Membrane Technology
ISSN (online): 1873-4049

# Lexicon-based Bidirectional Framework for Relational Triple Extraction

**Songtao Cai[1], Qicheng Ma[1], Yupeng Hou[1], Guangping Zeng[2,*]**

[1]*School of Computer & Communication Engineering, University of Science and Technology, Beijing, China*

[2]*Key Laboratory of Knowledge Engineering for Materials Science, University of Science and Technology, Beijing, China*

**Corresponding Author.*

**Abstract:**
A lexicon-based bidirectional extraction framework is proposed to solve the unidirectional frame sensitivity problem in relational triple extraction methods. The framework employs two complementary directions to extract entity pairs, effectively alleviating the overdependence on subject extraction results. A shared encoder is utilized to facilitate feature transfer between two directions, ensuring extractions in each direction are mutually reinforced and complementary. The shared structure leads to inconsistent convergence rates in training process, thus a shared-aware learning mechanism is introduced. Model's efficacy is confirmed by experiments demonstrating the adaptability and enhancement capabilities of bidirectional extraction framework and shared-aware mechanism for other labelling-based methods.

**Keywords:** relational triple extraction, joint extraction of entities and relations, bidirectional extraction framework, deep learning

## INTRODUCTION

Relational triple extraction (RTE) seeks to autonomously discern and extract entities and their relational dynamics from textual data. This undertaking is pivotal for numerous applications, including information retrieval, the assembly of knowledge graphs, and the functionality of question-and-answer (Q&A) systems.

Despite extensive research in RTE, current methodologies encounter several obstacles. Traditional techniques, predominantly unidirectional extraction frameworks as noted in [1], initiate by pinpointing subjects, followed by object and relation extraction. Although efficacious in certain scenarios, this method's effectiveness is heavily contingent on the precision of subject extraction, with inaccuracies significantly impeding overall performance.

Considering these constraints, we propose a lexicon-based bidirectional extraction framework, diverging from the conventional reliance on unidirectional entity extraction. Our approach commences by extracting potential subject-object pairs via dependent syntax analysis. Subsequently, it employs dual bi-directional flows for subject and object extraction, thereby diminishing the dependency on singular subject extraction outcomes. This bidirectional method enables a more holistic capture of relational data, improving both the accuracy and stability of the process. Additionally, we introduce a shared perceptual learning mechanism to address the issue of inconsistent convergence rates inherent in the shared structural framework. This mechanism promotes training stability and augments model performance.

Extensive testing across various datasets confirms the efficiency of our framework. The results consistently lay out that our framework not only achieves SOTA performance across these datasets but also enhances the efficacy of other labeling-based methodologies through the application of our bidirectional extraction framework and shared perceptual learning mechanism.

## RELATED WORK

Relational triple extraction has evolved significantly. Primal works in RTE predominantly adopted pipeline-based approaches, treating entity recognition and relation detection as distinct tasks. Zelenko et al. [2] and Chan and Roth [3] exemplified this approach. To address these challenges, previous methods typically faced issues like error propagation and slow processing because they handled entity and relation extraction separately. Recent research has moved towards integrated extraction models that concurrently manage both entities and relations to overcome these limitations. Such models, including RSAN as proposed by Yuan et al., have shown effectiveness in extracting overlapping triples and enhancing performance.

The advancement of neural architectures, especially deep learning models, has significantly impacted the development of more sophisticated extraction methods. Convolutional network models, recurrent nerual network models, attention models, and graph models have been at the forefront of this evolution [1]. These models excel in capturing the nuanced relationships between entities within texts and have set new benchmarks on publicly available datasets.

One of the main challenges in RTE is handling overlapping entities within the extracted triples. Different classes of overlaps, namely No entity overlap (NEO), entity pair over1ap (EPO), and single entity over1ap (SEO), present unique challenges that demand innovative solutions [1]. Approaches like L-RTE [4] and ETL-Span [5] have been developed to address these challenges by incorporating entity-guided joint learning methods.

Recent advancements in relational triple extraction highlight the evolution of various methodologies and approaches in this field. Yuan et al. [6] emphasized the significance of relation-specific features for enhancing triad extraction accuracy, suggesting a shift towards relation-centric approaches. Sun et al. [7] introduced a progressive multitask learning method, indicating the efficacy of multitask learning in complex RTE. Ye et al. [8] explored the generative aspects of NLP models, proposing a contrastive triple extraction method, which could improve the robustness of extraction techniques. Wei et al. [9] and Ren et al. [10] both underscored the importance of tagging in RTE, offering perspectives for refining accuracy and integration in extraction frameworks. Nayak and Ng [11] focused on optimizing encoder-decoder models, contributing to the shared perceptual learning of such systems. Hoffmann et al. [12] discussed weak supervision methods for extracting overlapping relations, vital for handling complex relational structures. A bidirectional framework for RTE was proposed by another study [13], critiquing unidirectional frameworks and suggesting an alternative. Additional research [9-13] introduced various methodologies, including span-based tagging, hierarchical boundary tagging [5], token pair linking problems [14], and innovative models like PRGC [15] and StereoRel [16], each addressing specific challenges such as redundancy, poor generalization, and information loss in the extraction process.

More recent studies have further expanded the scope of RTE. A novel model focusing on entity role attribute recognition addresses the overlapping problem in joint extraction models, employing an approach that integrates the segregation of low-level features with the unification of high-level concepts together for better prediction and interpretation [17]. Another research emphasizes the incorporation of syntactic information in RTE, proposing a hyper-relational knowledge graph embedding model named HINGE, to refine the interaction of triplets and associated key-value pairs, demonstrating its effectiveness on several link prediction tasks [18]. Additionally, a model focused on global features, which utilizes table filling, has been proposed to maximize the use of global relationships between relations and token pairs, thus enhancing the extraction process from complex sentences [19].

The studies above highlight the continuous evolution and diversification of methodologies in the field, collectively demonstrate the diverse and evolving landscape of methodologies in relational triple extraction, paving the way for more robust and efficient frameworks. Our proposed lexical-based bidirectional extraction framework and shared perceptual learning mechanism aim to synthesize these advancements, offering a comprehensive solution to the challenges faced in this domain. The adaptability and improved accuracy of our approach, evidenced by our experiments, mark a significant progression in RTE, building upon these foundational works.

## METHODOLOGY

This section outlines the specifics of our framework. An overview is depicted in Fig. 1. Our framework comprises three primary components: Feature extractor based on BERT and dependency syntax, Subject and object extractor, and Relational extractor. We will introduce these three components in three separate subsections.

### Feature Extractor

Firstly, a feature extractor is designed to obtain rich key features from the model for text and text syntax dimensions. Specifically, since different types of items in the triad have their own characteristics. We employ a pre-trained BERT model to create a sentence representation, and then generate initial representations of the subject, object, and relation through three different fully connected layers $V_s$, $V_o$, and $V_r$, respectively. The specific formulas are as follows:

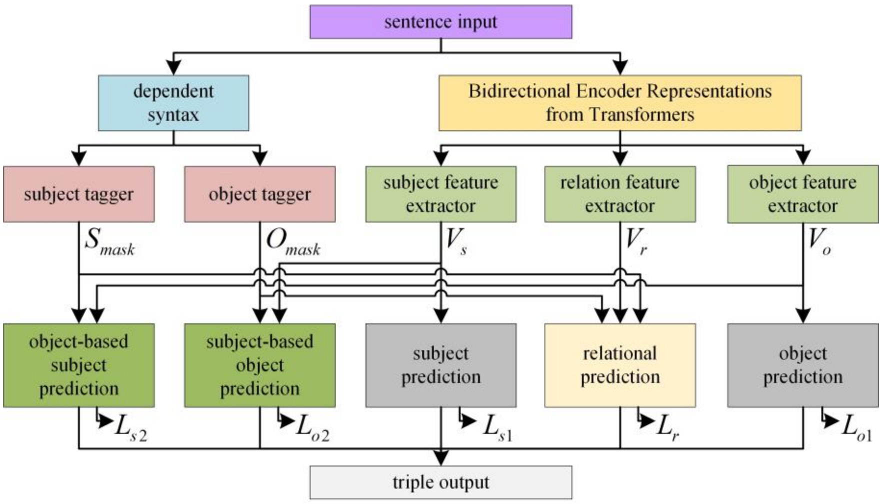


Figure 1. An overview of our proposed framework

$$V_s = W_s F + b \tag{1}$$

$$V_o = W_o F + b \tag{2}$$

$$V_r = W_r F + b \tag{3}$$

where $W_s$, $W_o$ and $W_r$ denote the initialized weights of fully connected layers respectively, $F$ is the feature representation of the sentence extracted by BERT, and $b$ denotes bias.

Subsequently, we find the lexical properties of all the words in the sentence through dependent syntax. Based on the lexical properties (e.g., nouns, proper nouns and pronouns), we find the words or phrases in the sentence that are the most probable to be the subject and object, and finally mask-tagging them, as shown in Figure 1.

**Subject and Object Extractor**

The subject and object extractor operates as a bidirectional framework, simultaneously pulling subjects and objects from two directions: one direction prioritizes the subject, using it as a condition to extract the object, while the reverse direction extracts in the opposite order.

Specifically, subject prediction module serves to extract all subjects from the given input sentences. In this process, each token within the sentence is assigned to a pair of probabilities, indicating the likelihood of being the begin and end token of a subject, respectively. The specific formulas are:

$$p_s^{begin} = f\left(W_s^{begin} V_s + b\right) \tag{4}$$

$$p_s^{end} = f\left(W_s^{end} V_s + b\right) \tag{5}$$

where $p_s^{begin}$ and $p_s^{end}$ indicate the probability of being the beginning marker and ending marker of the subject, respectively. $W_s^{begin}$ and $W_s^{end}$ are trainable matrices, and $b$ is bias vector. $f()$ denotes Sigmoid function.

Similarly, formulas for the object prediction module are as follows:

$$p_o^{begin} = f\left(W_o^{begin} V_o + b\right) \tag{6}$$

$$p_o^{end} = f\left(W_o^{end} V_o + b\right) \tag{7}$$

where $p_o^{begin}$ and $p_o^{end}$ denote the probability of begin marker and end marker of object, respectively. $W_o^{begin}$ and $W_o^{end}$ are trainable matrices.

Subject-based object prediction identifies objects using subjects extracted earlier. This method employs an iterative token structure, which processes each subject sequentially and extracts the related object through dependent syntax. Here, after a subject is chosen, every token in the sentence is assessed for its probability of being either the begin or end of the object linked to that subject, with the relevant formula detailed below:

$$P_o^{begin} = f\left(w_o^{begin}(V_o t_s\right)+b) \tag{8}$$

$$P_o^{end} = f\left(w_o^{end}(V_o t_s\right)+b) \tag{9}$$

where $P_o^{begin}$ and $P_o^{end}$ denote the probability of the begin marker and end marker of object, respectively. $t_s$ denotes the representation of the tagged subject. $w_o^{begin}$ and $w_o^{end}$ are trainable matrices.

Similarly, the specific formulas for the Object-based subject prediction module are as follows:

$$P_s^{begin} = f\left(w_s^{begin}(V_s t_o\right)+b) \tag{10}$$

$$P_s^{end} = f\left(w_s^{end}(V_s t_o\right)+b) \tag{11}$$

where $P_s^{begin}$ and $P_s^{end}$ denote the probability of the begin marker and end marker of subject, respectively. $t_o$ denotes the representation of the tagged object. $w_s^{begin}$ and $w_s^{end}$ are trainable matrices.

The cross-entropy loss function takes into account the predicted probabilities of each category to assess the difference between the model's output and the ground truth label, thereby offering precise error feedback to the model. Cross-entropy is particularly effective in multi-categorization tasks because it not only penalizes incorrect predictions, but also adjusts the strength of the penalty based on the uncertainty of the prediction. The losses of the above four labeling modules are denoted as $L_{s1}$, $L_{o2}$, $L_{o1}$ and $L_{s2}$ respectively. the specific formulas for $L_{s1}$ and $L_{o1}$ are given here as an example:

$$\sigma(p,t) = -\left[tlogp + (1-t)\log(1-p)\right] \tag{12}$$

$$L_{s1} = \frac{1}{2\times u}\sum_{n\in\{start,end\}}\sum_{i=1}^{n}\sigma\left(p_s^{i,n}, t_s^{i,n}\right) \tag{13}$$

$$L_{o1} = \frac{1}{2\times u}\sum_{n\in\{start,end\}}\sum_{i=1}^{n}\sigma\left(p_o^{i,n}, t_o^{i,n}\right) \tag{14}$$

where $\sigma(p,t)$ denotes binary cross entropy loss, $p\in(0,1)$ denotes the predicted probability, and $t$ is the true tag. $u$ denotes number of tokens in input sentence.

**Relational Extractor**

The relation extractor is designed to identify the type of relation present in a sentence. The main idea is to identify interconnections between subjects and objects. To achieve this goal, we identify subject-object pairs through dependent syntax and introduce attention mechanism to represent the relationship between subject and object, which essentially calculates the contextual importance score of each entity to achieve effective classification through the fully connected (FC) layer. Formulas are as follows:

$$SubjectScore = Softmax\left(W_s X_s + b\right) \tag{15}$$

$$ObjectScore = Softmax\left(W_o X_o + b\right) \quad (16)$$

where $X_s$ and $X_o$ denote subject and object, respectively. $W_s$ and $W_o$ are trainable matrices. $Softmax()$ denotes the activation function.

Then, the subject score and object score are combined with the vector $X_s$ and $X_o$, respectively. The final classification result is obtained after full connection and activation function.

## EXPERIMENTAL PROCESS AND RESULTS ANALYSIS

### Experimental Data

Our framework is evaluated on following benchmark datasets: NYT [12], WebNLG [20]. In addition, we follow the recent studies [7,9,14] with the same datasets. Statistics for datasets are listed in Table 1.

Keep in mind that both the NYT and WebNLG datasets have two different versions, each following particular annotation guidelines. The first version focuses on marking only the last token of entities, while the second encompasses the annotation of the entire span of the entities. Researchers often select between these versions based on their specific needs. For clarity, we refer to the datasets using the first annotation style as NYT* and WebNLG*, while those following the second style simply as NYT and WebNLG. It's evident that datasets with comprehensive annotations provide a more accurate measure of a model's true capabilities.

Table 1. Experimental datasets

| Dataset | NYT | | WebNLG | |
|---|---|---|---|---|
| | Test | Train | Test | Train |
| Normal | 3266 | 37013 | 246 | 1596 |
| SEO | 1297 | 14735 | 457 | 3406 |
| EPO | 978 | 9782 | 26 | 227 |
| ALL | 5000 | 56195 | 703 | 5019 |

### Experimental Setup

In our model, we initially set the following key hyperparameters: Adam optimizer is used with learning rate of $3\times10^{-5}$, batch size of 6, training period of 100, and maximum length to be 100 words. We utilize the SoftMax activation function, which effectively converts vectors into probability values.

### Performance Comparison

Table 2 outlines the performance of other models in comparison to our framework. Comparative models include ETL-Span [5], WDec [11], RSAN [6], RIN [10], CasRelLSTM [9], PMEIlLSTM [7], TPLinkerLSTM [14], R-BPtrNetLSTM [21], CGTUniLM [8], CasRelBERT [9], PMEIBERT [7], TPLinkerBERT [14], StereoRelBERT [16], PRGCBERT [15], R-BPtrNetlBERT [21], BiRTELSTM [13], BiRTEBERT [13].

As indicated in Table 2, our framework consistently surpasses the other models in all metrics as well as in overall performance.

The experimental results demonstrate performance of lexicon-based bidirectional extraction framework compared to other models. The framework consistently achieves higher precision, recall, and F1-scores across different datasets, including NYT and WebNLG. Improvement can be attributed to bidirectional extraction mechanism, which effectively mitigates the traditional reliance on subject extraction, enhancing the robustness and accuracy of the model. Additionally, the integration of the shared perceptual learning mechanism helps address convergence inconsistencies, ensuring more stable and reliable training outcomes. Overall, the results validate the effectiveness in achieving SOTA performance across various relational extraction tasks.

### Ablation Experiments

A series of ablation experiments are designed to acquire a deeper understanding of the roles and importance of each constituent in the proposed bidirectional extraction model. Through these experiments, the change in model performance when removing or modifying specific features can be evaluated, which not only validates the design choices, but also helps to identify critical parts of the model, which is crucial for future optimization and iteration.

To evaluate the effectiveness of the two-way extraction mechanism, the following comparison models are set up:

**Model A**: The complete two-way extraction model, which serves as the baseline control group.

**Model B**: Using only the subject-to-bin extraction path and ignoring the information flow from the bin to the subject.

**Model C**: Use only the bin-to-subject extraction path and ignore the subject-to-bin information flow.

Table 2. Performance comparison

| Model | NYT* | | | WebNLG* | | | NYT | | | WebNLG | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | P | R | F1 | P | R | F1 | P | R | F1 | P | R | F1 |
| R-BPtrNet$_{LSTM}$ | 90.9 | 91.3 | 91.1 | 90.7 | 94.6 | 92.6 | - | - | - | - | - | - |
| WDec | - | - | - | - | - | - | 88.1 | 76.1 | 81.7 | -- | - | - |
| RIN | 87.2 | 87.3 | 87.3 | 87.6 | 87.0 | 87.3 | 83.9 | 85.5 | 84.7 | 77.3 | 76.8 | 77.0 |
| CasRel$_{LSTM}$ | 84.2 | 83.0 | 83.6 | 86.9 | 80.6 | 83.7 | - | - | - | - | - | - |
| RSAN | - | - | - | - | - | - | 85.7 | 83.6 | 84.6 | 80.5 | 83.8 | 82.1 |
| PMEIl$_{LSTM}$ | 88.7 | 86.8 | 87.8 | 88.7 | 87.6 | 88.1 | 84.5 | 84.0 | 84.2 | 78.8 | 77.7 | 78.2 |
| TPLinker$_{LSTM}$ | 83.8 | 83.4 | 83.6 | 90.8 | 90.3 | 90.5 | 86.0 | 82.0 | 84.0 | 91.9 | 81.6 | 86.4 |
| ETL-Span | 84.9 | 72.3 | 78.1 | 84.0 | 91.5 | 87.6 | 85.5 | 71.7 | 78.0 | 84.3 | 82.0 | 83.1 |
| CGT$_{UniLM}$ | 94.7 | 84.2 | 89.1 | 92.9 | 75.6 | 83.4 | - | - | - | - | - | - |
| CasRel$_{BERT}$ | 89.7 | 89.5 | 89.6 | 93.4 | 90.1 | 91.8 | 89.8 | 88.2 | 89.0 | 88.3 | 84.6 | 86.4 |
| PMEI$_{BERT}$ | 90.5 | 89.8 | 90.1 | 91.0 | 92.9 | 92.0 | 88.4 | 88.9 | 88.7 | 80.8 | 82.8 | 81.8 |
| TPLinker$_{BERT}$ | 91.3 | 92.5 | 91.9 | 91.8 | 92.0 | 91.9 | 91.4 | 92.6 | 92.0 | 88.9 | 84.5 | 86.7 |
| StereoRel$_{BERT}$ | 92.0 | 92.3 | 92.2 | 91.6 | 92.6 | 92.1 | 92.0 | 92.3 | 92.2 | - | - | - |
| PRGC$_{BERT}$ | 93.3 | 91.9 | 92.6 | 94.0 | 92.1 | 93.0 | 93.5 | 91.9 | 92.7 | 89.9 | 87.2 | 88.5 |
| R-BPtrNetl$_{BERT}$ | 92.7 | 92.5 | 92.6 | 93.7 | 92.8 | 93.3 | - | - | - | - | - | - |
| BiRTE$_{LSTM}$ | 86.5 | 89.0 | 87.7 | 90.5 | 91.6 | 91.0 | 86.4 | 87.1 | 86.7 | 85.2 | 87.3 | 86.2 |
| BiRTE$_{BERT}$ | 92.2 | 93.8 | 93.0 | 93.2 | 94.0 | 93.6 | 91.9 | 93.7 | 92.8 | 89.0 | 89.5 | 89.3 |
| Our framework | 93.4 | 94.0 | 93.0 | 94.0 | 94.5 | 93.8 | 93.2 | 94.1 | 93.2 | 90.0 | 90.4 | 90.8 |

With this setup, the performance difference between unidirectional and bidirectional extraction can be compared to validate the effectiveness of bidirectional mechanism.

Considering the application of dependent syntactic analysis in subject-object pair extraction, the following models are designed for comparison:

**Model D**: Remove the dependency syntactic analysis function and do not use syntactic features to guide the identification of subject-object pairs.

This model helps to evaluate the specific contribution of dependent syntactic analysis to model performance.

A version of the model without this mechanism was designed to test the impact of the shared perceptual learning mechanism:

**Model E**: Each extractor is trained independently during the training process without sharing perceptual information.

The utility of the shared perceptual learning mechanism in multi-task learning is evaluated by comparing Model E with the full Model A.

Attention mechanism acts a vital role in relation extraction. An experiment was designed to assess its importance:

**Model F**: The attention mechanism is removed from the relationship extractor to test its impact on performance.

Each ablation model will be evaluated on the same standard dataset using precision, recall, and F1 score as the three metrics.

The full model (Model A) shows the best overall performance, demonstrating the effectiveness of the integration of the bi-directional extraction framework, dependent syntactic analysis, shared perceptual learning mechanism, and attention mechanism.

Models B and C show lower performance than Model A, indicating that bidirectional extraction captures relational data more comprehensively than unidirectional extraction, improving accuracy and robustness. The slightly better performance of Model C over Model B suggests that the object-to-subject path may provide additional contextual information, aiding in more accurate subject identification.

Table 3. Component ablation experiment results

| Model | Precision | Recall | F1 |
|---|---|---|---|
| Model A | 93.4 | 94 | 93.7 |
| Model B | 90.2 | 89.5 | 89.8 |
| Model C | 91 | 90.3 | 90.6 |
| Model D | 88.9 | 87.7 | 88.3 |
| Model E | 92.1 | 91.5 | 91.8 |
| Model F | 90.5 | 91.2 | 90.8 |

Model D performed the worst in the absence of dependent syntactic analysis. This demonstrates the important role of dependent syntactic analysis in correctly identifying subject-object pairs, especially when dealing with sentences with complex grammatical structures or containing nested and long-distance dependencies. Dependency syntactic analysis provides key structural information that is essential for the model to understand the various relationships in a sentence.

The relatively small performance degradation of Model E suggests that although the shared perceptual learning mechanism does help in optimizing model training and enhancing the synergy of the individual extraction tasks, its core functionality is inferior to that of the dependent syntactic analysis and the bidirectional extraction mechanism. The shared perceptual learning mechanism may mainly affect the efficiency of the training process and the generalization ability, rather than directly improving the basic recognition ability.

The performance of model F decreases after the removal of the attention mechanism, but the decrease is moderate. This indicates that the attention mechanism improves the capability to recognize the relationship between subjects and objects, helping to filter out irrelevant information by providing the model with the ability to be able to focus on key information. Nevertheless, the impact of the attention mechanism does not appear to be decisive, but it does effectively improve the extraction and categorization of relationships.

The analysis shows that while each component contributes to model performance, their influence varies. Dependency syntactic analysis and the bidirectional extraction mechanism are the most critical factors, emphasizing their central role in improving accuracy and robustness. Shared perceptual learning and the attention mechanism also have positive impacts, mainly optimizing training efficiency and handling complex data.

In summary, the design's reasonableness and effectiveness are verified through comparative experiments on multiple standard datasets. The proposed model not only reaches optimal performance but also improves the performance of other labeling-based methods by integrating the bidirectional extraction framework and shared perception learning mechanism.

## CONCLUSION

We present a lexicon-based bidirectional extraction framework designed to address the unidirectional frame sensitivity problem in traditional relational triple extraction methods. Extensive testing across various benchmark datasets confirms the efficiency of our approach, outperforming several well-established models. The bidirectional extraction mechanism, combined with the shared perceptual learning mechanism, enhances both accuracy and robustness, particularly by reducing dependency on subject extraction and improving training stability. Nonetheless, despite these encouraging outcomes, some limitations persist. The framework's performance remains influenced by the quality of the input data, and its effectiveness on more complex, domain-specific datasets may need additional improvement.

Future work could explore incorporating more advanced attention mechanisms or graph-based neural networks to enhance the model's ability to handle overlapping entities. Additionally, expanding the framework to adapt to more diverse and domain-specific datasets will help generalize its application in real-world scenarios. We also suggest investigating unsupervised or semi-supervised learning methods to lessen dependence on extensive annotated datasets, which could expand the framework's usability in low-resource environments.

**ACKNOWLEDGEMENT**

This research is supported by the National Natural Science Foundation of China (General Program), grant number 62072031. This research is additionally supported by the Second Department of the Second Institute of China Aerospace Science and Industry Corporation, with the grant number classified.

**REFRENCES**

[1] T. Nayak, N. Majumder, P. Goyal, S. Poria, Deep Neural Approaches to Relation Triplets Extraction: A Comprehensive Survey, Cogn Comput, Vol. 13, 2021, pp. 1215–1232, doi:10.1007/s12559-021-09917-7.

[2] D. Zelenko, C. Aone, A. Richardella, Kernel Methods for Relation Extraction, Journal of Machine Learning Research, Vol. 3, 2003, pp. 1083–1106.

[3] Y.S. Chan, D. Roth, Exploiting Syntactico-Semantic Structures for Relation Extraction, in Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1, Association for Computational Linguistics, USA, June 19, 2011, pp. 551–560.

[4] A. Hari, P. Kumar, Automated Relational Triple Extraction from Unstructured Text Using Transformer, in Innovations in Electrical and Electronic Engineering, Springer Nature Singapore, 2022, pp. 472–480.

[5] B. Yu, Z. Zhang, X. Shu, Y. Wang, T. Liu, B. Wang, S. Li, Joint Extraction of Entities and Relations Based on a Novel Decomposition Strategy, 2020.

[6] Y. Yuan, X. Zhou, S. Pan, Q. Zhu, Z. Song, L. Guo, A Relation-Specific Attention Network for Joint Entity and Relation Extraction, International Joint Conference on Artificial Intelligence, 2021, ISBN 978-0-9992411-6-5.

[7] K. Sun, R. Zhang, S. Mensah, Y. Mao, X. Liu, Progressive Multi-Task Learning with Controlled Information Flow for Joint Entity and Relation Extraction, Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 35, 2021, pp. 13851–13859, doi:10.1609/aaai. v35i15.17632.

[8] H. Ye, N. Zhang, S. Deng, M. Chen, C. Tan, F. Huang, H. Chen, Contrastive Triple Extraction with Generative Transformer, Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 35, 2021, pp. 14257–14265, doi:10.1609/aaai. v 35i16.17677.

[9] Z. Wei, J. Su, Y. Wang, Y. Tian, Y. Chang, A Novel Cascade Binary Tagging Framework for Relational Triple Extraction, 2020.

[10] K. Sun, R. Zhang, S. Mensah, Y. Mao, X. Liu, Recurrent Interaction Network for Jointly Extracting Entities and Classifying Relations, 2020.

[11] T. Nayak, H.T. Ng, Effective Modeling of Encoder-Decoder Architecture for Joint Entity and Relation Extraction, Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 34, 2020, pp. 8528–8535, doi:10.1609/aaai. v34i05.6374.

[12] R. Hoffmann, C. Zhang, X. Ling, L. Zettlemoyer, D.S. Weld, Knowledge-Based Weak Supervision for Information Extraction of Overlapping Relations, in Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, Eds. D. Lin, Y. Matsumoto, R. Mihalcea, Association for Computational Linguistics, Portland, Oregon, USA, June 2011, pp. 541–550.

[13] F. Ren, L. Zhang, X. Zhao, S. Yin, S. Liu, B. Li, A Simple but Effective Bidirectional Framework for Relational Triple Extraction, in Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining, Association for Computing Machinery, New York, NY, USA, February 15, 2022, pp. 824–832.

[14] Y. Wang, B. Yu, Y. Zhang, T. Liu, H. Zhu, L. Sun, TPLinker: Single-Stage Joint Extraction of Entities and Relations Through Token Pair Linking, 2020.

[15] H. Zheng, R. Wen, X. Chen, Y. Yang, Y. Zhang, Z. Zhang, N. Zhang, B. Qin, M. Xu, Y. Zheng, PRGC: Potential Relation and Global Correspondence Based Joint Relational Triple Extraction, 2021.

[16] X. Tian, L. Jing, L. He, F. Liu, StereoRel: Relational Triple Extraction from a Stereoscopic Perspective, in Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers), Eds. C. Zong, F. Xia, W. Li, R. Navigli, Association for Computational Linguistics, Online, August 2021, pp. 4851–4861.

[17] B.D. Trisedya, G. Weikum, J. Qi, R. Zhang, Neural Relation Extraction for Knowledge Base Enrichment, in Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Florence, Italy, 2019, pp. 229–240, DOI: 10.18653/v1/P19-1023.

[18] P. Rosso, D. Yang, P. Cudré-Mauroux, Beyond Triplets: Hyper-Relational Knowledge Graph Embedding for Link Prediction, Proceedings of The Web Conference 2020, 2020, pp. 1885–1896, DOI: 10.1145/3366423.3380257.

[19] J. Luzuriaga, E. Muñoz, H. Rosales-Méndez, A. Hogan, Merging Web Tables for Relation Extraction With Knowledge Graphs, IEEE Trans. Knowl. Data Eng., Vol. 35, No. 2, 2023, pp. 1803–1816, DOI: 10.1109/TKDE.2021.3101479.

[20] C. Gardent, A. Shimorina, S. Narayan, L. Perez-Beltrachini, Creating Training Corpora for NLG Micro-Planners, in Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Eds. R. Barzilay, M.-Y. Kan, Association for Computational Linguistics, Vancouver, Canada, July 2017, pp. 179–188.

[21] Y. Chen, Y. Zhang, C. Hu, Y. Huang, Jointly Extracting Explicit and Implicit Relational Triples with Reasoning Pattern Enhanced Binary Pointer Network, 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, R. Cotterell, T. Chakraborty, Y. Zhou, Association for Computational Linguistics, Online, June 2021, pp. 5694–5703.