




A survey on narrative extraction from textual data

Brenda Santana^{1,2}  · Ricardo Campos^{2,3} · Evelin Amorim² · Alípio Jorge^{2,4} · Purificação Silvano⁵ · Sérgio Nunes^{2,6}

Published online: 6 January 2023
© The Author(s) 2022

Abstract

Narratives are present in many forms of human expression and can be understood as a fundamental way of communication between people. Computational understanding of the underlying story of a narrative, however, may be a rather complex task for both linguists and computational linguistics. Such task can be approached using natural language processing techniques to automatically extract narratives from texts. In this paper, we present an in depth survey of narrative extraction from text, providing a establishing a basis/framework for the study roadmap to the study of this area as a whole as a means to consolidate a view on this line of research. We aim to fulfill the current gap by identifying important research efforts at the crossroad between linguists and computer scientists. In particular, we highlight the importance and complexity of the annotation process, as a crucial step for the training stage. Next, we detail methods and approaches regarding the identification and extraction of narrative components, their linkage and understanding of likely inherent relationships, before detailing formal narrative representation structures as an intermediate step for visualization and data exploration purposes. We then move into the narrative evaluation task aspects, and conclude this survey by highlighting important open issues under the domain of narratives extraction from texts that are yet to be explored.

Keywords Narrative extraction · Natural language processing · Computational linguistics · Computational narratology

1 Introduction

Recent years have shown a growing interest in the application of natural language processing techniques for extracting, summarizing, and creating new data from text. Today's digital media ecosystem generates massive streams of unstructured data, such as news articles,

✉ Brenda Santana
brenda.s.santana@inesctec.pt

✉ Ricardo Campos
ricardo.campos@inesctec.pt

✉ Alípio Jorge
amjorge@fc.up.pt

Extended author information available on the last page of the article

web pages, blog posts and tweets, thus posing a set of challenges related to their automatic understanding. Much of this content is communicated using natural language containing narratives that refer to stories involving multiple actors and events, occurring in varied locations according to a timeline. The composition of a narrative is given by the entities participating in the story, the events in which they are agents, and the temporal data that defines the sequence of narrated events. In this sense, as stated by Toolan (2013), a narrative can be seen as a perceived sequence of non-randomly connected events, a sequence of interconnected facts that are observed over a period of time involving basic elements such as organizations, persons, locations or time.

Although several attempts have been conducted in the domain of narrative extraction, some problems remain unsolved, which are not easily solved by the currently available Natural Language Processing (NLP) and Information Extraction (IE) techniques. Important challenges in this domain involve defining annotation schemes that are comprehensive enough to include the relevant features of the narrative elements, but, at the same time, not too cumbersome to avoid overloading the process of extraction; automatically extracting cohesive narratives from unstructured data with high effectiveness; detecting and characterizing storylines over single and multiple documents; identifying narrative segments within large portions of text; dealing with informal language and with different types of documents (e.g., social media and micro-blogs); coping with different human languages and with combinations of languages; defining meaningful formal representation of narratives as an intermediate step to visualization schemes; devising a standard evaluating narrative framework (made of datasets, baselines and metrics) as an entry-point for researchers interested on properly evaluating their methods; and contributing with the development of specialized linguistic resources for low-resource languages, such as narrative annotated data, to avoid creating niches.

To help the research community address these challenges, different techniques can be applied depending on the type and scale of the narrative, e.g. single news articles versus whole fiction books. In this paper, we aim to survey existing approaches and techniques related to the process of extracting narratives from a text, an emergent theme of research within the field of computational linguistics and artificial intelligence. To the best of our knowledge, there is currently no survey addressing the full process of narrative extraction from textual data. Hence, collecting, organizing, documenting and describing fundamental concepts of this process is of paramount importance for different related areas, where the extraction of narratives plays an important role. Domains such as journalism (Caselli et al. 2016), finance (El-Haj 2022), health (Sheikhalishahi et al. 2019), information access in digital libraries (Kroll et al. 2022), and other fields that require analyzing narratives over a common topic using NLP approaches, are among the main beneficiaries of this survey. Such communities are very active in the area, which makes it quite hard for a researcher to be aware of all the relevant contributions. It should be noted that in the different contexts where narratives can be extracted, there are common problems; however, there is also terminology, notation and specificities for each one. While in a financial narrative, there is a need for structure extraction (El-Haj et al. 2019) given the format in which the data are generally available, in narratives related to historical events (Lai et al. 2021), there is no such need. In this work, such particularities will not be discussed; instead, we will seek to focus on tasks that are common to narrative extraction from a general point of view.

In an attempt to provide a thorough account of relevant research developed within the field of computational linguistics and artificial intelligence (AI), we conduct a survey where the different approaches are organized according to the pipeline of tasks one can find in the extraction of narratives from text. The development of this study began with the

selection of fundamental articles understood as key to the subareas that make up the extraction of narratives. From these articles, we apply the snowball procedure (Wohlin 2014) to expand the pool of articles by considering relevant research published in high-quality conferences and journals in the fields of natural language processing and artificial intelligence.

The remainder of this paper is organized as follows. Section 2 gives an overview of the process of extracting narratives from a text by introducing the fundamental concepts, the data acquisition and annotation effort, and the narrative extraction pipeline. Section 3 presents the initial step of narrative extraction that comprises the pre-processing and parsing of a text. Section 4 introduces research developed on the identification and extraction of narrative components, namely lexical and syntactical components. Section 5 presents the techniques found in the literature to establish the linkage between the identified narrative components. Section 6 refers to research work focused on the representation of narratives structures. Section 7 provides a snapshot of the metrics and the datasets behind the evaluation efforts. Section 8 promotes a discussion of some important aspects of narratives research including open issues. Finally, Sect. 9 concludes this paper by presenting its final remarks.

2 Computational narrative extraction

In this section, we begin by presenting a detailed definition of the concept of narratives and its associated terminology. Following, we discuss important steps related to data acquisition and annotation. Finally, we describe the general pipeline behind the process of extracting narratives from textual data through the lens of a computational scope. The pipeline here introduced will define the structure of the rest of this paper.

2.1 Narrative definition

Narratives have long been studied in the field of linguistics. One of the first authors to introduce a formal definition was Adam (1992), who considers narratives to be a prototypical sequence obeying a thematic unity, regarding a chronological succession of events involving characters. The events are linked by causality relations, and compose a story with an initial situation, followed by a complication, reactions, resolution and final situation. From a structuralist theory point of view, narratives can be defined, as stated by Chatman (1980), as structures consisting of two parts: (1) a story, the content or chain of events (actions, happenings), plus what may be called the existents (characters, items of setting); and (2) a discourse, that is, the expression, how the content is communicated, for instance, by word (verbal language: oral and written), image (visual language), representation (theatrical language), etc.

The definition of what a narrative is has been dissected by several other authors over the years, reflecting the difficulty in reaching a broad consensus within the community. Riedl (2004) considers narratives as a cognitive tool for situated understanding, i.e., a structure designed to better understand the world around us. Motta (2005), instead, understands narratives as forms of relationships that are established due to culture, the coexistence between living beings that have interests, desires, and that are under the constraints and social conditions of hierarchy and power. As the author states, whoever narrates has some purpose in narrating, therefore no narrative is naive. The same narrative can be seen from a different point of view, i.e., from a focal point for seeing, hearing, smelling, and experiencing the

story's environments, characters, and events in the narrator's way (Al-Alami 2019). The role of characters in narratives has been surveyed to a huge extent by Labatut and Bost (2019).

Broader understandings define narratives as a sequence of events that need to have a "continuant subject and constitute a whole" (Prince 2019) so that the "significance of each event can be understood through its relation to that whole" (Elliott 2005). Zacks and Tversky (2001) go one step further by defining events in terms of their temporal structure and how they connect between each other. In this regard, Mostafazadeh et al. (2016b) shows that the order in which an event is described in a text, that is, the narrative, does not comply with its chronological sequence in time in 23% of the cases, meaning that simply looking at the sequence of the story in the text may not be enough to determine its temporal path.

The term story is often used interchangeably with narrative, though they are not synonyms. A story consists of events that are related by a narrator. A narrative is how a story is told or interpreted. On this account, a new event order means a new narrative of the same story, that is, a new perspective given by different observers (Zhang et al. 2019a).

The study of narratives from the computational perspective is carried out through a study area called Computational Narratology (Mani 2014). Its purpose is to study narrative from the computational and information processing point of view focusing on the algorithmic processes involved in creating and interpreting narratives, and the modeling of narrative structure in terms of formal, computable representations (Mani 2012). In parallel, computational narrative extraction, or simply Narrative Extraction can be defined as the use of computational tools for identifying, linking and visualizing narrative elements from textual sources. A closely related term is Computational Narrative Understanding (Piper et al. 2021), which broadens the perspective to social, personal and cultural dimensions. Research on this topic is very recent and open to debate. In this survey paper, we focus exclusively on the textual representation of a narrative. Broadly speaking, Narrative Extraction is framed as a sub-field of AI that makes heavy use of: Information Retrieval—to help users access information; Text Summarization—to summarize relevant and complementary information to narratives; Natural Language Processing—to identify, extract and relate the narrative elements; and Natural Language Generation—to produce text from structured data.

The list of potential applications is endless, going from virtual assistants, to chatbots, or improving information access and exploration in search tasks, to help uncover and interpret patterns in complex informational contexts, or applications that automatically generate alternative and customized representations of the source data (Wu 2019).

2.2 Data acquisition and annotation

The quality of the automation of tasks related to natural language processing is directly associated with the preparation and the annotation of data used to train NLP algorithms.

Data acquisition, despite being a relatively simple process, can raise several problems, mainly due to the lack of available datasets (Ide et al. 2002; Ide 2017) and to copyright issues (McEnery et al. 2006; Zeldes 2018). The same happens with data preparation for annotation, which requires choosing an appropriate format, such as JSON, XML, or CSV. Some can be more problematic and time-consuming than others (Ide 2017), and, so, one has to weigh in the advantages and disadvantages of each, taking into consideration the data original format, the annotation and extraction tools that will be used during the following phases, and how the corpus is going to be made

accessible. Stripping the data of unnecessary information and making sure that relevant metadata is kept is also key during this process. A case in point related to narrative extracted from news is the publication date, which is relevant to determine its timeline.

What follows is designing a suitable annotation scheme that is simultaneously tailored to encompass all the particularities of the target language(s), and comprehensive enough to be applied to other datasets or even be broadened. Bearing in mind, on the one hand, the diversity of annotation frameworks, and, on the other hand, the usefulness of establishing comparisons between annotated corpora from different genres in the same language, but also across languages, many proposals try to achieve this balance utilizing acknowledged standards, which have resulted in, for e.g., Ontologies of Linguistic Annotation (OLiA) (Chiarcos 2014), and ISO 24617—Language resource management—Semantic annotation framework. The decision about the annotation framework, and about the different layers of annotation, depends necessarily on a variety of factors, like the annotation purpose, the texts genre, among others (Pustejovsky et al. 2017). In the case of the narratives annotation, since it is relevant to feature participants, events, time and space, as well the relationships between them, the annotation scheme can be designed to include several intertwined semantic layers enabling temporal, referential, thematic, and spatial annotations (see for example Silvano et al. (2021)).

The adequacy of the annotation tool is also of great relevance for the efficiency of different tasks, namely creation, browsing, visualization, and querying of linguistic annotations. Although one can choose to tailor a tool to the specific features of one's project, it may be labor-saving to resort to the existing ones and, if necessary, proceed with some modifications. Some of the existing annotation tools that enable annotating different markables with several attributes, as well as establishing links between those markables are the following: MMAX2 (Müller and Strube 2006), MAE and MAI Stubbs (2011), BRAT (Stenetorp et al. 2012), or ANNIS (Krause and Zeldes 2014).

During the process of narrative extraction, some degree of human participation is necessary, either to verify the feasibility of the annotation scheme, to annotate linguistic aspects for which automatic models are insufficient, or to supervise the automatic annotation. Different strategies can be adopted depending, namely, on the type of annotators that the project wants or needs, with or without linguistic training, specifically: crowdsourcing (Estellés-Arolas and de Guevara 2012), class sourcing (Christopher Blackwell 2009) or gamification (Stieglitz et al. 2016). When using human annotation, calculating inter-annotator agreement (IAA) is crucial to ensure its reliability and accuracy. The two most used metrics in computational and corpus linguistics are: Cohen's Kappa (and its variation, Fleiss's Kappa) and Krippendorff's Alpha (Pustejovsky and Stubbs 2012).

2.3 Narrative extraction pipeline

The study of narratives from the perspective of natural language texts can be summarized into five major stages: (1) Pre-Processing and Parsing; (2) Identification and Extraction of Narrative Components; (3) Linking Components; (4) Representation of Narratives and (5) Evaluation. Each of these tasks gives rise to the structure adopted in the rest of this survey. An overall picture is shown in Fig. 1, and a detailed description of each task can be found in the following sections.

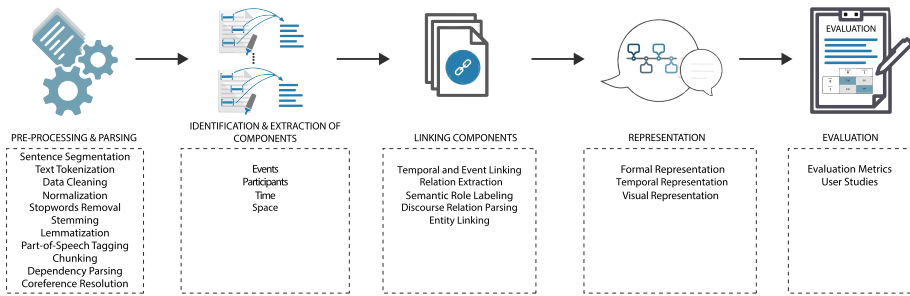


Fig. 1 The narrative extraction pipeline

3 Narratives pre-processing and parsing

Pre-processing and parsing of a text comprise a set of lexical and syntactic tasks. Lexical activities aim to split the text into basic units, called tokens, and to normalize them into different forms; syntactical analysis identifies the grammar class of tokens, and identifies the dependency between the tokens, producing trees or chunks that relate the tokens to each other. The following subsections detail both tasks.

3.1 Lexical tasks

Lexical tasks seek to standardize the input text as a way to better prepare the data content for the following steps (Sun et al. 2014). Despite being a simple procedure, they are considered to be a key factor to achieve effective results (Denny and Spirling 2018).

The main lexical tasks required for the narrative extraction process are: (1) Sentence segmentation, i.e., dividing the whole text into sentences and tokens to gather the lexicon of the language (Palmer 2007); for this task, using deep learning (DL) methods Schweter and Ahmed (2019) achieved SOTA results through the use of a multi-lingual system based on three different architectures of neural networks. (2) Text tokenization, applied to break sentences into tokens, the smallest unit of a text (Vijayarani et al. 2016); and (3) Text Cleaning and Normalization, which may involve a number of optional cleaning steps, including removing numbers, punctuation marks, accent marks, or stopwords, as well as apply stemming (Jabbar et al. 2020) and lemmatization (Bergmanis and Goldwater 2018) to observe other possible discrepancies between equal or similar words. These are foundational NLP tasks; hence they will not be explored in detail in this article. More details regarding these tasks can be found in Raina and Krishnamurthy (2022).

3.2 Syntactic tasks

The lexical analysis precedes a more profound observation, which aims to identify the grammar class of the words and to infer the relationship between words in a sentence to help understand the meaning of a text. To go over this process, we use the text presented in Fig. 2 as a running example in the remainder of this survey.

John, the magician, pulled a rabbit out of a hat at the show in Edinburgh this year. The audience loved his presentation at the XYZ contest, even more than the last one presented.

Fig. 2 Running example

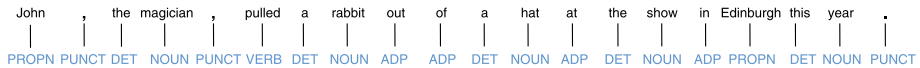


Fig. 3 Results of applying the stanza PoS tagging processor to the sentence: “John, the magician, pulled a rabbit out of a hat at the show in Edinburgh this year.”

3.2.1 Part-of-speech tagging

The first step in this stage is to assign parts-of-speech tags to each word of a given text (e.g., nouns, verbs, adjectives) based on its definition and its context, which is called the Part-of-Speech (PoS) tagging task. The current state-of-the-art is settled by Bohnet et al. (2018) in several languages. In this work, the authors used recurrent neural networks with sentence-level context for initial character and word-based representations, achieving an accuracy of 97.96% on top of the Penn Treebank dataset Marcus et al. (1993). With respect to narrative texts, since their elements can be typically associated to different word classes (for instance, nouns to participants, verbs to events), POS tagging plays a fundamental role (Palshikar et al. 2019; Quaresma et al. 2019; Yu and Kim 2021). Figure 3 shows the result of applying the Stanza¹ library (Qi et al. 2020) PoS tagger to the first sentence of our running example. In the figure, PROPON refers to a proper noun, PUNCT to punctuation, DET to determiner, VERB to a verb, and ADP to adposition. These tags are defined under the Universal Dependencies guidelines (de Marneffe et al. 2021).

3.2.2 Chunking and Dependency Parsing

Following PoS tagging, the parsing of the text can be conducted. The chunking task (a.k.a, shallow parsing) is responsible for identifying constituent parts of the sentences (nouns, verbs, adjectives, etc.) and linking them to higher-order units that have discrete grammatical meanings (noun groups or phrases, verb groups, and others). Figure 4 illustrates the results of applying chunking on top of the first sentence of our running example through the use of the benepar library (Kitaev et al. 2019). By looking at the figure, we can observe that both “a” (determiner - DT), as well as “rabbit” (noun-NN), belong to a higher group of noun phrases (NP). Note that adpositions (tagged as “ADP” in Fig. 3) are now represented in Fig. 4 with the “IN” tag. As in other tasks, the application of deep learning approaches (Hashimoto et al. 2017; Zhai et al. 2017; Akbik et al. 2018) has brought important improvements to this particular task.

The next step of the pipeline is to understand how all the words of the sentence relate to each other, which is done by a dependency parsing analysis. The objective, as referred by Jurafsky and Martin (2009), is to assign a single headword to each dependent word in the sentence through labeled arcs. The root node of the tree, that is, the head of the entire

¹ Available in: <https://stanfordnlp.github.io/stanza/>

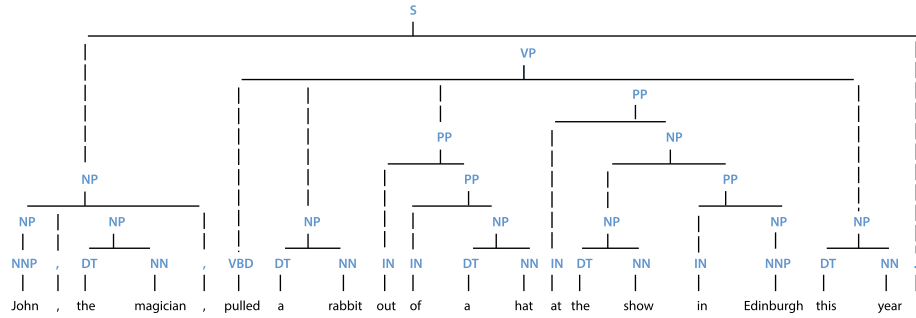


Fig. 4 Results of applying the benepar chunking model to the sentence: “John, the magician, pulled a rabbit out of a hat at the show in Edinburgh this year.”

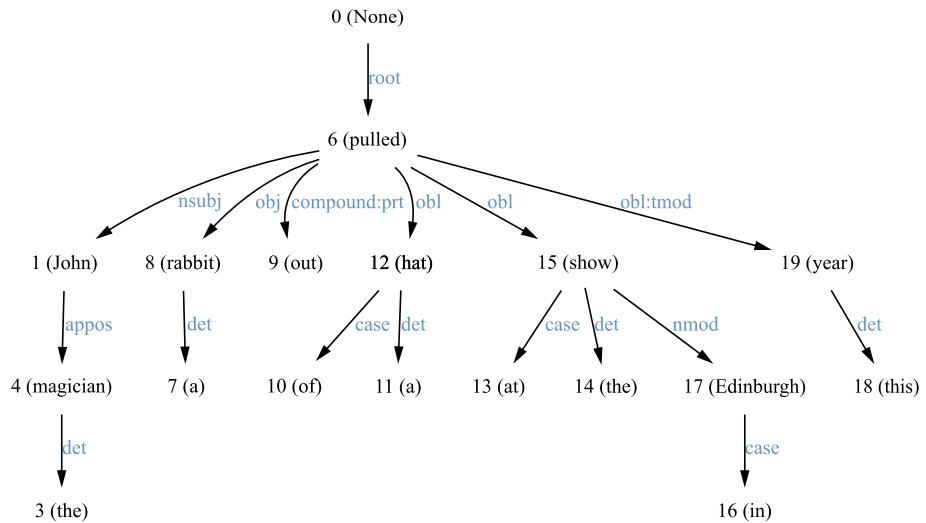


Fig. 5 Results of applying the Stanford CoreNLP Parser Dependency to the sentence: “John, the magician, pulled a rabbit out of a hat at the show in Edinburgh this year.”

structure, will be the main verb in the sentence (e.g., “pulled” in our running example). This task can enhance the performance of named entity recognition (further discussed in Sect. 4.2) models and be used to increase information extraction (de Oliveira et al. 2022). Figure 5 shows a dependency analysis of the first sentence of our running example alongside the type of relationship predicted to occur between two tokens. The results obtained stem from applying the Stanford CoreNLP Parser (Chen and Manning 2014), one of the most well-known tools in this regard. A quick look at the figure emphasizes the association between several terms, among which we highlight the relationship hierarchy between “John” and “magician”. The current state-of-the-art of this task, in the English and Chinese languages, is settled by Mrini et al. (2020). The model proposed by the authors used a combination of a label attention layer with Head-driven Phrase Structure Grammar (HPSG), and pre-trained XLNet embeddings Yang et al. (2019).

In a narrative context, the application of these tasks can be particularly useful to identify relations between narrative components (further discussed in Sect. 4).

3.2.3 Coreference resolution

Coreference resolution, also known as anaphora resolution, uses information about the semantic relation of coreferent expressions to infer the reference between them (Jiang and Cohn 2021). In addition to that, all the terms that reference the same real-world entity are marked (Mitkov 2014). For example, the word “presentation” (found in sentence 2 of our running example) refers to “show” (which can be found in sentence 1). Coreference resolution is essential for deep language understanding, and has shown its potential in various language processing problems (Poesio et al. 2016). Bringing it to the context of narrative extraction, some works highlight the importance of this task in the most different scenarios, such as clinical narratives (Jindal and Roth 2013a), newswire (Do et al. 2015), and also in violent death narratives from the USA’s Centers for Disease Control’s (CDC) National Violent Death Reporting System (Uppunda et al. 2021). In a general domain, several approaches have been proposed over the years (Pradhan et al. 2012; Mitkov 2014; Lee et al. 2013), but recently there has been a shift towards adopting transformers models (Ferreira Cruz et al. 2020; Kantor and Globerson 2019) to improve the connection between textual elements and, as such, enhance the coreference task. The current SOTA of coreference resolution is marked by Attree (2019) and Kirstain et al. (2021), both sticking to deep learning methods.

3.3 State-of-the-art and summary

Table 1 presents the current status of the tasks covered in this section. This includes lexical tasks (sentence segmentation, text tokenization, data cleaning, normalization, stopwords removal, stemming, lemmatization) and syntactic (PoS tagging, chunking, dependency parsing, coreference resolution). Herein, however, we only list those that have been undergoing constant improvements over the most recent years. In this table, the effectiveness of each study is shown according to the measures reported by the authors.

4 Identification and extraction of narrative components

The next step in the narrative extraction pipeline is to identify and extract the main elements that compose a narrative. This step comprises tasks like the detection and classification of events, the recognition and the classification of named entities (participants), the extraction of temporal information and of spatial data.

4.1 Events

Finding the events mentioned in the text is an essential step towards the extraction of narratives from a pre-processed text. Formally, an event may be defined (Allan et al. 1998) as something significant happening at a specific time and place with consequences. In the real world, this can be an explosion caused by a bomb, the birth of an heir, or the death of a famous person. Xie et al. (2008) went a little bit further, and suggested describing real-world events through the use of the 5W1H interrogatives: when, where, who, what, why, and how, anchored on journalistic practices. Based on this reference, an event such as the one referenced in our running example might be depicted along with the six aspects: *who*—John, the magician; *when*—2022 (taking into account the date of writing the text);

Table 1 State-of-the-art in pre-processing and parsing tasks

Task	Ref.	Method	Effectiveness	Benchmark dataset
Sentence segmentation	Schweter and Ahmed (2019)	BiLSTM	98.62 F1-score (English) 99.70 F1-score (Greek) 99.86 F1-score (Arabic)	Europarl Corpus Koehn (2005) SETimes Alperen and Tyers (2010) Leipzig Corpus Goldhahn et al. (2012)
Text tokenization	Graÿn et al. (2018)	Language-specific and language independent token identification rules	1% Error rate	SMULTRON Treebank Volk et al. (2015)
Lemmatization	Bergmanis and Goldwater (2018)	Standard encoder-decoder architecture + character-level sentence context	92.4 Accuracy (ambiguous words) 84.30 Accuracy (unseen words)	Universal Dependency Treebank v2.0 Nivre et al. (2020)
Part-of-speech tagging	Bohnet et al. (2018)	Recurrent network model	97.96 Accuracy	Penn Treebank Marcus et al. (1993)
Chunking	Akbik et al. (2018)	Neural character-level language modeling	96.72 F1-score 96.34 F1-Score	Penn Treebank Marcus et al. (1993)
Dependency parsing	Mrini et al. (2020)	Label attention layer + HPSG + XLNet	97.39 UAS 96.29 LAS	Penn Treebank Marcus et al. (1993)
Coreference resolution	Attree (2019)	Neural attention model	92.50 F1-Score	GAP Dataset Webster et al. (2018)

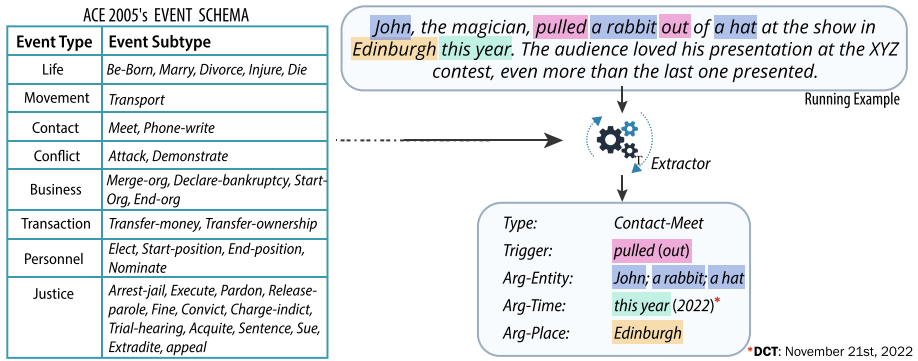


Fig. 6 Event extraction in the closed-domain of our running example

where - XYZ contest realized in Edinburgh; *what*—the magician pleased the audience during his presentation; *how*—taking a rabbit out of his hat; *why*—his presentation was better than the last one.

The task of event detection has its roots in the late 90's when the Topic Detection and Tracking (TDT) project emerged (Allan et al. 1998) as the first attempt to cope with the rising of huge volumes of data. Among the entire project, the detection of new events, also known as the first story detection (FSD) (Kontostathis et al. 2004), was a subtask of the TDT project, concerned with the detection and subsequent tracking of the first and coming stories of a given news event. One way to detect events is to rely on event triggers, cues that express an event's occurrence, most often single verbs or phrasal verbs, as referred by Boruş (2018), but also nouns, noun phrases, pronouns, adverbs, and adjectives. According to Araki (2018), 95% of the ACE (Automatic Content Extraction) event triggers consist of a single token. This is the case of our running example, where the events *pulled*, *loved*, *presented* are triggered by verbs, and *show*, *contest* and *presentation* are also events represented by nouns. Other works consider the case where multiple events appear in the same sentence (Balali et al. 2020). Following ACE (LDC 2005) terminology, we consider an event structure based on four subtasks: (1) Event mention (i.e., a sentence or expression that explains an event, including a cause and multiple arguments); (2) Event trigger (i.e., the key term that demonstrates the occurrence of an event most clearly, usually a verb or a noun); (3) Event argument (i.e., a reference to an entity, a temporal expression, or a value that works as an attribute or individual with a particular role in an event); and (4) Argument role (i.e., the link between an argument and the event in which it is involved). In its annotation guidelines for events, ACE 2005 (LDC 2005) defined 8 event types and 33 subtypes, where each event subtype corresponds to a set of argument roles. Figure 6 illustrates the event extraction process for our running example. The ACE 2005's predefined event scheme is presented in the left-hand part of the figure. The event trigger ("pulled"), the 3 argument roles ("Arg-Entity"), and 2 modifiers ("Arg-Time" and "Arg-Place") identified for the type "Contact-Meet" type are illustrated on the right-hand side of the figure. In this visualization, the time-reference "this year" is normalized to 2022 based on the document creation time (DCT).

The current state-of-the-art in event extraction task for the English language is provided by Feng et al. (2018), who developed a language-independent neural network. The authors tested on top of the ACE 2005 English event detection task (Walker et al. 2006) data,

achieving an F1-score of 73.40%. More detailed information about the event detection task can be found in Saeed et al. (2019) and a recent survey from Xiang and Wang (2019). In recent years, the task of monitoring streams of data and detection events has made a shift towards microblogs platforms, such as Twitter (Atefeh and Khreich 2015), a very different scenario when compared to the TDT era, when algorithms were developed to track news stories over time from traditional media like newspapers (Kalyanam et al. 2016).

The detection of events also plays a fundamental role in different kinds of applications, domains, and languages, as is the case of clinical narratives (Jindal and Roth 2013b; Adams et al. 2021), and historical events (Lai et al. 2021). Despite the significant advances in the last few years, the application of event extraction techniques in the context of narrative representation has been quite limited (Metilli et al. 2019), and only recently a few works began to emerge. Metilli et al. (2019), for example, made use of numerous novel discourse and narrative features, besides common relations, such as the time at which the event takes place, or coreference events, which detects whether two mentions of events refer to the same event (Araki 2018). Aldawsari and Finlayson (2019) presented a supervised model to automatically identify when one event is a sub-event of another, a problem known as sub-event detection or event hierarchy construction. Another strand tackles the problem of nested event structures, a common occurrence in both open domain (not limited to a single topic or subject) and domain specific (dedicated to a particular problem representation or solution) extraction tasks (McClosky et al. 2011). For instance, a “crime” event can lead to an “investigation” event, which can lead to an “arrest” event (Chambers and Jurafsky 2009).

The research of Metilli et al. (2019) applied a recurrent neural network model to event detection applied to biography texts obtained from Wikipedia. The model used a set of event categories (e.g., birth, conflict, marriage, and others) to depict a narrative. In the context of news stories, Zahid et al. (2019) proposed the development of heuristics to segment news according to a scheme that defines the organization and order of the events. The strategy followed in that research employed mechanisms that journalists usually exploit to compose news.

4.2 Participants

Another essential category of narrative elements is the participants. They are the “who” of the story, relevant to the “what” and the “why”. Participants, sometimes also referred to as actors, are involved in events in a varied number of ways and often correspond to entities in NLP. These, are usually identified through Named Entity Recognition (NER), a generic task aiming at seeking, locating and categorizing the entities mentioned in the text into pre-defined categories. Although the three most common named entities are person, organization, and location, further types such as numeric expressions (e.g., time, date, money, and percent expressions) can also be considered (Nadeau and Sekine 2007). In narrative terms, participants are more frequently restricted to the person and organization categories, though other categories such as animals may also be used. In our running example this would result in annotating “John” and “audience” as a person, and “rabbit” as an animal as can be observed in Fig. 7.

Overall, research on named entity recognition can be categorized into four main categories: rule-based approaches, unsupervised learning approaches, feature-based supervised learning approaches, and deep-learning-based approaches (Li et al. 2022). Beyond this, other aspects, such as textual genres, entity types, and language, are also considered by

John **PERSON**, the magician, pulled a rabbit **ANIMAL** out of a hat at the show in Edinburgh this year.
The audience **PERSON** loved his presentation at the XYZ contest, even more than the last one presented.

Fig. 7 Manual annotation of participants in our running example

research conducted on named entity identification (Goyal et al. 2018). The current state-of-the-art in named entity recognition was obtained by Wang et al. (2021) on the CoNLL 2003 NER task (Tjong Kim Sang and De Meulder 2003). To automate the process of finding better concatenations of embeddings for structured prediction tasks, the authors proposed Automated Concatenation of Embeddings (ACE) based on a formulation inspired by recent progress on neural architecture search. The method achieved an F1-score of 94.60%.

In narrative contexts, Lee et al. (2021) present a proposal that goes beyond the identification of entities in narratives, proposing a multi-relational graph contextualization to capture the implicit state of the participants in the story (i.e., characters' motivations, goals, and mental states). Piper et al. (2021) point out that NLP works focusing on agents have emphasized broadening the understanding of characters beyond named entities. This has been done through the concept of animacy detection, as in agents like "the coachman" or "the frog", while also distinguishing characters from other named referents (Piper et al. 2021). According to Oza and Dietz (2021)'s proposal, selecting a set of relevant entities for a story construction could be best achieved by using entity co-occurrences in retrieved text passages - especially when the relative relevance of passages is incorporated as link strength. In line with what the authors claim, their proposal is between 80% and 30% more effective than the best link-based approach.

4.3 Time

Temporal references are of the utmost importance for the understanding of the narrative and its timeline by anchoring each event or scene at a point in time. The process of identifying temporal references is commonly referred to as Temporal Tagging, and can be split into two sub-tasks: extraction and normalization (Strötgen and Gertz 2013). The former aims to correctly identify temporal expressions, and can be seen as a classification problem. The latter aims to normalize the identified temporal expressions, and can be seen as a more challenging process where different temporal expressions, carrying out the same meaning, need to be anchored at the same time-point. In this regard, three temporal taggers take the lead, Heideltime (Strötgen and Gertz 2013), SuTime (Chang and Manning 2012) and GuTime (Mani and Wilson 2000), which support the four basic types of temporal objects defined by Pustejovsky et al. (2005) in TimeML, the standard markup language for temporal annotation containing TIMEX3 tags for temporal expressions: Dates (e.g., "December 3, 2021"), Time (e.g., "5:37 a.m."), Durations (e.g., "four weeks", "several years"), and Sets (e.g., "every day", "twice a month"). Alongside with these, other realizations of temporal expressions can be found in a text (Strötgen et al. 2012; Campos et al. 2017), which pose additional challenges. We refer to explicit (e.g., "April 14, 2020"), implicit (e.g., "Christmas day 2019"), and relative temporal expressions (e.g., "yesterday"), a kind of temporal expression that requires further knowledge to be normalized. A discussion of the challenges associated with each one of them can be found in Strötgen and Gertz (2013, 2016). Other researchers (Jatowt et al. 2013) devised methods concerning document dating of non-timestamped documents.

The rise and continuous development of temporal taggers has paved the way to the emergence of upstream research at the intersection between information extraction and several NLP tasks, where temporal expressions play an important role. For example, in the information retrieval domain, temporal information (Campos et al. 2014) can be used for temporal clustering of documents (Campos et al. 2012) or temporal ranking (Berberich et al. 2010; Campos et al. 2016); in question answering to query knowledge bases (Sun et al. 2018); in document summarization to construct timeline summaries (Campos et al. 2021); and in web archives to estimate the relevance of past news (Sato et al. 2021).

At this stage of identification and extraction of narrative components, several works address the intersections between the tasks of identifying and extracting events, entities, and temporal information. Strötgen and Gertz (2012) explored event-centric aspects, considering related temporal information. In particular, they introduced the concept of event sequences, a set (or sequence) of chronological ordered events extracted from several documents. In their work, the geographic dimension played an important role by mapping event-sequences onto a map as well. In contrast, studies such as Agarwal et al. (2018) and Rijhwani and Preotiuc-Pietro (2020) explored the task of recognizing named entities considering time-aware aspects.

The incipient annotation of documents to support related research has also led to the development of markup languages (e.g., TimeML (Pustejovsky et al. 2005), a formal specification language for temporal and event expressions anchored on TIMEX3 tags), annotated corpora (e.g., TimeBank corpus (Pustejovsky et al. 2006), a standard set of English news articles annotated with temporal information under the TimeML 1.2.1 (Sauri et al. 2006) guidelines) and research competitions [e.g., TempEval series (Lim et al. 2019)].

4.4 Space

The problem of extracting geographic references from texts is longstanding (Martins et al. 2008), and is highly related to the field of Geographic Information Retrieval (GIR) (Purves et al. 2018).

Much of this information can be found in unstructured texts through references to places and locations (Purves et al. 2018).

Many difficulties arise when attempting to comprehend geographic information in natural language or free text (e.g., under-specified and ambiguous queries) (Purves et al. 2018). Challenges related to space involve the detection and resolution of references to locations, typically, but not exclusively, in the form of place names, or more formally toponyms, from unstructured text documents (Jones and Purves 2008). In our running example, to extract the location of the main event, one can consider “contests near Edinburgh”, which consists of three important parts, a theme (contests), a spatial relationship (near), and a location (Edinburgh). However, in the example, it is unclear whether there is only one contest in the city or multiple ones, and, if so, to which one the text is pointing to. The spatial relationship “near” also brings more questions than answers, making it difficult to understand whether it refers to downtown Edinburgh or some constrained space within the state of Edinburgh.

Spatial relationships such as these may be both geometric (obtained using coordinate systems imposed on the real world, such as latitude and longitude) and topological (spatially related, but without a measurable distance or complete direction) (Larson 1996). The current SOTA in toponym detection and disambiguation was achieved by Wang et al. (2019) on the SemEval 2019 Task 12 (Weissenbacher et al. 2019), taking as a basis

PubMed articles. In their approach, called DM_NLP, the authors made use of an ensemble of multiple neural networks. In a different approach, Yan et al. (2017) proposed a solution called augmented spatial contexts that learns vector embeddings and uses them to reason about place type similarity and relatedness.

When considering space in a narrative context, other issues arise. Geographic maps, landscape paintings, and other representations of space are not always narratives, yet all narratives presuppose a world with spatial extension, even when spatial information is concealed (Ryan 2014 [Online]). Generally speaking, a narrative space is a physically existing environment in which characters live and move. This space is usually identified as a setting, that is, the general socio-historic-geographical environment in which the action takes place (Ryan 2014 [Online]), i.e., grounded by referents in the real world (e.g., the Edinburgh city in our running example) or entirely fictional (Rohan kingdom in *The Lord of Rings*). Aiming to identify places, some works (Bamman et al. 2019; Brooke et al. 2016) have been exploring entity recognition and toponym resolution in the context of narratives, making it possible to recognize named locations, facilities, and geopolitical entities (Edinburgh, Rohan) within this kind of text (Piper et al. 2021). Another problem is related to coreference resolution. Unlike coreference resolution of named locations, long-document coreference of common items (e.g., the house and the room), which form the narrative universe for many fictional creations, can be tough. As indicated by Piper et al. (2021), many questions can only be addressed by calculating the distance between locations described in the text: how far do Frodo and Sam go on their journey? Systems that allow better inference on spatial information within narratives could provide important insights to get the reader closer to the narrative universe.

4.5 State-of-the-art and summary

Table 2 brings a summary of the current status of the tasks covered at the identification and extraction stage discussed in this section. In this table, the effectiveness of each study is shown according to the measures reported by the authors. The type column specifies the type of task addressed in the works. All the works discussed here refer to a semantic analysis of the text. This contrasts with Table 1 where semantic tasks, but, above all, lexical and syntactic, were addressed.

In our supplementary material, made available with this survey, one can find additional benchmark datasets relating to the topics covered in this section. This material has datasets available that can be used as a reference for carrying out the tasks described here. Some of these datasets are available in other languages and might help identify and extract narrative components also in these languages.

5 Linking components

After the process of identifying and extracting narrative components, it is crucial to extract the link relations between such pieces. Linking narrative components comprises the core of the extraction of narratives from the text. More than extracting separate elements from texts, establishing relations and detecting structures from them becomes essential to fully understand their meaning. In this stage, extracted pieces of information are connected, structuring the narrative at a global level. Thus, temporal and event linking, entity and

Table 2 State-of-the-art of Identification and Extraction Tasks

Task	Ref.	Method	Effectiveness	Benchmark Dataset
Event detection	Feng et al. (2018)	Language-independent neural network	73.40 F1-score	ACE 2005 multilingual training corpus Walker et al. (2006)
Named entity recognition	Wang et al. (2021)	Automated concatenation of embeddings + document-context	94.60 F1-score	CoNLL 2003 NER task Tjong Kim Sang and De Meulder (2003)
Named entity recognition	Shahzad et al. (2021)	Neural character-level language modeling	50.52 F1-score 90.71 F1-score	WNUT 2017 Emerging Entities task Derczynski et al. (2017)
Temporal information extraction	Strötgen and Gertz (2013)	Rule-based	91.94 F1-score	TempEval-3 Platinum English UzZaman et al. (2013) TempEval-3 TimeBank UzZaman et al. (2013)
Temporal information extraction	Chang and Manning (2012)	Deterministic rule-based system designed for extensibility.	0.92 F1-score (extends)	TempEval-2 Verhagen et al. (2010)
Temporal information extraction	Mani and Wilson (2000)	Based on both hand-crafted and machine-learn rules	83.2% accuracy	TDT-2 Corpus (part of) Cieri et al. (1999)
Toponym detection	Wang et al. (2019)	Neural networks ensemble	0.9346 F1-score (overlap macro) 0.9161 F1-score (strict macro)	SemEval 2019 Task 12 Weissenbacher et al. (2019)
Toponym disambiguation	Wang et al. (2019)	Neural networks ensemble	0.8234 F1-score (strict macro) 0.7821 F1-score (strict micro)	SemEval 2019 Task 12 Weissenbacher et al. (2019)

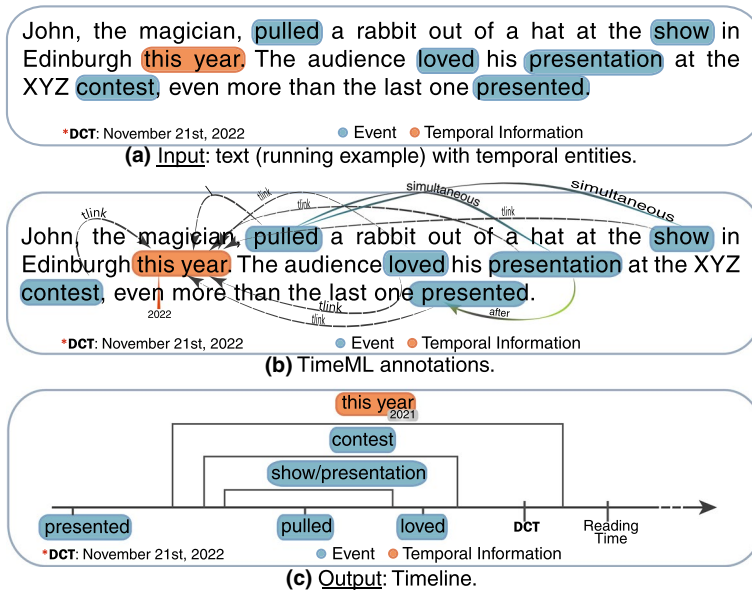


Fig. 8 Overview of temporal information extraction steps

event linking, and the relation extraction are considered. They will be addressed in this section.

5.1 Temporal reasoning

Despite concentrating the research community’s attention, predicting temporal and causal relations between events and temporal references, and investigating the evolution of an event over time as a whole (Kalyanam et al. 2016), remains a major challenge of the text understanding task (Leeuwenberg and Moens 2018) that goes beyond merely identifying temporal expressions in text documents. Ultimately, understanding a text with regards to its temporal dimension requires the following three steps: (1) identification of temporal expressions and events; (2) identification of temporal relations between them; and (3) timeline constructions. An overview of each one of them is shown in Fig. 8 for our running example.

Given an input text, the system determines the temporal expressions and the events (top). Next, it assigns TimeML annotations (middle). Finally, it produces a timeline as an output (bottom). Figure 8a begins by showing the input text. By looking at it, one can easily detect “this year” temporal expression (normalized to 2022 assuming that the Document Creation Date—DCD— refers to November 21st, 2022) and the chronology of the events. In our example, the event *pulled* is carried out by a magician called John. During the show, he pulled a rabbit out of a hat. The presentation was more welcomed by the public than another previous presentation. Despite all the information collected, it is not clear whether the “last one presented” (event) was held during the present edition or last year’s contest, or even if the “show” consisted of only “pulling a rabbit out of a hat”. Ambiguity in the text might also raise doubts about whether the previous presentation refers to John or another

magician. In Fig. 8b, we can observe a possible interpretation of the relation phase, where temporal links are established between the temporal expressions and the events. Finally, in Fig. 8c, we can see the generated timeline. In this visualization, one can observe a reference to the document creation time, the present time (reading time), and the ordering of the events according to the time-reference “this year” normalized to 2022. This apparently simple example shows how difficult it may be to understand a text in detail. In the following subsections, we present research tackling the identification of temporal expressions, events and their relations, before exploring timeline constructions.

5.1.1 Identification of temporal expressions, events and their relations

Mostafazadeh et al. (2016b), in their work, proposed a pipeline strategy for extracting events and linking them temporally, without building a timeline. The authors demonstrated that events that follow a protocol or script could be related to each other through temporal associations. These relationships are based on temporal reasoning, and provide a broad understanding of language. In another study, Roemmele and Gordon (2018) applied a neural encoder-decoder to detect events before predicting the relations between adjacent events in stories. The authors assumed that subsequent events are temporally related, and as such the extraction of events and their respective relations can be inferred. The goal is to model events that are the cause of other events and also events that are the effect of other events. Similarly, Yao and Huang (2018) deemed that the order of events in a text is the same as that of the temporal order of them in a narrative. Based on this assumption, the authors used a weakly supervised technique to propose a novel strategy to extract events and temporal relations between events in news articles, blogs, and novel books across sentences in narrative paragraphs. In another proposal, Han et al. (2019a) introduced a strategy where the extraction of events and temporal relations was jointly learned. Hence, the information about one task is employed to leverage the learning of the other. This strategy selected shared representation learning and structured prediction to solve the tasks. The results achieved were superior to the previous works. Ning et al. (2017) suggested a structured learning approach to identifying temporal relations in natural language text. Their solution was evaluated on top of the TempEval-3 data (UzZaman et al. 2013), achieving a temporal awareness [metric provided by UzZaman and Allen (2011)] of 67.2%. In the health domain, Tang et al. (2013) developed a temporal information extraction system capable of identifying events, temporal expressions, and their temporal relations found in clinical texts. In this same domain, Leeuwenberg and Moens (2020) proposed an annotation scheme for extraction of implicit and explicit temporal information for clinical reports identified events, which provides probabilistic absolute event timelines by modeling temporal uncertainty with information bounds. A comprehensive overview of the research conducted over the years on identifying events, time references, and properly connected them is given by Derczynski (2016).

5.1.2 Timeline construction

The next step, after establishing the relationships between events and the temporal entities, is to arrange temporal information in such a manner that the remaining narrative components can be organized with regards to time (Leeuwenberg and Moens 2019). One important aspect of this, as stated by Leeuwenberg and Moens (2019), is the temporal reasoning task, which refers to the process of combining different temporal cues into a coherent

temporal view. This is of the utmost importance as previous works have shown that the information presented in the text may not match the elapsed order (Mostafazadeh et al. 2016b). Getting to know the correct order of the events may also enable to dive deeply in the analysis of particular components. Antonucci et al. (2020) employed this notion to investigate the evolution of characters of books over time. The authors trained word embeddings related to different characters in distinct parts of literary texts (e.g., chapters). These embeddings are also called dynamic or temporal embeddings (Bamler and Mandt 2017) since, from them, it is possible to analyze relations between concepts over time. This work demonstrated that characters with strong friendships have similar behavior; therefore, such characters evolve similarly over time. Another outcome of the temporal reasoning task is building the network of events in chronological order, i.e., the text's narrative. Over the years, different researches (Bethard et al. 2007; Mani and Schiffman 2005) have been proposed with one such goal. In Leeuwenberg and Moens (2018), the authors proposed two models that predict relative timelines in linear complexity and new loss functions for the training of timeline models using TimeML-style annotations. On the sidelines, Jia et al. (2021) proposed the use of temporal information as a way to provide complex question answering related to temporal information through knowledge graphs. An overview of the temporal reasoning task, on the extraction, and on how to combine temporal cues from text into a coherent temporal view is presented in a survey of Leeuwenberg and Moens (2019).

5.2 Relation extraction

Capturing the linking between the pairs of identified entities (or participants) is an important aspect of the narrative understanding. Known as relation extraction task (Qin et al. 2021), it attempts to determine a semantic link between two named entities in a sentence. Performing this task on top of our running example would result in identifying an instance of a *pulled* relation between “John” and “a rabbit”. Such semantic relations are usually structured in the form of $\langle e_1, rel, e_2 \rangle$ triples, where e_1 and e_2 are named-entities, and *rel* is a relationship type (Batista 2016 [Online]). In our case, it could be exemplified by the triple: $\langle John, pulled, a rabbit \rangle$. Considering that in narratives, events are often directly linked, extracting such relationships is also of paramount importance.

Semantic relations also encompass objectal relations (i.e., relations between discourse entities seen as extra-linguistic concepts). These relations aim to state how two discourse entities are referentially related to one another (ISO 24617-9:2019 (E) 2000). Other NLP tasks, like Question-Answering Systems (Li et al. 2019), and the creation of Knowledge Graphs (Zhang et al. 2019b), can benefit from identifying such a relation. An example of entity relation extraction employed to build Knowledge Graphs is provided in Han et al. (2019b). In this work, the authors present OpenNRE, an open-source and expandable toolkit for implementing neural models for relation extraction. Using this tool, one can train custom models to extract structured relational facts from the plain text, which can be later used to expand a knowledge graph.

The traditional relation extraction can be solved by applying one of the following approaches: (1) rule-based; (2) weakly-supervised; (3) supervised; (4) distantly supervised; and (5) unsupervised. In addition to this type of relation extraction, one can also extract semantic relationships by following Open Information Extraction approaches (OIE) (Batista 2016 [Online]), a research field that obtains domain-independent relations from a text. This task is supposed to extract all kinds of n-ary relations in the text (Xavier et al. 2015). Two methods can be applied by the Open Information approaches (1) rule-based;

and (2) data-based. The former relies on hand-crafted patterns derived from PoS-tagged text or dependency parse tree rules. The latter generates patterns based on training data represented as a dependency tree or PoS-tagged text.

Over the last few years, neural networks have also made their path in the particular domain of relation extraction (Qin et al. 2021), through different neural network approaches (Baldini Soares et al. 2019; Hendrickx et al. 2010; Nadgeri et al. 2021). Exploiting the narrative structure, (Tang et al. 2021) proposed a Multi-tier Knowledge Projection Network (MKPNet). The strategy was designed to leverage multi-tier discourse knowledge and present a knowledge projection paradigm for event relation extraction. According to the authors, such a paradigm can effectively leverage the commonalities between discourses and narratives for event relation extraction. This study had a focus on the projection of knowledge from discourses to narratives. (Lv et al. 2016) also proposed a strategy that takes advantage of the narrative structure. The authors applied an autoencoder to a set of features, including word embeddings, in a clinical narrative dataset, the 2010 I2B2 relation challenge (Uzuner et al. 2011). Using the same dataset, the 2010 I2B2 relation challenge, and the 2018 National NLP Clinical Challenges (N2C2) dataset (Henry et al. 2020), Wei et al. (2019) proposed two different models to extract relations from clinical narratives. The models used a language model based on transformers to extract the features, and BI-LSTM neural network attention to detect the relations. Another example comes from finances related narratives, in this context a shared task was proposed by Mariko et al. (2022a) with the intend to extract causality between existing relations. In this shared task experiments were done on top of FinCausal dataset (Mariko et al. 2022b), a dataset extracted from different 2019 financial news. The best results were achieve by a team that developed an ensemble of sequence tagging models based on the BIO scheme using the RoBERTa-Largemodel, which achieved an F1 score of 94.70 to win the FinCausal 2022 challenge.

In addition to these types of connections, entities can also be also linked to other kinds of data. Knowledge databases can furnish a unique identity to entities, which can enrich the narrative with information, and aid to disambiguate the entities as well. The task of linking knowledge base-entities and narrative entities is called entity linking. In the following section, we will discuss this task.

5.3 Entity linking

Over the years, efforts have been developed to explore relations between entities in texts (Freitas et al. 2009; Hasegawa et al. 2004). This willingness to recognize individuals in a document is a significant move towards understanding what the document is all about (Balog 2018). Truly understanding a text requires, however, linking these individuals with other pieces of information. The temporal connection between events is one way to achieve this. Another way is to relate the events of the narrative with entities from a knowledge base, which stores data about entities using ontological schemes (Ehrlinger and Wöß 2016). According to Gruber (1993), ontologies are formal and explicit specifications of a shared conceptualization that can serve as a modeling basis for different purposes of representation. Representational primitives defined by ontologies are typically classes (or sets), attributes (or properties), and relationships (or relations among class members) (Gruber 2008).

Formally, the process of linking entities can be understood as the task of recognizing (named entity recognition), disambiguating (named entity disambiguation [Eshel et al. 2017]), and linking entities (named entity linking) with unique entity identifiers from a

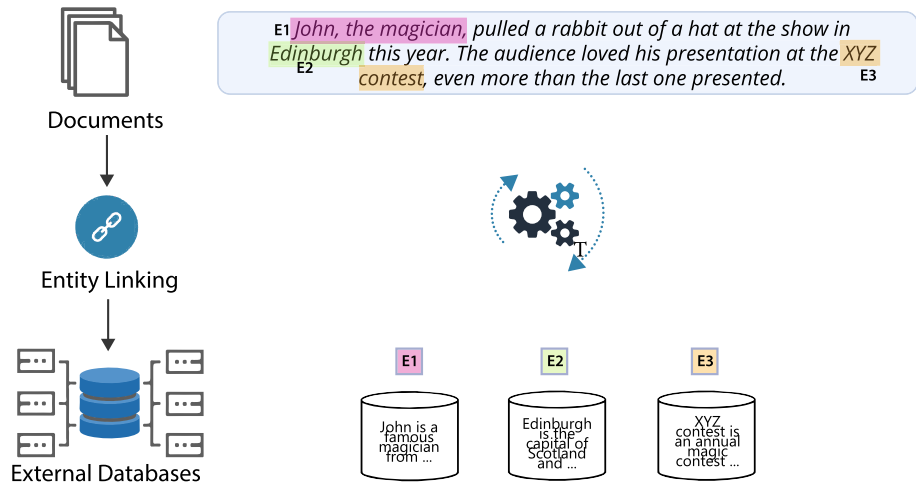


Fig. 9 Overview of entity linking task result when applied to our running example

given reference knowledge-base, such as DBpedia (Auer et al. 2007), or Yago (Suchanek et al. 2007). Given the sentence “Bush was a former president of the USA from 1989 to 1993”, the idea is to determine that “Bush” refers to George H. Bush, an American Politician who served as the 41st president of the United States from 1989 to 1993, and not to “Bush” a British rock band formed in London, England in 1992. When connecting an entity with Wikipedia data, such as Wikidata (Vrandečić and Krötzsch 2014), this task is known as Wikification. As emphasized by Szymański and Naruszewicz (2019), when a document is Wikified, the reader can better understand it because related topics and enriched knowledge from a knowledge base are easily accessible. From a system-to-system view, the meanings of a Wikified document’s core concepts and entities are conveyed by anchoring them in an encyclopedia or a structurally rich ontology. In the context of a narrative, it might mean a more knowledge-rich narrative.

Figure 9 illustrates the results of applying entity linking on top of our running example. By looking at the figure, one can observe the linkage of the detected entities or events in the text with their presence on external sources, such as knowledge bases. One can get insights into who is “John”, “Edinburgh” or the “XYZ contest”. This information might be helpful not only to add more knowledge but also to disambiguate the case of entities that are associated with more than one concept.

In the scope of entity linking efforts, Raiman and Raiman (2018) achieved the state-of-the-art focusing on a cross-lingual approach, a type system on an English dataset supervised with French data. In particular, they constructed a type system and used it to constrain a neural network’s outputs to respect the symbolic structure.

5.4 Semantic role labeling

Understanding narrative stories involves the ability to recognize events and their participants, as well as to recognize the role that a participant plays in an event, as “‘who’ did ‘what’ to ‘whom’”, and ‘when’, ‘where’, ‘why’, and ‘how’ (ISO 24617-4:2014). Generally, this can be achieved by Semantic Role Labeling (SRL). This task assigns semantic roles to the constituents of the sentence (Aher et al. 2010). Contrary to syntactic analysis

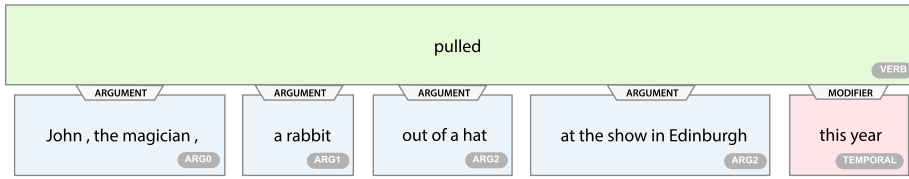


Fig. 10 Semantic role labeling result when applied to the sentence: “John, the magician, pulled a rabbit out of a hat at the show in Edinburgh this year.”

(such as the ones conducted on chunking, dependency parsing, etc.), it acts on a semantic level, and is responsible for capturing predicate-argument relations, such as “who did what to whom” (He et al. 2018), towards making sense of a sentence’s interpretation. SRL recovers the latent predicate-argument structure of a sentence, providing representations that answer basic questions about a sentence’s meaning. Figure 10 illustrates the results of applying SRL on top of the first sentence of our running example, through the use of AllenNLP² model, which is the implementation of Shi and Lin (2019).

The relation established between the different parts of the sentence can be observed by looking at the figure. In this relation, all the parts are linked by the verb “pulled (out)”: “John, the magician”, “a rabbit” and “of a hat” are the core arguments, while “at the show in Edinburgh” and “this year” are modifiers, the former conveying spatial information and the latter temporal information. The current state-of-the-art of this task is settled by He et al. (2018) with an F1-score of 85.5%, using a deep learning approach based on a BiLSTM neural network to predict predicates and arguments on top of the OntoNotes benchmark (Pradhan et al. 2013).

In a narrative-focused approach, Mihaylov and Frank (2019) proposed the use of linguistic annotations as a basis of a discourse-aware semantic self-attention encoder for reading comprehension on narrative texts. In this work, the authors adopted the 15 fine-grained discourse relation sense types (see section 5.5 for more information about discourse relations) from the Penn Discourse Tree Bank (PDTB). According to what is mentioned by the author, combining such an annotation scheme with self-attention yields significant improvements. Following this approach, this study’s results indicate that SRL significantly improves *who* and *when* questions and that discourse relations also improve the performance on *why* and *where* questions. These results are illustrated on top of the NarrativeQA (Kočíský et al. 2018) reading comprehension.

5.5 Discourse relation parsing

When reading or listening to a text, the reader/listener establishes relations of meaning between the different parts, be they clauses, sentences, or paragraphs. These relations are discourse relations (DReIs)—also known as rhetorical relations or coherence relations—and are crucial to explain how discourse is organized. For that reason, they have been the basis of several frameworks, such as Rhetorical Structure Theory (RST) (MANN and Thompson 1988), Segmented Discourse Representation Theory (Asher et al. 2003).

² Available in: <https://demo.allennlp.org/semantic-role-labeling>

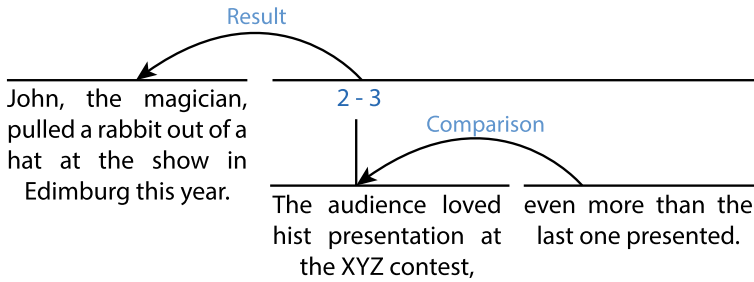


Fig. 11 The scheme application to our running example

Looking at our running example, since no connective is present, the identification of the DRel would be very challenging using automatic methods. However, taking into consideration lexical and other semantic information, as well as our world knowledge, one can infer that the two sentences are related by the DRel Result, because loving the magician's presentation is a consequence of him pulling a rabbit out of the hat. The second sentence can be divided into two arguments/ text unities related by Comparison. Figure 11 illustrates the annotation of this example using an RST markup tool (O'Donnell 2000).

In NLP, the process of uncovering the DRels between text units is called discourse parsing, a very complex task, despite a few advances over the past years. One of the most recognized discourse parsers is the one used by the news corpus Penn Discourse Tree Bank (PDTB) (Prasad et al. 2018). This end-to-end discourse parser (LIN et al. 2014) follows the same steps as PDTB human annotators: identification of the discourse connectives, arguments segmentation, labeling them Arg1 and Arg2, and recognition of the explicit DRel. When no explicit relation is extracted, the second step labels the implicit DRels. The last step consists in labeling the attribution spans (see Potter (2019) for more information about attribution relations), which are text units that reveal if the content should be attributed to the writer or to another participant.

The 2021 edition of the DISRPT Shared Task (Zeldes et al. 2021) added the task of Discourse Relation Classification across RST, SDRT and PDTB, to the two existing tasks from the previous edition (Zeldes et al. 2019), Elementary Discourse Unit Segmentation and Connective Detection. The Shared Task was performed in relation to 16 datasets in 11 languages. Overall, the system with the best results was DisCoDisCo (Gessler et al. 2021) with a Transformer-based neural classifier, which was able to surpass state-of-the-art scores from 2019 DISRPT regarding the first two tasks and to obtain a solid score on the 2021 benchmark for the third task.

Establishing DRels between events, such as cause, result, or temporal sequence, is of paramount importance to understand the narrative. Much has been achieved with manual annotation of datasets, which has been the basis for the development of some Discourse Parsers, namely within shared tasks (see LI et al. (2021), for a review and future trends). However, many issues need further research so that Discourse Parsing can fully contribute to narrative extraction, in particular concerning the identification of implicit DRels and the specification of the high-level discourse structures.

Table 3 State-of-the-art in information linking tasks

Task	Ref.	Method	Effectiveness	Benchmark dataset
Temporal and event linking	Ning et al. (2017)	Structured Learning	67.20 Temporal Awareness	TempEval-3 UzZaman et al. (2013)
Named entity linking	Raiman and Raiman (2018)	(1) Type system, and (2) use it to constrain the outputs of a neural network to respect the symbolic structure.	94.88 micro-precision 90.85 micro-precision F1-Score	CoNLL 2003 NER task Tjong Kim Sang and De Meulder (2003)
Relation extraction	Baldini Soares et al. (2019)	BERT-based Model	89.5 F1-score	SemEval-2010 Task 8 Hendrickx et al. (2010)
Semantic role labeling	He et al. (2018)	BiLSTM	85.50 F1-score	OntoNote Pradhan et al. (2013)
Discourse relation parsing	Gessler et al. (2021)	Transformer-based neural classifier	EDU segmentation on Treebanked data - 91.48 F1-score (average) Connective detection on Treebanked data - 91.22 F1-score (average) Relation classification on Treebanked data - 61.82 F1-score (average)	DISRPT 2021 Zeldes et al. (2021)

5.6 State-of-the-art and summary

Table 3 brings a summary of the current status of five of the tasks (temporal reasoning, entity relation extraction, entity linking, semantic role labeling, and discourse relation parsing) covered at the information linking stage discussed in this section. The type column specifies the type of task addressed in those works. In this table, each study's effectiveness is displayed according to the measures reported by the authors. All the discussed tasks are related to semantic analysis at this pipeline stage.

In our supplementary material, made available with this survey, one can find a list of additional benchmark datasets relating to the topics covered in this section. The datasets mentioned there can be used as a reference for carrying out the tasks described here. Some of these datasets are available in languages other than English and are useful for the tasks of linking narrative components in these other languages.

6 Representation of narratives

The representation of narratives can be categorized in two levels, the conceptual and the visual level. The first one considers that the elements in a narrative are codified as concepts that allow abstracting the meaning. This codification allows the exploration and detection of the elements of a narrative, smoothing the process of analysis of a narrative. The second level of representation is built using visual elements like lines, graphs, icons, pictures, or other graphical resources used to depict the narrative. This level of representation is more accessible for people of different degrees of general expertise. Next, we detail both levels of representation.

6.1 Narrative ontologies

Ontologies are a flexible and complex framework in computer science that allows building schemes to represent several kinds of concepts. They aid the multi-layered meanings of a narrative to be pictured as faithfully as possible (Ciotti 2016) since the main elements are conceptually identified.

Over the years, some ontologies have been proposed as a means to representing narratives. Khan et al. (2016), for example, proposed an ontology for narratives applying it to Homer's *Odyssey*. Damiano and Lieto (2013), in turn, described an ontology for narratives for the hero's journey, which is a traditional archetype of the histories of the western culture. The ontology is applied in digital archives for artwork, and then to aid the browsing through the digital artifacts. Tests were made in a small dataset, and some information about artworks was obtained by reasoning. In another proposal for cultural and historic events, Meghini et al. (2021) depicted the Narrative Ontology (NOnt). The authors built an ontology that comprises some rules, like laws of physics. Then, they applied it to heritage crafts, and case studies were conducted to validate the proposed model.

Although, ontology is a flexible framework, building an ontology that encompasses all the elements in a narrative is a cumbersome task. Thus, few proposals undertake this challenge.

6.2 Formal semantic representation

Representing and learning common sense knowledge for the interpretation of a narrative is one of the fundamental problems in the quest for a profound understanding of language (Mostafazadeh et al. 2016a). Logic-based representations, like DRT (Kamp and Reyle 1993) and Abstract Meaning Representation (AMR) (Banarescu et al. 2013), express recursive formal meaning structures that have a model-theoretic interpretation, although following different formalizations. DRT adopts a dynamic and compositional perspective, which entails determining how the meaning of a sentence can change the context, and enables a straightforward and elegant representation of not only anaphoric relations, both nominal and temporal, but also of other linguistic phenomena. Within DRT's framework, the processing of the discourse is performed one sentence at a time in Discourse Representation Structures (DRSs) in an incremental manner. Each sentence is represented in the DRSs by discourse referents, which represent entities in the discourse, always displayed at the top of the box, and by conditions, which establish a relation of identity between the discourse referents and the corresponding element of the sentence, typically displayed below the universe constituted by the discourse referents. These DRSs are recursive formal meaning structures that have a model-theoretic interpretation, and can be translated into first-order logic (FOL). Asher (1993) and Asher et al. (2003) extended DRT with the inclusion of discourse relations and with contributions from other dynamic semantic proposals and pragmatics, naming it Segmented Discourse Representation Theory (SDRT).

The aforementioned semantic representations are used by the Groningen Meaning Bank (GMB) (Basile et al. 2012; Bos et al. 2017) and the Parallel Meaning Bank (PMB) (Abzianidze et al. 2017). While PMB adopts DRT as a single semantic formalism to fully represent the different annotated linguistic phenomena, GMB goes further and enhances DRT semantics with discourse relations between DRSs, using Boxer (Bos 2015; Curran et al. 2007), a system which utilizes λ -calculus to generate the meaning representation of texts.

Another logic-based representation, but for sentences, is the AMR (Banarescu et al. 2013). The sentence is represented as a single-rooted graph (Damonte et al. 2017). The syntax of AMRs can be defined recursively, and it is possible to specify a systematic translation to first-order logic. AMRs without recurrent variables are in the decidable two-variable fragment of FOL. The AMR Bank (Banarescu et al. 2013), based on this form of representation, is a set of English sentences paired with simple, readable semantic representations in some cases manually constructed by human annotators, but also generated in a semi-automatic manner (Gruzitis et al. 2018). It is possible to represent narratives using AMR as well. Droog-Hayes et al. (2018), for instance, employ AMR to represent Russian folktales and extract the narrative structure.

6.3 Visual representation

Narrative visualization addresses the interplay between narrative and visualization. This includes perspectives that range from the use of visualization elements to enrich traditional narratives, to the exploration of narrative techniques in visualization-rich artefacts. In this section, we focus on the visual and structured representation of a narrative as an important step, not only as a final output of the narrative extraction pipeline, but also as a machine-readable representation that can be used as a tool for human inspection and validation of each of the pipeline steps.

Broadly speaking, narrative visualization can be mapped into 7 genres (Segel and Heer 2010): magazine-style, annotated chart, partitioned poster, flow chart, comic strip, slide show, and film/video/animation. As stated by the authors, such genres vary primarily in terms of the number of frames - distinct visual scenes, multiplexed in time and/or space - that each contains, and the ordering of their visual elements. None of these genres are mutually exclusive: they can act as components and be combined to create more complex visual genres.

A less explored area is the visualization of narratives themselves, a perspective of particular interest in the context of this survey, where the focus is on the representation of the extracted narrative as closely as possible to its original form. Metilli et al. (2019) developed a semi-automatic software that can import knowledge from Wikidata to allow users to construct and visualize narratives, based on a proposed ontology to annotate narratives. Baikadi et al. (2011), in turn, proposed a framework for visualizing a narrative, and introduced an environment designed to explore narrative visualization to support novice writers. Milon-Flores et al. (2019), instead, applied algorithms to extract emotions and characters involved in literary work to propose a methodology that is able to generate audiovisual summaries by the combination of emotion-based music composition and graph-based animation.

The representation of narrative participants and events over time can be accomplished by projecting these elements in two-dimensional layouts. Munroe (2009) introduced the concept of a narrative chart, a visual representation that encodes narrative elements and interactions in a representation where lines flow from left (past) to right (future) representing each participant, and events are depicted using ellipsis to which participants are connected. Kim et al. (2018) presented the Story Explorer, a visualization tool to explore and communicate nonlinear narratives through the representation of story curves, which contrast story order (y-axis) with narrative order (x-axis). Story Explorer is used to analyze and discuss narrative patterns in 10 popular nonlinear movies.

Less elaborate visual schemes can also be used to represent the elements from a narrative and provide an accessible understanding of it. A simpler, but powerful resource to represent the elements of narrative, their relationships, and the chronological order of them is based on the use of knowledge graphs (Ehrlinger and Wöß 2016) and their linkage to knowledge-bases (von Landesberger et al. 2011). Li et al. (2018), for instance, employed a graph structure to extract and link narrative events. Amorim et al. (2021) in their work proposed the Brat2Viz tool to build a visual representation of a knowledge graph from a DRT representation of a narrative. A useful visual representation, in this context, is also the Message Sequence Chart (MSC) (Harel and Thiagarajan 2003), a diagram created to specify systems requirements. Due to its flexibility, MSCs are also suitable to schematize other kinds of processes. In Brat2Viz tool (Amorim et al. 2021), the authors generate a MSC from a DRT as well. Further initiatives also explore MSC for narrative visualizations (Palshikar et al. 2019; Hingmire et al. 2020).

A discussion of the challenges associated with narrative visualization, and the general field of information visualization, can be found in de Ponte Figueiras (2016), and also in Tong et al. (2018). The work of Edmond and Bednarz (2021) also addresses the multidisciplinary nature of the field and outline possible trajectories based on the emphasis put on the narrative itself or the data visualization components used.

7 Evaluation

With the growing maturity and understanding of the narrative extraction process, new procedures have come into play to help formalizing the narrative extraction evaluation step. In particular, shared tasks were proposed to establish a common experimental setup and overcome the lack of datasets. In this section, we discuss the evaluation carried out by studies developed in the area.

Representations made of the extracted narratives need to be evaluated for their level of abstraction to the application (e.g., health care systems, games, news, etc.). The evaluation of the extracted narrative is an essential step of its understandability (Riedl and Young 2010) because it aims to observe the content's understanding by the final consumers and specialists, which can provide insights and possible improvements. However, computational evaluation presents hard challenges. In the literature, few studies evaluate narratives computationally extracted or generated since this is a subjective and application-specific task.

In cases where the evaluation is considered, a manual evaluation method is usually adopted. In this case, the narrative result is delivered to specific (usually hired) people, who evaluate the generated result from pre-defined domain criteria, as it is done in Motwani et al. (2019), and give their feedback and overall perception of it. However, ways of automatically evaluating the extracted narratives are sought. Goyal et al. (2010), for example, evaluated the representation of narratives texts through plot units measuring the F-score achieved by the model. Other works (Metilli et al. 2019; Zahid et al. 2019) also assessed the tasks developed in the context of narratives in terms of this measure. Research works that use this measure have the advantage of observing a given task's performance isolated, i.e., separately.

Narratives, however, need to be evaluated as a whole. Narrative Cloze Test is a common evaluation framework for script learning introduced by Chambers and Jurafsky (2008). Based on the cloze task (Taylor 1953), the narrative cloze evaluation approach consists of a sequence of narrative events in a document where one event has been removed. The evaluation is then done by predicting the missing verb and typed dependency. In the context of narratives, this approach became widely used to explore the commonsense reasoning of narratives outcome. In a similar approach, Mostafazadeh et al. (2016a) proposed the Story Cloze Test. The authors presented this measure as a generic story understanding evaluation framework that can also evaluate story generation models (e.g., by calculating the log-likelihoods assigned by the story generation model to the two ending alternatives), which does not necessarily indicate a requirement for explicit narrative knowledge learning. According to the authors, models that perform well in the Story Cloze Test reveal some deeper understanding of the story. In this work, Mostafazadeh et al. (2016a) also presented a corpus of 50k five-sentence commonsense stories (ROCStories (Mostafazadeh 2016)) developed to enable a brand new framework for evaluating story understanding - bringing a great contribution to the narratives field.

Another form of evaluating narratives was proposed by Kočiský et al. (2018) in the NarrativeQA Reading Comprehension Challenge. In this challenge, one is presented to a dataset and a respectively set of tasks in which questions about stories must be answered. As pointed out by the authors, these tasks are designed so that successfully answering their questions requires understanding of the underlying narrative rather than relying on shallow pattern matching or salience. Approaches like these are essential to the reasoning of narratives, in the sense that they aim to provide understandable narratives to the end-user

focusing on the outcome and in the causality between the events. Another aspect that should be considered when evaluating a narrative is from whose point of view the story is being told (Brahman et al. 2021).

As discussed in the beginning of this survey, having access to appropriate datasets is one of the most important steps behind any evaluation procedure. In this context, we make available a summarized list of the most important datasets in several languages that suit the evaluation tasks approached in this survey. We refer the interested reader to check the supplementary material of this paper.

8 Open issues in the narrative extraction pipeline

As exposed in the previous sections, the extraction of narratives comprehends a series of interrelated tasks related to different areas, which amount to an intricate and complicated enterprise. As such, and although much has been accomplished, different challenges need to be met. In this section, we refer to the most prominent, organizing them into general and narrative extraction-oriented.

8.1 General Issues

8.1.1 Complexity of narratives

Human statements frequently contain ambiguity, error, implicit information, and other sources of complexity. Therefore, the creation of cognitive agents with human-level natural language understanding capabilities involves mimicking human cognition (McShane 2017). As narratives are composed of human declarations, the study of narratives presents a high level cognitive challenge. Some of the problems inherent to narratives are coreference resolution, for instance, when the subject is null, polysemy (i.e., the multiplicity of meanings of a word or phrase), synonymy (i.e., the expression of the same idea with different terms), ambiguity, related, not only to polysemy, but also to syntactic structure, presuppositions or sarcasm. Natural Language Understanding, a research area of NLP, can aid in the resolution of some of these issues. However, a model able to deal with the aforementioned issues, and to perform at a level proximate to human reasoning is still far from being developed.

8.1.2 Narratives across documents

As demonstrated in this survey paper, narratives are formed by a series of links between involved participants and events, and organized according to their causality over time, which may mean that different documents that report the same event can compose a familiar narrative. However, few works found in the literature of the area explore the extraction of narratives from multiple documents. Future research should propose new methods to automatically identify, interpret, and relate the different elements of a narrative, which will likely come from various sources. This dilemma is related to the fact that current models are often centered on recurrent neural networks (which, although can represent contexts longer than other types of networks, still have limitations for this type of strategy). Working in vast environments (i.e., contexts) is closely connected to natural language and requires scaling up internal processes before larger documents can be handled.

8.1.3 Low-resource languages

Although narratives are present in a set of diverse languages, most of the data available, however, is mainly in English. One consequence of this is that low-resource languages end up receiving far less attention. For example, while the task of temporal information extraction has drawn much attention in recent years, research efforts have mostly focused in the English language (Mirza 2016), and the development of new solutions for less known languages ends up being compromised by the lack of properly annotated datasets.

8.2 Narrative extraction issues

Pre-processing and parsing, and the identification and extraction of narrative components, are two of the components that present the best results in the process of narrative extraction. This stems from the fact that the tasks covered by these processing stages have many applications and have been studied across several NLP tasks over the years. The remaining tasks of the narrative extraction pipeline, however, still present many challenges for further enhancements. In the following, we describe some of those challenges, highlighting future directions whenever appropriate.

8.2.1 Datasets annotation

The annotation process is a key element when working on narrative extraction. Since a significant number of the tasks from the narrative extraction pipeline hinge on the existence of large annotated datasets, the effort of annotation is huge. One of the biggest issue is precisely the lack of manually annotated datasets available with the necessary information. Checking the manual annotations can also be troublesome because, often, the multilayer annotation is so dense that the annotator can barely unravel what was annotated. A useful solution to this is to resort to visual representations such as knowledge graphs and message sequence charters to carry out the supervising task. Nonetheless, further developments on these visual representations are needed. Another problem is related to the inter-annotator agreement. As stated before, narratives are complex and not always straightforward, which reinforces the relevance of assessing the level of agreement between annotators, both during the process of creating the annotated datasets and evaluating the annotation performed by means of automatic methods.

8.2.2 Temporal reasoning and a cross-lingual approach

Several challenges related to the temporal domain may be understood as possible reasons that prevent the development of more elaborated solutions in different languages. For instance, the free-text temporal expression (Strötgen and Gertz 2016), which deals with general phrases not covered by standard definitions, the resolution of implicit or relative temporal expressions, or the problem posed in normalizing temporal expressions across different time-zones, are among some of the most well-known reasons. Thus, the most promising approaches rely on cross-lingual transformer language models and cross-lingual sentence embeddings that leverage universal language commonalities.

8.2.3 Incorporating domain knowledge into narrative reasoning systems

From the linguistic point of view, some texts have a stereotypical structure already well established in the area (as is the case of medical reports). The knowledge of the domain is essential to understand the structure and context of a document to improve the reasoning process. The integration of different knowledge might provide, for example, a hierarchical structure of terminology that can help to identify whether two different statements refer to the same event. For a semantic and pragmatic analysis, other computational tools, such as automated reasoning systems, are useful. These tools may help to explore how inference processes can clarify pragmatic phenomena such as conversational implications and more explicitly context-based understanding.

8.2.4 Semantic relations in narratives

Most text processing tools concentrate on extracting relatively simple constructs from the local lexical context and focus on the document as a unit or even smaller units such as sentences or phrases, rather than on relations of different elements within the document or even cross-document relations. Correctly inducing semantic relations, between participants, participants and events or between text spans, within a natural language story is also an open issue in narrative studies. For instance, the task of discourse parsing has still a long road to go (Morey et al. 2017), to achieve discourse relation identification beyond local dependencies.

8.2.5 Open information extraction for narratives construction

Implementing a combination of three relation extraction techniques (machine learning, heuristics, and a hybrid combination of both) may be an interesting approach to be explored in the context of narratives. As narratives result from a semantic structuring of information extracted from a text, which presents n-ary relations among themselves, the Open IE exploration might provide meaningful insights to narratives studies. As far as we know, nothing of this kind has yet been proposed.

8.2.6 Narratives perspective

The narrative perspective, also known as the point of view, is the vantage point through which the events of a story are filtered and then transmitted to the audience. Thus, the same story might have different point of views (POV) depending on the person (narrator/ character) who narrates the story or the angle from which one looks (Al-Alami 2019; Brahman et al. 2021). Extracting narratives from different points of view using the same dataset, combining them in a coherent representation of the story or deciding on which are the most relevant are still open challenges. Considering different narrative points of view also influences the task of evaluating narratives.

8.2.7 Narratives representations

Without using some form of a semantic representation that offers an abstraction from the details of lexical and syntactic realizations, comprehension is arguably unattainable.

Representing embedded stories - stories told inside a story—is also a complex task that remains open. Gervás (2021) proposed a simple model to represent them. Nevertheless, the author stated that there is substantial further work in the depicted model. Narrative visualization is an active area of research, mostly as a result of a strong community focused on interactive storytelling and the use of data visualization techniques to improve narrative understanding. In contrast, there is ample opportunity for further research in the visual representation of narratives themselves. Existing research is mostly focused on the representation of narrative elements such as participants and events, whereas work on commonly used narrative techniques—e.g., focalization, allegory, personification, amplification—is scarce and represents a challenging opportunity for future research. With the development of the field, research on the proposal of common visual vocabularies and patterns for recurring visual solutions also constitute a pertinent opportunity for further research. Another aspect where research opportunities exist is the study of user interaction with narratives—which degree of manipulation is useful to improve narrative understanding? Which elements and dimensions of a visual representation should be open to user interaction? Finally, evaluating the effectiveness of the visualization of a narrative typically resort to user studies involving ad-hoc tasks—an opportunity for the development of reference evaluation benchmarks and guidelines exists.

8.2.8 Evaluation

Coming up with a framework that evaluates the narrative extraction pipeline as a whole is a crucial step for further developments. One of the difficulties refers to the process of creating a gold standard dataset, a labor intensive task that is highly dependent on the subjective interpretation of several factors, such as the emphasis put on each narrative participant, or the level of detail to include. Raters may also have differing viewpoints on annotating data or disagree about what details should be kept, contributing to a lower consensus between raters and intra-raters. On the application level, the purposes of evaluation may vary according to the context addressed. In general, the proposed systems should answer questions like: What is this narrative about? Does the system accurately extract the required elements to tell a story from it properly? How accurately and efficiently does the present system represents the narrative in terms of clarity and amount of data to the application context? The processing of narrative in large-scale or from multiple documents also bumps the issue that supervision is expensive to obtain. When considering the evaluation metrics for natural language processing, another issue is related to how well these methods match and generalize to the real complexity of human languages and how many more interesting natural language inference datasets can be generated.

9 Conclusion

Narratives are an essential tool for communication, representation, and understanding information. Computational systems that are able to identify narrative elements and structures can naturally interact with human users. Such systems understand collaborative contexts as an emerging narrative and can express themselves through storytelling (Riedl 2004). This survey paper provides simultaneously an account of the study of narrative extraction and a roadmap for future research. To this end, we propose a narrative extraction

pipeline, defining the key tasks involved in this process identified in the literature. By doing this, we set a common ground for further studies, highlighting the different stages of the narrative process and the most prominent approaches. During the course of this survey paper, we also pointed out extensive literature focused on extracting narratives and supporting NLP tasks. Nonetheless, and despite several recent advances, there are still important open issues demonstrating that narrative extraction is a rich and promising research area that requires multidisciplinary knowledge in the crossroads between linguistics and computation.

Acknowledgements The authors Alípio Jorge, Ricardo Campos, and Evelin Amorim are financed by the project Text2Story, financed by the ERDF European Regional Development Fund through the North Portugal Regional Operational Programme (NORTE 2020), under the PORTUGAL 2020 and by National Funds through the Portuguese funding agency, FCT - Fundação para a Ciência e a Tecnologia within project PTDC/CCI-COM/31857/2017 (NORTE-01-0145-FEDER-03185). All the authors are also financed by National Funds through the Portuguese funding agency, FCT - Fundação para a Ciência e a Tecnologia, within project LA/P/0063/2020. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES)—Finance Code 001.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abzianidze L, Bjerva J, Evang K, et al. (2017) The Parallel Meaning Bank: Towards a multilingual corpus of translations annotated with compositional meaning representations. In: Proceedings of the 15th Conference of the European Chapter of the ACL: Volume 2, Short Papers. ACL, Valencia, Spain, pp 242–247
- Adam J (1992) Les textes: types et prototypes : récit, description, argumentation, explication et dialogue. Fac. : série Linguistique/collection créée par Henri Mitterand, Nathan, Paris
- Adams G, Alsentzer E, Ketenci M, et al. (2021) What's in a summary? laying the groundwork for advances in hospital-course summarization. In: Proceedings of the 2021 Conference of the North American Chapter of the ACL: Human Language Technologies. ACL, Online, pp 4794–4811
- Agarwal P, Strötgen J, del Corro L, et al. (2018) diaNED: Time-aware named entity disambiguation for diachronic corpora. In: Proceedings of the 56th Annual Meeting of the ACL (Volume 2: Short Papers). ACL, Melbourne, Australia, pp 686–693
- Aher A, Jain K, Dan A, et al. (2010) A survey on semantic role labeling and dependency parsing. Semantic Role Labeling Task p 42
- Akbik A, Blythe D, Vollgraf R (2018) Contextual string embeddings for sequence labeling. In: Proceedings of the 27th International Conference on Computational Linguistics. ACL, Santa Fe, New Mexico, USA, pp 1638–1649
- Al-Alami S (2019) Point of view in narrative. Theory and Practice in Language Studies
- Aldawsari M, Finlayson M (2019) Detecting subevents using discourse and narrative features. In: Proceedings of the 57th Annual Meeting of the ACL. ACL, Florence, Italy, pp 4780–4790
- Allan J, Carbonell JG, Doddington G, et al. (1998) Topic Detection and Tracking Pilot Study Final Report <https://doi.org/10.1184/R1/6626252.v1>, https://kilthub.cmu.edu/articles/journal_contribution/Topic_Detection_and_Tracking_Pilot_Study_Final_Report/6626252
- Alperen MS, Tyers F (2010) South-east european times: A parallel corpus of balkan languages. In: Proceedings of the LREC workshop on exploitation of multilingual resources and tools for Central and (South-) Eastern European Languages, pp 49–53

- Amorim E, Ribeiro A, Cantante I, et al. (2021) Brat2viz: a tool and pipeline for visualizing narratives from annotated texts. In: 4th International Workshop on Narrative Extraction from Texts, pp 49–56
- Antonucci A, Vani K, Mellace S (2020) Temporal embeddings and transformer models for narrative text understanding. In: Proceedings of Text2Story—3rd Workshop on Narrative Extraction From Texts co-located with ECIR'20, pp 71–77
- Araki J (2018) Extraction of event structures from text. PhD thesis, Ph. D. thesis, Carnegie Mellon University
- Asher N (1993) Reference to abstract objects in discourse. Kluwer, Dordrecht
- Asher N, Asher NM, Lascarides A (2003) Logics of conversation. Cambridge University Press, Cambridge
- Atefeh F, Khreich W (2015) A survey of techniques for event detection in twitter. *Comput Intell* 31(1):132–164
- Attree S (2019) Gendered ambiguous pronouns shared task: boosting model confidence by evidence pooling. [arXiv:1906.00839](https://arxiv.org/abs/1906.00839)
- Auer S, Bizer K, Kobilarov G et al (2007) Dbpedia: a nucleus for a web of open data. In: Aberer K, Choi KS, Noy N et al (eds) The semantic web. Springer, Berlin, pp 722–735
- Baikadi A, Goth J, Mitchell CM, et al. (2011) Towards a computational model of narrative visualization. In: Workshops at the 7th Artificial Intelligence and Interactive Digital Entertainment Conference
- Balali A, Asadpour M, Campos R et al (2020) Joint event extraction along shortest dependency paths using graph convolutional networks. *Knowl Based Syst* 210(106):492
- Baldini Soares L, FitzGerald N, Ling J, et al. (2019) Matching the blanks: distributional similarity for relation learning. In: Proceedings of the 57th Annual Meeting of the ACL. ACL, Florence, Italy, pp 2895–2905
- Balog K (2018) Entity linking. Springer, Cham, pp 147–188
- Bamler R, Mandt S (2017) Dynamic word embeddings. In: Precup D, Teh YW (eds) Proceedings of the 34th International Conference on Machine Learning, Proceedings of Machine Learning Research, vol 70. PMLR, Sydney, NSW, Australia, pp 380–389
- Bamman D, Popat S, Shen S (2019) An annotated dataset of literary entities. In: Proceedings of the 2019 Conference of the North American Chapter of the ACL: Human Language Technologies, Volume 1 (Long and Short Papers) . ACL, Minneapolis, Minnesota, pp 2138–2144
- Banarescu L, Bonial C, Cai S, et al. (2013) Abstract meaning representation for sembanking. In: Proceedings of the 7th Linguistic Annotation Workshop and Interoperability with Discourse. ACL, Sofia, Bulgaria, pp 178–186
- Basile V, Bos J, Evang K, et al. (2012) Developing a large semantically annotated corpus. In: Proceedings of the LREC'12. European Language Resources Association (ELRA) , Istanbul, Turkey, pp 3196–3200
- Batista DS (2016 [Online]) Large-scale semantic relationship extraction for information discovery. Ph.D. thesis, Universidade de Lisboa, Lisboa, Portugal
- Berberich K, Bedathur S, Alonso O et al (2010) A language modeling approach for temporal information needs. *Advances in Information Retrieval*. Springer, Berlin, pp 13–25
- Bergmanis T, Goldwater S (2018) Context sensitive neural lemmatization with lematius. In: Proceedings of the 2018 Conference of the North American Chapter of the ACL: Human Language Technologies, Volume 1 (Long Papers) . ACL, New Orleans, Louisiana, pp 1391–1400
- Bethard S, Martin JH, Klingenstein S (2007) Timelines from text: identification of syntactic temporal relations. In: International Conference on Semantic Computing (ICSC 2007) , pp 11–18
- Bohnet B, McDonald R, Simões G, et al. (2018) Morphosyntactic tagging with a meta-BiLSTM model over context sensitive token encodings. <https://aclanthology.org/P18-1246>
- Boroş E (2018) Neural methods for event extraction. PhD thesis, Université Paris-Saclay
- Bos J (2015) Open-domain semantic parsing with boxer. In: Proceedings of the 20th Nordic Conference of Computational Linguistics (NODALIDA'15) . Linköping University Electronic Press, Sweden, Vilnius, Lithuania, pp 301–304
- Bos J, Basile V, Evang K et al (2017) The groningen meaning bank. *Handbook of linguistic annotation*. Springer, Dordrecht, pp 463–496
- Brahman F, Huang M, Tafjord O, et al. (2021) “Let your characters tell their story”: a dataset for character-centric narrative understanding. In: Findings of the Association for Computational Linguistics: EMNLP 2021. Association for Computational Linguistics, Punta Cana, Dominican Republic, pp 1734–1752. <https://doi.org/10.18653/v1/2021.findings-emnlp.150>, <https://aclanthology.org/2021.findings-emnlp.150>
- Brooke J, Hammond A, Baldwin T (2016) Bootstrapped text-level named entity recognition for literature. In: Proceedings of the 54th Annual Meeting of the ACL (Volume 2: Short Papers) . ACL, Berlin, Germany, pp 344–350

- Campos R, Jorge AM, Dias G, et al. (2012) Disambiguating implicit temporal queries by clustering top relevant dates in web snippets. In: Proceedings of the The 2012 IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technology - Volume 01. IEEE Computer Society, USA, WI-IAT '12, p 1–8
- Campos R, Dias G, Jorge AM et al (2014) Survey of temporal information retrieval and related applications. *ACM Comput Surv* 47(2):1–41
- Campos R, Dias G, Jorge A et al (2016) Gte-rank: a time-aware search engine to answer time-sensitive queries. *Inform Process Management* 52(2):273–298
- Campos R, Dias G, Jorge A et al (2017) Identifying top relevant dates for implicit time sensitive queries. *Inform Retr J* 20(4):363–398
- Campos R, Pasquali A, Jatowt A et al (2021) Automatic generation of timelines for past-web events. Springer, Cham, pp 225–242
- Caselli T, van Erp Anne-Lyse Minard M, Finlayson M, et al. (2016) Computing news storylines. <https://sites.google.com/site/newsstorylines2016/>
- Chambers N, Jurafsky D (2008) Unsupervised learning of narrative event chains. In: Proceedings of ACL-08: HLT. ACL, Columbus, Ohio, pp 789–797
- Chambers N, Jurafsky D (2009) Unsupervised learning of narrative schemas and their participants. In: Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Vol. 2. ACL, USA, ACL'09, p 602–610
- Chang AX, Manning C (2012) SUTime: a library for recognizing and normalizing time expressions. In: Proceedings of the (LREC'12) . European Language Resources Association (ELRA) , Istanbul, Turkey, pp 3735–3740
- Chatman S (1980) Story and discourse: narrative structure in fiction and film. Cornell University Press, Ithaca
- Chen D, Manning CD (2014) A fast and accurate dependency parser using neural networks. EMNLP'14. ACL, Doha, pp 740–750
- Chiaros C (2014) Towards interoperable discourse annotation. discourse features in the ontologies of linguistic annotation. In: Proceedings of the LREC'14. European Language Resources Association (ELRA) , Reykjavik, Iceland, pp 4569–4577
- Christopher Blackwell TRM (2009) Technology, collaboration, and undergraduate research
- Cieri C, Graff D, Liberman M, et al. (1999) The tdt-2 text and speech corpus. In: Proceedings of the DARPA Broadcast News workshop
- Ciotti F (2016) Toward a formal ontology for narrative. *MATLIT: Materialidades da Literatura* 4(1):29–44
- Curran J, Clark S, Bos J (2007) Linguistically motivated large-scale NLP with C & C and boxer. In: Proceedings of the 45th Annual Meeting of the ACL Companion Volume Proceedings of the Demo and Poster Sessions. ACL, Prague, Czech Republic, pp 33–36
- Damiano R, Lieto A (2013) Ontological representations of narratives: a case study on stories and actions. In: *CMN*, pp 76–93
- Damonte M, Cohen SB, Satta G (2017) An incremental parser for Abstract Meaning Representation. In: Proceedings of the 15th Conference of the European Chapter of the ACL: Volume 1, Long Papers. ACL, Valencia, Spain, pp 536–546
- de Marneffe MC, Manning CD, Nivre J et al (2021) Universal dependencies. *Comput Linguist* 47(2):255–308
- de Oliveira LFA, Pagano A, de Oliveira LES et al (2022) Challenges in annotating a treebank of clinical narratives in Brazilian Portuguese. In: Pinheiro V, Gamallo P, Amaro R et al (eds) *Computational processing of the Portuguese Language*. Springer International Publishing, Cham, pp 90–100
- de Ponte Figueiras AR (2016) How to tell stories using visualization: strategies towards narrative visualization. PhD thesis, Universidade Nova de Lisboa, Lisboa, Portugal
- Denny MJ, Spirling A (2018) Text preprocessing for unsupervised learning: why it matters, when it misleads, and what to do about it. *Political analysis* 26(2):168–189
- Derczynski L (2016) Automatically ordering events and times in text. *Studies in computational intelligence*. Springer, Cham
- Derczynski L, Nichols E, van Erp M, et al. (2017) Results of the WNUT2017 shared task on novel and emerging entity recognition. In: Proceedings of the 3rd Workshop on Noisy User-generated Text. ACL, Copenhagen, Denmark, pp 140–147
- Do QNT, Bethard S, Moens MF (2015) Adapting coreference resolution for narrative processing. In: Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, Lisbon, Portugal, pp 2262–2267. <https://aclanthology.org/D15-1271>

- Droog-Hayes M, Wiggins G, Purver M (2018) Automatic detection of narrative structure for high-level story representation. In: The 5th AISB Computational Creativity Symposium, pp 26–33
- Edmond C, Bednarz T (2021) Three trajectories for narrative visualisation. *Visual Inform* 5(2):26–40
- Ehrlinger L, Wöß W (2016) Towards a definition of knowledge graphs. *SEMANTiCS (Posters, Demos, SuCESS)* 48(1–4):2
- El-Haj M (2022) Financial narrative processing workshops. <http://wp.lancs.ac.uk/cfie/>. Accessed 17 Oct 2022
- El-Haj M, Rayson P, Alves P et al (2019) Multilingual financial narrative processing: analyzing annual reports in English, Spanish, and Portuguese. *Multilingual text analysis: challenges, models and approaches*. World Scientific, New Jersey, pp 441–463
- Elliott J (2005) Using narrative in social research: qualitative and quantitative approaches. Sage Publications, Thousand Oaks
- Eshel Y, Cohen N, Radinsky K, et al. (2017) Named entity disambiguation for noisy text. In: Proceedings of the 21st Conference on Computational Natural Language Learning (CoNLL 2017). ACL, Vancouver, Canada, pp 58–68
- Estellés-Arolas E, de Guevara FGL (2012) Towards an integrated crowdsourcing definition. *J Inf Sci* 38(2):189–200. <https://doi.org/10.1177/0165551512437638>
- Feng X, Qin B, Liu T (2018) A language-independent neural network for event detection. *Sci China Inf Sci* 61(9):92–106
- Ferreira Cruz A, Rocha G, Lopes Cardoso H (2020) Coreference resolution: toward end-to-end and cross-lingual systems. *Information* 11(2):74
- Freitas C, Santos D, Mota C, et al. (2009) Relation detection between named entities: Report of a shared task. In: Proceedings of the Workshop on Semantic Evaluations: Recent Achievements and Future Directions. ACL, Boulder, Colorado, SEW'09, p 129–137
- Gervás P (2021) A model of interpretation of embedded stories. In: *Text2Story@ ECIR*, pp 23–31
- Gessler L, Behzad S, Liu YJ, et al. (2021) DisCoDisCo at the DISRPT2021 shared task: a system for discourse segmentation, classification, and connective detection. In: Proceedings of the 2nd Shared Task on Discourse Relation Parsing and Treebanking (DISRPT 2021). ACL, Punta Cana, Dominican Republic, pp 51–62
- Goldhahn D, Eckart T, Quasthoff U (2012) Building large monolingual dictionaries at the leipzig corpora collection: from 100 to 200 languages. In: Proceedings of the LREC'12. European Language Resources Association (ELRA), Istanbul, Turkey, pp 759–765
- Goyal A, Riloff E, Daumé H III (2010) Automatically producing plot unit representations for narrative text. *EMNLP'10*. ACL, Cambridge, MA, pp 77–86
- Goyal A, Gupta V, Kumar M (2018) Recent named entity recognition and classification techniques: a systematic review. *Comput Sci Rev* 29:21–43
- Graën J, Bertamini M, Volk M, et al. (2018) Cutter—a universal multilingual tokenizer. In: *CEUR Workshop Proceedings, CEUR-WS*. CEUR-WS, Aachen, 2226, pp 75–81
- Gruber T (2008) *Ontology*. Encyclopedia of database systems
- Gruber TR (1993) A translation approach to portable ontology specifications. *Knowl Acquis* 5(2):199–220
- Gruzitis N, Pretkalnina L, Saulite B, et al. (2018) Creation of a balanced state-of-the-art multilayer corpus for NLU. In: Proceedings of the LREC'18. European Language Resources Association (ELRA), Miyazaki, Japan
- Han R, Ning Q, Peng N (2019a) Joint event and temporal relation extraction with shared representations and structured prediction. In: Proceedings of the EMNLP'19 and IJCNLP'19 Conferences. ACL, Hong Kong, China, pp 434–444
- Han X, Gao T, Yao Y, et al. (2019b) OpenNRE: an open and extensible toolkit for neural relation extraction. In: Proceedings of the EMNLP'19 and the IJCNLP'19 Conferences. ACL, Hong Kong, China, pp 169–174
- Harel D, Thiagarajan PS (2003) *Message sequence charts*. Kluwer Academic Publishers, USA, pp 77–105
- Hasegawa T, Sekine S, Grishman R (2004) Discovering relations among named entities from large corpora. In: Proceedings of the 42nd Annual Meeting on ACL. ACL, Barcelona, Spain, ACL '04, p 415–es
- Hashimoto K, Xiong C, Tsuruoka Y, et al. (2017) A joint many-task model: growing a neural network for multiple NLP tasks. In: Proceedings of the EMNLP'17. ACL, Copenhagen, Denmark, pp 1923–1933
- He L, Lee K, Levy O, et al. (2018) Jointly predicting predicates and arguments in neural semantic role labeling. In: Proceedings of the 56th Annual Meeting of the ACL (Volume 2: Short Papers). ACL, Melbourne, Australia, pp 364–369
- Hendrickx I, Kim SN, Kozareva Z, et al. (2010) SemEval-2010 task 8: multi-way classification of semantic relations between pairs of nominals. In: Proceedings of the 5th International Workshop on Semantic Evaluation. ACL, Uppsala, Sweden, pp 33–38

- Henry S, Buchan K, Filannino M et al (2020) 2018 n2c2 shared task on adverse drug events and medication extraction in electronic health records. *J Am Med Inf Assoc* 27(1):3–12
- Hingmire S, Ramrakhiani N, Singh AK, et al. (2020) Extracting message sequence charts from Hindi narrative text. In: *Proceedings of the 1st Joint Workshop on Narrative Understanding, Storylines, and Events*, pp 87–96
- Ide N (2017) Case study: the manually annotated sub-corpus. Springer, Netherlands, pp 497–519
- Ide N, Reppen R, Suderman K (2002) The American national corpus: more than the web can provide. In: *Proceedings of the LREC'02. European Language Resources Association (ELRA)*, Las Palmas, Canary Islands - Spain
- ISO 24617-4:2014, (2014) Language resource management—semantic annotation framework—part 4: semantic roles (semaf-sr). Standard, International Organization for Standardization, Geneva, CH
- ISO 24617-9:2019 (E), (2000) Language resource management—semantic annotation framework—Part 9: reference annotation framework. Standard, International Organization for Standardization, Geneva, CH
- Jabbar A, Iqbal S, Tamimy MI et al (2020) Empirical evaluation and study of text stemming algorithms. *Artif Intell Rev* 53(8):5559–5588
- Jatowt A, Au Yeung CM, Tanaka K (2013) Estimating document focus time. In: *Proceedings of the 22nd ACM International Conference on Information & Knowledge Management*. ACM, New York, NY, USA, CIKM'13, p 2273–2278
- Jia Z, Pramanik S, Saha Roy R, et al. (2021) Complex temporal question answering on knowledge graphs. *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*
- Jiang F, Cohn T (2021) Incorporating syntax and semantics in coreference resolution with heterogeneous graph attention network. In: *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Association for Computational Linguistics, Online, pp 1584–1591, <https://doi.org/10.18653/v1/2021.naacl-main.125>, <https://aclanthology.org/2021.naacl-main.125>
- Jindal P, Roth D (2013a) End-to-end coreference resolution for clinical narratives. In: *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence*. AAAI Press, Beijing, China, IJCAI'13, p 2106–2112
- Jindal P, Roth D (2013b) Extraction of events and temporal expressions from clinical narratives. *J Biomed Inf* 46:S13–S19. 2012 i2b2 NLP Challenge on Temporal Relations in Clinical Data
- Jones CB, Purves RS (2008) Geographical information retrieval. *Int J Geograph Inf Sci* 22(3):219–228
- Jurafsky D, Martin JH (2009) *Speech and language processing*, 2nd edn. Prentice-Hall Inc, Upper Saddle River
- Kalyanam J, Velupillai S, Conway M, et al. (2016) From event detection to storytelling on microblogs. In: *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, IEEE, pp 437–442
- Kamp H, Reyle U (1993) *From discourse to logic: introduction to modeltheoretic semantics of natural language, formal logic and discourse representation theory*, 1st edn. Springer, Netherlands
- Kantor B, Globerson A (2019) Coreference resolution with entity equalization. In: *Proceedings of the 57th Annual Meeting of the ACL*. ACL, Florence, Italy, pp 673–677
- Khan F, Bellandi A, Benotto G, et al. (2016) Leveraging a narrative ontology to query a literary text. In: *CMN. Schloss Dagstuhl - Leibniz-Zentrum für Informatik*, Germany
- Kim NW, Bach B, Im H et al (2018) Visualizing nonlinear narratives with story curves. *IEEE Trans Visualiz Comput Gr* 24(1):595–604
- Kirstain Y, Ram O, Levy O (2021) Coreference resolution without span representations. [arXiv:2101.00434](https://arxiv.org/abs/2101.00434)
- Kitaev N, Cao S, Klein D (2019) Multilingual constituency parsing with self-attention and pre-training. In: *Proceedings of the 57th Annual Meeting of the ACL*. ACL, Florence, Italy, pp 3499–3505
- Kočiský T, Schwarz J, Blunsom P et al (2018) The NarrativeQA reading comprehension challenge. *Trans ACL* 6:317–328
- Koehn P (2005) Europarl: a parallel corpus for statistical machine translation. In: *MT summit*, Citeseer, pp 79–86
- Kontostathis A, Galitsky LM, Pottenger WM et al (2004) A survey of emerging trend detection in textual data mining. *Survey of text mining*. Springer, New York, pp 185–224
- Krause T, Zeldes A (2014) ANNIS3: a new architecture for generic corpus query and visualization. *Digi Scholarsh Humanit* 31(1):118–139
- Kroll H, Plötzky F, Pirklbauer J, et al. (2022) What a publication tells you—benefits of narrative information access in digital libraries. In: *Proceedings of the 22nd ACM/IEEE Joint Conference on Digital Libraries*. Association for Computing Machinery, New York, NY, USA, JCDL'22, <https://doi.org/10.1145/3529372.3530928>

- Labatut V, Bost X (2019) Extraction and analysis of fictional character networks: a survey. *ACM Comput Surv.* <https://doi.org/10.1145/3344548>
- Lai VD, Van Nguyen M, Kaufman H et al (2021) Event extraction from historical texts: a new dataset for black rebellions. *Find Assoc Comput Linguist: ACL-IJCNLP 2021*:2390–2400
- Larson RR (1996) Geographic information retrieval and spatial browsing. *Geographic information systems and libraries: patrons, maps, and spatial information* [papers presented at the 1995 Clinic on Library Applications of Data Processing, April 10–12, 1995]
- LDC (2005) Ace (automatic content extraction) English annotation guidelines for events
- Lee H, Chang A, Peirsman Y et al (2013) Deterministic coreference resolution based on entity-centric, precision-ranked rules. *Comput Linguist* 39(4):885–916
- Lee IT, Pacheco ML, Goldwasser D (2021) Modeling human mental states with an entity-based narrative graph. *Proceedings of the 2021 Conference of the North American Chapter of the ACL: Human Language Technologies*
- Leeuwenberg A, Moens MF (2018) Temporal information extraction by predicting relative time-lines. In: *Proceedings of the EMNLP'18. ACL, Brussels, Belgium*, pp 1237–1246
- Leeuwenberg A, Moens MF (2019) A survey on temporal reasoning for temporal information extraction from text. *J Artif Intell Res* 66:341–380
- Leeuwenberg A, Moens MF (2020) Towards extracting absolute event timelines from English clinical reports. *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 28:2710–2719
- Li J, Liu M, Qin B et al (2021) A survey of discourse parsing. *Front Comput Sci* 16:165329
- Li J, Sun A, Han J et al (2022) A survey on deep learning for named entity recognition. *IEEE Trans Knowl Data Eng* 34(1):50–70. <https://doi.org/10.1109/TKDE.2020.2981314>
- Li X, Yin F, Sun Z, et al. (2019) Entity-relation extraction as multi-turn question answering. *arXiv preprint arXiv:1905.05529*
- Li Z, Ding X, Liu T (2018) Constructing narrative event evolutionary graph for script event prediction. In: *Proceedings of the 27th International Joint Conference on Artificial Intelligence. AAAI Press, Stockholm, Sweden, IJCAI'18*, p 4201–4207
- Lim CG, Jeong YS, Choi HJ (2019) Survey of temporal information extraction. *JIPS* 15(4):931–956
- Lin Z, Ng HT, Kan MY (2014) A pdtb-styled end-to-end discourse parser. *Nat Lang Eng* 20(2):151–184
- Lv X, Guan Y, Yang J et al (2016) Clinical relation extraction with deep learning. *Int J Hybrid Inf Technol* 9(7):237–248
- Mani I (2012) Computational modeling of narrative. *Synth Lectures Human Lang Technol* 5(3):1–142
- Mani I (2014) *Computational Narratology*, De Gruyter, Berlin, Munchen, Boston, pp 84–92
- Mani I, Schiffman B (2005) Temporally anchoring and ordering events in news. *Time and Event Recognition in Natural Language* John Benjamins
- Mani I, Wilson G (2000) Robust temporal processing of news. In: *Proceedings of the 38th Annual Meeting of the ACL. ACL, USA*, pp 69–76
- Mann W, Thompson S (1988) Rhetorical structure theory: toward a functional theory of text organization. *Text* 8:243–281
- Marcus MP, Marcinkiewicz MA, Santorini B (1993) Building a large annotated corpus of English: the Penn treebank. *Comput Linguist* 19(2):313–330
- Mariko D, Abi-Akl H, Trottier K, et al. (2022a) The financial causality extraction shared task (FinCausal 2022). In: *Proceedings of the 4th Financial Narrative Processing Workshop @LREC2022. European Language Resources Association, Marseille, France*, pp 105–107. <https://aclanthology.org/2022.fnp-1.16>
- Mariko D, Abi-Akl H, Trottier K, et al. (2022b) Fincausal. <https://github.com/yseop/YseopLab>
- Martins B, Manguinhas H, Borbinha J (2008) Extracting and exploring the geo-temporal semantics of textual resources. In: *2008 IEEE International Conference on Semantic Computing*, pp 1–9
- McClosky D, Surdeanu M, Manning CD (2011) Event extraction as dependency parsing. In: *Proceedings of the 49th Annual Meeting of the ACL: Human Language Technologies—Vol.1. ACL, Portland, Oregon, USA, HLT'11*, p 1626–1635
- McEnery T, Xiao R, Tono Y (2006) *Corpus-based language studies: an advanced resource book. Routledge applied linguistics*, Routledge
- McShane M (2017) Natural language understanding (nlu, not nlp) in cognitive systems. *AI Magazine* 38(4):43–56
- Meghini C, Bartalesi V, Metilli D (2021) Representing narratives in digital libraries: the narrative ontology. *Semantic Web* 12(2):241–264
- Metilli D, Bartalesi V, Meghini C (2019) Steps towards a system to extract formal narratives from text. In: *Text2Story@ ECIR*, pp 53–61

- Mihaylov T, Frank A (2019) Discourse-aware semantic self-attention for narrative reading comprehension. In: Proceedings of the EMNLP'19 and the IJCNLP'19 Conference. ACL, Hong Kong, China, pp 2541–2552
- Milon-Flores DF, Ochoa-Luna J, Gomez-Nieto E (2019) Generating audiovisual summaries from literary works using emotion analysis. In: 2019 32nd SIBGRAPI Conference on Graphics, Patterns and Images, pp 31–38
- Mirza P (2016) Extracting temporal and causal relations between events. PhD thesis, University of Trento
- Mitkov R (2014) Anaphora resolution. Routledge, London
- Morey M, Muller P, Asher N (2017) How much progress have we made on RST discourse parsing? a replication study of recent results on the RST-DT. In: Proceedings of the EMNLP'17. ACL, Copenhagen, Denmark, pp 1319–1324
- Mostafazadeh N (2016) Story cloze test and rocstories corpora. <https://www.cs.rochester.edu/nlp/rocstories/>. Accessed 20 Apr 2021
- Mostafazadeh N, Chambers N, He X, et al. (2016a) A corpus and evaluation framework for deeper understanding of commonsense stories. [arXiv:1604.01696](https://arxiv.org/abs/1604.01696)
- Mostafazadeh N, Grealish A, Chambers N, et al. (2016b) Caters: Causal and temporal relation scheme for semantic annotation of event structures. In: Proceedings of the 4th Workshop on Events, pp 51–61
- Motta LG (2005) Análise pragmática da narrativa jornalística. In: Congresso Brasileiro de Ciências da Comunicação, Intercom, pp 1–16
- Motwani A, Naik A, Karlapalem K (2019) Extracting evidence summaries from detective novels. In: Text2Story@ ECIR, pp 5–11
- Mrini K, Deroncourt F, Tran QH, et al. (2020) Rethinking self-attention: Towards interpretability in neural parsing. In: Findings of the Association for Computational Linguistics: EMNLP 2020. Association for Computational Linguistics, Online, pp 731–742. <https://aclanthology.org/2020.findings-emnlp.65>
- Munroe R (2009) xkcd: Movie Narrative Charts
- Müller C, Strube M (2006) Multi-level annotation of linguistic data with mmax2. In: Corpus Technology and Language Pedagogy: New Resources, New Tools, New Methods
- Nadeau D, Sekine S (2007) A survey of named entity recognition and classification. *Linguisticae Investigationes* 30(1):3–26
- Nadgeri A, Bastos A, Singh K, et al. (2021) KGPool: dynamic knowledge graph context selection for relation extraction. In: Findings of the ACL: ACL-IJCNLP 2021. ACL, pp 535–548
- Ning Q, Feng Z, Roth D (2017) A structured learning approach to temporal relation extraction. In: Proceedings of the EMNLP'17. ACL, Copenhagen, Denmark, pp 1027–1037
- Nivre J, de Marneffe MC, Ginter F, et al. (2020) Universal Dependencies v2: An evergrowing multilingual treebank collection. In: Proceedings of the Twelfth Language Resources and Evaluation Conference. European Language Resources Association, Marseille, France, pp 4034–4043, <https://aclanthology.org/2020.lrec-1.497>
- O'Donnell M (2000) RSTTool 2.4—a markup tool for Rhetorical Structure Theory. In: INLG'2000 Proceedings of the 1st International Conference on Natural Language Generation. ACL, Mitzpe Ramon, Israel, pp 253–256
- Oza P, Dietz L (2021) Which entities are relevant for the story? In: Text2Story@ECIR
- Palmer DD (2007) Chapter 2: Tokenisation and sentence segmentation. In: Handbook of natural language processing
- Palshikar G, Pawar S, Patil S, et al. (2019) Extraction of Message Sequence Charts from Narrative History Text. In: Proceedings of the 1st Workshop on Narrative Understanding, pp 28–36
- Piper A, So RJ, Bamman D (2021) Narrative theory for computational narrative understanding. In: Proceedings of the EMNLP'21. ACL, Online and Punta Cana, Dominican Republic, pp 298–311
- Poesio M, Stuckardt R, Versley Y (2016) Anaphora resolution. Springer, Berlin
- Potter A (2019) The rhetorical structure of attribution. In: Proceedings of the Workshop on Discourse Relation Parsing and Treebanking 2019. ACL, Minneapolis, MN, pp 38–49
- Pradhan S, Moschitti A, Xue N et al (2012) CoNLL-2012 shared task: modeling multilingual unrestricted coreference in OntoNotes. Joint Conference on EMNLP and CoNLL—Shared Task. ACL, Jeju Island, Korea, pp 1–40
- Pradhan S, Moschitti A, Xue N, et al. (2013) Towards robust linguistic analysis using OntoNotes. In: Proceedings of the Seventeenth Conference on Computational Natural Language Learning. ACL, Sofia, Bulgaria, pp 143–152
- Prasad R, Webber B, Lee A (2018) Discourse annotation in the PDTB: the next generation. Proceedings 14th Joint ACL—ISO Workshop on Interoperable Semantic Annotation. ACL, Santa Fe, New Mexico, USA, pp 87–97


- Prince G (2019) Narratology
- Purves RS, Clough P, Jones CB, et al. (2018) Geographic information retrieval: progress and challenges in spatial search of text. Now Publishers
- Pustejovsky J, Stubbs A (2012) Natural Language Annotation for Machine Learning. No. v. 9, p. 878 in *A Guide to corpus-building for applications*. O'Reilly Media, Incorporated, Sebastopol, CA
- Pustejovsky J, Knippen R, Littman J et al (2005) Temporal and event information in natural language text. *Lang Resour Eval* 39(2–3):123–164
- Pustejovsky J, Verhagen M, Sauri R, et al. (2006) Timebank 1.2. <https://catalog.ldc.upenn.edu/LDC2006T08>. Accessed 07 July 2020
- Pustejovsky J, Bunt H, Zaenen A (2017) Designing annotation schemes: from theory to model. Springer, Netherlands, Dordrecht, pp 21–72
- Qi P, Zhang Y, Zhang Y, et al. (2020) Stanza: a python natural language processing toolkit for many human languages. In: *Proceedings of the 58th Annual Meeting of the ACL: System demonstrations*
- Qin Y, Yang W, Wang K et al (2021) Entity relation extraction based on entity indicators. *Symmetry* 13(4):539
- Quaresma P, Nogueira VB, Raiyani K et al (2019) Event extraction and representation: a case study for the Portuguese language. *Information* 10(6):205
- Raiman J, Raiman O (2018) Deeptype: multilingual entity linking by neural type system evolution. [arXiv:1802.01021](https://arxiv.org/abs/1802.01021)
- Raina V, Krishnamurthy S (2022) *Natural Language Processing*, Apress, Berkeley, pp 63–73. https://doi.org/10.1007/978-1-4842-7419-4_6
- Riedl MO (2004) Narrative planning: balancing plot and character. PhD thesis, North Carolina State University
- Riedl MO, Young RM (2010) Narrative planning: balancing plot and character. *J Artif Intell Res* 39:217–268
- Rijhwani S, Preotiuc-Pietro D (2020) Temporally-informed analysis of named entity recognition. In: *Proceedings of the 58th Annual Meeting of the ACL*. ACL, Online, pp 7605–7617
- Roemmele M, Gordon A (2018) An encoder-decoder approach to predicting causal relations in stories. In: *Proceedings of the 1st Workshop on Storytelling*, pp 50–59
- Ryan ML (2014) [Online] Space. In: Hühn P, Meister JC, Pier J et al (eds) *The living handbook of narratology*. Hamburg University, Hamburg
- Saeed Z, Abbasi RA, Maqbool O et al (2019) What's happening around the world? A survey and framework on event detection techniques on twitter. *J Grid Comput* 17(2):279–312
- Sato M, Jatowt A, Duan Y, et al. (2021) Estimating contemporary relevance of past news. In: *2021 ACM/IEEE Joint Conference on Digital Libraries (JCDL)*. IEEE Computer Society, Los Alamitos, CA, USA, pp 70–79. <https://doi.ieeecomputersociety.org/10.1109/JCDL52503.2021.00019>
- Sauri R, Littman J, Knippen B, et al. (2006) Timeml annotation guidelines version 1.2.1
- Schweter S, Ahmed S (2019) Deep-EOS: general-purpose neural networks for sentence boundary detection. In: *Proceedings of the 15th Conference on Natural Language Processing (KONVENS)*
- Segel E, Heer J (2010) Narrative visualization: telling stories with data. *IEEE Trans Visualiz Comput Gr* 16(6):1139–1148
- Shahzad M, Amin A, Esteves D, et al. (2021) Inferner: an attentive model leveraging the sentence-level information for named entity recognition in microblogs. *The International FLAIRS Conference Proceedings* 34
- Sheikhalishahi S, Miotto R, Dudley JT et al (2019) Natural language processing of clinical notes on chronic diseases: systematic review. *JMIR Med Inf* 7(2):e12239
- Shi P, Lin J (2019) Simple bert models for relation extraction and semantic role labeling. [arXiv:1904.05255](https://arxiv.org/abs/1904.05255)
- Silvano P, Leal A, Silva F, et al. (2021) Developing a multilayer semantic annotation scheme based on ISO standards for the visualization of a newswire corpus. In: *Proceedings of the 17th Joint ACL—ISO Workshop on Interoperable Semantic Annotation*. Association for Computational Linguistics, Groningen, The Netherlands (online), pp 1–13. <https://aclanthology.org/2021.isa-1.1>
- Stenetorp P, Pyysalo S, Topić G, et al. (2012) Brat: a web-based tool for NLP-assisted text annotation. In: *Proceedings of the Demonstrations at the 13th Conference of the European Chapter of the ACL*. ACL, Avignon, France, pp 102–107
- Stieglitz S, Lattemann C, Robra-Bissantz S et al (2016) *Gamification: using game elements in serious contexts*. Springer, Cham
- Strötgen J, Gertz M (2012) Event-centric search and exploration in document collections. In: *Proceedings of the 12th ACM/IEEE-CS Joint Conference on Digital Libraries*. ACM, New York, NY, USA, JCDL'12, pp 223–232
- Strötgen J, Gertz M (2013) Multilingual and cross-domain temporal tagging. *Lang Resour Eval* 47(2):269–298

- Strötgen J, Gertz M (2016) Domain-sensitive temporal tagging. *Synth Lectures Human Lang Technol* 9(3):1–151
- Strötgen J, Alonso O, Gertz M (2012) Identification of top relevant temporal expressions in documents. In: *Proceedings of the 2nd Temporal Web Analytics Workshop*. ACM, New York, NY, USA, TempWeb'12, pp 33–40
- Stubbs A (2011) MAE and MAI: lightweight annotation and adjudication tools. In: *Proceedings of the 5th Linguistic Annotation Workshop*. ACL, Portland, Oregon, USA, pp 129–133
- Suchanek FM, Kasneci G, Weikum G (2007) Yago: a core of semantic knowledge. In: *Proceedings of the 16th International Conference on World Wide Web*. ACM, New York, NY, USA, WWW'07, pp 697–706
- Sun X, Liu X, Hu J, et al. (2014) Empirical studies on the nlp techniques for source code data preprocessing. In: *Proceedings of the 2014 3rd International Workshop on Evidential Assessment of Software Technologies*. ACM, New York, NY, USA, EAST 2014, pp 32–39
- Sun Y, Cheng G, Qu Y (2018) Reading comprehension with graph-based temporal-casual reasoning. In: *Proceedings of the 27th International Conference on Computational Linguistics*. ACL, Santa Fe, New Mexico, USA, pp 806–817
- Szymański J, Naruszewicz M (2019) Review on wikification methods. *AI Commun* 32(3):235–251
- Tang B, Wu Y, Jiang M et al (2013) A hybrid system for temporal information extraction from clinical text. *J Am Med Inf Assoc* 20(5):828–835
- Tang J, Lin H, Liao M, et al. (2021) From discourse to narrative: knowledge projection for event relation extraction. [arXiv:2106.08629](https://arxiv.org/abs/2106.08629)
- Taylor WL (1953) Cloze procedure: a new tool for measuring readability. *Journal Q* 30(4):415–433
- Tjong Kim Sang EF, De Meulder F (2003) Introduction to the CoNLL-2003 shared task: Language-independent named entity recognition. In: *Proceedings of the Seventh Conference on Natural Language Learning at HLT-NAACL 2003*, pp 142–147. <https://aclanthology.org/W03-0419>
- Tong C, Roberts R, Borgo R et al (2018) Storytelling and visualization: an extended survey. *Information* 9(3):65
- Toolan MJ (2013) *Narrative: a critical linguistic introduction*, 2nd edn. Interface, Routledge, London
- Uppunda A, Cochran S, Foster J, et al. (2021) Adapting coreference resolution for processing violent death narratives. In: *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Association for Computational Linguistics, pp 4553–4559. <https://doi.org/10.18653/v1/2021.naacl-main.361>, <https://aclanthology.org/2021.naacl-main.361>
- Uzuner Ö, South BR, Shen S et al (2011) 2010 i2b2/va challenge on concepts, assertions, and relations in clinical text. *J Am Med Inf Assoc* 18(5):552–556
- UzZaman N, Allen J (2011) Temporal evaluation. In: *Proceedings of the 49th Annual Meeting of the ACL: Human Language Technologies*. ACL, Portland, Oregon, USA, pp 351–356
- UzZaman N, Llorens H, Derczynski L, et al. (2013) SemEval-2013 task 1: TempEval-3: evaluating time expressions, events, and temporal relations
- Verhagen M, Sauri R, Caselli T, et al. (2010) Semeval-2010 task 13: Tempeval-2. In: *Proceedings of the 5th International Workshop on Semantic Evaluation*. ACL, Uppsala, Sweden, pp 57–62
- Vijayarani S, Janani R et al (2016) Text mining: open source tokenization tools-an analysis. *Adv Comput Intell Int J (ACII)* 3(1):37–47
- Volk M, Göhring A, Rios A, et al. (2015) SMULTRON (version 4.0)—The Stockholm MULTilingual parallel TReebank. http://www.cl.uzh.ch/research/parallelcorpora/paralreetreebanks_en.html, an English-French-German-Quechua-Spanish-Swedish parallel treebank with sub-sentential alignments
- von Landesberger T, Kuijper A, Schreck T et al (2011) Visual analysis of large graphs: state-of-the-art and future research challenges. *Comput Gr Forum* 30:1719–1749
- Vrandečić D, Krötzsch M (2014) Wikidata: a free collaborative knowledgebase. *Commun ACM* 57(10):78–85
- Walker C, Medero SSJ, Maeda K (2006) Ace 2005 multilingual training corpus. <https://catalog.ldc.upenn.edu/LDC2006T06>. Accessed 18 May 2020
- Wang X, Ma C, Zheng H, et al. (2019) DM_NLP at SemEval-2018 task 12: A pipeline system for toponym resolution. In: *Proceedings of the 13th International Workshop on Semantic Evaluation*. ACL, Minneapolis, Minnesota, USA, pp 917–923
- Wang X, Jiang Y, Bach N, et al. (2021) Automated concatenation of embeddings for structured prediction. [arXiv:2010.05006](https://arxiv.org/abs/2010.05006)
- Webster K, Recasens M, Axelrod V, et al. (2018) Mind the gap: a balanced corpus of gendered ambiguous. In: *Transactions of the ACL*, p to appear

- Wei Q, Ji Z, Si Y, et al. (2019) Relation extraction from clinical narratives using pre-trained language models. In: AMIA Annual Symposium Proceedings, American Medical Informatics Association, p 1236
- Weissenbacher D, Magge A, O'Connor K, et al. (2019) SemEval-2019 task 12: Toponym resolution in scientific papers. In: Proceedings of the 13th International Workshop on Semantic Evaluation. ACL, Minneapolis, Minnesota, USA, pp 907–916
- Wohlin C (2014) Guidelines for snowballing in systematic literature studies and a replication in software engineering. In: Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering. Association for Computing Machinery, New York, NY, USA, EASE'14. <https://doi.org/10.1145/2601248.2601268>
- Wu Y (2019) Is automated journalistic writing less biased? An experimental test of auto-written and human-written news stories. *Journalism Practice* 0 (0) :1–21
- Xavier CC, Strube de Lima VL, Souza M (2015) Open information extraction based on lexical semantics. *J Braz Comput Soc* 21(1):4
- Xiang W, Wang B (2019) A survey of event extraction from text. *IEEE Access* 7:173111–173137
- Xie L, Sundaram H, Campbell M (2008) Event mining in multimedia streams. *Proceedings of the IEEE* 96(4):623–647
- Yan B, Janowicz K, Mai G, et al. (2017) From itdl to place2vec: Reasoning about place type similarity and relatedness by learning embeddings from augmented spatial contexts. In: Proceedings of the 25th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. ACM, New York, NY, USA, SIGSPATIAL'17
- Yang Z, Dai Z, Yang Y, et al. (2019) Xlnet: Generalized autoregressive pretraining for language understanding. In: *Advances in neural information processing systems*, pp 5754–5764
- Yao W, Huang R (2018) Temporal event knowledge acquisition via identifying narratives. In: Proceedings of the 56th Annual Meeting of the ACL (Volume 1: Long Papers) . ACL, Melbourne, Australia, pp 537–547
- Yu HY, Kim MH (2021) Automatic event extraction method for analyzing text narrative structure. In: 2021 15th International Conference on Ubiquitous Information Management and Communication (IMCOM) , IEEE, pp 1–4
- Zacks JM, Tversky B (2001) Event structure in perception and conception. *Psychol Bull* 127(1):3
- Zahid I, Zhang H, Boons F, et al. (2019) Towards the automatic analysis of the structure of news stories. In: *Text2Story@ ECIR*, pp 71–79
- Zeldes A (2018) *Multilayer corpus studies*. Routledge, London
- Zeldes A, Das D, Maziero EG, et al. (2019) Introduction to discourse relation parsing and treebanking (DISRPT) : 7th workshop on Rhetorical Structure Theory and related formalisms. In: Proceedings of the Workshop on Discourse Relation Parsing and Treebanking 2019. ACL, Minneapolis, MN, pp 1–6
- Zeldes A, Liu YJ, Iruskieta M, et al. (2021) The DISRPT 2021 shared task on elementary discourse unit segmentation, connective detection, and relation classification. In: Proceedings of the 2nd Shared Task on Discourse Relation Parsing and Treebanking (DISRPT 2021) . ACL, Punta Cana, Dominican Republic, pp 1–12
- Zhai F, Potdar S, Xiang B, et al. (2017) Neural models for sequence chunking. In: 31st AAAI Conference on Artificial Intelligence
- Zhang H, Boons F, Batista-Navarro R (2019) Whose story is it anyway? Automatic extraction of accounts from news articles. *Inf Process Management* 56(5):1837–1848
- Zhang N, Deng S, Sun Z, et al. (2019b) Long-tail relation extraction via knowledge graph embeddings and graph convolution networks. In: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers) . Association for Computational Linguistics, Minneapolis, Minnesota, pp 3016–3025. <https://doi.org/10.18653/v1/N19-1306>, <https://aclanthology.org/N19-1306>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Brenda Santana^{1,2}  · **Ricardo Campos**^{2,3} · **Evelin Amorim**² · **Alípio Jorge**^{2,4} · **Purificação Silvano**⁵ · **Sérgio Nunes**^{2,6}

Evelin Amorim
evelin.f.amorim@inesctec.pt

Purificação Silvano
msilvano@letras.up.pt

Sérgio Nunes
sergio.nunes@inesctec.pt

¹ Institute of Informatics, Federal University of Rio Grande do Sul, Av. Bento Gonçalves, 9500, Porto Alegre, Rio Grande do Sul 91509-900, Brazil

² LIAAD, INESC TEC, Rua Dr. Roberto Frias, Porto 4200-465, Porto, Portugal

³ Ci2 - Smart Cities Research Center, Polytechnic Institute of Tomar, Av. Dr. Aurélio Ribeiro 3, Tomar 2300-313, Ribatejo, Portugal

⁴ Faculty of Sciences, University of Porto, Rua do Campo Alegre, Porto 4169-007, Porto, Portugal

⁵ Faculty of Arts and Humanities, University of Porto, Via Panorâmica Edgar Cardoso, Porto 4150-564, Porto, Portugal

⁶ Faculty of Engineering, University of Porto, Rua Dr. Roberto Frias, Porto 4200-465, Porto, Portugal