WikiWars-UA: Ukrainian corpus annotated with temporal expressions

Natalia Grabar¹, Thierry Hamon^{2,3}

¹ CNRS, Univ. Lille, UMR 8163 - STL - Savoirs Textes Langage, F-59000 Lille, France;
² LIMSI, CNRS, Université Paris-Saclay, F-91405 Orsay, France;

³ Université Paris 13, Sorbonne Paris Cité, F-93430 Villetaneuse, France natalia.grabar@univ-lille.fr, hamon@limsi.fr

Abstract. Reliability of tools and reproducibility of study results are important features of modern Natural Language Processing (NLP) tools and methods. The scientific research is indeed increasingly coming under criticism for the lack of reproducibility of results. First step towards the reproducibility is related to the availability of freely usable tools and corpora. In our work, we are interested in automatic processing of unstructured documents for the extraction of temporal information. Our main objective is to create reference annotated corpus with temporal information related to dates (absolute and relative), periods, time, etc. in Ukrainian, and to their normalization. The approach relies on the adaptation of existing application, automatic pre-annotation of WikiWars corpus in Ukrainian and its manual correction. The reference corpus permits to reliably evaluate the current version of the automatic temporal annotator and to prepare future work on this topics. The corpus is freely available for the research on https://github.com/thhamon/WikiWarsUA

Keywords: Temporality · Information Extraction · Ukrainian · Wiki-Wars · HeidelTime · Reference Corpus

1 Introduction

Unstructured documents are the most common source of information, and they may represent the majority of information available in different sources and domains. Yet, the work on unstructured narrative texts is very demanding on automatic methods for detecting, extracting, formalizing and organizing information contained in these documents. If information extraction (IE), which is part of Natural Language Processing (NLP), proposes such methods and aims at detecting and extracting relevant pieces of information from textual data, the question on availability of corpora, resources and reference data is very important. Indeed, such data are crucial for designing, testing and evaluating the automatic methods. Another important issue is related to the reliability of tools and to the reproducibility of study results across similar data from different sources. The scientific research is indeed increasingly coming under criticism for the lack of

reproducibility of results [5, 7, 6]. First step towards the reproducibility of results is the availability of freely usable tools and corpora.

In our work, we focus on detection and extraction of temporal information, such as it occurs in these sentences:

- Корабель Аполлон-11 стартував <u>16 липня 1969</u> о <u>13 годині 32 хви- лини</u> за Грінвічем. (The Apollo-11 ship took off at 1:32 pm GMT on 7/16/1969.)
- Протягом <u>трьох годин</u>, поки налагоджували зв'язок із Москвою, Гагарін давав інтерв'ю і фотографувався. (<u>During three hours</u>, while establishing communication with Moscow, Gagarin was interviewed and photographed.)
- Корейська війна збройний конфлікт між Корейською Народно-Демократичною Республікою та Південною Кореею, який тривав з <u>25 червня</u> <u>1950 року</u> до <u>27</u> липня <u>1953 р</u>. (Korean war is an armed conflict between Democratic People's Republic of Korea and South Korea, which lasted from <u>25th of June</u> <u>1950</u> up to 27th of July 1953.)
- В екваторіальному та тропічному поясі припливи і відпливи здебільшого повторюються <u>двічі на добу</u>. (In the equatorial and tropical areas, high and low tides mostly occur twice a day.)
- Тривали <u>118 років</u>, з примиренням. (Lasted for <u>118 years</u>, including armistices.)
- До <u>середини 260-х до н. е.</u> Римська республіка остаточно підпорядкувала собі Апеннінський півострів. (By <u>the mid of 260 BC</u>, the Roman Republic had gained control of the Italian peninsula.)
- Основним джерелом з історії греко-перських воєн є «Історія» Геродота, що містить опис подій до <u>478</u> до н. е. включно. ("The Histories" by Herodotus, which contains description of events <u>up to 478 BC</u>, is the main source on history of the Greco-Persian Wars.)

Temporal information is important for several tasks and areas, as it allows to structure the entities and events according to their chronological occurrence. This is important in several situations. For instance, in historical studies, the events are usually ordered and then taught and studied in this order. Temporality has become an important research field and several challenges addressed this task up to now: ACE [1], SemEval [22, 23, 21], I2B2 2012 [20]. Yet, the main work is done on texts written in English. We propose to work with texts written in Ukrainian.

In what follows, we first present some related work (Sec. 2). We then precise our objectives (Sec. 3), introduce the material used (Sec. 4) and the proposed method (Sec. 5). Our results and their discussion are presented in Section 6. Finally, we conclude with some directions for future work (Sec. 7).

2 Related Work

Work on temporal information relies on three important steps when processing unstructured narrative documents: identification of linguistic expressions that are indicative of the temporality and their normalization [22, 4, 17, 11], and modeling and chaining of temporal information and events [2, 13, 14, 20, 9]. Identification of temporal expressions, which corresponds to the first step, provides basic knowledge for further tasks aiming at the processing of the temporality. The existing available automatic systems, such as *HeidelTime* [17] or *SUTIME* [4], exploit rule-based approaches, which makes them adaptable to new data, areas, and languages. Such tools usually encode temporal information with the TimeML standard.

TimeML¹ [14] is an annotation standard for temporal expressions proposed in 2010. Since then, it has became the reference for encoding temporal information in different languages. For instance, it has been used in several contexts: for encoding temporal data in challenge corpora such as TempEval [23, 21, 3] and I2B2 [20], for preparing corpora² annotated with temporal expressions such as TimeBank, TempEval, I2B2 and Clinical TempEval corpora.

TimeML offers the possibility to encode several types of temporal information and expressions (i.e. TIMEX3 tags):

- 1. Expressions of dates, time, durations or sets (attribute types). Dates and time are represented according to the ISO-8601 norm;
- 2. ISO-normalized forms of the expressions (attribute value), such as in (from examples above):
 - 16 липня 1969 о 13 годині 32 хвилини ⇒ 1969-07-16Т13:32:00
 - *трьох годин* \Rightarrow РЗН
 - -двічі на добу \Rightarrow P1D
- Quantity and frequency of the set expressions (attributes quant or freq), such as in this expression of frequency:

 $- \partial вічі на \partial oбу \Rightarrow 2X$

- 4. Begin and end anchors for durations (beginpoint and endpoint attributes). For instance, in example Корейська війна - збройний конфлікт між Корейською Народно-Демократичною Республікою та Південною Кореєю, який тривав з 25 червня 1950 року до 27 липня 1953 р. (Korean war is an armed conflict between Democratic People's Republic of Korea and South Korea, which lasted from 25th of June 1950 up to 27th of July 1953.), the begin anchor is 25th of June 1950 and the end anchor is 27th of July 1953. The implicit duration is 3 years, 1 month and 2 days, which is normalized in P3Y1M2D.
- 5. Temporal modifiers, which have been introduced in order to annotate changed or clarified temporal expressions. For instance, in example До <u>середини 260-х</u> <u>до н. е.</u> Римсъка республіка остаточно підпорядкувала собі Апеннінський півострів. (By <u>the mid of 260 BC</u>, the Roman Republic had gained control of the Italian peninsula.), the date 260 до н. е. is changed by середини, which is the date modifier attribute MID.

In addition to the annotation of temporal expressions, TimeML also allows to describe events as well as relations between temporal expressions and/or events.

¹ http://www.timeml.org

² http://timexportal.wikidot.com/

In this paper, we only focus on the annotation of temporal expressions (TIMEX3) related to dates, durations and time.

There is quite few available corpora with temporal annotations. In addition to corpora mentioned above and created as part of challenges, there is the WikiWars corpus³ [12] which provides a collection of texts issued from Wikipedia articles. These texts describe the course of the most famous wars in history, including the biggest wars that happened in the 20th century. The corpus contains 22 articles (such as WW1, WW2, Vietnamese war, Russo-Japanese war, or Punic wars). The main interest in working with these Wikipedia articles is that they contain several dates, as they are typically associated with battles, meetings, armistices, etc. The initial project contains articles in English. It has been extended to three other languages (German, Vietnamese and Croatian) [16,10]. Hence, another interest in working with this corpus is that is contains comparable information and data in several languages.

3 Objectives

The purpose of our work is to build reference corpus in Ukrainian language annotated with temporal information. Temporal information is detected and normalized in the ISO format with respect of the TIMEX3 norm. By comparison with the previous work, which focused on the adaptation of *HeidelTime* to Ukrainian [8], we currently propose the manually verified reference corpus annotated with temporal information. The corpus is freely available for the research⁴. The availability of this corpus may permit a more thorough evaluation of the automatically obtained results.

4 Material

Encyclopedic articles are obtained from the Wikipedia resource in Ukrainian⁵, which is a free and collaborative resource. This encyclopedia contains information on a great variety of topics. For our work, we created the WikiWars corpus [12] in Ukrainian, which contains Wikipedia articles describing the most famous wars in history, including the biggest wars of the 20th century. Since the articles related to the Riffian War, the Sudanese Civil War and the Chaco War are available and contain enough temporal expressions, we also included them in the corpus. Overall, the corpus contains 25 articles (such as WW1, WW2, Vietnamese war, Russo-Japanese war, or Punic wars), and 66,479 word occurrences. The articles have been collected similarly to the building of the original WikiWars corpus [12].

³ http://timexportal.wikidot.com/wikiwars

⁴ https://github.com/thhamon/WikiWarsUA

⁵ https://uk.wikipedia.org

5 Methods

The methods are composed of three main steps: pre-annotation of texts with *Hei-delTime* adapted to Ukrainian, manual correction, and evaluation of the current version of automatic annotations obtained with *HeidelTime*. We also propose a comparison with the English version of the annotations.

5.1 Pre-annotation

For the pre-annotation, we use the *HeidelTime* application. During a preliminary study, *HeidelTime* [17] has been extended to over 200 other languages [18] using existing multilingual resources such as Wiktionary ⁶, which provides data for 170 languages. The test of this version provided no results for Ukrainian (lexical ambiguity and polysemy, missing translations, Wiktionary resources not suitable for the purpose...). Hence, *HeidelTime* was first adapted to the Ukrainian language [8].

HeidelTime is a cross-domain temporal tagger that extracts temporal expressions from documents and normalizes them according to the TIMEX3 annotation standard, which is part of the markup language TimeML [14]. This is a rule-based system. Because the source code and the resources (patterns, normalization information, and rules) are strictly separated, it is possible to develop and implement resources for additional languages and areas using *HeidelTime* rule syntax. Three kinds of resources have been improved:

- linguistic patterns, which describe linguistic elements of the temporality (days of the week, months, numbers, etc.). This type of resources is used for the detection of temporality in texts;
- normalization resources, which are created to permit the normalization of the detected elements. In this way, all the detected units are normalized. Thanks to these resources, normalization can be performed for absolute (Example (1)) and relative (Example (2)) dates, durations and sets. Thus, the normalized values of Examples (1) and (2) are 2015-05-07 and 2017-05-09, respectively if we consider that these two dates are related;
- rules for composing more sophisticated detection of temporality, such as periods, intervals and specific expressions.
- (1) 7 травня 2015 року. (May 7th, 2015.)
- (2) Через два дні. (Two days later.)

HeidelTime was adapted to Ukrainian language on newspaper articles from $\mathcal{Y}\kappa pa\"incoka$ npas ∂a^7 . Then, it was used for the automatic annotation of the WikiWars corpus, which provided 2,226 annotations. Because of the diversity of Wikipedia articles and topics, the detection of temporal information must cover a great variety of dates and formats, several of which do not occur in modern newspaper articles, on which the system in Ukrainian has been developed.

⁶ https://www.wiktionary.org/

⁷ https://www.pravda.com.ua/news/

5.2 Manual correction of annotations

le <u>E</u> dit F <u>o</u> r	rmat <u>T</u> ools	<u>H</u> elp								
	имом і Карфа	геном за п	ануванн	я над Захід	ним С				, Третя Пунічна віі огою Риму і знищ	
	о зіткнення :	з Карфаген	юм — мо	огутньою д					льша експансія Рі имувала велику ча	
ерша Пунічна	війна (<mark>264</mark> - <mark>2</mark>	41 до н. е.								
ого, як цар су рийняли рішен ротидію карф евдало. У 263 айважливішу с цо безкарно сі	усідніх Сираку ння про вклю рагенян, зайн дон.е. рим. фортецю Ка пустошував (уз Гієрон II і чення Месс няла місто. ляни розби рфагена в береги Італ	точав вій ани в Іт У відпов ли Гієроя Сицилії; ії, римля	йну з мамер алійський с відь Карфаг на II і змуси карфагенц ани побудув	отинця союз; геног лийог цібули залидя	ями, ті звернул навесні 264 до олосив війну Рі о вступити з н витиснені в за о 260 до н. е. 1	ися в <mark>265 до н. е</mark> н. е. римська арм иму. Сиракузці ра ними в союз. У <mark>26</mark> ахідну частину ос	по допомогу иія переправи зом з карфаго 2 до н. е. вони трова. Щоб вг блів. У 260 до	кідному узбережж до Риму. Римські пася до Сицилії і, і нянами обложилі і узяли Акрагант (поратися з флото <mark>н. е,</mark> флот Карфа	народні збори незважаючи на и Мессану, але (Агригент), им Карфагена,
о Африки. У <mark>2</mark> сля кількох не жкими, і Кар	2 <mark>56 до н. е.</mark> , з евдач карфаі фаген, мобіл	авдавши п генці зверн ізувавши в	оразки ф кулися до сі ресурс	рлоту Карф римського и, зібрав во	фагена о кома елике	а біля миса Екн андувача Атілія наймане війсы	юм, вони висадил і Регула з прохані ко під командуваї	ися в Клупейс ням про мир, а нням спартан	вирішили перенес ькій бухті (на схід але римські умови ця Ксантіппа. Вео	від Карфагена) виявилися дуже сною 255 до н. е
сантіпп нагол астка загинул			експеди	иційну арміі	ю. Хоч	а римський фл	от взяв гору над	ескадрою Ка	офагена біля м. Ге	ерм, велика його
							_		ю фортецею Кар	
а північно-захі	ідному побер	ежжі Сици					е, в наступному			
арфагена — Ј кертвою штор 247 до н. е. в абіги на бере	Лілібей і Дрег рму. зійська Карф: эги Італії. Сит	пану. Але сг агена на Сі зуація зміни	проба уз нцилії оч лася тіл	яти Лілібей олив енерг ьки тоді, ко	иляни г 1 в <mark>249</mark> - ійний оли ри	поступово підп до н. е. прова Гамількар Барі мляни великок	юрядкували собі і алилася, а в <mark>248 д</mark> ка. Він розвернув о напругою сил (в	асю Сицилію і он.е. римськ активні дії пр ведення надз	блокували дві ост ий флот в чергов оти римлян, здійс вичайного подат	анні твердині ий раз став :нюючи постійні ку) побудували
арфагена — Ј кертвою штор абіги на бере овий флот. У І репани, Карф обов'язавшиск ахідного Сере руга Пунічна в	Лілібей і Дрег эму. зійська Карф ги Італії. Сит березні 241 фаген був ви ь сплатити в едземномор'я війна (219–20	пану. Але сг агена на Си уація зміни до н. е. цей мушений піт іелику конт я.)1 до н. е.).	іроба уз іцилії оч лася тіл і флот р ги на уки рибуцію.	яти Лілібей олив енерг ьки тоді, ки озбив еска падення ми . В результа	иляни г 1 в 249 оли ри адру Ка ару, по аті Пе	поступово підп до н. ел прова Гамількар Барі мляни великок арфагена побл ступившись на ршої Пунічної в	юрядкували собі алилася, а в 248 д ка. Він розвернув о напругою сил (в лизу Егатських ос а користь Риму св	асю Сицилію і он.е. римськ активні дії пр ведення надз тровів. Розум оїми сицилійс	блокували дві ост ий флот в чергов оти римлян, здійс	анні твердині ий раз став нюючи постійні ку) побудували падіння Лілібея і и і
арфагена — Ј ертвоко штор 247.дон.е. в ббіги на бере звий флот. У репани, Карф Бов'язавшисі ахідного Сере руга Пунічна в 9 Events с	Лілібей і дрег ому. зійська Карф ти Італії. Сит барезні 241 фаген був вин ь сплатите був війна (218–20 9 Signals	пану. Але ст агена на Си уація зміни до н. е. цей мушений пі елику конт я. О1 до н. е.). Ф Timex3	проба уз ицилії оч лася тіл і флот р ги на уки рибуцію.	яти Лілібей оливенерг ьки тоді, к озбивеска адення ми . В результ пks результ	иляни́г 1 в 249 ійний оли ри адру Ка адру Ка ару, по аті Пе iLinks	поступово підп до нь ез прова Гамількар Барі мляни великох арфагена побл ступившись на ршої Пунічної в Ф TLinks	орядкували собі і лилася, а в 240 д ка. Він розвернув о напругою сил (в низу Егатських ос користь Риму св ійни Римська рес	асю Сицилію і о н.е. римськ активні дії пр ведення нада тровів. Розум оїми сицилійс публіка перет	блокували дві ост ий флот в чергов оти римлян, здійс вичайного поданного подиникучість кими володінням ворилася на силь	анні твердині ий раз став нюючи постійні ку) побудували падіння Лілібея I и I ну державу
арфагена — Ј ертвою штор 247. до н. е.) в бабги на бере эвий флот. У репани, Карф бов'язавшиса жидного Сере руга Пунічна в 9 Events с ТІD	Лілібей і дрег ому. зійська Карф. зійська Карф. ти Італії. Сит березні 241 фаген був вик ь сплатити в даземномор'я війна (218–20 Війна (218–20 Війна (218–20 Війна (218–20) Війна (218–20)	пану. Але ст агена на Си уація зміни до н. е. цей мушений пі елику конт я. О1 до н. е.). Ф Timex3	іроба уз іцилії оч лася тіл і флот р ги на уки рибуцію.	яти Лілібей олив енерг ьки тоді, ки озбив еска падення ми . В результа	иляни́г 1 в 249 ійний оли ри адру Ка адру Ка ару, по аті Пе iLinks	поступово підп до н. е прова Гамількар Барі мляни великок арфагена поб. ступившись на ршої Пунічної в Ф TLinks Value	юрядкували собі алилася, а в 248 д ка. Він розвернув о напругою сил (в лизу Егатських ос а користь Риму св	асю Сицилію і он.е. римськ активні дії пр ведення надз тровів. Розум оїми сицилійс	блокували дві ост ий флот в чергов оти римлян, здійс вичайного подат ючи неминучість ькими володінням	анні твердині ий раз став нюючи постійні ку) побудували падіння Лілібея і и і
арфагена — Ј ертвою штору 247 до н. е. в ббіги на бере заий флот. У Ј оепани, Карф бов'язавшисц жідного Сере оуга Пунічна в ру Events с TID 50) 264	Липібей і Дрег ому. зійська Карф: ти Італії. Сит березні 241 фаген був вий ь сплатити в едземномор'я війна (218–20 війна (218–20 війна (218–20 війна (218–20 війна (218–20 війна (218–20 війна (218–20) війна (218–20) війна (218–20) війна (218–20) війна (218–20) війна (218–20) війна (218–20) війна (218–20) війна (218–20)	пану. Але ст агена на Си уація зміни до н. е. цей мушений пі елику конт я. О1 до н. е.). Ф Timex3	проба уз ицилії оч лася тіл і флот р ги на уки рибуцію.	яти Лілібей оливенерг ьки тоді, к озбивеска адення ми . В результ пks результ	иляни́г 1 в 249 ійний оли ри адру Ка адру Ка ару, по аті Пе iLinks	поступово підп до нь ез прова Гамількар Барі мляни великох арфагена побл ступившись на ршої Пунічної в Ф TLinks	орядкували собі і лилася, а в 240 д ка. Він розвернув о напругою сил (в низу Егатських ос користь Риму св ійни Римська рес	асю Сицилію і о н.е. римськ активні дії пр ведення нада тровів. Розум оїми сицилійс публіка перет	блокували дві ост ий флот в чергов оти римлян, здійс вичайного поданного подиникучість кими володінням ворилася на силь	анні твердині ий раз став нюючи постійні ку) побудували падіння Лілібея I и I ну державу
арфагена — J ертвою штор забіги на бере звий флот. у I репани, Карф обов'язавшисса ахідного Сере руга Пунічна в <u>9 Events с б</u> <u>TID</u> 60) 264 60) 264 61) 241 рг	Липібей і Дрег ому. зійська Карф: ти Італії. Сит березні 241 фаген був вий ь сплатити в едземномор'я війна (218–20 війна (218–20 війна (218–20 війна (218–20 війна (218–20 війна (218–20 війна (218–20) війна (218–20) війна (218–20) війна (218–20) війна (218–20) війна (218–20) війна (218–20) війна (218–20) війна (218–20)	пану. Але ст агена на Си уація зміни до н. е. цей мушений пі елику конт я. О1 до н. е.). Ф Timex3	проба уз ицилії оч лася тіл і флот р ги на уки рибуцію.	яти Лілібей оливенерг ьки тоді, к озбивеска адення ми . В результ пks результ	иляни́г 1 в 249 ійний оли ри адру Ка адру Ка ару, по аті Пе iLinks	тоступово підп до н. с. прова Гамількар Барі мляни великою арфагена побл ступившись на ршої Пунічної в Гар TLinks Value ВС0264	орядкували собі і лилася, а в 240 д ка. Він розвернув о напругою сил (в низу Егатських ос користь Риму св ійни Римська рес	асю Сицилію і о н.е. римськ активні дії пр ведення нада тровів. Розум оїми сицилійс публіка перет	блокували дві ост ий флот в чергов оти римлян, здійс вичайного поданного подиникучість кими володінням ворилася на силь	анні твердині ий раз став нюючи постійні ку) побудували падіння Лілібея I и I ну державу
247 до н. е. в ертвою штору 247 до н. е. в абіги на бере звий флот. У репани, Кар репани, Кар рога Пунічна в 9 Events / d TID 60) 264 61) 241 pp	Липібей і Дрег ому. зійська Карф: зійська Карф: сти Італії. Сит березні 243. фаген був вий ь сллатити в едземномор'я війна (216–20 війна (216–20 війна (216–21) війна (216	пану. Але ст агена на Си уація зміни до н. е. цей мушений пі елику конт я. О1 до н. е.). Ф Timex3	проба уз ицилії оч лася тіл і флот р ги на уки рибуцію.	яти Лілібей оливенерг ьки тоді, к озбивеска адення ми . В результ пks результ	иляни́г 1 в 249 ійний оли ри адру Ка адру Ка ару, по аті Пе iLinks	тоступово підп до н. е. прова Гамількар Барі мляни великою ступившись на ршої Пунічної в Value BC0264 BC0241	орядкували собі і лилася, а в 240 д ка. Він розвернув о напругою сил (в низу Егатських ос користь Риму св ійни Римська рес	асю Сицилію і о н.е. римськ активні дії пр ведення нада тровів. Розум оїми сицилійс публіка перет	блокували дві ост ий флот в чергов оти римлян, здійс вичайного поданного подиникучість кими володінням ворилася на силь	анні твердині ий раз став нюючи постійні ку) побудували падіння Лілібея I и I ну державу
арфагена — Ј ертвою штор абіги на бере езвий флот. У репани, Карф бов'язавшист оруга Пунічна в 9 Events 0 7 ID 50) 264 51) 241 pp 52) 218 53) 201 pp 53) 201 pp	Лілібей ї Дрег ому. війська Карф. еги Італії. Сит Баревліг 248 фаген бувар Барен бувар Барен бувар війна (218–20 9 Signals) 7 Уре DATE DATE DATE DATE DATE DATE DATE	пану. Але ст агена на Си уація зміни до н. е. цей мушений пі елику конт я. О1 до н. е.). Ф Timex3	проба уз ицилії оч лася тіл і флот р ги на уки рибуцію.	яти Лілібей оливенерг ьки тоді, к озбивеска адення ми . В результ пks результ	иляни́г 1 в 249 ійний оли ри адру Ка адру Ка ару, по аті Пе iLinks	тоступово підп до н. сі прова Гамількар Барі шляни великой арфагена побл ступившись ступившись ршої Пунічної в Value ВСО264 ВСО264 ВСО264 ВСО218 ВСО218 ВСО218	орядкували собі і лилася, а в 240 д ка. Він розвернув о напругою сил (в низу Егатських ос користь Риму св ійни Римська рес	асю Сицилію і о н.е. римськ активні дії пр ведення нада тровів. Розум оїми сицилійс публіка перет	блокували дві ост ий флот в чергов оти римлян, здійс вичайного поданного подиникучість кими володінням ворилася на силь	анні твердині ий раз став нюючи постійні ку) побудували падіння Лілібея I и I ну державу
арфагена — J ертвою штор 247. до н. еl в абіги на бере звий флот. У репани, Карф бов'язавшисс ажідного Сере руга Пунічна в 9 Events 2 6 01 264 601 264 601 241 pp 62) 218 633 201 pp 64) 149 65) 146 pp	Ліпібей і Дрег ому. зійська Карфі- ти талії. Сит (горезні 241) даген бувешкі 241) даген бувешкі 241 едземномор'я війна (212-20 війна (212-20) війна (212-20) с Signals (Туре ДАТЕ ДАТЕ ДАТЕ ДАТЕ ДАТЕ ДАТЕ	пану. Але ст агена на Си уація зміни до н. е. цей мушений пі елику конт я. О1 до н. е.). Ф Timex3	проба уз ицилії оч лася тіл і флот р ги на уки рибуцію.	яти Лілібей оливенерг ьки тоді, к озбивеска адення ми . В результ пks результ	иляни́г 1 в 249 ійний оли ри адру Ка адру Ка ару, по аті Пе iLinks	тоступово підп 2 о н. сі прова Гамількар Барі мляни великою арфагена побл ступившись на ршої Пунічної є 9 ТLinks 19 ТLinks 10 TLinks 10 TLinks	орядкували собі і лилася, а в 240 д ка. Він розвернув о напругою сил (в низу Егатських ос користь Риму св ійни Римська рес	асю Сицилію і о н.е. римськ активні дії пр ведення нада тровів. Розум оїми сицилійс публіка перет	блокували дві ост ий флот в чергов оти римлян, здійс вичайного поданного подиникучість кими володінням ворилася на силь	анні твердині ий раз став нюючи постійні ку) побудували падіння Лілібея I и I ну державу
арфаѓена — Јербалена — Јербалена — Јербалена — Јертвою штор ертвою штор Јевбиги на бере јевий флот. У ј релани, Карборелани, Карборелани, Карборелани, Карборелани, Карборелани, Карбореланиј 9 Events d TID 60) 264 63) 2018 рг. 63) 2018 рг. 63) 2018 рг. 63) 2018 рг. 63) 2018 рг. 63) 2018 рг.	Лілібей і Дрегорму. мійскак Карфінгік Карфінги Галії ситі алаген буз війскак Карфінги Салії ситі алаген буз війскак Салії ситі салаген Салії ситі салаген Салії ситі війна (200-20 війна (200-20) війна (200-20) вій	пану. Але сг агена на Си уація зміни до н. е. цей мушений пі лелику конт я. О1 до н. е.). Ф Timex3	проба уз ицилії оч лася тіл і флот р ги на уки рибуцію.	яти Лілібей оливенерг ьки тоді, к озбивеска адення ми . В результ пks результ	иляни́г 1 в 249 ійний оли ри адру Ка адру Ка ару, по аті Пе iLinks	тоступово підп 2 о н. с. прова Гамількар Барі імляни великох арфагена побу арфагена побу аршись на ступившись на ршої Пунічної в Value ВС0241 ВС0241 ВС0241 ВС0221 ВС02421 ВС0244 ВС0241 ВС0244 ВС0241 ВС0244 ВС0241 ВС0244 ВС0241 ВС0244 ВС0244 ВС0244 ВС0244 ВС0244 ВС0244 ВС0244 ВС0244 ВС0244 ВС0244 ВС0	орядкували собі і лилася, а в 240 д ка. Він розвернув о напругою сил (в низу Егатських ос користь Риму св ійни Римська рес	асю Сицилію і о н. с. римськ активні дії пр ведення нада тровів. Розум оіми сицилійс публіка перет	блокували дві ост ий флот в чергов оти римлян, здійс вичайного поданного подиникучість кими володінням ворилася на силь	анні твердині ий раз став нюючи постійні ку) побудували падіння Лілібея I и I ну державу
арфагена — / арфагена — / арфагена — / ертвою штор ертвою штор 26/7 дон и сі в абіги на бере змий флот. У Іро Бов'язавшись, Карф бов'язавшись, Карф бов'язавшись, Карф 50/264 В. 10/241 рр 50/264 В. 2218 В. 30/201 рр 50/264 В. 50/264 В.	Ліпбей і Дрег ому. зійська Карф. Гелазлі сит белазлі сит белазлі сит белазлі сит белазлі сит белазлі сит белазлі сит війна (212-50 війна (212-50) війна (212-50 війна (212-50) війна	пану. Але сг агена на Си уація зміни до н. е. цей мушений пі лелику конт я. О1 до н. е.). Ф Timex3	проба уз ицилії оч лася тіл і флот р ги на уки рибуцію.	яти Лілібей оливенерг ьки тоді, к озбивеска адення ми . В результ пks результ	иляни́г 1 в 249 ійний оли ри адру Ка адру Ка ару, по аті Пе iLinks	тоступово підп 2 о н. сі прова Гамількар Барі мляни великом арфагена побл ступившись на ршої Пунічної є Value ВС0264 ВС0241 ВС02218 ВС02218 ВС02218 ВС02218 ВС02218 ВС02218 ВС02218 ВС02218 ВС02218 ВС0224	орядкували собі і лилася, а в 240 д ка. Він розвернув о напругою сил (в низу Егатських ос користь Риму св ійни Римська рес	асю Сицилію і о н.е. римськ активні дії пр ведення нада тровів. Розум оїми сицилійс публіка перет	блокували дві ост ий флот в чергов оти римлян, здійс вичайного поданного подиникучість кими володінням ворилася на силь	анні твердині ий раз став нюючи постійні ку) побудували падіння Лілібея I и I ну державу
арфаѓена — Јерфаѓена — Јерфаѓена — Јертако штор ертако штор 247. гон и е 16 веј вобити на бере вий фоло. тУ Јерба и Карф Беленки, Карф Обов'язавшиск уруга Пунјчна г 9 Events d 10 9 Events d 10 10 241 ре 55) 146 рр 55) 146 рр 56) гердин. 56) гердин.	Лілібей і Дрег ому. зійскак Карф. уги Італії. Сит фален був вим ь сплатити в саразні асій війна (200-20 g Signals) Туре DATE DATE DATE DATE DATE DATE DATE DATE	пану. Але сг агена на Си уація зміни до н. е. цей мушений пі лелику конт я. О1 до н. е.). Ф Timex3	проба уз ицилії оч лася тіл і флот р ги на уки рибуцію.	яти Лілібей оливенерг ьки тоді, к озбивеска адення ми . В результ пks результ	иляни́г 1 в 249 ійний оли ри адру Ка адру Ка ару, по аті Пе iLinks	тоступово підп 2016 г. прова Гамількар Барі Іміляни великох арфагена побл ступившись на ршої Пунічної в Уаіце ВС0241 ВС0244 ВС0241 ВС0241 ВС0244 ВС024	орядкували собі і лилася, а в 240 д ка. Він розвернув о напругою сил (в низу Егатських ос користь Риму св ійни Римська рес	асю Сицилію і о н. с. римськ активні дії пр ведення нада тровів. Розум оіми сицилійс публіка перет	блокували дві ост ий флот в чергов оти римлян, здійс вичайного поданного подиникучість кими володінням ворилася на силь	анні твердині ий раз став нюючи постійні ку) побудували падіння Лілібея I и I ну державу
арфагена — / ертвою штор ертвою штор 247 дон н. ей в вбиги на бере зюй флот. У Чренани, Карф бов'язавшись, Карф бов'язавшись, Карф бов'язавшись, Карф 500 264 301 261 др 540 149 р 540 149 р 541 149 р 551 146 рр 551 146 рр 551 146 рр 551 146 рр 571 264 33 241 до 4 372 64	Ліпбей і Дрег ому. ўйскак Карф. раген був вик валазії. Сит блоззії. 421 валазії. Сит блоззії. 421 валазії. Сит валазії. Сит валазії. Сит валазії. Сит валазії. Сит валазії. сита тити в валазії. сита тити в сита тити в сити в сит	пану. Але сг агена на Си уація зміни до н. е. цей мушений пі лелику конт я. О1 до н. е.). Ф Timex3	проба уз ицилії оч лася тіл і флот р ги на уки рибуцію.	яти Лілібей оливенерг ьки тоді, к озбивеска адення ми . В результ пks результ	иляни́г 1 в 249 ійний оли ри адру Ка адру Ка ару, по аті Пе iLinks	тоступово підп до н. сі прова Гамількар Барі міляни великок арфагена побл ступлившись на ршої Пунічної є Value ВСО264 ВСО218 ВСО201 ВСО2018 ВСО218 ВСО218 ВСО214 ВСО218 ВСО218 ВСО214 ВСО218 ВСО218 ВСО264 ВСО264 ВСО264	орядкували собі і лилася, а в 240 д ка. Він розвернув о напругою сил (в низу Егатських ос користь Риму св ійни Римська рес	асю Сицилію і о н. с. римськ активні дії пр ведення нада тровів. Розум оіми сицилійс публіка перет	блокували дві ост ий флот в чергов оти римлян, здійс вичайного поданного подиникучість кими володінням ворилася на силь	анні твердині ий раз став нюючи постійні ку) побудували падіння Лілібея I и I ну державу
арфаѓена — Јерфаѓена — Јерфаѓена — Јертвою штор сертвою штор 247 дон и сеј в биги на берее звиќ диол ту Урепани, Карф бов'язавшиси мијдного Сере 9 Events је и 7 ID 50) 264 51) 241 рр 52) 218 53) 201 рр 53) 241 рр 53) 241 рр 53) 241 рр 53) 241 рр 53) 241 рр 53) 241 рр 53) 242 рр 53) 1240 рр 53) 1241 рр 53) 1240 рр 54) 1240 рр 55) 1240 рр 55) 1240 рр 55) 1240 рр 56) 1240 рр 57)	Лілібей і Дрег ому. зійскак Карф. ги Гталії. Сит борезні 421 фаген був вик ь сплатити в в слаятити в в сла	пану. Але сг агена на Си уація зміни до н. е. цей мушений пі лелику конт я. О1 до н. е.). Ф Timex3	проба уз ицилії оч лася тіл і флот р ги на уки рибуцію.	яти Лілібей оливенерг ьки тоді, к озбивеска адення ми . В результ пks результ	иляни́г 1 в 249 ійний оли ри адру Ка адру Ка ару, по аті Пе iLinks	тоступово підп 2 о н. с. прова Гамількар Барі шляни великох арфагена побл ступившись на ршої Пунічної в Содана ВСО241 ВСО241 ВСО241 ВСО241 ВСО241 ВСО241 ВСО241 ВСО241 ВСО241 ВСО241 ВСО241 ВСО241 ВСО241 ВСО241	орядкували собі і лилася, а в 240 д ка. Він розвернув о напругою сил (в низу Егатських ос користь Риму св ійни Римська рес	асю Сицилію і о н. с. римськ активні дії пр ведення нада тровів. Розум оіми сицилійс публіка перет	блокували дві ост ий флот в чергов оти римлян, здійс вичайного поданного подиникучість кими володінням ворилася на силь	анні твердині ий раз став нюючи постійні ку) побудували падіння Лілібея I и I ну державу
apdparenta — Jepdarenta — Jepdarenta — Jepdarenta — Jeptasou urtop ceptasou urtop ceptasou urtop solit A depension solit A	Ліпбей і Дрег ому. «) () () () () () () () () () () () () ()	пану. Але сг агена на Си уація зміни до н. е. цей мушений пі- телику конт я. О1 до н. е.). Ф Timex3	проба уз ицилії оч лася тіл і флот р ги на уки рибуцію.	яти Лілібей оливенерг ьки тоді, к озбивеска адення ми . В результ пks результ	иляни́г 1 в 249 ійний оли ри адру Ка адру Ка ару, по аті Пе iLinks	тоступово підп до н. сі прова Гамількар Барі шляни великой арфагена побл ступившись на ршої Пунічної в Value ВСО264 ВСО264 ВСО218 ВСО218 ВСО218 ВСО218 ВСО218 ВСО2146 Р1189 ВСО264 ВСО264 ВСО264 ВСО264 ВСО284 РСО284 РСО284 РСО284 РСО284	орядкували собі і лилася, а в 240 д ка. Він розвернув о напругою сил (в низу Егатських ос користь Риму св ійни Римська рес	асю Сицилію і о н. с. римськ активні дії пр ведення нада тровів. Розум оіми сицилійс публіка перет	блокували дві ост ий флот в чергов оти римлян, здійс вичайного поданного подиникучість кими володінням ворилася на силь	анні твердині ий раз став нюючи постійні ку) побудували падіння Лілібея I и I ну державу
apфařenia — Jepčařenia — Jepčař	Лілібей і Дрег ому. зійскак Карф. ги Італії. Сит (отразні 421) раген був вик ведземномор'я війна (212-52) війна (212-52) війна (212-52) отте рате	пану. Але сг агена на Си уація зміни до н. е. цей мушений пі- телику конт я. О1 до н. е.). Ф Timex3	проба уз ицилії оч лася тіл і флот р ги на уки рибуцію.	яти Лілібей оливенерг ьки тоді, к озбивеска адення ми . В результ пks результ	иляни́г 1 в 249 ійний оли ри адру Ка адру Ка ару, по аті Пе iLinks	тоступово підп 2 о н. с. прова Гамількар Барі шляни великох арфагена побл ступившись на ршої Пунічної в Содана ВСО241 ВСО241 ВСО241 ВСО241 ВСО241 ВСО241 ВСО241 ВСО241 ВСО241 ВСО241 ВСО241 ВСО241 ВСО241 ВСО241	орядкували собі і лилася, а в 240 д ка. Він розвернув о напругою сил (в низу Егатських ос користь Риму св ійни Римська рес	асю Сицилію і о н. с. римськ активні дії пр ведення нада тровів. Розум оіми сицилійс публіка перет	блокували дві ост ий флот в чергов оти римлян, здійс вичайного поданного подиникучість кими володінням ворилася на силь	анні твердині ий раз став нюючи постійні ку) побудували падіння Лілібея I и I ну державу
арфагена — Ј кертвою штор 247 до н. е. в обіги на бере овий флот. У репани, Карф обов'язавшисі ахідного Сере 	Лілібей і Дрег ому. війскак Карф. ги Італії. Сит альген був війскак Карф. війна (200-42) війна	пану. Але сг агена на Си уація зміни до н. е. цей мушений пі- телику конт я. О1 до н. е.). Ф Timex3	проба уз ицилії оч лася тіл і флот р ги на уки рибуцію.	яти Лілібей оливенерг ьки тоді, к озбивеска адення ми . В результ пks результ	иляни́г 1 в 249 ійний оли ри адру Ка адру Ка ару, по аті Пе iLinks	100 тупово підп 2 ло н. сі прова Гамількар Барі мляни великою верфагена побл ступившись на ршої Пунічної є 9 ТLinks 10 ТLinks 10 ТLinks 10 ТLinks 10 ТLinks 10 ТLinks 10 TLinks 10 TLin	орядкували собі і лилася, а в 240 д ка. Він розвернув о напругою сил (в низу Егатських ос користь Риму св ійни Римська рес	асю Сицилію і о н. с. римськ активні дії пр ведення нада тровів. Розум оіми сицилійс публіка перет	блокували дві ост ий флот в чергов оти римлян, здійс вичайного поданного подиникучість кими володінням ворилася на силь	анні твердині ий раз став нюючи постійні ку) побудували падіння Лілібея I и I ну державу
арфагена — J кертвою штор кертвою штор сертвою штор овий флот. У репани, Карф обов'язавшисс видного Сере круга Пунічна и Середна – Сере круга Пунічна и Середна – Сере (со) 264 (со) 2	Ліпбей і Дрег ому. зійскак Карф. ги Італії. Сит балагна буз війскак Карф. ги Італії. Сит балагна буз відагна буз війна (200-42) війна	пану. Але сг агена на Си уація зміни до н. е. цей мушений пі- телику конт я. О1 до н. е.). Ф Timex3	проба уз ицилії оч лася тіл і флот р ги на уки рибуцію.	яти Лілібей оливенерг ьки тоді, к озбивеска адення ми . В результ пks результ	иляни́г 1 в 249 ійний оли ри адру Ка адру Ка ару, по аті Пе iLinks	тоступово підп 2 о н. с. прова Гамількар Барі шляни великой арфагена побу арфагена побу арфагена побу арфагена побу Ступившись на ступившись на ступившись на вслаза всла вслаза	орядкували собі і лилася, а в 240 д ка. Він розвернув о напругою сил (в низу Егатських ос користь Риму св ійни Римська рес	асю Сицилію і о н. с. римськ активні дії пр ведення нада тровів. Розум оіми сицилійс публіка перет	блокували дві ост ий флот в чергов оти римлян, здійс вичайного поданного подиникучість кими володінням ворилася на силь	анні твердині ий раз став нюючи постійні ку) побудували падіння Лілібея и ну державу
арфагена — J кертвою штор сертвою штор 247 до н. е в абіги на бере овий флот. У репани, Карф обов'язавшисс амідного Сере руга Пунічна с то то сб) 264 (61) 241 pp (62) 218 (63) 201 pp (63) 201 pp (63) 201 pp (65) 149 (55) 146 pp (65) 149 (55) 140 pp (65) 264 (3) 241 до н (63) 264 (3) 241 до н (57) 264 (3) 241 до н (57) 265 до н (57) 265 до н (57) 265 до н (57) 263 до н (57) 263 до н	Ліпбей і Дрег ому. зійскак Карф. еги Італії. Сит (блезні -24) ваген був вик вадземномор'ї війна (112-14) ваген був вик вадземномор'ї війна (112-14) ваген був вик вадземномор'ї війна (112-14) вала вала вала вала одате	пану. Але сг агена на Си уація зміни до н. е. цей мушений пі- телику конт я. О1 до н. е.). Ф Timex3	проба уз ицилії оч лася тіл і флот р ги на уки рибуцію.	яти Лілібей оливенерг ьки тоді, к озбивеска адення ми . В результ пks результ	иляни́г 1 в 249 ійний оли ри адру Ка адру Ка ару, по аті Пе iLinks	тоступово підп 2 о н. сі прова Гамількар Бари мляни великор афагена побл ступившись на ршої Пунічної є Уаіце ВС0264 ВС0241 ВС0218 ВС0218 ВС0218 ВС02149 ВС0241 ВС0241 ВС0241 ВС0241 ВС0264 ВС02241 ВС0264 ВС02241 ВС0264 ВС02264	орядкували собі і лилася, а в 240 д ка. Він розвернув о напругою сил (в низу Егатських ос користь Риму св ійни Римська рес	асю Сицилію і о н. с. римськ активні дії пр ведення нада тровів. Розум оіми сицилійс публіка перет	блокували дві ост ий флот в чергов оти римлян, здійс вичайного поданного подиникучість кими володінням ворилася на силь	анні твердині ий раз став нюючи постійні ку) побудували падіння Лілібея I и I ну державу

Fig. 1. Reading and annotation of temporal information with Callisto on example of Punic wars.

The results provided by the automatic annotation of temporality are then corrected and completed manually. We use the Callisto application developed by the MITRE Corp.⁸. As can be seen in Figure 1, this application permits to visualize the existing annotations in the text (upper part) together with the

⁸ https://mitre.github.io/callisto/manual/use.html

detected linguistic units and their normalizations (lower part). Then, each annotation can be modified or deleted, and new annotations can be created. The same operations are available for the normalization information. Modifiers can also be added to the annotations, such as *mid*, *end*, *after*.

5.3 Evaluation

After the manual correction of the annotations and the creation of the reference data, the results from the previously obtained automatic annotation [8] are evaluated against these reference data with classical evaluation measures [15]:

- true positives TP: number of correctly extracted or normalized temporal expressions;
- precision \mathcal{P} : percentage of the relevant temporal expressions extracted and normalized divided by the total number of the temporal expressions extracted and normalized.
- recall \mathcal{R} : percentage of the relevant temporal expressions extracted divided by the number of the expected temporal expressions;
- F-measure \mathcal{F} : the harmonic mean of the precision and recall values $\frac{\mathcal{P}*\mathcal{R}}{\mathcal{P}+\mathcal{P}}$.

The evaluation is done with scripts available from previous work [17]. The evaluation measures are computed with strict and relaxed values, according to whether the boundaries of the temporal expressions are detected correctly (strict boundaries) or not (intersection between reference and automatically extracted linguistic units).

6 Results and Discussion

During the manual correction of the results, we observed two main difficulties in the results of the current version of *HeidelTime* in Ukrainian:

- Ambiguity of *nishiu* meaning both *midnight* and *north*. Currently, every occurrence of this marker is automatically annotated as temporal information, like in this example:

У спробах полегшити тиск з <TIMEX3 type="TIME" value="1967-06-12T24:00">півночі</TIMEX3>, <TIMEX3 type="DATE" value="1967-08-09">9 серпня</TIMEX3> мобільна бригада армії Біафри у складі 3000 осіб за підтримки артилерії та бронемашин переправилася на західний берег Нігера. (Trying to ease the pressure from north, 9th of August mobile group of the Biafra army composed of 3,000 people and supported by artillery and armoured cars crossed the Niger river and reached its western side.)

The disambiguation of this marker may rely on prepositions it subsumes or on additional analysis of texts, as pre-processing or post-processing step. The native tokenization of HeidelTime is unable to tokenize text in several situations, for which reason, several dates have been missed by the automatic annotation. For instance, all the dates have been missed in this example: Пунічні війни — три війни (Перша Пунічна війна 264–241 рр. до н. е., Друга Пунічна війна 218–201 рр. до н. е., Третя Пунічна війна 149–146 рр. до н. е.) між Римом і Карфагеном за панування над Західним Середземномор'ям. (Punic wars - three wars (First Punic war 264-241 BC, Second Punic war 218-201 BC, Third Punic war 149-149 BC) between Rome and Carthage for the for the conquest of the Western Mediterranean.))

Further to the manual correction, the reference annotations amount up to 2,719 temporal units. The automatic extraction provides 2,116 temporal units, among which 2,018 are correct (True Positives). For comparison, in the English version of the WikiWars corpus (19 files), the reference data contain 1,858 temporal expressions.

Table 1 indicates global strict and relaxed values of Precision \mathcal{P} , Recall \mathcal{R} and F-measure \mathcal{F} of the extracted results. As expected, relaxed values are better because they accept intersection between reference data and automatically extracted results. Yet, the strict values are high as well, which means that the system is quite successful in the extraction of temporal linguistic expressions.

Table 1. TIMEX3: strict and relaxed values for Precision, Recall and F-measure.

		\mathcal{R}	
strict match	0.85	0.66	0.75
relaxed match	0.95	0.74	0.83

Table 2 indicates performance of the system for the detection of types of temporal expressions (which may be of three natures: date, duration, time) and for their normalization. As in other works, we can see that evaluation values obtained for the extraction of linguistic temporal expressions are higher than values for their normalization [19]. In several situations, it may indeed be complicated to compute the normalized values, and typically when past of future events are just mentioned in the texts. Improvement of the computing of the normalization values is one of the challenges for future work.

Table 2. TIMEX3: extraction and normalization values (Precision, Recall, F-measure).

	TP	\mathcal{P}	\mathcal{R}	\mathcal{F}
type	$1,\!897$	0.90	0.70	0.80
normalization	$1,\!651$	0.78	0.60	0.69

⁸ N Grabar, T Hamon

7 Conclusion and Future Work

The main purpose of this work is the creation of the reference corpus with annotations of temporal expressions and their normalizations. We proposed to build such corpus in Ukrainian as part of the WikiWars corpora. The pre-annotation is done automatically with the Ukrainian version of *HeidelTime*. The annotations are then verified and completed manually. Overall, on 25 Wikipedia articles, we count 2,719 reference temporal units. This reference corpus permits to make the evaluation of the current version of *HeidelTime*: up to 0.80 F-measure for the detection of temporal expressions and up to 0.69 F-measure for their normalization. This corpus is freely available for the research and accessible on https://github.com/thhamon/WikiWarsUA. For people working on the automatic detection of temporal expressions in Ukrainian. One of the challenges is related to the normalization of the temporal expressions and to their link with the events described.

References

- ACE challenge: The ACE 2004 evaluation plan. evaluation of the recognition of ace entities, ace relations and ace events. Tech. rep., ACE challenge (2004), http://www.itl.nist.gov/iad/mig/tests/ace/2004
- Batal, I., Sacchi, L., Bellazzi, R., Hauskrecht, M.: A temporal abstraction framework for classifying clinical temporal data. In: Ann Symp Am Med Inform Assoc (AMIA). pp. 29–33 (2009)
- Bethard, S., Savova, G., Palmer, M., Pustejovsky, J.: Semeval-2017 task 12: Clinical tempeval. In: Int Workshop on Semantic Evaluation (SemEval-2017). pp. 565–572. Association for Computational Linguistics, Vancouver, Canada (August 2017)
- Chang, A.X., Manning, C.D.: SUTIME: A library for recognizing and normalizing time expressions. In: LREC. pp. 3735–3740 (2012)
- Chapman, W.W., Nadkarni, P.M., Hirschman, L., D'Avolio, L.W., Savova, G.K., Uzuner, O.: Overcoming barriers to nlp for clinical text: the role of shared tasks and the need for additional creative solutions. J Am Med Inform Assoc 18(5), 540–543 (2011)
- Cohen, K.B., Xia, J., Roeder, C., Hunter, L.E.: Reproducibility in natural language processing: A case study of two R libraries for mining PubMed/MEDLINE. In: LREC Int Conf Lang Resour Eval. pp. 6–12 (2016)
- Collins, F., Tabak, L.: Nih plans to enhance reproducibility. Nature 505, 612–613 (2014)
- Grabar, N., Hamon, T.: Automatic detection of temporal information in ukrainian general-language texts. In: COLINS 2018. pp. 1–11 (2018)
- Grouin, C., Grabar, N., Hamon, T., Rosset, S., Tannier, X., Zweigenbaum, P.: Hybrid approaches to represent the clinical patient's timeline. J Am Med Inform Assoc 20(5), 820–7 (2013)
- Jeong, Y.S., Joo, W.T., Do, H.W., Lim, C.G., Choi, K.S., Choi, H.J.: Korean timeml and korean timebank. In: Chair), N.C.C., Choukri, K., Declerck, T., Goggi, S., Grobelnik, M., Maegaard, B., Mariani, J., Mazo, H., Moreno, A., Odijk, J., Piperidis, S. (eds.) Proceedings of the Tenth International Conference on Language

Resources and Evaluation (LREC 2016). European Language Resources Association (ELRA), Paris, France (may 2016)

- Kessler, R., Tannier, X., Hagège, C., Moriceau, V., Bittar, A.: Finding salient dates for building thematic timelines. In: Annual Meeting of the Association for Computational Linguistics. pp. 730–739 (2012)
- Mazur, P., Dale, R.: WikiWars: A new corpus for research on temporal expressions. In: Int Conf on Empirical Methods in Natural Language Processing. pp. 913–922 (2010)
- 13. Moskovitch, R., Shahar, Y.: Medical temporal-knowledge discovery via temporal abstraction. In: Ann Symp Am Med Inform Assoc (AMIA). pp. 452–456 (2009)
- Pustejovsky, J., Lee, K., Bunt, H., Romary, L.: ISO-TimeML: An international standard for semantic annotation. In: Chair), N.C.C., Choukri, K., Maegaard, B., Mariani, J., Odijk, J., Piperidis, S., Rosner, M., Tapias, D. (eds.) Int Conf Language Resources and Evaluation (LREC'10). European Language Resources Association (ELRA), Valletta, Malta (may 2010)
- Sebastiani, F.: Machine learning in automated text categorization. ACM Computing Surveys 34(1), 1–47 (2002)
- Strötgen, J., Gertz, M.: Wikiwarsde: A german corpus of narratives annotated with temporal expressions. In: Conf of the German Society for Comp Linguistics and Language Technology (GSCL 2011). pp. 129–134. Hamburg, Germany (September 2011)
- Strötgen, J., Gertz, M.: Temporal tagging on different domains: Challenges, strategies, and gold standards. In: Int Conf on Language Resources and Evaluation (LREC'12). pp. 3746–3753. ELRA (2012)
- Strötgen, J., Gertz, M.: A baseline temporal tagger for all languages. In: Int Conf on Empirical Methods in Natural Language Processing. pp. 541–547. ACL (2015)
- Strötgen, J., Armiti, A., Canh, T.V., Zell, J., Gertz, M.: Time for more languages: Temporal tagging of Arabic, Italian, Spanish, and Vietnamese. ACM Transactions on Asian Language Information Processing 13(1), 1–21 (2014)
- Sun, W., Rumshisky, A., Uzuner, Ö.: Evaluating temporal relations in clinical text: 2012 i2b2 challenge. JAMIA 20(5), 806–813 (2013)
- UzZaman, N., Llorens, H., Derczynski, L., Allen, J., Verhagen, M., Pustejovsky, J.: Semeval-2013 task 1: Tempeval-3: Evaluating time expressions, events, and temporal relations. In: Int Workshop on Semantic Evaluation (SemEval 2013). pp. 1–9. Atlanta, Georgia, USA (June 2013), http://www.aclweb.org/anthology/ S13-2001
- Verhagen, M., Gaizauskas, R., Schilder, F., Hepple, M., Katz, G., Pustejovsky, J.: Semeval-2007 task 15: Tempeval temporal relation identification. In: Int Workshop on Semantic Evaluations (SemEval-2007). pp. 75-80. Prague, Czech Republic (June 2007), http://www.aclweb.org/anthology/S/S07/S07-1014
- Verhagen, M., Sauri, R., Caselli, T., Pustejovsky, J.: Semeval-2010 task 13: Tempeval-2. In: Int Workshop on Semantic Evaluation. pp. 57–62. Uppsala, Sweden (July 2010), http://www.aclweb.org/anthology/S10-1010