, Bing L, Cambria E, Si L, Miao C (2022) MELM: data augmentation with masked entity language modeling for low-resource NER. In: ACL, pp 2251–2262