

# Fine-Grained Event Trigger Detection

Duong Minh Le

VinAI Research, Vietnam  
v.duonglm1@vinai.io

Thien Huu Nguyen\*

Department of Computer and Information Science  
University of Oregon, Eugene, OR 97403, USA  
thien@cs.uoregon.edu

## Abstract

Most of the previous work on Event Detection (ED) has only considered the datasets with a small number of event types (i.e., up to 38 types). In this work, we present the first study on fine-grained ED (FED) where the evaluation dataset involves much more fine-grained event types (i.e., 449 types). We propose a novel method to transform the Semcor dataset for Word Sense Disambiguation into a large and high-quality dataset for FED. Extensive evaluation of the current ED methods is conducted to demonstrate the challenges of the generated datasets for FED, calling for more research effort in this area.

## 1 Introduction

Understanding events in text is an important aspect of Natural Language Processing (NLP). Toward this end, Event Detection (ED), a task of Information Extraction (IE), aims to identify event triggers in sentences and classify them into some predefined types of interest. Event triggers represent the most important words (usually single verbs or nominalizations) in the sentences that evoke the events. The current state-of-the-art methods for ED feature the deep learning models where many new network architectures are introduced in the last couple of years (Nguyen and Grishman, 2015; Chen et al., 2015; Liu et al., 2017, 2019a; Lai et al., 2020b).

Among others, the rapid development of the deep learning models for ED can be partly attributed to the availability of the large datasets to evaluate the models (e.g., the ACE 2005 and TAC KBP 2015 datasets (Walker et al., 2006; Mitamura et al., 2015)). Unfortunately, a major issue in these existing datasets for ED is that they tend to only focus on a limited set of event types. For example, the popular ACE 2005 dataset is only annotated

for 33 event subtypes (e.g., *Attack*, *Start-Position*, *Elect*) while the number of events in the TAC KBP dataset (Mitamura et al., 2015) is 38. On the one hand, the limited numbers of types are unable to cover a wide range of possible events in practice (Araki and Mitamura, 2018). On the other hand, the small label sets often amount to the coarse-grained event types in the existing datasets that cannot capture the slightly different nuances (i.e., fine-grained distinction) of the events. For instance, both the words “quit” and “fired” in the two sentences “He decided to **quit** the job.” and “He was **fired** due to a policy violation.” (respectively) would be considered as the trigger words of the same event type of *End-Position* in the ACE 2005 dataset. However, the nuances in these two events are quite different (i.e., in term of the willingness of the job termination) and the ability to characterize such subtle distinction would be useful for the downstream applications (Choi et al., 2018).

In order to address these problems, we propose to explore the problem of Fine-grained Event Detection (FED) that seeks to solve ED with much larger and finer-grained sets of event types (motivated by the fine-grained entity typing task (Ling and Weld, 2012; Choi et al., 2018)). To our knowledge, this is the first work to explicitly study FED in the literature. A major challenge in this research direction is the creation of the evaluation datasets to enable effective model development and analysis. In particular, it is non-trivial to design a large set of fine-grained event types to be applied to annotate the datasets. In addition, with such a large number of fine-grained event types (i.e., 449 in this work), the traditional labeling procedure with human involvement might be too expensive and error-prone when it comes to the generation of large datasets for FED. To this end, we introduce a novel method to address these challenges and produce a large dataset for FED based on WordNet and Word Sense

\*Corresponding author.

Disambiguation (WSD) datasets. Our method involves two major steps where we first leverage the synset typology in WordNet to formulate the fine-grained event types and then convert the annotated datasets for WSD to establish the datasets for our FED problem. This novel data generation procedure minimizes the human effort and allows us to create a large and high-quality dataset with 449 fine-grained event types for FED. Finally, we extensively evaluate the state-of-the-art ED models on the proposed FED dataset. The experiments show that the performance of the current ED models is not yet satisfactory for FED and further research is needed to advance the performance in this area. We will publicly release the proposed dataset to promote the future research on FED.

## 2 Data Generation Procedure

The goal of this section is to generate a large dataset for ED with many fine-grained event types to evaluate the FED models. Our proposed procedure to achieve this goal involves two major steps. First, we identify the eventive synsets/senses in WordNet 3.0 (Miller, 1995) and group them into classes with similar eventive meanings. These classes would serve as the fine-grained event types in the resulting FED dataset. As the result, we obtain a mapping from the set of WordNet synsets to the set of the fine-grained event types for our problem (some WordNet synsets might not be mapped to any event type in our case). Afterward, we leverage the Semcor dataset for WSD (Miller et al., 1994) and map the synsets annotated for the words in this dataset into the event types in our setting. This conversion process produces a dataset whose words are assigned with the fine-grained event types in our FED problem. As the Semcor dataset is manually annotated, the resulting FED dataset would be large and have high quality if the synset-event type mapping is constructed well.

In particular, for eventive synset/sense identification, we first start with nouns. Following (Araki and Mitamura, 2018), we assume that any synset for a noun subsumed by one of the three following synsets via the WordNet hyponyms would be considered as eventive:  $state_n^2$  (i.e., the way something is with respect to its main attributes),  $process_n^6$  (i.e., a sustained phenomenon or one marked by gradual changes through a series of states), and  $event_n^1$  (i.e., something that happens at a given place and time). In this way, we find 13,166 even-

tive synsets over 82,115 synsets for nouns. We

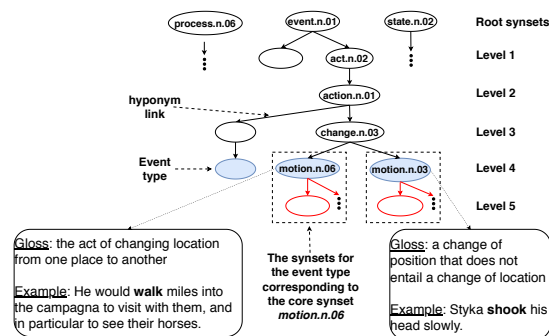


Figure 1: The synsets at different levels and the examples for some event types and their core synsets.

call these three general synsets as the eventive root synsets in the following. Starting from these root synsets, we traverse the synset graph in WordNet by following the hyponym links. The graph traversal procedure will generate three different trees whose nodes are the eventive synsets and roots correspond to the three selected synsets. For convenience, we call the synsets in WordNet that can be reached by one of the three synsets above after  $n$  hyponym links as the synsets at the  $n$ -th level<sup>1</sup> (so the root synsets are at the zero level). In order to form the fine-grained event types for FED, we select the WordNet synsets at the 4<sup>th</sup> level as the core meanings (called the core synsets) for the event types in our dataset (there are 2,637 core eventive synsets found in this way). We empirically choose the synsets at the 4<sup>th</sup> level to balance two factors. On the one hand, the synsets at the shallower levels lead to too general event types that cannot achieve the expected fine-grained property. On the other hand, going deeper for the core event meanings reduces the numbers of examples per event type in the final FED dataset converted from Semcor.

Given a core synset  $A$ , we identify the other synsets with similar meaning to  $A$  and combine them to represent a fine-grained event type (called  $E$ ) in our dataset (i.e., the event type  $E$  will involve several semantically similar synsets in WordNet). In this work, we include two following classes of synsets in the event type  $E$  for  $A$ :

- The synsets for nouns that can be reached from  $A$  with the hyponym links: Intuitively, the synsets subsumed by  $A$  would exhibit the general eventive

<sup>1</sup>It is possible that some eventive synsets in WordNet might reside at more than one level as they can be reached from the three root nodes with multiple paths. We resolve this conflict by putting these synsets on the closest level to the roots.

meanings of  $A$  with some certain distinctions.

- The synsets of the derivationally related forms of the lemmas/senses  $l$  in  $A$  and their synset descendants (via the hyponym links): In WordNet, the derivationally related forms of a lemma  $l$  in the synset  $A$  involve the lemmas from different syntactic categories (e.g., verbs and adjectives) that have the same root form as  $l$  and are semantically related to  $l$  and  $A$  (e.g., destruction  $\rightarrow$  destroy) (Miller, 1995). Due to such semantic similarity, we expect that the synsets of the derivationally related forms of the lemmas in  $A$  and their descendants also express the same eventive meaning as  $A$ , thereby enriching the synsets for  $E$  with the syntactic categories beyond nouns.

Up to this point, we obtain a set of 2,637 event types, each represented by a core synset and a set of related synsets. We combine all the other synsets (i.e., the ones that do not appear in any of the 2637 types) to create a single event type called *Other* as in the traditional ED task. With these grouping information, we can now create a mapping from the synsets/senses in WordNet to the 2638 established event types for our dataset (called the *M-WordNet-Event* mapping). Based on this mapping, we transform each example in the WSD Semcor dataset, which involves a sentence and a word of interest, into an example in the new dataset for FED (called FedSemcor) where the synset/sense label for the word in the original example of Semcor is mapped into the corresponding event type in FedSemcor. As the final processing step, we remove from FedSemcor any event types that have less than 10 examples to ensure that the event types are adequately represented in our dataset. This step significantly reduces the number of event types in FedSemcor, leaving us 449 event types (not including *Other*) with 34,666 examples. These are called the positive examples where we have 77.2 examples per event type in average. As a result, the synsets for the removed event types are also included in the synset set for the *Other* type in the mapping. The number of examples with the *Other* type in the final FedSemcor dataset (called the negative examples) is 98,309. Figure 1 illustrates the synset levels and some examples for the event types.

**Implementation Details:** In the actual implementation of the data generation procedure for FedSemcor, given a core synset  $A$ , we do not include all the descendants of  $A$  and the synsets of the derivationally related forms of  $A$ 's lemmas into the

	ACE	TAC KBP	FedSemcor
# event types	33	38	449
# positive examples	4,907	11,975	34,666
# negative examples	104,217	126,934	98,309

Table 1: Statistics for the FedSemcor, ACE 2005 and TAC KBP 2015 datasets. Negative examples refer to the non-trigger words while positive examples are annotated trigger words for the event types of interest.

synset set for the corresponding event type  $E$  for  $A$ . Instead, we only include the descendants that are at at most 2 hyponym links away from  $A$  and the synsets of the derivationally related forms of  $A$ 's lemmas in  $E$ . This is based on our empirical investigation of the data where the descendants with more than 2 links away tend to have semantic drifts from  $A$ , potentially introducing noise into the event type  $E$ . For example with the core synset  $motion_n^6$  (i.e., *the act of changing location from one place to another*) at the 4<sup>th</sup> level, the descendants at the 5<sup>th</sup>, 6<sup>th</sup> and 7<sup>th</sup> levels include: level 5:  $approach_n^2$  (i.e., *the act of drawing spatially closer to something*), level 6:  $access_n^6$  (i.e., *the act of approaching or entering*), and level 7:  $back_door_n^6$  (i.e., *a secret or underhand means of access (to a place or a position)*). As we can see, while the synsets at the 5<sup>th</sup> and 6<sup>th</sup> levels are related to the original core synset, the synset at the 7<sup>th</sup> level already involves some semantic departure from the one at the 4<sup>th</sup> level that should be avoided to improve the precision.

**Dataset Statistics:** Table 1 reports some statistics for FedSemcor and some prior popular datasets for ED (i.e., ACE 2005 (Walker et al., 2006) and TAC KBP 2015 (Mitamura et al., 2015)) to facilitate the comparison. As we can see from the table, FedSemcor has more positive examples, but less negative examples than ACE 2005 and TAC KBP 2015, making FedSemcor a more balanced dataset than the other two. In addition, we show the distribution of 50 event types with the highest numbers of examples in FedSemcor in Figure 2. Finally, Figure 3 illustrates the distribution of the sentence lengths for the examples in FedSemcor.

**Evaluation of FedSemCor:** As we rely on the manual annotation in Semcor for the synsets for the words, the main bottleneck in the data generation procedure is the mapping from the WordNet synsets to the 450 event types in FedSemCor (including *Other*). In order to evaluate the quality of this mapping, we sample 500 synsets from WordNet that are different from the core synsets of the 449 positive event types. Two experienced NLP re-

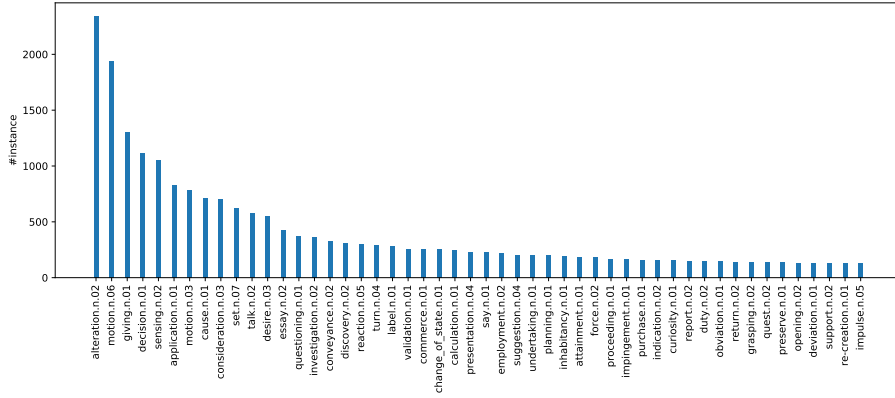


Figure 2: The distribution of the 50 event types with the highest numbers of examples/instances in FedSemcor.

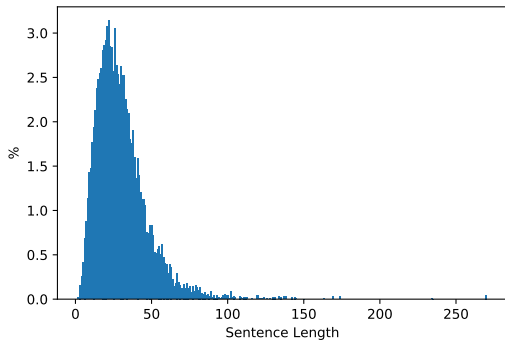


Figure 3: The distribution of the sentence lengths for the examples in FedSemcor.

searchers then independently examine each of these 500 sampled synsets to determine the appropriate event type for it (among the 450 types). In doing so, they examined the glossaries of the synsets as well as the examples provided by WordNet. The two annotators achieved 79.8% agreement for which the synsets with conflicts are resolved by a third NLP researcher. Afterward, we apply the synset-event type mapping obtained in the data generation procedure to annotate the 500 sampled synsets for the event types. The event types provided by the mapping are then compared with those from the annotators, leading to 83.6%, 78.6% and 81.0% as the precision, recall, and F1 scores respectively.

### 3 Evaluation

**Models and Data:** In order to understand the complexity of the FedSemcor dataset for FED, this section evaluates the performance of the state-of-the-art models for the traditional ED problem on this dataset. In particular, we first split FedSemcor into the training, development and test data

Dataset	Training		Development		Test	
	#Pos	#Neg	#Pos	#Neg	#Pos	#Neg
FedSemcor	20,799	58,985	5,546	21,038	8,321	18,286

Table 2: The size of the dataset portions.

using the 6:2:2 ratio over the entire dataset. Table 2 presents the statistics about these data portions. Note that similar to some prior ED work (Nguyen and Grishman, 2015; Chen et al., 2015), our FED problem is formulated as a word classification problem where given a word in an input sentence, the models need to predict the event type for the word.

Afterward, we consider the following representative models for ED: **CNN** (Nguyen and Grishman, 2015), **DMCNN** (Chen et al., 2015), **SupAtt** (Liu et al., 2017) (i.e., supervised attention), **GCN** (Nguyen and Grishman, 2018), and **MOGANED** (Yan et al., 2019) (i.e., a Multi-Order Graph Convolution model). MOGANED is the state-of-the-art model with uncontextualized word embeddings in traditional ED (i.e., on ACE 2005). For these models, we use both the traditional word embeddings `word2vec` and the recent contextualized word embeddings BERT (i.e., the uncased base model) (Devlin et al., 2019) as the pre-trained word embeddings. For BERT, we further evaluate the ED models in (Wang et al., 2019) (called **DMBERT**) and (Yang et al., 2019) (called **BERT-ED**) that have the best-reported performance on ACE 2005 for ED.

For the experiments in this work, we re-tune the hyper-parameters of the models on the development set of FedSemcor. In particular, depending on which components each model has, we use the following bounds to search for the hyperparameters: [100, 200, 300, 400, 500] for the dimensionality of the hidden vectors in the layers of all

the feed-forward, BiLSTM, and GCN networks, [1, 2, 3] for the numbers of layers for BiLSTM and GCN, [16, 32, 64] for the mini-batch size, [1e-5, 1e-4, 1e-3, 1e-2, 1e-1] for the learning rate of the Adam optimizer, and [10, 20, 30, 40, 50] for the dimensions of the feature embeddings, i.e., position embeddings in CNN (Nguyen and Grishman, 2015; Chen et al., 2015).

Finally, in order to demonstrate the benefit of the conversion of Semcor (i.e., a WSD dataset) into FedSemcor for FED, we consider a WSD-based baseline for FED where the state-of-the-art WSD model in (Hadiwinoto et al., 2019) is trained on the training data of FedSemcor. As this is a WSD model, instead of using the mapped event types as the labels for the examples (i.e., the 450 types) in the training data, we employ the original word senses of the words as the labels to train this WSD model. Afterward, we apply the trained WSD model on the test data of FedSemcor, producing a word sense for each example. In the last step, the mapping *M-WordNet-Event* is utilized to convert the predicted word senses for the test set examples into the event types for FedSemcor that would be evaluated to obtain the FED performance for this baseline (called **WSD-based**). Note that this WSD model also uses the BERT embeddings.

**Results:** Table 3 shows the performance of the models on the test set of FedSemcor. From the table, we see that GCN has the best performance among the models with `word2vec` while BERT-ED outperforms all the BERT-based models. However, the best performance on FedSemcor (i.e., 65.0% F1 score with BERT-ED) is still far behind the typical performance (i.e., up to 80.7% in (Yang et al., 2019)) of the models on the traditional ED datasets (i.e., ACE 2005). This suggests the more challenging nature of FedSemcor and FED over traditional ED, presenting a challenge for the future research in this area. Importantly, the performance of the ED models (i.e., with the BERT embeddings) is significantly better than the WSD-based baseline (i.e., up to 9% performance gap with BERT-ED), clearly testifying to the advantages of the conversion from Semcor into FedSemcor for FED.

## 4 Related Work

ED has been studied extensively in the last decade, featuring feature-based models (Ahn, 2006; Ji and Grishman, 2008; Li et al., 2013, 2015), deep learning models (Chen et al., 2015; Nguyen et al.,

Model	word2vec			BERT		
	P	R	F1	P	R	F1
WSD-based	-	-	-	47.2	68.7	56.0
CNN	48.8	53.6	51.1	52.8	69.8	60.1
DMCNN	43.7	51.6	47.4	56.9	71.5	63.4
SupAtt	58.5	46.0	51.5	59.5	71.1	64.8
GCN	53.7	60.0	<b>56.7</b>	58.9	71.5	64.6
MOGANED	48.6	61.8	54.4	55.8	71.2	62.6
DMBERT	-	-	-	57.7	63.2	60.3
BERT-ED	-	-	-	59.2	72.1	<b>65.0</b>

Table 3: The performance on the FedSemcor test set.

2016b,a; Nguyen and Grishman, 2016; Chen et al., 2017; Liu et al., 2018; Yan et al., 2019; Ngo et al., 2020; Lai et al., 2020b), and few/zero-shot learning models (Huang et al., 2018; Lai and Nguyen, 2019; Lai et al., 2020a). The rapid development of such models has been facilitated by the availability of the ED datasets in different domains, including the general domain with the popular ACE and TAC KBP datasets (Walker et al., 2006; Mitamura et al., 2015, 2016), the biomedical domain (Kim et al., 2009, 2011), literature (Sims et al., 2019), cybersecurity (Satyapanich et al., 2020; Man Duc Trong et al., 2020), and the open domain (Araki and Mitamura, 2018; Liu et al., 2019b). However, these datasets only involve a small number of event types and none of them has considered ED with many fine-grained event types as we do.

Our FED task is also related to fine-grained entity typing that aims to classify entity mentions into a fine-grained set of types (Karn et al., 2017; Shimaoka et al., 2016; Lin and Ji, 2019). The techniques to generate datasets for fine-grained entity typing include distant supervision (Ling and Weld, 2012; Abhishek et al., 2017) and manual annotation (Murty et al., 2018; Choi et al., 2018). Notably, (Del Corro et al., 2015) also uses WordNet to establish the fine-grained entity types, applying different entity mention extractors over external corpus. Our work is different as we focus on fine-grained event types using the manually annotated corpus Semcor to generate data.

## 5 Conclusion

We study a new task of FED, featuring 449 fine-grained event types in the dataset for ED. A novel method to generate the evaluation dataset for FED is introduced, leveraging manually annotated WSD datasets (i.e., Semcor) and the eventive synsets in WordNet. We evaluate the state-of-the-art ED models on the new dataset to show the opportunities for the future research on FED.

## References

- Abhishek Abhishek, Ashish Anand, and Amit Awekar. 2017. [Fine-grained entity type classification by jointly learning representations and label embeddings](#). In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers*, pages 797–807, Valencia, Spain. Association for Computational Linguistics.
- David Ahn. 2006. [The stages of event extraction](#). In *Proceedings of the Workshop on Annotating and Reasoning about Time and Events*, pages 1–8, Sydney, Australia. Association for Computational Linguistics.
- Jun Araki and Teruko Mitamura. 2018. [Open-domain event detection using distant supervision](#). In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 878–891, Santa Fe, New Mexico, USA. Association for Computational Linguistics.
- Yubo Chen, Shulin Liu, Xiang Zhang, Kang Liu, and Jun Zhao. 2017. [Automatically labeled data generation for large scale event extraction](#). In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 409–419, Vancouver, Canada. Association for Computational Linguistics.
- Yubo Chen, Liheng Xu, Kang Liu, Daojian Zeng, and Jun Zhao. 2015. [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 (Volume 1: Long Papers)*, pages 167–176, Beijing, China. Association for Computational Linguistics.
- Eunsol Choi, Omer Levy, Yejin Choi, and Luke Zettlemoyer. 2018. [Ultra-fine entity typing](#). In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 87–96, Melbourne, Australia. Association for Computational Linguistics.
- Luciano Del Corro, Abdalghani Abujabal, Rainer Gemulla, and Gerhard Weikum. 2015. [FINET: Context-aware fine-grained named entity typing](#). In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 868–878, Lisbon, Portugal. Association for Computational Linguistics.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 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, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.
- Christian Hadiwinoto, Hwee Tou Ng, and Wee Chung Gan. 2019. [Improved word sense disambiguation using pre-trained contextualized word representations](#). 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)*, pages 5297–5306, Hong Kong, China. Association for Computational Linguistics.
- Lifu Huang, Heng Ji, Kyunghyun Cho, Ido Dagan, Sebastian Riedel, and Clare Voss. 2018. [Zero-shot transfer learning for event extraction](#). In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2160–2170, Melbourne, Australia. Association for Computational Linguistics.
- Heng Ji and Ralph Grishman. 2008. [Refining event extraction through cross-document inference](#). In *Proceedings of ACL-08: HLT*, pages 254–262, Columbus, Ohio. Association for Computational Linguistics.
- Sanjeev Karn, Ulli Waltinger, and Hinrich Schütze. 2017. [End-to-end trainable attentive decoder for hierarchical entity classification](#). In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*, pages 752–758, Valencia, Spain. Association for Computational Linguistics.
- Jin-Dong Kim, Tomoko Ohta, Sampo Pyysalo, Yoshinobu Kano, and Jun’ichi Tsujii. 2009. [Overview of BioNLP’09 shared task on event extraction](#). In *Proceedings of the BioNLP 2009 Workshop Companion Volume for Shared Task*, pages 1–9, Boulder, Colorado. Association for Computational Linguistics.
- Jin-Dong Kim, Yue Wang, Toshihisa Takagi, and Akinori Yonezawa. 2011. [Overview of Genia event task in BioNLP shared task 2011](#). In *Proceedings of BioNLP Shared Task 2011 Workshop*, pages 7–15, Portland, Oregon, USA. Association for Computational Linguistics.
- Viet Duc Lai, Franck Dernoncourt, and Thien Huu Nguyen. 2020a. [Exploiting the matching information in the support set for few shot event classification](#). In *Proceedings of the 24th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD)*, Singapore.
- Viet Duc Lai and Thien Nguyen. 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)*, pages 243–248, Hong Kong, China. Association for Computational Linguistics.
- Viet Duc Lai, Tuan Ngo Nguyen, and Thien Huu Nguyen. 2020b. [Event detection: Gate diversity and syntactic importance scores for graph convolution neural networks](#). In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language*

- Processing (EMNLP)*, pages 5405–5411, Online. Association for Computational Linguistics.
- Qi Li, Heng Ji, and Liang Huang. 2013. [Joint event extraction via structured prediction with global features](#). In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 73–82, Sofia, Bulgaria. Association for Computational Linguistics.
- Xiang Li, Thien Huu Nguyen, Kai Cao, and Ralph Grishman. 2015. [Improving event detection with abstract meaning representation](#). In *Proceedings of the First Workshop on Computing News Storylines*, pages 11–15, Beijing, China. Association for Computational Linguistics.
- Ying Lin and Heng Ji. 2019. [An attentive fine-grained entity typing model with latent type representation](#). 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)*, pages 6197–6202, Hong Kong, China. Association for Computational Linguistics.
- Xial Ling and Daniel Weld. 2012. Fine-grained entity recognition. In *Proceedings of the Twenty-Sixth Conference on Artificial Intelligence (AAAI)*, page 94–100, Toronto, Canada. Association for the Advancement of Artificial Intelligence.
- Jian Liu, Yubo Chen, and Kang Liu. 2019a. [Exploiting the ground-truth: An adversarial imitation based knowledge distillation approach for event detection](#). In *Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence (AAAI)*, Honolulu, Hawaii. Association for the Advancement of Artificial Intelligence.
- Jian Liu, Yubo Chen, Kang Liu, and Jun Zhao. 2018. Event detection via gated multilingual attention mechanism. In *Thirty-Second AAAI Conference on Artificial Intelligence (AAAI)*, New Orleans, Louisiana. Association for the Advancement of Artificial Intelligence.
- Shulin Liu, Yubo Chen, Kang Liu, and Jun Zhao. 2017. [Exploiting argument information to improve event detection via supervised attention mechanisms](#). In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1789–1798, Vancouver, Canada. Association for Computational Linguistics.
- Xiao Liu, Heyan Huang, and Yue Zhang. 2019b. [Open domain event extraction using neural latent variable models](#). In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 2860–2871, Florence, Italy. Association for Computational Linguistics.
- Hieu Man Duc Trong, Duc Trong Le, Amir Pouran Ben Veyseh, Thuat Nguyen, and Thien Huu Nguyen. 2020. [Introducing a new dataset for event detection in cybersecurity texts](#). In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 5381–5390, Online. Association for Computational Linguistics.
- George A. Miller. 1995. Wordnet: a lexical database for english. In *Communications of the ACM*, 38(11):39–41.
- George A. Miller, Martin Chodorow, Shari Landes, Claudia Leacock, and Robert G. Thomas. 1994. [Using a semantic concordance for sense identification](#). In *HUMAN LANGUAGE TECHNOLOGY: Proceedings of a Workshop held at Plainsboro, New Jersey, March 8-11, 1994*.
- Teruko Mitamura, Zhengzhong Liu, and Eduard Hovy. 2015. Overview of tac kbp 2015 event nugget track. In *Proceedings of the Text Analysis Conference (TAC)*, Gaithersburg, Maryland. National Institute of Standards and Technology (NIST).
- Teruko Mitamura, Zhengzhong Liu, and Eduard Hovy. 2016. Overview of tac kbp 2016 event nugget track. In *Proceedings of the Text Analysis Conference (TAC)*, Gaithersburg, Maryland. National Institute of Standards and Technology (NIST).
- Shikhar Murty, Patrick Verga, Luke Vilnis, Irena Radovanovic, and Andrew McCallum. 2018. [Hierarchical losses and new resources for fine-grained entity typing and linking](#). In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 97–109, Melbourne, Australia. Association for Computational Linguistics.
- Nghia Trung Ngo, Tuan Ngo Nguyen, and Thien Huu Nguyen. 2020. Learning to select important context words for event detection. In *Proceedings of the 24th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD)*, Singapore.
- Thien Huu Nguyen, , Adam Meyers, and Ralph Grishman. 2016a. New york university 2016 system for kbp event nugget: A deep learning approach. In *Proceedings of the Text Analysis Conference (TAC)*, Gaithersburg, Maryland. National Institute of Standards and Technology (NIST).
- Thien Huu Nguyen, Kyunghyun Cho, and Ralph Grishman. 2016b. [Joint event extraction via recurrent neural networks](#). In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 300–309, San Diego, California. Association for Computational Linguistics.
- Thien Huu Nguyen and Ralph Grishman. 2015. [Event detection and domain adaptation with 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 (Volume 2: Short*

- Papers*), pages 365–371, Beijing, China. Association for Computational Linguistics.
- Thien Huu Nguyen and Ralph Grishman. 2016. [Modeling skip-grams for event detection with convolutional neural networks](#). In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 886–891, Austin, Texas. Association for Computational Linguistics.
- Thien Huu Nguyen and Ralph Grishman. 2018. Graph convolutional networks with argument-aware pooling for event detection. In *Thirty-Second AAAI Conference on Artificial Intelligence (AAAI)*, New Orleans, Louisiana. Association for the Advancement of Artificial Intelligence.
- Taneeya Satyapanich, Francis Ferraro, and Tim Finin. 2020. [Casie: Extracting cybersecurity event information from text](#). In *Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI)*, New York, New York. Association for the Advancement of Artificial Intelligence.
- Sonse Shimaoka, Pontus Stenetorp, Kentaro Inui, and Sebastian Riedel. 2016. [An attentive neural architecture for fine-grained entity type classification](#). In *Proceedings of the 5th Workshop on Automated Knowledge Base Construction*, pages 69–74, San Diego, CA. Association for Computational Linguistics.
- Matthew Sims, Jong Ho Park, and David Bamman. 2019. [Literary event detection](#). In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 3623–3634, Florence, Italy. Association for Computational Linguistics.
- Christopher Walker, Stephanie Strassel, Julie Medero, and Kazuaki Maeda. 2006. Ace 2005 multilingual training corpus. In *Technical report, Linguistic Data Consortium*.
- Xiaozhi Wang, Xu Han, Zhiyuan Liu, Maosong Sun, and Peng Li. 2019. [Adversarial training for weakly supervised event detection](#). 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)*, pages 998–1008, Minneapolis, Minnesota. Association for Computational Linguistics.
- Haoran Yan, Xiaolong Jin, Xiangbin Meng, Jiafeng Guo, and Xueqi Cheng. 2019. [Event detection with multi-order graph convolution and aggregated attention](#). 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)*, pages 5766–5770, Hong Kong, China. Association for Computational Linguistics.
- Sen Yang, Dawei Feng, Linbo Qiao, Zhigang Kan, and Dongsheng Li. 2019. [Exploring pre-trained language models for event extraction and generation](#). In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 5284–5294, Florence, Italy. Association for Computational Linguistics.