



OPEN ACCESS

EDITED BY

Jennifer D'Souza,
Technische Informationsbibliothek (TIB),
Germany

REVIEWED BY

Hamed Babaei Giglou,
Technische Informationsbibliothek (TIB),
Germany
Azanzi Jiomekong,
University of Yaounde I, Cameroon

*CORRESPONDENCE

Pakpoom Buabthong
✉ pakpoom.b@nrru.ac.th

RECEIVED 26 December 2023

ACCEPTED 08 April 2024

PUBLISHED 29 April 2024

CITATION

Saetia C, Thonglong A, Amornchaiteera T,
Chalothorn T, Taerungruang S and
Buabthong P (2024) Streamlining event
extraction with a simplified annotation
framework. *Front. Artif. Intell.* 7:1361483.
doi: 10.3389/frai.2024.1361483

COPYRIGHT

© 2024 Saetia, Thonglong, Amornchaiteera,
Chalothorn, Taerungruang and Buabthong.
This is an open-access article distributed
under the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other forums is
permitted, provided the original author(s) and
the copyright owner(s) are credited and that
the original publication in this journal is cited,
in accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Streamlining event extraction with a simplified annotation framework

Chanatip Saetia¹, Areeya Thonglong²,
Thanpitcha Amornchaiteera², Tawunrat Chalothorn¹,
Supawat Taerungruang² and Pakpoom Buabthong^{3*}

¹Kasikorn Labs, Kasikorn Business-Technology Group, Nonthaburi, Thailand, ²Department of Thai, Faculty of Humanities, Chiangmai University, Chiang Mai, Thailand, ³Faculty of Science and Technology, Nakhon Ratchasima Rajabhat University, Nakhon Ratchasima, Thailand

Event extraction, grounded in semantic relationships, can serve as a simplified relation extraction. In this study, we propose an efficient open-domain event annotation framework tailored for subsequent information extraction, with a specific focus on its applicability to low-resource languages. The proposed event annotation method, which is based on event semantic elements, demonstrates substantial time-efficiency gains over traditional Universal Dependencies (UD) tagging. We show how language-specific pretraining outperforms multilingual counterparts in entity and relation extraction tasks and emphasize the importance of task- and language-specific fine-tuning for optimal model performance. Furthermore, we demonstrate the improvement of model performance upon integrating UD information during pre-training, achieving the F1 score of 71.16 and 60.43% for entity and relation extraction respectively. In addition, we showcase the usage of our extracted event graph for improving node classification in a retail banking domain. This work provides valuable guidance on improving information extraction and outlines a methodology for developing training datasets, particularly for low-resource languages.

KEYWORDS

event extraction, annotation guideline, Universal Dependencies, generative model, event graph

1 Introduction

The advent of large language models (LLMs) has enabled significant progress in the field of natural language processing (NLP) and has helped provide promising results for various tasks (Brown et al., 2020). Many types of LLM have been proposed to solve both language-specific and domain-specific tasks (Lewis et al., 2019; Chung et al., 2022; Touvron et al., 2023). However, LLMs primarily favored well-resourced language with large updated training corpora, which may lead to hallucination problems, especially in lower-resource languages, in which the training corpora are not abundantly available (Ji et al., 2023). Extracting knowledge from these low-resource languages is not only beneficial as it helps include more available data. It could also provide deeper insight into the model's behavior across linguistic variations.

To mitigate the hallucination problem, researchers have explored augmenting LLMs with external structured data sources, such as knowledge graphs (Guu et al., 2020; Asai et al., 2021; Mialon et al., 2023). Integrating structured information graphs with the LMs has been one of the common approaches (Yao et al., 2019; Kang et al., 2022), as graphs can be constructed in a domain-specific fashion, such as finance (Yang et al., 2018; Elhammadi et al., 2020).

Event graphs, which store event information from unstructured plain texts that describe “who, when, where, what, why” and “how” of the action, can provide a simplified version of a more generalized knowledge graph (Xiang and Wang, 2019; Li et al., 2022). Focusing on event extraction is particularly promising for enhancing NLP in low-resource settings because it involves parsing relationships within the narrow scope of particular events, thus requiring less extensive linguistic understanding for the model. Although close-domain event extraction, which follows specific domain schema, may provide better results in downstream retrieval tasks (Chambers et al., 2014; Björne and Salakoski, 2018; Han et al., 2018), this specialization often results in complex annotation systems that can be cumbersome and domain-restrictive, especially for low-resource languages. Moreover, while the use of additional syntactic information for extraction tasks has been studied in English (Fader et al., 2011; Wang C. et al., 2023; Wang Z. et al., 2023), it remains under-explored in low-resource languages.

In this work, we propose a methodology that streamlines the process for open-domain event extraction for corporate documents written in Thai and demonstrates its utility in a downstream task. Our guideline aims to make structured information extraction more accessible, by reducing the complexity of the annotation process. We also utilize Universal Dependencies [UD; Nivre et al., 2016] during the pre-training step to help the extraction model better understand the structural information of the sentences.

The main contributions of this work are as follows:

- **Annotation framework:** We offer a simplified annotation guideline that streamlines the event extraction process and presents a comparative analysis with the traditional Universal Dependency (UD) framework.
- **Event extraction models:** We explore the impact of language-specific and task-specific pre-training as well as the incorporation of UD on the improvement of the overall extraction performance.
- **Applications:** We demonstrate that the extracted event graph can be utilized to improve a downstream task, namely, node classification in a retail banking domain.

The rest of this paper is organized as follows. Section 2 analyzes previous work. Section 3 describes our methodology. Section 4 reports on our experiments. Section 5 provides a discussion of the results. Section 6 elaborates on the application of the event graphs. Section 7 concludes with a summary. By simplifying the initial extraction process, our method could allow for a more straightforward transition into an extraction task for other types of relations, such as, part-of or causal relations, which often require a deeper understanding of the interconnectedness of entities beyond their basic semantic relationships.

2 Related work

In this section, the background of the paper is explained along with literature reviews, outlining the previous work on event extraction, and Universal Dependencies. First, the definition and prior works of event extraction are explained. Second, the Universal Dependencies are described including the definition and its advantages.

2.1 Event extraction

Event extraction typically aims to extract event attributes from a raw, answering the 5W1H (who, what, when, where, why, and how) questions (Xiang and Wang, 2019). In earlier work, event extraction is considered a sequence labeling-based task (Gupta and Manning, 2014; Chen et al., 2020). The event trigger and its arguments are extracted as a span of words with an inside-outside-beginning (BIO) tagging system (Li et al., 2022). However, multiple events may be found in a given sentence, thus later necessitating the classification of the relation between each argument with its trigger.

The event extraction task can generally be categorized into two groups: close domain and open domain (Xiang and Wang, 2019; Liu et al., 2021, 2023). The close-domain extraction aims to extract a pre-defined structure based on supervised datasets. Most approaches first identify the event trigger, followed by its corresponding attributes (Huang et al., 2017; Xiang and Wang, 2019). Each event attributed is connected to the trigger with a pre-defined relation.

Various methods were proposed to address close-domain extraction (Chen et al., 2015a; Huang et al., 2017; Li et al., 2020). Some treat the event extraction as a sequence of sub-tasks: trigger identification, Trigger classification, argument identification, and argument role classification (Chen et al., 2015b; Yang et al., 2019; Li et al., 2022). However, this technique could lead to error propagation during the process (Li et al., 2019; Zhang et al., 2019). To minimize this error propagation, joint-trained models were proposed (Hsu et al., 2021; Lu et al., 2021). Many approaches adopt deep learning model architecture to train an end-to-end event extraction (Nguyen and Nguyen, 2019; Wadden et al., 2019). Recently, conditional generations from language models yield promising accuracy among many NLP tasks. Such models have been adopted for event extraction, achieving state-of-the-art accuracy over the complex classification models (Hsu et al., 2021; Lu et al., 2021). Nevertheless, the learning for these deep learning model approaches is supervised, necessitating a large amount of training data, which is not practical for low-resource languages.

Although the accuracy of the close-domain models is promising, most datasets are still limited to specific domains like medical data, historical documents, or specific types of news (Vanegas et al., 2015; Björne and Salakoski, 2018; Han et al., 2018). Thus, to extract a generic event from more generalized corpora, open-domain event extraction was developed (Chau et al., 2019; Liu et al., 2019). The early model considers the headline phrase as an event and disambiguates the events using Wordnet (Miller, 1995) and word sense disambiguation (Chau et al., 2019). This method leads to suboptimal performance as the

arguments of an event are not necessarily positioned next to an event trigger keyword. To address this limitation, another model utilizes an unsupervised method using a neural latent variable model to extract an event (Liu et al., 2019). However, because of its unsupervised architecture, this method is not controllable and can extract an inaccurate event entity.

2.2 Low-resource event extraction

Similar to other tasks under a low-resource setting, the development of event extraction for low-resource languages generally focuses on methods that require less amount of training data. Zero-shot learning is one of the most common approaches to help the model perform tasks without additional training samples. Previous work on zero-shot event extraction has explored the use of representation in other latent spaces such as semi-Markov conditional random fields (Lu and Roth, 2012), Abstract Meaning Representation (AMR; Huang et al., 2018), pre-defined ontological structure (Zhang et al., 2021). Alternatively, event extraction tasks may be formulated as different tasks such as question-answering (Lyu et al., 2021). However, these techniques necessitate that proficient models already exist in the target language.

On the other hand, few-shot learning can be utilized to minimize the amount of new training data that is specific to the extraction tasks, while improving the overall performance of the models. Early models use a prototypical network to classify the extracted token (Snell et al., 2017; Lai and Nguyen, 2019), or minimize the supervised training data by providing the trigger terms in the annotation guideline as seeds for each event type (Bronstein et al., 2015). More recent work addresses the issue of low sample diversity by introducing Adaptive Knowledge-Enhanced Bayesian Meta Learning (AKE-BML) that uses a prior knowledge distribution to generate the posterior distribution for each event type (Shen et al., 2021). Techniques used in a few-shot setting typically work well when there exists a known distribution within a given task followed by model refinement through additional examples in the target tasks. For example, in Thai, we can pre-train the model with a syntactic structure such as UD, then fine-tune the model with a small number of labels for event extraction.

Furthermore, cross-lingual transfer may be employed when both languages have well-established parallel corpus. Recent methods have proposed transferring the entire universal structures across languages (Li et al., 2016; Subburathinam et al., 2019; Lou et al., 2022), or leveraging multilingual embedding when training the extraction model (M'hamdi et al., 2019). However, the cross-lingual approach typically requires extensive lexical mapping which may not be suitable for this initial stage of the development.

2.3 Relation extraction

In addition to models specific to event extraction, other relation extraction models may be utilized. End-to-end deep learning models have been proposed to concurrently extract entities and relation (Bekoulis et al., 2018; Eberts and Ulges, 2019; Hang et al., 2021). SpERT (Eberts and Ulges, 2019), in particular, has shown

promising results on both entity and relation extraction evaluated over the SciERC dataset (Luan et al., 2018).

Moreover, generative pre-trained language models have been reported to achieve high performance on many NLP tasks (Brown et al., 2020; Touvron et al., 2023). Structured prediction using generative LMs, in particular, has recently attracted interest, due to their flexibility and applicability to new datasets. Most models are trained to generate structured output for named entities recognition or relation extraction from unstructured texts (Eberts and Ulges, 2019; Lu et al., 2021; Paolini et al., 2021). DeepStruct (Wang C. et al., 2023), for example, offers state-of-the-art performance when predicting the triplet from various domains, namely T-REx (Elsahar et al., 2018), TEKGEN, KELM (Agarwal et al., 2021), WebNLG (Colin et al., 2016), and ConceptNet (Speer et al., 2017).

Nevertheless, the models may not perform well in other languages that are not primarily present in the pre-training dataset. Other syntactic or semantic information, such as Universal Dependencies (UD) may assist in cross-lingual transfer of the extraction capabilities.

2.4 Universal Dependencies

Universal Dependencies (UD; Nivre et al., 2016) is a cross-language framework that allows for consistency in the annotation of syntactic grammatical structure (parts of speech, morphological features, and syntactic dependencies). Given this UD, a reliable graph can be created to represent the syntactic structure of an arbitrary text. Some event extraction models have been reported to benefit from the incorporation of UD (Björne and Salakoski, 2018; Chau et al., 2019). Unsupervised techniques can extract phrases and their relation from the UD graphs (Chau et al., 2019). Other work used the output of the UD as a graph feature along with a graph neural network to improve an event extraction model (Liu et al., 2018; Ahmad et al., 2021). Nevertheless, developing extraction models that rely too heavily on UD may pose similar limitations to those with languages that have low annotated training data, since the models may learn to capture only the explicit syntactic relationship and not the generalized semantic structure of the sentences.

3 Methodology

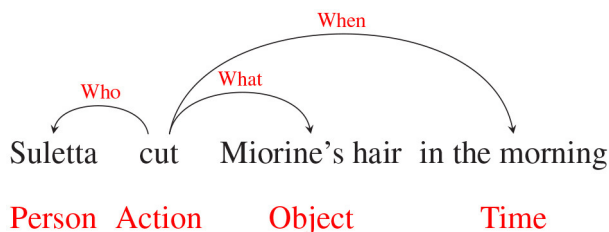
This section outlines the annotation process and the event extraction models used in this work.

3.1 Annotation framework

Frameworks for annotating text typically have two distinct aspects: (1) the practical means of how to annotate, and (2) the rules governing the annotation process (Pyysalo et al., 2012; Stenetorp et al., 2012; Cassidy et al., 2014). For (1), in this work, we configured INCEpTION (Klie et al., 2018) for entity and relation tagging. For (2), the complete annotation guideline is provided in the

Supplementary material, while the abbreviated version, along with the design reasoning, is presented as follows.

Briefly, instead of the traditional event annotation where the trigger verb is identified first, the events are tagged based on 5W1H questions. The annotation guideline proposed two-stage tagging, which first labels entity spans and then links the relations among them. An example of a fully annotated sentence is shown below.



Entities, which are graph nodes of an event graph, are extracted as triggers and their corresponding arguments, represented as word spans. These entity spans are categorized into seven types to include the semantic meaning of an entity. One of the types is denoted as *Action*, which is similar to the trigger of an event. Other types are the semantic type of the argument, like *Person*, *Object*, and *Location*.

After getting entity spans, the subsequent step is establishing and classifying the relations among the spans. The classified relation types are designed to primarily address WH questions, which are *what*, *who*, *when*, and *where*. The *how* and *why* are not included, since the phrase that describes these two relations can be highly subjective depending on the annotator. Nevertheless, we also include additional relations, namely, same-unit, benefit, and value, in the guideline as these relations are not semantically ambiguous and can be potentially useful for downstream information extraction tasks.

3.2 Event extraction models

Two candidate models are selected for the event extraction task based on their inference settings: generative and span-based classification.

Span-based joint entity and relation extraction. The two models, SpERT (Eberts and Ulges, 2019) for span-based classification and DeepStruct (Wang C. et al., 2023) for the generative approach, were selected based on their demonstrated state-of-the-art performance in their respective tasks. SpERT has shows superior performance in span-based classification tasks, benchmarked on CoNLL-2003 (Tjong Kim Sang and De Meulder, 2003). Similarly Deepstruct has exhibited strong performance using generative approach on ACE-2005 (Walker and Consortium, 2005) corpus due to their superior performance in their respective tasks.

3.2.1 Span-based classification model

For the baseline model, SpERT is used to represent a relatively more straightforward approach to the event extract task. In this approach, the model first recognizes the spans of the token of interest (entity extraction), then, with each pair of spans, learns to

classify the relation types (relation extraction). Nevertheless, both entity extraction and relation extraction are trained jointly.

3.2.2 Generative model

To study the effect of incorporating UD structure into the model, a separate model based on DeepStruct is used. The model is trained in a generative setting using a short prompt and the text of interest as the input, with the event triplets as the output (shown in Figure 1) In the UD pre-training, two tasks are trained jointly but with different prompts: part-of-speech (POS) tagging, and dependency (DEP) tagging. In contrast to the previous work where the extracted triplets are only constrained to a few important relations, each word in the input sentence of our approach will result in its own POS and DEP triplets. Note that in this generative setting, both entity extraction and relation extraction are inferred simultaneously from the model.

Since the original model is pre-trained only with the English dataset, herein we pre-trained the model on our Thai dataset (UD). After the pre-training process, the model is fine-tuned on the annotated Thai event dataset. Similar to the pre-training stage, the three tasks are trained jointly using different prompts and outputs.

4 Experiments and results

In this section, we compare the annotation time between event annotation using our proposed guideline and the traditional UD annotation. The annotated data was then used in a comparative study between different approaches to event extraction tasks.

4.1 Time for annotation

Our proposed guideline was used to annotate news articles and internal corporate documents written in Thai. To measure the time for annotation, two annotators were tasked to label the documents according to our guidelines as well as the standard Thai UD annotation for 1 month. Afterward, the number of annotated sentences for each task was divided to calculate the daily average from both annotators was averaged per day and divided by the number of days in that month. The statistics of the resulting annotated data are shown in Table 1.

The number of sentences annotated using our event extraction guideline compared to using the typical UD guideline are 292.77 and 19.2 sentences per day, respectively, indicating ~10 times faster annotation speed.

4.2 Event extraction model

The annotated event dataset was used to evaluate event extraction models described in this section. First, the dataset is split into a train, validation, and test dataset with ratios of 0.78, 0.11, and 0.11, respectively (we allocated 2,000 sentences each to the validation and test split and used the remaining for the training). To evaluate the model, the micro-average F1 score, calculated separately between the entities F1 score and

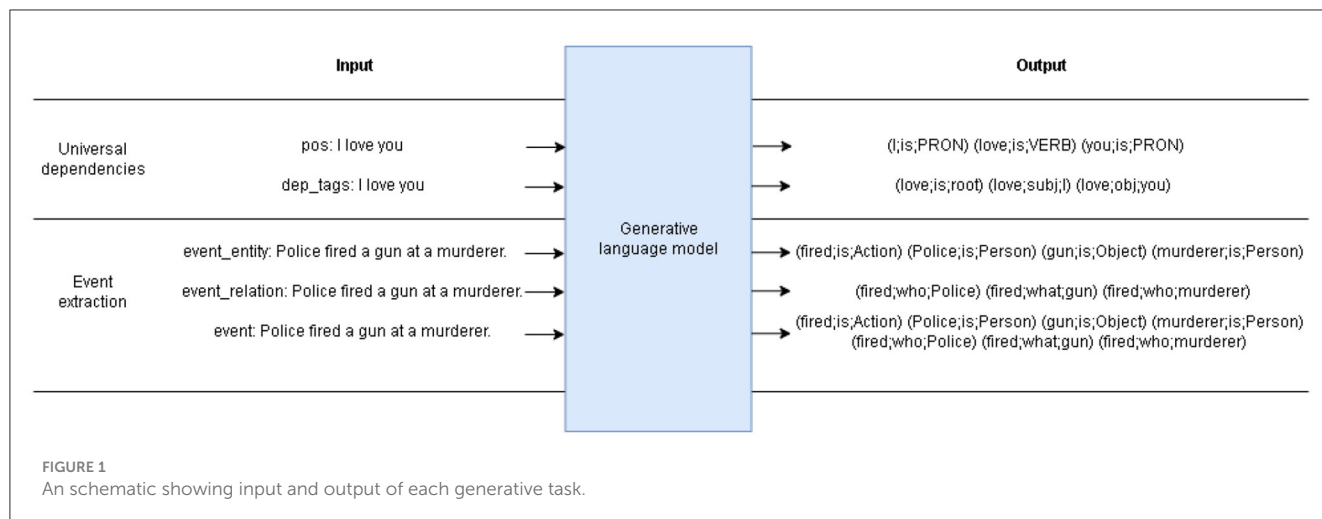


FIGURE 1 An schematic showing input and output of each generative task.

TABLE 1 The data statistics of the annotated dataset of event extraction.

Statistic	Train	Validation	Test
# sentences	13,566	2,000	2,000
# entities	15,306	2,173	2,073
# relations	9,117	1,295	1,223

the relation F1 score (Eberts and Ulges, 2019), is used. Models based on SpERT (Eberts and Ulges, 2019) and DeepStruct (Wang C. et al., 2023) are employed to compare the performance between a span-based classification model and a generative model. To study the effect of the language-specific pre-training, a multilingual BERT (Devlin et al., 2019) and a Thai-specific WangchanBERTa (Lowphansirikul et al., 2021) are used in the span-based model. Lastly, in the generative settings, the pre-training model with mT5 (Xue et al., 2021) is compared to pre-training with our Thai UD dataset. All models are fine-tuned with the annotated event training set for the event extraction task.

Table 2 shows the micro-average F1 score for entity and relation extraction. For the span-based model, using language-specific pre-training substantially outperforms the multilingual one for both entity (66.97 vs. 44.58) and relation extraction (59.20 vs. 33.86). In our settings, generative models yield better results than the span-based ones. Notably, for entity extraction, the generative model trained with the multilingual pre-training can still outperform the language-specific span-based model (68.87 vs. 66.97). Finally, the best result in both entity and relation extraction is achieved when using language-specific UD pre-training (71.16 for entity extraction and 60.43 for relation extraction).

5 Discussion

Compared to the baseline UD tagging, event annotation following our guideline is substantially faster. The decrease can be attributed not only to the fewer number of relations but also to the less complex annotation scheme that the annotators need to process. Annotating using our proposed guideline mostly

follows the semantic structure of the sentence, eliminating the need to recognize minor syntactic relations like “case,” or “disclose.” The more complicated relations between clauses like “acl,” “advcl,” “csubj,” or “xcomp” are also omitted. In addition, event annotation treats multiple-word phrases as single units, eliminating the need to understand the intraterm connection. As a result, when developing the data for structural information extraction models, starting from semantic relations similar to the proposed event extraction could be more practical and time-efficient, especially for languages with no pre-existing structural training data.

From the subsequent span-based classification result, the model using language-specific pretraining outperforms the multilingual one in both entity and relation extraction, likely attributed to both language-specific and task-specific fine-tuning. Previous work has reported that using multilingual BERT performs substantially worse for low-resource languages, like Thai, as it does not benefit from cross-lingual transfer (Wu and Dredze, 2020) and shows that monolingual BERT-based models perform even worse for NER, POS, DEP tagging. In our case, we show that fine-tuning using task- and language-specific data offers an option to improve upon the monolingual BERT-based models.

When comparing the models in different settings, although the generative model with multilingual pretraining outperforms most of the span-based ones, it still lags behind the monolingual SpERT on the relation extraction task. This discrepancy is likely because the entity recognition task can benefit from the encoder-decoder architecture used in this work. A similar observation has also been previously reported (Wu et al., 2023). Nevertheless, specific downstream tasks must be taken into account when selecting candidate baseline models, as other types, such as masked LMs, could be computationally cheaper for domain-specific training.

In contrast to entity extraction, the relation extraction task could benefit more from the span-based two-step classification architecture. While SpERT inherently approaches relation extraction as a direct classification task, the generative-based method necessitates the simultaneous learning of relation generation with the identification of the entities of interest.

TABLE 2 The result of entity and relation extraction for event extraction of each model.

	Model	P (%)	R (%)	F1 (%)
Entity	SpERT (multilingual BERT)	37.05	55.93	44.58
	SpERT (Wangchanberta)	68.27	65.71	66.97
	DeepStruct (mT5)	78.22	61.52	68.87
	DeepStruct (UD)	78.96	64.76	71.16
Relation	SpERT (multilingual BERT)	25.44	50.58	33.86
	SpERT (Wangchanberta)	60.21	58.22	59.20
	DeepStruct (mT5)	50.69	65.49	57.15
	DeepStruct (UD)	53.56	69.33	60.43

Lastly, when UD is included during the pre-training stage, the generative model outperforms in both tasks. Using UD information allows the model to learn the syntactic structure of the language, potentially aiding in the semantic inference of the subsequent relation extraction.

This result motivates the use of UD in conjunction with a more simplified event annotation framework when developing models for structure extraction, especially for low-resource languages. Although UD annotation is substantially more time-consuming, our work shows that including such information is likely beneficial to the subsequent semantic-related tasks.

6 Applications of event graphs

After obtaining the list of event attributes from the event extraction model, these sets of structured event information can be adopted to enhance other downstream tasks. In this section, we demonstrate the application of the extracted event graph to improve node classification in the retail banking product domain. Additionally, we explore the potential of transforming our event graph into a more generic knowledge graph where the types of relations are not constrained to only those present in our event annotation guideline.

The event graph in this experiment was constructed from the list of event triplets extracted using the UD-pretrained model from a set of 6,024 internal documents written in Thai, describing the details of financial products and services. This results in 69,801 nodes and 168,964 relations. Out of the total entity nodes, 500 nodes were selected and labeled into one of the 15 categories: “Process,” “Debit,” “Credit,” “Loan,” “Service,” “Promotion,” “System,” “Right,” “Fee,” “Insurance,” “Document,” “Contact,” “Account,” “Statement,” and “RewardPoint.” These nodes were selected such that the resulting 500-node sub-graphs were sufficiently connected (no disconnected graphs), and the numbers of each label were balanced. The averaged F1-score of 5-fold cross-validation of these 500-node sub-graph was then used to assess the performance of the model.

In the baseline model, only the text embedding derived from a pre-trained Thai language model, Wangchanberta (Lowphansirikul et al., 2021), was used. For our model, the node embedding derived from the event graph using Hash-GNN (Tan et al., 2020) was

TABLE 3 The comparison between models with and without node embedding as a feature.

Model	F1-macro (%)
W/o node embedding	75.87
W/ node embedding	77.71

concatenated with the original text embedding as an additional feature.

Table 3 shows the averaged F1 scores of the model using text embedding or text+node embedding as features. The result shows an ~2 percentage point improvement (77.71% from 75.87%) when the model uses node embedding in conjunction with text embedding. This improvement underscores the significance of the relational information provided by our event graph using the simple Hash-GNN. To achieve further improvement, one could employ more advanced (though computationally more expensive) node embedding techniques, namely, GCN (Kipf and Welling, 2017) or GAN (Veličković et al., 2018). In addition to the improved performance, our node classification approach adaptable to other domains and can assist organizations in processing large textual data. A similar technique could be employed to categorize entity names present in internal documents, by labeling small subset samples and then using a classification model with the extracted event graph to incorporate contextual information.

Moreover, our extracted event graph can also be merged and reformatted to construct a more generic knowledge graph. Briefly, the procedure involves finding a pair of triplets such that the head entity of one pair is the same as the tail entity of the other pair. For example, the sentence “A criminal, previously exonerated, stole a car” would be converted into {subj, rel, obj} = {A criminal, stole, a car}. By merging the triplets afterward, the model is allowed to be trained under the constraint of recognizing only seven predefined relation types, yet allowing the extracted triplets to be rearranged to cover more generalized relations. Such a generalized knowledge graph can then be applied to assist in other domain-specific or language-specific information retrieval tasks, such as question answering on knowledge graphs (KGQA; Khongcharoen et al., 2022), or KG-enhanced LLMs (Pan et al., 2023).

7 Conclusion

In this paper, we introduced a streamlined event annotation framework that allows for substantially faster labeling over the baseline UD tagging. We propose that initiating the development of data for structural information extraction models with simple semantic relations, akin to event extraction, proves more practical, particularly for languages with no pre-existing structural training data.

Language-specific pretraining helps achieve better performance over the multilingual counterparts in both entity and relation extraction tasks. Notably, we underscored the importance of fine-tuning using task- and language-specific data to improve upon monolingual BERT-based models.

Under different settings, while the generative model with multilingual pretraining generally performs well, the span-based two-step classification architecture of SpERT shows a particular advantage for relation extraction tasks. The integration of UD information during the pre-training stage further improved the performance in both tasks, indicating a potential synergistic relationship between syntactic structure understanding and subsequent semantic inference.

Moreover, we leveraged the structured event information obtained from the event extraction model to improve node classification in the retail banking product domain. We also proposed a simple method for converting our event graph into a more generic knowledge graph that expands beyond our event relation types.

In conclusion, our research underscores the value of semantic-based event extraction, language-specific pretraining, and the integration of syntactic structure understanding through UD for improved performance in structural information extraction tasks. The methods we propose are not only efficient but also versatile, with potential applications in other domains, especially for developing similar structural training data for low-resource languages.

Data availability statement

The data supporting the conclusions of this article will be made available by the authors, upon reasonable request.

Author contributions

CS: Formal analysis, Investigation, Methodology, Validation, Writing – original draft. AT: Data curation, Formal analysis,

Writing – original draft. TA: Data curation, Investigation, Writing – original draft. TC: Supervision, Writing – original draft, Writing – review & editing. ST: Conceptualization, Formal analysis, Supervision, Writing – original draft, Writing – review & editing. PB: Investigation, Supervision, Validation, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

This work was supported by Innovation Research and Development at Kasikorn Business-Technology Group (KBTG).

Conflict of interest

CS and TC were employed at Kasikorn Labs, Kasikorn Business-Technology Group. The authors declare that this study received funding from Kasikorn Business-Technology Group. The funder had the following involvement in the study: study design, data collection and analysis.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/frai.2024.1361483/full#supplementary-material>

References

Agarwal, O., Ge, H., Shakeri, S., and Al-Rfou, R. (2021). "Knowledge graph based synthetic corpus generation for knowledge-enhanced language model pre-training," in *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 3554–3565.

Ahmad, W. U., Peng, N., and Chang, K.-W. (2021). "GATE: graph attention transformer encoder for cross-lingual relation and event extraction," in *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 35, 12462–12470.

Asai, A., Yu, X., Kasai, J., and Hajishirzi, H. (2021). One question answering model for many languages with cross-lingual dense passage retrieval.

- Adv. Neural Inform. Process. Syst. 34, 7547–7560. doi: 10.48550/arXiv.2107.11976
- Bekoulis, G., Deleu, J., Demeester, T., and Develder, C. (2018). Joint entity recognition and relation extraction as a multi-head selection problem. *Expert Syst. Appl.* 114, 34–45. doi: 10.48550/arXiv.1804.07847
- Björne, J., and Salakoski, T. (2018). “Biomedical event extraction using convolutional neural networks and dependency parsing,” in *Proceedings of the BioNLP 2018 Workshop* (Melbourne, VIC), 98–108.
- Bronstein, O., Dagan, I., Li, Q., Ji, H., and Frank, A. (2015). “Seed-based event trigger labeling: How far can event descriptions get us?,” in *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Vol. 2: Short Papers)*, eds. C. Zong and M. Strube (Beijing: Association for Computational Linguistics), 372–376.
- Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., et al. (2020). Language models are few-shot learners. *Adv. Neural Inform. Process. Syst.* 33, 1877–1901. doi: 10.48550/arXiv.2005.14165
- Cassidy, T., McDowell, B., Chambers, N., and Bethard, S. (2014). “An annotation framework for dense event ordering,” in *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Vol. 2: Short Papers)*, eds. K. Toutanova and H. Wu (Baltimore, MD: Association for Computational Linguistics), 501–506.
- Chambers, N., Cassidy, T., McDowell, B., and Bethard, S. (2014). Dense event ordering with a multi-pass architecture. *Trans. Assoc. Comput. Linguist.* 2, 273–284. doi: 10.1162/tacl_a_00182
- Chau, M. T., Esteves, D., and Lehmann, J. (2019). Open-domain event extraction and embedding for natural gas market prediction. *arXiv preprint arXiv:1912.11334*. doi: 10.48550/arXiv.1912.11334
- Chen, L., Ruan, W., Liu, X., and Lu, J. (2020). “SeqVAT: Virtual adversarial training for semi-supervised sequence labeling,” in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 8801–8811.
- Chen, Y., Xu, L., Liu, K., Zeng, D., and Zhao, J. (2015a). “Event extraction via dynamic multi-pooling convolutional neural networks,” in *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Vol. 1: Long Papers)* (Beijing), 167–176.
- Chen, Y., Xu, L., Liu, K., Zeng, D., and Zhao, J. (2015b). “Event extraction via dynamic multi-pooling convolutional neural networks,” in *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Vol. 1: Long Papers)*, eds. C. Zong and M. Strube (Beijing: Association for Computational Linguistics), 167–176.
- Chung, H. W., Hou, L., Longpre, S., Zoph, B., Tay, Y., Fedus, W., et al. (2022). Scaling instruction-finetuned language models. *arXiv preprint arXiv:2210.11416*. doi: 10.48550/arXiv.2210.11416
- Colin, E., Gardent, C., Mrabet, Y., Narayan, S., and Perez-Beltrachini, L. (2016). “The webnlg challenge: generating text from dbpedia data,” in *Proceedings of the 9th International Natural Language Generation Conference* (Edinburgh), 163–167.
- Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2019). “BERT: pre-training of deep bidirectional transformers for language understanding,” in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Vol. 1 (Long and Short Papers)*, eds. J. Burstein, C. Doran, and T. Solorio (Minneapolis, MN: Association for Computational Linguistics), 4171–4186.
- Eberts, M., and Ulges, A. (2019). Span-based joint entity and relation extraction with transformer pre-training. *arXiv preprint arXiv:1909.07755*. doi: 10.48550/arXiv.1909.07755
- Elhammedi, S., Lakshmanan, L. V., Ng, R., Simpson, M., Huai, B., Wang, Z., et al. (2020). “A high precision pipeline for financial knowledge graph construction,” in *Proceedings of the 28th International Conference on Computational Linguistics*, 967–977.
- Elsahar, H., Vougiouklis, P., Remaci, A., Gravier, C., Hare, J., Laforest, F., et al. (2018). “T-REx: a large scale alignment of natural language with knowledge base triples,” in *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)* (Miyazaki).
- Fader, A., Soderland, S., and Etzioni, O. (2011). “Identifying relations for open information extraction,” in *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, eds. R. Barzilay and M. Johnson (Edinburgh: Association for Computational Linguistics), 1535–1545.
- Gupta, S., and Manning, C. D. (2014). “Improved pattern learning for bootstrapped entity extraction,” in *Proceedings of the Eighteenth Conference on Computational Natural Language Learning* (Ann Arbor, MI), 98–108.
- Guu, K., Lee, K., Tung, Z., Pasupat, P., and Chang, M. (2020). “Retrieval augmented language model pre-training,” in *International Conference on Machine Learning (PMLR)*, 3929–3938.
- Han, S., Hao, X., and Huang, H. (2018). An event-extraction approach for business analysis from online chinese news. *Electr. Commerce Res. Appl.* 28, 244–260. doi: 10.1016/j.elerap.2018.02.006
- Hang, T., Feng, J., Wu, Y., Yan, L., and Wang, Y. (2021). Joint extraction of entities and overlapping relations using source-target entity labeling. *Expert Syst. Appl.* 177:114853. doi: 10.1016/j.eswa.2021.114853
- Hsu, I., Huang, K. H., Boschee, E., Miller, S., Natarajan, P., Chang, K. W., et al. (2021). DEGREE: a data-efficient generative event extraction model. *arXiv preprint arXiv:2108.12724*. doi: 10.48550/arXiv.2108.12724
- Huang, L., Ji, H., Cho, K., Dagan, I., Riedel, S., and Voss, C. (2018). “Zero-shot transfer learning for event extraction,” in *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Vol. 1: Long Papers)*, eds. I. Gurevych and Y. Miyao (Melbourne, VIC: Association for Computational Linguistics), 2160–2170.
- Huang, L., Ji, H., Cho, K., and Voss, C. R. (2017). Zero-shot transfer learning for event extraction. *arXiv preprint arXiv:1707.01066*. doi: 10.48550/arXiv.1707.01066
- Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., et al. (2023). Survey of hallucination in natural language generation. *ACM Comput. Surv.* 55, 1–38. doi: 10.48550/arXiv.2202.03629
- Kang, M., Baek, J., and Hwang, S. J. (2022). KALA: knowledge-augmented language model adaptation. *arXiv preprint arXiv:2204.10555*. doi: 10.48550/arXiv.2204.10555
- Khongcharoen, W., Saetia, C., Chalothorn, T., and Buabthong, P. (2022). “Question answering over knowledge graphs for thai retail banking products,” in *Proceeding of The 17th International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP 2022)* (Chiangmai: IEEE).
- Kipf, T. N., and Welling, M. (2017). Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*. doi: 10.48550/arXiv.1609.02907
- Klie, J.-C., Bugert, M., Boulosa, B., Eckart de Castilho, R., and Gurevych, I. (2018). “The INCEpTION platform: machine-assisted and knowledge-oriented interactive annotation,” in *Proceedings of the 27th International Conference on Computational Linguistics: System Demonstrations* (Santa Fe), 5–9.
- Lai, V. D., and Nguyen, T. (2019). “Extending event detection to new types with learning from keywords,” in *Proceedings of the 5th Workshop on Noisy User-generated Text (W-NUT 2019)*, eds. W. Xu, A. Ritter, T. Baldwin, and A. Rahimi (Hong Kong: Association for Computational Linguistics), 243–248.
- Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., et al. (2019). BART: denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. *arXiv preprint arXiv:1910.13461*. doi: 10.48550/arXiv.1910.13461
- Li, D., Huang, L., Ji, H., and Han, J. (2019). “Biomedical event extraction based on knowledge-driven tree-LSTM,” in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Vol. 1 (Long and Short Papers)*, eds. J. Burstein, C. Doran, and T. Solorio (Minneapolis, MN: Association for Computational Linguistics), 1421–1430.
- Li, F., Huang, R., Xiong, D., and Zhang, M. (2016). “Learning event expressions via bilingual structure projection,” in *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, eds. Y. Matsumoto and R. Prasad (Osaka: The COLING 2016 Organizing Committee), 1441–1450.
- Li, F., Peng, W., Chen, Y., Wang, Q., Pan, L., Lyu, Y., et al. (2020). “Event extraction as multi-turn question answering,” in *Findings of the Association for Computational Linguistics: EMNLP 2020 (EMNLP)*, 829–838.
- Li, Q., Li, J., Sheng, J., Cui, S., Wu, J., Hei, Y., et al. (2022). A survey on deep learning event extraction: approaches and applications. *IEEE Trans. Neural Netw. Learn. Syst.* 2022, 1–21. doi: 10.1109/TNNLS.2022.3213168
- Liu, J., Min, L., and Huang, X. (2021). An overview of event extraction and its applications. *arXiv preprint arXiv:2111.03212*. doi: 10.48550/arXiv.2111.03212
- Liu, Q., Luan, Z., Wang, K., Zou, Y., Liu, B., and Zhang, Y. (2023). Document-level event extraction—a survey of methods and applications. *J. Phys.* 2504:e012008. doi: 10.1088/1742-6596/2504/1/012008
- Liu, X., Huang, H., and Zhang, Y. (2019). Open domain event extraction using neural latent variable models. *arXiv preprint arXiv:1906.06947*. doi: 10.18653/v1/P19-1276
- Liu, X., Luo, Z., and Huang, H. (2018). Jointly multiple events extraction via attention-based graph information aggregation. *arXiv preprint arXiv:1809.09078*. doi: 10.18653/v1/D18-1156
- Lou, C., Gao, J., Yu, C., Wang, W., Zhao, H., Tu, W., et al. (2022). “Translation-based implicit annotation projection for zero-shot cross-lingual event argument extraction,” in *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '22* (New York, NY: Association for Computing Machinery), 2076–2081.
- Lowphansirikul, L., Polpanumas, C., Jantrakulchai, N., and Nutanong, S. (2021). Wangchanberta: pretraining transformer-based thai language models. *arXiv preprint arXiv:2101.09635*. doi: 10.48550/arXiv.2101.09635

- Lu, W., and Roth, D. (2012). "Automatic event extraction with structured preference modeling," in *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (Vol. 1: Long Papers)*, eds. H. Li, C. Y. Lin, M. Osborne, G. G. Lee, and J. C. Park (Jeju Island: Association for Computational Linguistics), 835–844.
- Lu, Y., Lin, H., Xu, J., Han, X., Tang, J., Li, A., et al. (2021). Text2Event: controllable sequence-to-structure generation for end-to-end event extraction. *arXiv preprint arXiv:2106.09232*. doi: 10.18653/v1/2021.acl-long.217
- Luan, Y., He, L., Ostendorf, M., and Hajishirzi, H. (2018). "Multi-task identification of entities, relations, and coreference for scientific knowledge graph construction," in *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, eds. E. Riloff, D. Chiang, J. Hockenmaier, and J. Tsujii (Brussels: Association for Computational Linguistics), 3219–3232.
- Lyu, Q., Zhang, H., Sulem, E., and Roth, D. (2021). "Zero-shot event extraction via transfer learning: challenges and insights," in *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Vol. 2: Short Papers)*, eds. C. Zong, F. Xia, W. Li, and R. Navigli (Brussels: Association for Computational Linguistics), 322–332.
- M'hamdi, M., Freedman, M., and May, J. (2019). "Contextualized cross-lingual event trigger extraction with minimal resources," in *Proceedings of the 23rd Conference on Computational Natural Language Learning (CoNLL)*, eds. M. Bansal and A. Villavicencio (Hong Kong: Association for Computational Linguistics), 656–665.
- Mialon, G., Dessi, R., Lomeli, M., Nalmpantis, C., Pasunuru, R., Raileanu, R., et al. (2023). Augmented language models: a survey. *arXiv preprint arXiv:2302.07842*. doi: 10.48550/arXiv.2302.07842
- Miller, G. A. (1995). Wordnet: a lexical database for english. *Commun. ACM* 38, 39–41.
- Nguyen, T. M., and Nguyen, T. H. (2019). "One for all: neural joint modeling of entities and events," in *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33 (Hawaii), 6851–6858.
- Nivre, J., De Marneffe, M. C., Ginter, F., Goldberg, Y., Hajic, J., Manning, C. D., et al. (2016). "Universal dependencies v1: a multilingual treebank collection," in *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16) (Portorož)*, 1659–1666.
- Pan, S., Luo, L., Wang, Y., Chen, C., Wang, J., and Wu, X. (2023). Unifying large language models and knowledge graphs: a roadmap. *arXiv preprint arXiv:2306.08302*. doi: 10.48550/arXiv.2306.08302
- Paolini, G., Athiwaratkun, B., Krone, J., Ma, J., Achille, A., Anubhai, R., et al. (2021). Structured prediction as translation between augmented natural languages. *arXiv preprint arXiv:2101.05779*. doi: 10.48550/arXiv.2101.05779
- Pyysalo, S., Ohta, T., Rak, R., Sullivan, D., Mao, C., Wang, C., et al. (2012). Overview of the ID, EPI and REL tasks of bionlp shared task 2011. *BMC Bioinform.* 13(Suppl.11):S2. doi: 10.1186/1471-2105-13-S11-S2
- Shen, S., Wu, T., Qi, G., Li, Y. F., Haffari, G., and Bi, S. (2021). "Adaptive knowledge-enhanced Bayesian meta-learning for few-shot event detection," in *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, eds. C. Zong, F. Xia, W. Li, and R. Navigli (Brussels: Association for Computational Linguistics), 2417–2429.
- Snell, J., Swersky, K., and Zemel, R. S. (2017). Prototypical networks for few-shot learning. *arXiv preprint arXiv:1703.05175*. doi: 10.48550/arXiv.1703.05175
- Speer, R., Chin, J., and Havasi, C. (2017). "ConceptNet 5.5: an open multilingual graph of general knowledge," in *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 31 (San Francisco, CA).
- Stenetorp, P., Pyysalo, S., Topić, G., Ohta, T., Ananiadou, S., and Tsujii, J. (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 Association for Computational Linguistics*, ed. F. Segond (Avignon: Association for Computational Linguistics), 102–107.
- Subburathinam, A., Lu, D., Ji, H., May, J., Chang, S.-F., Sil, A., et al. (2019). "Cross-lingual structure transfer for relation and event extraction," in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, eds. K. Inui, J. Jiang, V. Ng, and X. Wan (Hong Kong: Association for Computational Linguistics), 313–325.
- Tan, Q., Liu, N., Zhao, X., Yang, H., Zhou, J., and Hu, X. (2020). "Learning to hash with graph neural networks for recommender systems," in *Proceedings of The Web Conference 2020 (Taipei)*, 1988–1998.
- Tjong Kim Sang, E. F., and 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 (Edmonton, AB)*, 142–147.
- Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., et al. (2023). LLaMA: open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*. doi: 10.48550/arXiv.2302.13971
- Vanegas, J. A., Matos, S., González, F., and Oliveira, J. L. (2015). An overview of biomolecular event extraction from scientific documents. *Computat. Math. Methods Med.* 2015:571381. doi: 10.1155/2015/571381
- Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lió, P., and Bengio, Y. (2018). Graph attention networks. *arXiv preprint arXiv:1710.10903*. doi: 10.48550/arXiv.1710.10903
- Wadden, D., Wennberg, U., Luan, Y., and Hajishirzi, H. (2019). Entity, relation, and event extraction with contextualized span representations. *arXiv preprint arXiv:1909.03546*. doi: 10.48550/arXiv.1909.03546
- Walker, C., and Consortium, L. D. (2005). *ACE 2005 Multilingual Training Corpus*. Linguistic Data Consortium. Available online at: <https://catalog.ldc.upenn.edu/LDC2006T06> (accessed February 1, 2023).
- Wang, C., Liu, X., Chen, Z., Hong, H., Tang, J., and Song, D. (2023). DeepStruct: pretraining of language models for structure prediction. *arXiv preprint arXiv:2205.10475*. doi: 10.48550/arXiv.2205.10475
- Wang, Z., Li, T., and Li, Z. (2023). Unsupervised numerical information extraction via exploiting syntactic structures. *Electronics* 12:1977. doi: 10.3390/electronics12091977
- Wu, S., and Dredze, M. (2020). "Are all languages created equal in multilingual BERT?," in *Proceedings of the 5th Workshop on Representation Learning for NLP*, eds. S. Gella, J. Welbl, M. Rei, F. Petroni, P. Lewis, E. Strubell, et al. (Hong Kong: Association for Computational Linguistics), 120–130.
- Wu, S., Irsoy, O., Lu, S., Dabrovolski, V., Dredze, M., Gehrmann, S., et al. (2023). BloombergGPT: a large language model for finance. *arXiv preprint arXiv:2303.17564*. doi: 10.48550/arXiv.2303.17564
- Xiang, W., and Wang, B. (2019). A survey of event extraction from text. *IEEE Access* 7, 173111–173137. doi: 10.1109/ACCESS.2019.2956831
- Xue, L., Constant, N., Roberts, A., Kale, M., Al-Rfou, R., Siddhant, A., et al. (2021). "mT5: a massively multilingual pre-trained text-to-text transformer," in *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, eds. K. Toutanova, A. Rumshisky, L. Zettlemoyer, D. Hakkani-Tur, I. Beltagy, S. Bethard, et al. (Hong Kong: Association for Computational Linguistics), 483–498.
- Yang, H., Chen, Y., Liu, K., Xiao, Y., and Zhao, J. (2018). "DCFEE: a document-level Chinese financial event extraction system based on automatically labeled training data" in *Proceedings of ACL 2018, System Demonstrations (Melbourne, VIC)*, 50–55.
- Yang, S., Feng, D., Qiao, L., Kan, Z., and Li, D. (2019). "Exploring pre-trained language models for event extraction and generation," in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, eds. A. Korhonen, D. Traum, and L. Márquez (Florence: Association for Computational Linguistics), 5284–5294.
- Yao, L., Mao, C., and Luo, Y. (2019). KG-BERT: BERT for knowledge graph completion. *arXiv preprint arXiv:1909.03193*. doi: 10.48550/arXiv.1909.03193
- Zhang, H., Wang, H., and Roth, D. (2021). "Zero-shot label-aware event trigger and argument classification," in *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, eds. C. Zong, F. Xia, W. Li, and R. Navigli (Florence: Association for Computational Linguistics), 1331–1340.
- Zhang, J., Qin, Y., Zhang, Y., Liu, M., and Ji, D. (2019). "Extracting entities and events as a single task using a transition-based neural model," in *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19 (Macao)*, 5422–5428.