A Comprehensive Survey on Relation Extraction: Recent Advances and New Frontiers

XIAOYAN ZHAO, The Chinese University of Hong Kong, Hong Kong, Hong Kong
YANG DENG, Singapore Management University, Singapore, Singapore
MIN YANG, Shenzhen Institutes of Advanced Technology Chinese Academy of Sciences, Shenzhen, China
LINGZHI WANG, The Chinese University of Hong Kong, Hong Kong, Hong Kong
RUI ZHANG, Huazhong University of Science and Technology, Wuhan, China
HONG CHENG, The Chinese University of Hong Kong, Hong Kong, Hong Kong
WAI LAM, The Chinese University of Hong Kong, Hong Kong, Hong Kong
YING SHEN, Sun Yat-Sen University, Shenzhen, China
RUIFENG XU, Computer Science and Technology, Harbin Institute of Technology Shenzhen, Shenzhen, China

Relation extraction (RE) involves identifying the relations between entities from underlying content. RE serves as the foundation for many natural language processing (NLP) and information retrieval applications, such as knowledge graph completion and question answering. In recent years, deep neural networks have dominated the field of RE and made noticeable progress. Subsequently, the large pre-trained language models (PLMs) have taken the state-of-the-art RE to a new level. This survey provides a comprehensive review of existing deep learning techniques for RE. First, we introduce RE resources, including datasets and evaluation metrics. Second, we propose a new taxonomy to categorize existing works from three perspectives, i.e., text representation, context encoding, and triplet prediction. Third, we discuss several important challenges faced by RE and summarize potential techniques to tackle these challenges. Finally, we outline some promising future directions and prospects in this field. This survey is expected to facilitate researchers' collaborative efforts to address the challenges of real-world RE systems.

CCS Concepts: • **Computing methodologies → Natural language processing**; **Neural networks;**

This research is supported in part by grants from the Research Grant Council of the Hong Kong Special Administrative Region, China (No. CUHK 14217622). Min Yang was supported by National Key Research and Development Program of China (2022YFF0902100), National Natural Science Foundation of China (62376262), the Natural Science Foundation of Guangdong Province of China (2024A1515030166), Shenzhen Science and Technology Innovation Program (KQTD20190929172835662), Shenzhen Basic Research Foundation (JCYJ20210324115614039).

Authors' Contact Information: Xiaoyan Zhao, The Chinese University of Hong Kong, Hong Kong, Hong Kong; e-mail: xzhao@se.cuhk.edu.hk; Deng Yang, Singapore Management University, Singapore, Singapore; e-mail: ydeng@nus.edu.sg; Min Yang (Corresponding author), Shenzhen Institutes of Advanced Technology Chinese Academy of Sciences, Shenzhen, Guangdong, China; e-mail: min.yang@siat.ac.cn; Lingzhi Wang, The Chinese University of Hong Kong, Hong Kong, Hong Kong; e-mail: lzwang1120@gmail.com; Rui Zhang, Huazhong University of Science and Technology, Wuhan, China; e-mail: rayteam@yeah.net; Hong Cheng, The Chinese University of Hong Kong, Hong Kong, Hong Kong; e-mail: hcheng@se.cuhk.edu.hk; Wai Lam, The Chinese University of Hong Kong, Hong Kong, Hong Kong; e-mail: wlam@se.cuhk.edu.hk; Ying Shen (Corresponding author), Sun Yat-Sen University, Shenzhen, Guangdong, China; e-mail: sheny76@mail.sysu.edu.cn; Ruifeng Xu, Computer Science and Technology, Harbin Institute of Technology Shenzhen, Shenzhen, GuanDong, China; e-mail: xuruifeng@hit.edu.cn.

ACM 0360-0300/2024/07-ART293
https://doi.org/10.1145/3674501

Additional Key Words and Phrases: Relation extraction, deep learning, pre-trained language models, low-resource relation extraction

ACM Reference Format:
Xiaoyan Zhao, Yang Deng, Min Yang, Lingzhi Wang, Rui Zhang, Hong Cheng, Wai Lam, Ying SHEN, and Ruifeng Xu. 2024. A Comprehensive Survey on Relation Extraction: Recent Advances and New Frontiers. *ACM Comput. Surv.* 56, 11, Article 293 (July 2024), 39 pages. https://doi.org/10.1145/3674501

1 Introduction

Relation extraction (**RE**) is an essential task in **natural language processing** (**NLP**), which involves extracting entities and relations between them from underlying content. In this article, we primarily focus on binary relations as the main unit of analysis for RE tasks. Each relation is represented as a triplet ⟨*head_entity, relationship, tail_entity*⟩, consisting of two entities and the relation between them. RE facilitates the extraction of structured information from vast troves of unstructured texts, thereby unlocking the value hidden within such data. It can be used for many downstream applications [111, 186], such as **knowledge graph** (**KG**) completion [23] and alignment [229], **question answering** (**QA**) [100], and information retrieval [202]. In the era of **Large Language Models** (**LLMs**), RE methods continue to demonstrate significant advantages. LLMs struggle to accurately retain all the knowledge implied within the text, especially in handling long-tail texts where errors in judging relationships between entities are prone to occur. Thus, RE techniques serve as a potent technical complement in enhancing the accuracy of LLMs. Furthermore, in rapidly evolving domains where new entities, relationships, and concepts frequently emerge, RE methods offer the flexibility to effectively adapt to and incorporate new information, offering scalable solutions to the daunting task of mining structured insights from the vast expanse of unstructured data. Therefore, designing automatic approaches to extract the relations between entities contained in unstructured texts becomes increasingly necessary, leading to the booming development of RE.

In recent years, advances in **deep neural networks** (**DNNs**) and PLMs have significantly improved the performance of RE. These approaches can be categorized into two main types: **the pipeline-based RE approaches** [108] and **joint RE approaches** [112, 219, 238]. Pipeline-based approaches extract entities and relations from unstructured text through two separate stages, which first identify entities from the text and then detect the relation between any pairs of entities. For example, as illustrated in Figure 1, given the sentence "*ChatGPT is a chatbot launched by OpenAI*", pipeline-based approaches first identify the entities "*ChatGPT*" and "*OpenAI*", and then predict the relation "*product*" between the two entities. In the early stage, pipeline-based RE approaches [68, 108, 114] primarily use **Named Entity Recognition** (**NER**) tools to extract entities, and then classify the relations of entity pairs using supervised learning algorithms with feature engineering. Pipeline-based RE methods [12, 24, 179] often assume that the target entities are already identified, and the RE models merely need to predict the relations between any pair of entities. However, because the entity and RE processes are separated, pipeline-based approaches tend to suffer from error propagation, where relation classification can be affected by errors introduced during entity recognition.

Joint (non-pipeline) approaches, on the other hand, aim at addressing this challenge by jointly modeling entity recognition and relation classification tasks within a unified framework. Taking the second example in Figure 1 as an example, the sentence "*Sam Altman is the co-founder and CEO of OpenAI*" contains two relationships (i.e., "*co-founder*" and "*CEO*") with overlapping entities. RE systems must be able to accurately identify and distinguish between overlapping entities and

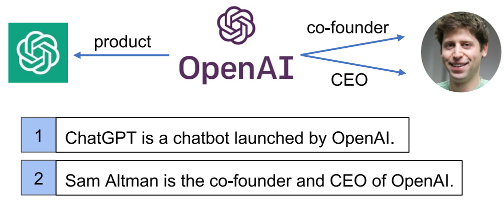


Fig. 1. Examples of RE.

relationships. Joint RE approaches tend to be less susceptible to error propagation due to several key reasons. First, this holistic training approach allows the model to learn optimal representations for both tasks concurrently, minimizing the impact of errors in one aspect on the other. Second, joint models can optimize objectives directly related to the overall task, such as maximizing the likelihood of correct entity pairs and relations. Third, by jointly learning entity recognition and RE tasks, joint models can adapt to errors in one task by leveraging information from the other, thus compensating for potential mistakes made at earlier stages. So far, many joint RE approaches have been proposed to extract entity and relation simultaneously. We generally divide them into four categories: span-based approaches [198, 241], sequence-to-sequence (Seq2Seq) approaches [223], MRC-based approaches [90, 238], and sequence labeling approaches [65, 219].

Despite the advances of **deep learning** (**DL**) for RE, several challenging problems still need to be solved in real-world scenarios. For example, many relations are "long-tail", where only a few frequent relations receive sufficient training examples. In contrast, the remaining infrequent relations usually suffer from a lack of labeled training data. However, DL requires massive amounts of training corpus, which is difficult to obtain in many real-life applications, especially in *low-resource settings*. **Distant supervision relation extraction (DSRE)** [217] is particularly appealing as it leverages existing structured information, such as KGs and databases, to generate labeled training data automatically. Nevertheless, distant supervision techniques may suffer from the wrong labeling problem and fail to handle long-tail relations with limited labeled instances. Therefore, **few-shot relation extraction (FSRE)** with limited labeled training samples has become a hot research topic [24, 121].

In addition, most existing studies focus on extracting relational facts from individual sentences. However, many real-life applications require the RE systems to identify entities and relations from a long document with multiple sentences. Following this direction, some recent studies [32, 82] have been proposed to solve *cross-sentence RE*, which attempt to identify relations that are mentioned across multiple sentences. Generally, there are two main research lines on cross-sentence RE. The first line of research is **document-level RE (DocRE)** [74], which has the potential to overcome the inherent limitations of sentence-level approaches and better capture the full range of relational information present in a document. The second line of research is **dialogue RE** [24, 144], which aims at discovering relation triplets appearing in multi-turn dialogues. Furthermore, the prosperity of RE in the general field motivates some works to focus on domain-specific RE from specialized articles. To the best of our knowledge, domain-specific RE approaches are still under-explored in previous surveys. In this article, we summarize the advanced RE approaches in specific domains (e.g., scientific, finance, medical, and biochemical).

Recently, some studies have also focused on other promising yet challenging RE problems, including **multi-modal relation extraction** [242], **cross-lingual relation extraction** [135], **temporal relation extraction** [151], and **evolutionary relation extraction** [237]. To facilitate building a comprehensive understanding of RE, we also review recent advances that address these challenging RE problems.

Overall, RE studies [36, 188] have been thriving in recent years. Although there have been several surveys on RE, they do not provide sufficient reviews of the above recent DL-based advances, current challenges, and future directions. In particular, the early surveys [9, 114, 220] emphasized traditional RE approaches (i.e., rule-based and machine learning-based approaches) in sentence-level settings. Detroja et al. [36] focused on both traditional RE approaches and DL-based approaches, but did not fully explore and omit the recent DL approaches. Moreover, Han et al. [59] reviewed existing RE approaches from four specific directions (i.e., utilizing more data, performing more efficient learning, handling more complicated context, and orienting more open domains). Xu et al. [188] focused on the low-resource RE problem. Bassignana and Plank [10] discussed RE datasets and scientific relation classification approaches. The most related survey to ours was proposed by Nayak et al. [111], which collectively introduced the general DL-based RE model architectures. However, several challenging issues and new frontiers in recent RE studies have not been discussed. Specifically, we argue that existing surveys mainly focused on limited aspects of RE and lack an in-depth sorting of the logical relationships among the classic approaches. Moreover, many emerging developments in this field have not yet been adequately explored. For example, PLMs and LLMs (e.g., BERT [38], GPT-3 [16], and ChatGPT [1]), which have been widely applied to enhance the outcomes of downstream RE in various scenarios, remain largely unexplored in previous RE surveys. In this survey, we first organize the general frameworks in the representative RE approaches and fully comb the recent studies into categories, illustrating the differences and connections between RE subtasks. In addition, we discuss the performance of RE on current solutions in diverse challenging settings (i.e., low-resource settings and cross-sentence settings) and specific domains (i.e., biomedical, finance, legal, and scientific fields), respectively. Furthermore, we present in-depth analyses that reveal the primary issues of RE with PLMs and discuss four main challenges (multi-modal RE, cross-lingual RE, temporal RE, and evolutionary RE) that need to be addressed. The holistic and multi-faceted views of RE methods discussed in our survey would allow readers to obtain a comprehensive landscape of available RE solutions and a good understanding of potential future directions. It is worth noting that to provide a comprehensive overview of RE, we selected bibliographical references based on criteria such as the significance of contributions (novel methods, datasets, metrics) and diversity of approaches. This approach ensures the inclusion of influential and representative works, reflecting the latest advancements and trends in the field across various techniques, datasets, and applications.

Contributions of this survey. This survey aims at providing a comprehensive overview of DL techniques in RE, which can provide researchers and practitioners with a comprehensive landscape of this area. First, we introduce representative RE datasets for verifying the RE methods. Second, we present a taxonomy classifying the representative RE approaches into several categories. Moreover, we explore and summarize the recent challenges and solutions faced by RE. Lastly, we present potential future directions in this field. This survey serves to facilitate collaborative efforts among researchers in tackling the challenges of RE.

In summary, we offer a comprehensive survey of RE techniques, analyzing the performance of RE models across various task settings and summarizing the limitations of existing models along with future directions for development. We begin with an overview of the research area, emphasizing the existing gaps in the literature concerning RE. Moving forward, Section 2 delves into datasets and evaluation metrics, establishing a foundational understanding for subsequent discussions. In Section 3, we explore DL techniques tailored for RE, with a taxonomy of text representation, context encoding, and triplet prediction. Section 4 discusses current challenging RE problems and solutions, including handling low-resource scenarios, cross-sentence extraction, and

[1]https://chat.openai.com/

Table 1. Statistics on RE Datasets

Corpus Name	General	Specific	Multi-lingual	Relation	Train/Test	Leaderboard
NYT [125]	✔			24	5.6 k/5 k	✔ [3]
WebNLG [51]	✔			171	5019/703	✔
WikiReading* [64]	✔			884	14.85 M/3.73 M	✔
WIKI-TIME [197]	✔			57	97.6 k/40 k	
SciERC [97]		Scientific		7	2,136/551	✔
FOBIE [81]		Scientific		3	1,238/300	✔
DialogRE [211]	✔			37	6 k/1.9 k	✔
FewRel 2.0 [50]		Medical		100+25	56 k/14 k	✔ [4]
ChemProt [115]		Biochemical		14	19.5 k/16.9 k	✔
DDI [63]		Biochemical		5	25.3 k/5.7 k	✔
DocRED* [204]	✔			96	4 k/1 k	✔
CUAD [62]		Legal		25	10.48 k/2.62 k	✔
FinRED [138]		Finance		29	5,699/1,068	✔
SMiLER [135]	✔		✔	36	733 k/15 k	
mLAMA [80]	✔		✔	5	–	
ACE 2023 [41]	✔		✔	24	100 k/50 k	✔
ACE 2024 [41]	✔		✔	24	300 k/50 k	✔

Document-level datasets are marked with*, while others are sentence-level datasets. The domains of the datasets are divided into general, specific, and multi-lingual categories. The ✔ mark in the leaderboard column indicates that the dataset has a leaderboard on the Articles with Code website[2].

adapting to domain-specific RE. Subsequently, in Section 5, we critically examine the integration of pre-trained language models and propose future directions, encompassing multi-modal, cross-lingual, temporal, evolutionary, and explainable RE approaches. Finally, in Section 6, we conclude the article by summarizing key findings and outlining promising directions for further research.

2 Preliminary

In this section, we first provide a formal problem definition of RE. Then, we introduce the recent benchmark corpora proposed for training deep RE models.[2] Finally, we present the evaluation metrics for evaluating the RE models.

2.1 Problem Definition

RE aims at automatically identifying the relations between entities in unstructured texts. Formally, given a natural language text x, the goal of the RE task is to predict a set of triplets, each consisting of a head entity e_1, a relation type r, and a tail entity e_2. The entities e_1 and e_2 can be words, phrases, or other syntactic units in the text, while the relation type r is a predefined type $r \in R$ that describes the relation between e_1 and e_2.

2.2 Datasets

Annotated datasets are crucial for the development of RE methods. We summarize the recently released and widely used benchmark datasets for RE in Table 1,[3,4,5] noting that the datasets listed are some representative examples and that many others also exist. Generally, these RE

[2]There are several online platforms for RE, including Google Cloud Natural Language https://cloud.google.com/natural-language, IBM Watson Natural Language Understanding https://www.ibm.com/products/natural-language-understanding and TextRazor https://www.textrazor.com/

[3]https://paperswithcode.com/datasets

[4]https://nlpprogress.com/english/relationship_extraction.html

[5]https://thunlp.github.io/2/fewrel2_da.html

Table 2. F1-scores (%) of Recent Representative Models on the NYT, WebNLG, and SciERC Datasets

Model	F1 (%)	Text Representation	Context Encoding	Triplet Prediction
		NYT		
UniRel [152]	93.7	Word-level, Position-enhanced	PLMs-based	Span-based
REBEL [17]	93.4	Word-level	PLMs-based	Seq2Seq
SPN [147]	92.5	Word-level, Position-enhanced	PLMs-based, Attention	Span-based
TDEER [89]	92.5	Word-level, Character-level	CNN&RNN, Attention	Sequence labeling
PFN [199]	92.4	Word-level, Position-enhanced	PLMs-based	Span-based
RIFRE [236]	92.0	Word-level, Position-enhanced	CNN&RNN, Attention	Sequence labeling
TPLinker [170]	91.9	Word-level, Position-enhanced	PLMs-based	Sequence labeling
RIN [148]	87.8	Word-level	RNN	Pipeline
		WebNLG		
UniRel [152]	94.7	Word-level, Position-enhanced	PLMs-based	Span-based
PFN [199]	93.6	Word-level, Position-enhanced	PLMs-based	Span-based
SPN [147]	93.4	Word-level, Position-enhanced	PLMs-based, Attention	Span-based
TDEER [89]	93.1	Word-level, Character-level	CNN&RNN, Attention	Sequence labeling
RIFRE [236]	92.6	Word-level, Position-enhanced	CNN&RNN, Attention	Sequence labeling
TPLinker [170]	91.9	Word-level, Position-enhanced	PLMs-based	Sequence labeling
RIN [148]	90.1	Word-level	RNN	Pipeline
CGT [207]	83.4	Word-level, Position-enhanced	PLMs-based	Sequence labeling
		SciERC		
PL-Marker [206]	53.2	Word-level, Position-enhanced	PLMs-based	Span-based
TriMF [140]	52.44	Word-level, Position-enhanced	PLMs-based	Span-based
SpERT.PL [131]	51.25	Word-level, Position-enhanced	PLMs-based, Attention	Span-based
SpERT [46]	50.84	Word-level, Position-enhanced	PLMs-based	Span-based
PURE [246]	50.1	Word-level	PLMs-based, Attention	Pipeline
DyGIE++ [165]	48.4	Word-level, Syntactic-enhanced	GNN, PLMs-based	Span-based
DyGIE [99]	41.6	Word-level, Syntactic-enhanced	GNN, PLMs-based	Span-based
SciIE [97]	39.3	Word-level, Syntactic-enhanced	GNN	Span-based

The models are categorized by their underlying architectures, including Transformer-based models, and LSTMs marked with †.

datasets can be roughly classified into four categories based on their data sources: (1) general corpora collected from news articles; (2) encyclopedic corpora collected from Wikipedia and Wikidata; (3) domain-specific corpora that contain scientific, finance, medical, and biochemical articles; (4) multi-lingual corpora that include input texts in multiple languages; and (5) multi-modal corpora that contain textual relations with visual information. Note that all datasets are manually annotated, except the NYT dataset [125], which is created by a distant supervision approach using the **knowledge base** (**KB**) Freebase. Some follow-up works calibrated subsets of the NYT dataset to obtain more accurate annotation, like WIKI-TIME [197]. Most existing datasets focus on sentence-level RE in the general domain. Recently, some works have started to focus on the annotation and evaluation setups in more complex scenarios, including document-level [64, 70, 204], low-resource [50], multi-modal [242], and multi-lingual [135] settings. To provide quantitative results, we include the corresponding leaderboard links in Table 1. Additionally, Table 2 lists representative methods for the popular benchmark datasets, offering a clear benchmark for future research in this domain. Table 2 highlights the effectiveness of Transformer-based models, which dominate the top performance across all datasets, reflecting their superior capability in handling complex language tasks.

2.3 Evaluation Metrics

The performance of supervised learning RE systems is typically measured by comparing the predicted labels to the corresponding ground-truth annotations. There are three main metrics [150]: precision (P), recall (R), and F1 score. Specifically, P measures the proportion of correctly

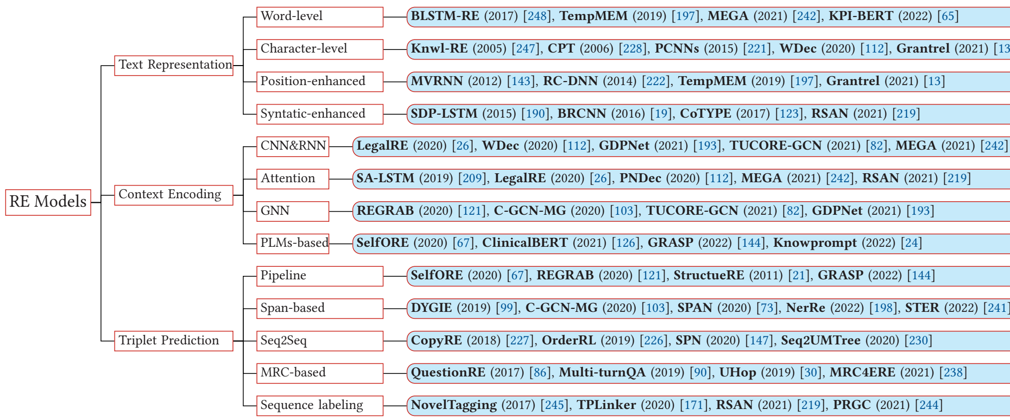


Fig. 2. The taxonomy of RE models and the corresponding representative methods in each category are listed.

recognized results, while R assesses the proportion of all correctly recognized entities. The F1 score, being the harmonic mean of precision and recall, offers a balanced reflection of the system's performance. For distant supervised RE tasks, labels are generated automatically from external KBs and may not be entirely accurate. As a result, metrics in supervised RE may not fully reflect the model's performance in real-world scenarios. Therefore, Precision@K, the **Precision-Recall Curve** (**PRC**), and its **Area Under Curve** (**AUC**) [250] are adopted as evaluation metrics for evaluating distantly supervised RE. Precision@K measures the proportion of correctly identified relations among the top-K predictions made by the system. For each instance, the RE model generates a ranked list of K predictions, prioritizing them based on certain relevance criteria. The top K predictions represent the subset of relations that the model deems most likely or relevant. However, evaluating Precision@K requires manual effort, as researchers must annotate the top-K output results of systems. The PRC and AUC enable us to understand the precision-recall tradeoff across various thresholds, comprehensively assessing the performance of distant-supervised RE models. While the metrics presented are valuable and commonly used for evaluating RE systems, it is important to note that they are not exhaustive. Other metrics may also be relevant, depending on the specific characteristics of the task and the goals of the evaluation.

3 DL Techniques for RE

Recent advances in RE have largely been driven by DL techniques. In this section, we propose a new taxonomy to summarize DL-based RE approaches from three perspectives: text representation, context encoding, and triplet prediction. For each part, we present a comprehensive review of approaches in the literature. The subset of representative models illustrated in Figure 2 serves as examples to illustrate this taxonomy.

3.1 Text Representation

For DL-based RE approaches, it is vital to learn powerful representations of the input data. Text representations encode each input token with a real-valued vector. Words that are similar in meaning are expected to be closer in vector space. The ability of such distributed representations to capture syntactic and semantic properties of words affects the language modeling performance of DL-based RE approaches. We review and discuss the various types of text representation learning approaches used in previous RE works, including word-level, character-level, position-level, and syntactic-level representations.

3.1.1 Word-Level Embeddings. Recent studies have demonstrated the importance of pre-trained word embeddings, which encode the meaning of input units into a real-valued continuous space. These word embeddings can be either fine-tuned or fixed during training.

Non-contextualized word embeddings, such as Word2Vec [105] and GloVe [116], are obtained by unsupervised algorithms, including **continuous bag-of-word** (**CBOW**) and continuous skip-gram models. These studies [158, 222] use high-dimensional distributed representations of words as input features for RE tasks, which encode the semantic information about entity words and help identify the relations among entities. For example, Zheng et al. [245] proposed an end-to-end model to jointly extract entities and relations in a single model, constructing the word embeddings trained on the NYT corpus through the Word2Vec toolkit. Zhou et al. [248] used the pre-trained 300-dimensional word vectors from Google in their proposed neural model for extracting relations.

Contextualized word embeddings of PLMs, such as BERT [38] and ELMo [117], have demonstrated the importance of pre-trained word embeddings. These PLMs can be further fine-tuned during RE model training. A significant advantage is that the embeddings are contextualized by their surrounding text, meaning the same word can have different embedding depending on its contextual use.

3.1.2 Character-Level Embeddings. To capture the sub-word level information, **character-level embeddings** [112] are introduced to encode fine-grained information such as n-gram, prefix, and suffix features. Previous methods [221, 228, 247] explore the utilization of both internal and external contexts. In these studies, the sentence is partitioned into three segments based on the two entities of interest: the internal context encompassing characters within these entities, and the external context encompassing characters surrounding them. Additionally, PLMs like BERT [38] inherently take subwords into account, further enriching the character-level representation and enabling the model to infer representations for unseen words. This characteristic is advantageous for handling out-of-vocabulary scenarios.

3.1.3 Position-Level Embeddings. In addition, Yan et al. [197] proposed using **position-enhanced embeddings** for text representations in RE, and experimental results demonstrated that adding position information could sufficiently exploit the relative distance of the target entity pairs. Yuan et al. [217] encoded the position information in sentences, which can be formulated as follows: first, for a sentence, transform the word at position i into a pre-trained word vector v_i [104]. Then, they calculate the relative distances to the target entities (i.e., d_1 and d_2) in the sentence and look up the position embedding table [222] to find their position embeddings p_{d_1} and p_{d_2}. Here, the position embedding table is randomly initialized and further updated during the processing of training. The word representation w_i is represented by concatenating v_i with p_{d_1} and p_{d_2}. After repeating these steps, each sentence is transformed into a fixed-sized matrix $C = [w_i, w_2, \ldots, w_m]^T$, where m is the maximum length of the sentence in whole input data, and w_i is a fixed-length vector. Sentences shorter than m are padded with zero vectors. Zeng et al. [221, 222] proposed using position embeddings for feature extraction in RE, and their results show that adding position information is superior to only using word information. Zeng et al. [222] exploited the position information to encode the relative distances to the target entity pairs. Zhang et al. [232] augmented the word representations with extra distributed representations of word position by combining the LSTM model with entity position-aware attention.

3.1.4 Syntactic-Level Embeddings. Moreover, another line of studies explored **syntactic-enhanced embeddings** to incorporate rich syntactic-related features into word embeddings, such as the **shortest dependency path** (**SDP**), **Part-of-Speech** (**POS**) tagging, WordNet hypernyms, and grammatical relations [26, 219]. For example, Zeng et al. [222] incorporated prior knowledge in texts, such as syntactic parsing and POS tagging, where the performance outperforms the baseline only using word-level representations. Xu et al. [190] employed rich features in addition to word embeddings, including the SDP, POS tags, WordNet hypernyms, and grammatical relations, jointly integrating syntax and semantics. Xu et al. [184] constructed a comprehensive word representation by concatenating the word representation and the syntactical representation, which contains dependency labels and dependency edge directions. Cai et al. [19] and Nayak and Ng [112] focused on the syntactic structures in the input sentences, which were obtained by a dependency parser and provided complementary evidence for relationships. Their boosted performances demonstrate that adding additional information may lead to improvements in RE performance. Ren et al. [123] proposed a domain-independent framework CoTYPE, which jointly embeds text features, type labels, entity and relation mentions. The entity and relation mentions with relevant candidate types are integrated into the model.

3.1.5 Summary. Word-level embeddings are commonly used as standalone representations, whereas other embeddings, such as character-level, position-level, and syntactic-level embeddings, are less frequently utilized on their own. While each type can represent specific aspects of textual information independently, hybrid embeddings combine multiple types of embeddings to capture a wider range of linguistic features, thereby enhancing the overall representation quality for RE tasks. However, blending features in hybrid embeddings can introduce complexity, potentially impacting the generality of neural RE models. The selection of external features depends on the specific application requirements.

3.2 Context Encoding

The word-level embeddings aim at extracting lexical-level features from the given input data. Context encoding is designed to learn sentence-level features by capturing contextualized information and filtering out irrelevant information in the text representation. Context encoding can be implemented with any popular neural network architecture, such as CNNs, RNNs, attention-based neural networks, PLMs, and prompt tuning. These methods aim at retaining almost all the information required to successfully predict the outputs.

3.2.1 Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs). CNNs [139, 217, 221, 222] effectively learn local and position-invariant contextual representations. For example, Zeng et al. [222] were among the first to use a convolutional deep neural network (CNN) for RE. It encodes the meaning of sentences not explicitly represented in the input representation. Zeng et al. [221] utilized a piecewise CNN model to scale hidden vectors for each word. The obtained feature vectors are then used to determine the relations through a feed-forward layer with a softmax function. Shen and Huang [139] incorporated a combination of the CNN model and an attention network, which extracts the global features and attentive features in the sentence. Yuan et al. [217] adopted a **piecewise-CNN** (**P-CNN**) to consider the specific situation in RE, consisting of a convolutional layer and a Piecewise Max-pooling layer [118]. Overall, CNNs are good at capturing local features within a sentence. However, CNNs may not capture long-distance dependencies efficiently, which is crucial in understanding complex sentence structures in RE tasks.

RNNs [68, 71, 71, 112, 232], including **long-short term memory** (**LSTM**) and **gated recurrent unit** (**GRU**), have shown remarkable achievements in modeling sequential data. This property provides an excellent way to compose long context-dependent representations of sequence [101]. Jat et al. [71] proposed a **bidirectional gated recurrent unit** (**Bi-GRU**) to extract the long-term dependency among the words in the input sequence. The text representations encoded by **bidirectional long-short term memory** (**Bi-LSTM**) [112] can efficiently incorporate the past and future text information [68, 71]. Zhang et al. [232] introduced a position-aware attention mechanism over

Bi-LSTM for the RE task, efficiently utilizing semantic similarity-based and position-based information. Overall, RNNs are designed to handle sequential data, making them more suitable than CNNs for capturing long-range dependencies in text. They sequentially process words and can theoretically remember all previous information. However, in practice, RNNs often struggle with long sequences due to vanishing or exploding gradient problems, making it hard to capture very long-distance dependencies.

3.2.2 Attention-Based Neural Networks. Attention-based neural networks [139, 209] enhance the correlations between relation representations and text representations, highlighting important information for RE. Earlier studies [71, 139] incorporated word-level attention with sentence-level RE. Recent works [92, 112, 209] combined attention networks with various models to capture multiple-grained entity and relation features. For example, Nayak and Ng [112] proposed a multi-focused attention model for RE, where dependency distance is incorporated to help identify the triplets in the input. The multi-factor attention helps focus on various pieces of evidence to determine the relationship. Yu et al. [209] introduced segment-level attention to select and model distributed representations of relational expressions. Li et al. [92] proposed a self-attention [161] enhanced model with entity-aware embeddings. Overall, attention mechanisms allow models to focus on relevant parts of the text when predicting relations, effectively overcoming the limitations of CNNs and RNNs in handling long-range dependencies. They can capture complex sentence structures and relationships between entities regardless of their position in the text. However, these models can be computationally expensive and require a significant amount of data to train effectively.

3.2.3 Graph Neural Networks (GNNs). GNNs [103, 129, 163, 177] attempt to capture the non-linear structure of the input sequence by constructing semantic graphs, empowering the RE models with relational reasoning ability on graphs. GNN-based methods offer several key advantages, such as the ability to capture the global structure of the graph, and the ability to learn the representations of nodes and edges simultaneously. Such graph-based models [56, 163, 177, 231] construct the non-linear structure of the input sequence via graphs, which provides a better way to represent the relationships between entities. In particular, Zhang et al. [231] utilized GCN and the syntactic dependency tree to construct the graph structure among the nodes. Then they built the adjacency matrix of the graph and included the edges from the SDP. Guo et al. [56] used a multi-head self-attention-based soft pruning strategy, which can identify the importance of edges in the graph. And some works [32, 103, 109, 129] used the shortest dependency tree path to create the connections among nodes. Overall, GNNs can capture the interconnectedness of entities and relations in a way that is difficult for purely sequential models. However, constructing such graph structures requires additional preprocessing, and requires that relationships are easy to accurately represent in a graph.

3.2.4 Pre-Trained Language Models (PLMs). Recently, PLMs [121, 142, 238] have shown remarkable achievements in modeling RE problems by eliciting rich knowledge from large pre-trained models. PLMs usually trained on large-scale corpora, such as BERT [38], ELMo [117], RoBERTa [95], and SpanBERT [78], intrinsically incorporate auxiliary embeddings (e.g., position and segment embeddings). PLMs [121, 238] provide rich semantic knowledge to the RE task, where the fine-tuning process is performed on annotated task-specific data to adapt semantic information for RE. As PLMs are pre-trained on large amounts of text data, fine-tuning for RE tasks [44, 142, 180] allows them better to understand the meaning and context of words and sentences. Particularly in scenarios with limited data availability, fine-tuning PLMs on the target task has proven to be an effective practice. Utilizing the information embedded in PLMs as the primary representation offers significant advantages in such cases. One of the challenges is that downstream RE tasks fine-tuned on PLMs usually have different objective forms, leading to performance degradation. Prompt tuning [24, 144] provides a new paradigm to stimulate the relation information of PLMs by bridging the format gap between the pre-training tasks and the downstream RE tasks. Recent works [60, 133] show that prompt learning can effectively leverage the knowledge encoded in the PLMs, especially for few-shot RE tasks in Section 4.1.2. Overall, PLMs offer substantial pre-trained knowledge that can capture intricate language patterns and dependencies. The main drawbacks of PLMs are their resource-intensive nature, requiring significant computational power for both training and inference and their tendency to overfit on smaller or domain-specific datasets.

3.2.5 Summary. Comparing the above encoders, there's a significant overlap in the application of these models, with many advanced systems combining their strengths. For example, earlier approaches combined RNNs or CNNs with attention mechanisms [139, 232] to capture both local features and global dependencies. Integrating GNNs with attention mechanisms [56] allows for dynamic focus on different parts of the graph, enhancing the model's ability to capture complex relationships. PLMs [38] inherently incorporate attention mechanisms like Transformers. The choice of model often depends on the specific requirements of the task, including the nature of the data, the computational resources available, and the desired level of accuracy.

3.3 Triplet Prediction

The triplet prediction involves detecting the entity boundaries and classifying the relation types in the input sentence. Various modeling paradigms have been proposed for decoding triplets in recent RE models. As illustrated in Figure 3, we group the existing triplet prediction paradigms into five categories, depending on the specific formulation of the RE task. The corresponding types of target triplets for RE models are listed, respectively.

3.3.1 The Pipeline-Based (Classification) Approaches. The pipeline-based (classification) approaches separate the extraction of entities and relations [108, 144]. As illustrated in Figure 3(a), the pipeline approaches carry out entity recognition and relation classification sequentially [108, 159]. In the first stage, all candidate entities in the sentence are annotated manually or identified via NER models. Then, a classifier is used to determine the relation between every possible pair of identified entities. The ultimate goal is to accurately and consistently identify and extract all relevant relationships from the input text. The pipeline approach assumes that entities are already identified, and models aim at identifying the relationship (relations R or *None*) between pairs of entities.

Different from the pipeline methods, joint-extraction approaches aim at finding both entities and relations in a sentence by extracting valid relation triplets. These models face a challenge when extracting triplets from sentences with overlapping entities, which can be divided into three categories: (i) **No Entity Overlap** (**NEO**), where triplets do not share any entities; (ii) **Single Entity Overlap** (**SEO**), where at least two triplets share exactly one entity; and (iii) **Entity Pair Overlap** (**EPO**), where at least two triplets share some entities in the same or reverse order. A sentence can belong to both the SEO and EPO categories. As shown in Table 3 [218], the overlapping entities are marked in bold. The triplets in the second example (SEO class) share one single entity, Donald Trump. The triplets in the third example (EPO class) have overlapping entity pairs (Japan, Tokyo). Joint RE aims at extracting all relevant relation triplets present in a given sentence.

3.3.2 The Span-Based Approaches. The span-based approaches [40, 103] process each sentence into spans and perform span classification to obtain predicted entities. Simultaneously, the detected entity pairs are regarded as candidate triplets for relation classification, as illustrated in Figure 3(b). Span-based approaches are shown with superior to previous pipeline-based methods [198]. These

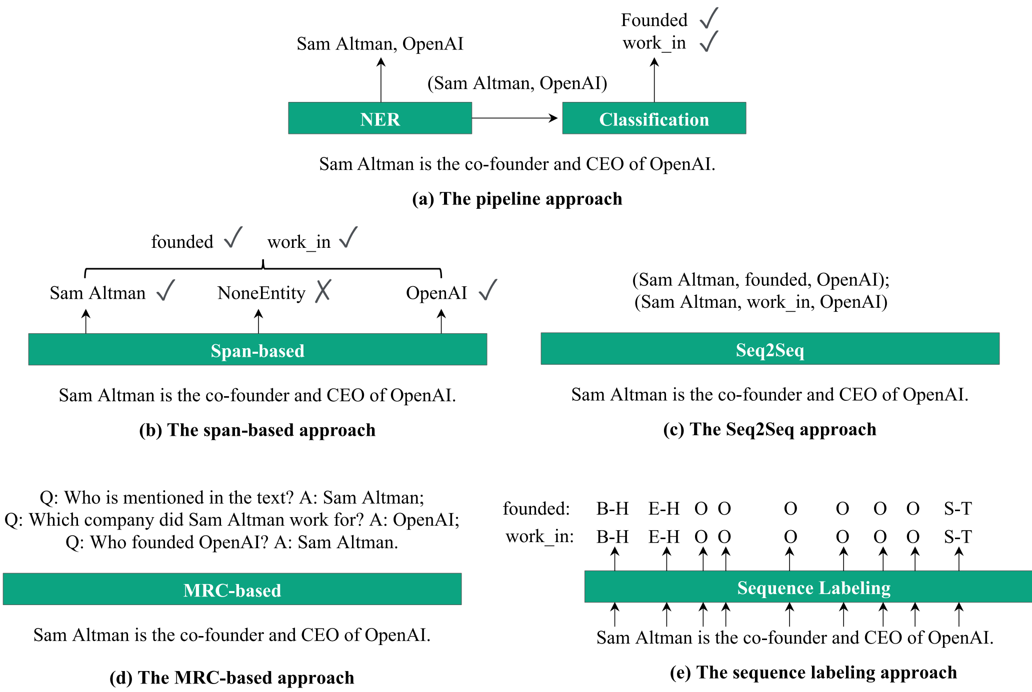


Fig. 3. Different RE modeling paradigms: (a) is the pipeline-based approach; (b)–(e) are the joint approaches. The target triplet types in different RE models are shown. Each paradigm uses the input sentence "Sam Altman is the co-founder and CEO of OpenAI." The target output triplets are <*Sam Altman, founded, OpenAI*> and <*Sam Altman, work_in, OpenAI*>.

Table 3. Examples of NEO, SEO, and EPO Cases

	Text	Triplets
NEO	The [United States] president [Donald Trump] will visit [Beijing], [China].	(Donald Trump, President_of, United States) (China, Contains, Beijing)
SEO	The [United States] president [Donald Trump] was born in [New York City].	(Donald Trump, President_of, United States) (Donald Trump, Born_in, New York City)
EPO	Martin went to [Tokyo] last week, which is the capital of [Japan].	(Japan, Contains, Tokyo) (Japan, Capital, Tokyo)

methods [45, 72, 73] utilize pre-trained Transformer blocks to map word embeddings into BERT embeddings, calculate span and relation representations, perform classification and filtration tasks, and generate contextual semantic representations using multiple attention variants. Zhao et al. [241] further define the privileged features in the RE task and propose a contrastive student-teacher learning framework to utilize the expert knowledge during training to enhance the performance of the model. Some works [97, 99, 165] utilize dynamically constructed span graphs to achieve high performance on various tasks such as entity recognition and RE. The most confident entity spans are selected and linked with confidence-weighted relation types and coreferences to construct the graphs, which iteratively optimize span representations.

3.3.3 The Seq2Seq-Based Approaches. The Seq2Seq-based approaches [112, 223, 226, 227, 230] receive unstructured text as input and directly generate ⟨*head_entity, relationship, tail_entity*⟩

triplets as a sequential output. Formally, a source sentence $S = \{x_1, x_2, \ldots, x_n\}$ is represented as a sequence of words, where x_i is the ith word in S and n is the length of S. Based on the text representation (in Section 3.1), the tag classifier predicts the relation types. The target sentence T is represented as a sequence of words $T = \{t_1, t_2, \ldots, t_m\}$, where t_j is the jth word in T. Figure 3(c) shows the target triplet types in Seq2Seq-based models, which are able to tackle the overlapping relations and reduce the excessive computations. However, when dealing with tasks involving multiple triplets within a single sentence, the inherent linearization process of Seq2Seq models may pose challenges in processing extracts from multiple triplets with overlapping entities. To address these challenges, recent studies [147, 230] have focused on strategies to avoid the limitations imposed by the sequential nature of Seq2Seq models. For example, to deal with the overlapping problem, recent works [171, 218, 244] design labeling strategies and perform the tagging process for multiple turns. These methods create specific sentence representations for each relation and then perform sequence labeling to extract the corresponding head and tail entities.

3.3.4 The MRC-Based Approaches. MRC-based approaches [90, 238] treat the entity RE task as a multi-turn QA task. For example, as shown in Figure 3(d), the relation type "work_in" between "Sam Altman" and "Open AI" can be formulated as "Question: Who is mentioned in the text? Answer: Sam Altman" and "Question: Which company did Sam Altman work for? Answer: Open AI". Therefore, the extraction of entities and relations in a sentence can be transformed into the QA task of identifying answer spans from the context. This transformation allows the RE task to exploit well-developed **machine reading comprehension** (**MRC**) models [90], which extract text spans in passages given queries. For example, Levy et al. [86] first formulated the RE task as a QA task, where the relations are defined by natural-language question templates. Li et al. [90] and Zhao et al. [238] further transformed the RE task into a multi-turn QA task, providing a natural way to identify the entities and relations in a sentence. The RE process is thus converted into extracting information from textual passages by answering questions posed about the text. Additionally, approaches [30, 76] treat the task as a series of questions and answers, where each turn corresponds to a step in the extraction process. Overall, the key idea of MRC-based approaches is to formulate questions that prompt the model to identify relevant entities and relations within the text.

3.3.5 The Sequence Labeling Approaches. Sequence labeling approaches solve RE task through shared parameters in an end-to-end manner, as illustrated in Figure 3(e). They perform joint RE by treating entity and relation types as well-designed tags [245] and predict a single tag for each token. Yu et al. [210] tackled the joint RE extraction using an end-to-end sequence labeling framework based on functional decomposition. By breaking down the original task into smaller components, the learning process is simplified, resulting in improved overall performance, as presented by the empirical analysis in [210]. To tackle the overlapping cases, some works [65, 171, 210, 219, 244] perform sequence labeling in multi-turn by generating a specific tag sequence for each given relation.

3.3.6 Summary. Previous RE surveys often overlook diverse decoding mechanisms. To bridge this gap, we provide a systematic survey of DL-based RE approaches focused on classifying the relations. Overall, both pipeline-based and joint RE approaches exhibit their pros and cons. The advantage of pipeline-based methods is that they are staged to detect named entities and classify relations, explicitly modeling the entity and relation information. However, the pipeline-based approaches assume that the entities are independent of relations, making them prone to accumulating errors and failing to capture the dependencies between entities and relations. In contrast, joint RE approaches are motivated by the fact that the entities and relations are closely related in real-world applications, thus avoiding error accumulation. Additionally, multiple relation triplets within an input text may share overlapping entities or relations.

Table 4. Overview of Input and Output for Each RE Task with Examples

	Task	Input	Example Input*	Output	Example Output
Low-resource RE	Distant Supervision RE	A bag of sentences S_b consisting of b sentences and an entity pair (e_1, e_2) presenting in all sentences.	Sentence-bag: #1: Barack Obama was born in the United States. #2: Barack Obama was the first African American to be elected to the president of the United States. #3: Barack Obama served as the 44th president of the United States from 2009 to 2017. ([136])	Bag relation r of the sentence-bag S_b	president_of
	Few-shot RE	(Train on small support set S) Predict the relation r for any given query instance x.	In 2001, he also published the "Khaki Shadows" that recounted the military history of Pakistan during the cold war. ([49])	Relation r	Facet_of
Cross-sentence RE	Document-level RE	Each sentence d_i in a document d	Lutsenko is a former minister of internal affairs. He occupied this post in the cabinets of Yulia Tymoshenko. The ministry of internal affairs is the Ukrainian police authority. ([109])	The relation r for each entity pair (e_1, e_2)	(Lutsenko, manage, internal affairs) (Lutsenko, work_with, Yulia Tymoshenko) (Yulia Tymoshenko, country_of_citizenship, Ukrainian)
	Dialogue RE	A dialogue $d = s_1 : t_1, s_2 : t_2, \ldots, s_m : t_m$ and entity piar (e_1, e_2)	S1: Hey Pheebs. | S2: Hey! | S1: Any sign of your brother? | S2: No, but he's always late. | S1: I thought you only met him once? | S2: Yeah, I did. I think it sounds y'know big sistery, y'know, 'Frank's always late.' | S1: Well relax, he'll be here.([211])	The relation r of (e_1, e_2) based on d	(Frank, per:siblings, S2) (S2, per:siblings, Frank) (S2, per:alternate_names, Pheebs)
Domain-specific RE	RE in Biomedical Field	A sentence s is inserted with four makers.	Patient was given e11 ibuprofen e12 for high e21 fever e22. ([126]) (Note: e11, e12, e21, and e22 at the beginning and end of the target entities (e_1, e_2).)	Relation r	may_treat
	RE in Finance Field	A sentence s	MEXICO CITY — State-owned oil company Pemex is reporting second quarter losses of $US5.2 billion ($A7.16 billion) due mainly lower petroleum prices. ([138])	(e_1, r, e_2) triplet set	(Pemex, product_or_material_produced, petroleum) (Pemex, headquarters_location, Mexico City)
	RE in Legal Field	A sentence s	On August 19, 2014, Mr. Su sold methamphetamine to Mr. Wang in Community A. ([145])	(e_1, r, e_2) triplet set	(Mr. Su, traffic_in, methamphetamine) (Mr. Su, sell_drug_to, Mr. Wang)
	RE in Scientific Field	A sentence s	MORPA is a fully implemented parser method developed for a text-to-speech system. ([97])	(e_1, r, e_2) triplet set	(MORPA, Used_for, text-to-speech system) (MORPA, Hyponym_of, parser)

4 Challenging RE Problems and Solutions

Section 3 summarizes the common practice for general RE problems. In this section, we review recent challenging RE problems and corresponding solutions. Table 4 shows the input and output for each challenging problem with examples.

4.1 Low-Resource RE

Supervised learning with DNNs requires a large-scale annotated training corpus which is difficult to obtain in real-world applications, especially in low-resource settings. Recently, many efforts have been made to address low-resource RE.

4.1.1 DSRE. DSRE aims at automatically leveraging the facts in large-scale KBs to generate the annotated triplets as weak supervision. This technique can be traced back to the early work of [107], which proposed obtaining relationships for entity pairs aligned in KBs, such as Wikidata [164], DBpedia [14], and Freebase [15]. Despite the large amount of training data obtained through distant supervision, DSRE suffers from noisy label problems because individual sentences may give incorrect cues. The noise present in this data mainly comes in two forms: (1) the obtained relations do not match the original meaning of the sentences, and (2) the relations and entities are missing due to incomplete KBs.

Existing DSRE studies mainly tackle the task at different granularities: (1) **Sentence-level**. These works [12, 44, 197] focus on finding accurate relational labels from the semantics of the input sentences. This approach is based on the strict assumption that if a pair of entities are found to share a relation in the KB, then any sentence containing that pair of entities is considered a positive instance of that relation. (2) **Bag-level**. This kind of approach is based on a slack assumption that at least one sentence in a "bag" of sentences should express the relation. There may exist several relations that can be chosen between specific entity pairs. To mitigate the effects of noisy samples and make them more robust, Zeng et al. [221] proposed a **Piecewise Convolutional Neural Networks** (**PCNNs**) model, which treats the distant supervised RE task as a multi-instance problem. The model avoids feature engineering and takes the uncertainty of instance labels into consideration. Yaghoobzadeh et al. [196] proposed to address two types of noise from DS and pipeline input features, respectively. They introduced multi-instance multi-label learning algorithms to learn fine-grained entity typing and integrated entity typing into RE to tackle the noise. To convert noisy labeling sentences into meaningful training data, Shang et al. [136] proposed an unsupervised deep clustering to produce new high-confidence relation labels for noisy sentences. Yu et al. [212] formulated the DSRE as a hierarchical classification task and constructed the hierarchical bag representation to extract relations in a top-down manner.

Additionally, some works [179, 217] that consider both sentence-level and bag-level information simultaneously, explore explicit cross-level interactions to further improve the performance of DSRE. For example, Yuan et al. [215] first used a linear attenuation simulation to reflect words' importance, then proposed a non-IID relevance embedding to capture the mutual information of instances in the bag. Ye and Ling [208] proposed intra-bag and inter-bag attention models to address the noisy bag problem in a multi-instance distant supervision setting. Yuan et al. [217] first employed sentence-level selective attention to reduce the effect of noise, then adopted cross-bag selective attention to capture the entity pairs with higher quality. Gou et al. [52] applied a dynamic parameter-enhanced network to DSRE, dynamically determining the sentence information to alleviate the style shift problem for predicting the long-tail relations. Zhao et al. [234] proposed context-aware based on frame semantics to combine the semantic knowledge within a hierarchical neural network to alleviate the noisy labels in DSRE. Dai et al.[34] employed a cross-stitch mechanism to capture the interaction between the text encoder and KG encoder, allowing the model to share the information thoroughly. Shang et al. [137] constructed a force-directed graph and introduced the attractive force to learn the correlation and mutual exclusion between different relations.

Summary. The development of DSRE has been characterized by continuous efforts to improve the accuracy and robustness of RE models in the face of noisy and incomplete data. Researchers have explored various methodologies and advanced models to enhance DSRE's performance in extracting relational information from large-scale KBs.

4.1.2 FSRE. FSRE aims at predicting the relationship between two entities in a sentence by training with a few labeled instances for each relation. In realistic scenarios, only common relationships can obtain enough labeled examples, while most other relationships have very limited relational facts. FSRE has the potential to handle "long-tail" relations that have limited relational facts. In this section, we systematically present advanced FSRE approaches by categorizing them into two groups: (1) Metric Learning and (2) Knowledge-enhanced Learning. Additionally, we will discuss the most recent prompt-based methods with PLMs for few-shot RE tasks in Section 5.

(1) Metric Learning. One popular approach for FSRE is metric learning [94, 121], where the term "metric" refers to the distance function used to measure the similarity or dissimilarity between samples in the embedding space. The model is optimized by minimizing the distances between query samples and their corresponding class prototypes, thereby improving its ability to assign new instances to the nearest relation class prototype. A relevant approach is learning prototypes of relations from the contextual information for capturing the semantics of relations, which significantly improves accuracy [39]. Some approaches [49, 213, 214, 240] also introduce external information to compensate for the limited information in FSRE. For example, Qu et al. [121] proposed a global relation graph with text descriptions of entities and relations collected from Wikidata. Gao et al. [49] proposed a hybrid attention-based prototypical network to tackle the noisy problem in FSRE. They designed instance-level and feature-level attention to highlight important instances and features. Yu et al. [213] proposed a multi-prototype embedding network to jointly extract relation triples. The prototype representations learned by specifical prototype-aware regularization can inject the implicit correlations between entities and relations. MapRE [42] considered both label-agnostic and label-aware semantic mapping information for FSRE. HCRP [58] learned relation label information by contrastive learning and allowed the model to adaptively learn to focus on hard work. To endow a new model with the ability to optimize rapidly, REGRAB [121] proposes a Bayesian meta-learning method by incorporating an external global relation graph. Overall, these approaches leverage metric learning strategies to learn a distance metric or similarity function, which can not only effectively distinguish between different types of relations but also facilitate better generalization from a limited number of labeled data for certain relations.

(2) Knowledge-enhanced Learning. Many FSRE works also employ external knowledge to enrich the auxiliary semantic information. According to the data structure, external knowledge can be divided into (1) unstructured text spans, including the descriptions of entity and relation, and (2) a structured KG. For unstructured text span, TD-proto [200] proposes a collaborative attention module to enhance the prototypical network with entity and relation descriptions. ConceptFERE [201] model introduces the inherent concepts of entities to provide appropriate clues for relation classification, bridging the gap between the representations of relation types and text. Wang et al. [166] proposed a discriminative rule-based knowledge method where a logic-aware inference module is adopted to avoid the adverse effect of text features. In comparison, some approaches explore the abundant KG information. Liu et al. [93] proposed to inject triples in KG into texts, which transforms the sentences into knowledge-enhanced sentence trees. Roy and Pan [127] incorporated entity-level KG into pre-trained BERT for clinical RE, integrating the medical knowledge by several techniques. Sainz et al. [130] reformulated the RE task as an entailment task with hand-made verbalizations of relation labels, which helps generalize to the unseen label. Knowledge-enhanced RE methods leverage external knowledge sources to deepen the understanding of entity relations. However, this integration presents challenges, including managing noise and ensuring effective incorporation. Errors may arise from mapping concepts from external sources to the RE task, mainly due to semantic mismatches. Additionally, the complexity of filtering relevant information while discarding noise impacts efficiency and adds to the difficulty of the algorithm design.

Summary. The development of FSRE has been driven by the need to extract relational information from text when only a limited number of labeled instances are available for training. FSRE approaches leverage metric learning and knowledge-enhanced learning strategies to address the challenge of limited labeled data for RE. These methods advance the field by effectively generalizing from a small number of labeled instances and enriching semantic understanding with external knowledge sources.

4.2 Cross-Sentence RE

While most existing works have focused on extracting relational facts from individual sentences, many relational facts are expressed across multiple sentences within a long context. As a result, many studies have shifted research attention from sentence-level to cross-sentence. Cross-sentence RE aims to overcome the inherent limitations of sentence-level approaches and identify all relations mentioned across multiple sentences. Generally, there are two main research lines in cross-sentence RE: DocRE and dialogue RE.

4.2.1 DocRE. DocRE aims at extracting the triplets mentioned in a document. Although sentence-level RE approaches have achieved impressive results [90, 109], they still fail to handle the DocRE, as the documents contain richer information and more complex structures than sentences. Unlike sentence-level RE, which aims at classifying the relations of one or several entity pairs, DocRE requires the RE model to identify and focus on the relevant context within the document for a specific entity pair. Additionally, one entity pair can appear multiple times within a document, with each appearance having a distinct relation, making DocRE more challenging than sentence-level RE. For instance, as shown in Figure 1, the relation "*worked_on*" between "*Sam Altman*" and "*ChatGPT*" can only be found in the long context of the document instead of a single sentence. Other sentences between these two sentences may contain irrelevant information. This requires the RE models to be capable of capturing long-distance dependency in relational information.

Recent RE methods [59, 219] effectively capture complex interactions requiring reasoning over long-distance entities across multiple sentences. There are mainly two methods to infer relations from multiple sentences:

(1) **The Graph-based Approaches.** These works [32, 109] construct document graphs with attention or dependency structures, bridging the entities spreading far apart in the document. Relational inference information is gathered from the GNNs. These methods perform multi-hop reasoning in the overall graph structure to obtain meaningful entity representations. For example, Nan et al. [110] automatically constructed a document-level graph to empower relational reasoning across sentences. To enable the model with multi-hop reasoning, they proposed a refinement strategy to incrementally aggregate relevant information. Christopoulou et al. [33] proposed an edge-oriented graph neural model to construct a document-level graph using multiple nodes and edges. Tran et al. [157] extended the edge-oriented model by incorporating explicit relation classification-related node representations. Li et al. [87] proposed a graph-enhanced dual attention network to characterize the complex interactions among potential relation instances. Zeng et al. [225] designed a graph aggregation-and-inference network featuring a double graph. There is a heterogeneous mention-level graph to capture the interactions among different mentions and an entity-level graph to infer relations between entities. To cover more logical reasoning chains, Zeng et al. [224] developed a logical reasoning module to represent intra- and inter-sentential relations.

(2) The Path-based (non-graph) Approaches. These works [26, 70] attempt to enrich the local contextual information surrounding the target entity pair. They extract paths connected to the given entities to retain semantic information for predicting the relations. For example, Yao et al. [205] utilized context-aware LSTM to encode sentences and infer relations in the document. Xu et al. [181] formulated the distinctive dependencies by incorporating the structural dependencies based on the self-attention mechanism. Moreover, some works explore synthesizing implicit long-distance information modeled by transformer-based methods and multi-scale neural architectures [74]. Xu et al. [185] proposed an encoder-classifier-reconstructor model for DocRE, where the reconstructor is used to reconstruct the path dependencies from the graph representation. To tackle the multi-label and multi-entity problem in DocRE, Zhou et al. [249] deployed adaptive thresholding and localized context pooling to transfer attention from PLMs to decide the context relevant to the relation. Wang et al. [168] constructed a unified positive-unlabeled learning method to tackle the incomplete labeling problem in DocRE. Chen et al. [25] introduced an iterative extraction for DocRE and proposed an imitation learning to cast the extraction problem as a Markov decision process. Guo et al. [55] discovered that the inadequate training paradigm leads to underwhelming performance instead of the model capacities. Therefore, they propose a generative framework for DocRE which generates a symbolic sequence from a relation matrix to help model learning.

Summary. Graph-based RE approaches construct document-level graphs and utilize graph structures to model the complicated relationships among entities. Such graph representations efficiently capture local and global contextual information, thereby facilitating the discovery of implicit relations. However, the construction and maintenance of graph structures can be computationally demanding, and the quality of the underlying graph representation plays a crucial role in determining the effectiveness of graph-based models. Conversely, path-based (non-graph-based) approaches concentrate on the sequential context and semantic patterns within candidate entity pairs, employing syntactic dependency structures or pre-trained models to connect target entities directly or through contextual tokens. These approaches generally exhibit greater computational efficiency and adaptability for various RE tasks. However, path-based approaches may be less proficient than graph-based approaches at capturing global relational information. This limitation arises due to the sequential nature of path-based approaches, which may struggle to comprehensively capture relationships spanning distant parts of the document. Overall, both graph-based and path-based approaches have shown effectiveness in DocRE tasks, with the potential for further enhancement through document understanding and the integration of multi-hop reasoning capabilities for inferring complex relationships.

4.2.2 Dialogue RE. In addition to extracting semantic relations from sentences and documents, recent RE research also explores dialogue scenarios. The relation triplets in dialogue usually have low information density and do not appear simultaneously. This suggests that dialogue RE should be aware of the multiple speakers and arguments within a conversation. The represented dialogue RE approaches can be divided into two categories: (1) Fine-tuning PLMs with specific dialogue RE objectives. To capture the diverse relational information between arguments in the dialogue, some strategies are applied to build an RE model that obtains the contextualized turn representations [82], such as constructing the latent multi-view graph and heterogeneous dialogue graph. Cai and Lam [18] proposed a graph transformer to explicitly encode relations and enable direct communication between distant node pairs. Yao et al. [203] proposed a heterogeneous graph transformer to model the different relations among individual subgraphs, including direct, indirect, and possible relations between nodes. (2) Prompt-based approaches. These utilize prompting exemplars constructed with trainable words to incorporate potential relational knowledge. For example, Chen et al. [24] injected knowledge among the relation labels into prompting. Son et al. [144] proposed an argument-aware prompting strategy to capture the relational clues.

Furthermore, due to the properties of low information density and high personal pronoun frequency [82] in dialogue, more research efforts are needed to capture such sparse semantics among multiple speakers. Albalak et al. [4] proposed a model-agnostic framework D-REX that focuses on dialogue RE and the explainability of methods. D-REX frames RE as a reranking task and incorporates relation- and entity-specific explanations in the intermediate steps. Yu et al. [211] defined the trigger words in dialogue RE which indicates the existence of a given relation. They showed such manually annotated text spans play a critical role in cross-sentence RE. Xue et al. [191] took a novel input format and utilized a BERT-based model to capture the interrelations among entity pairs. Additionally, some studies utilize token-graph models to track the speaker-related information for cross-sentence RE in dialogues. Chen et al. [22] deployed a token graph attention network. Xue et al. [192] proposed capturing relationships by generating a latent multi-view graph and selecting critical words for RE. Qiu et al. [120] proposed an $\alpha-\beta-\gamma$ strategy, an incremental parsing strategy for dynamic inference upon any incoming sentence, to infer social relations in dialogues. This strategy models the social network as a graph to ensure the consistency of relations.

Summary. Dialogues typically encompass complex discourse structures, implicit relations behind conversations, and dynamic interactions between speakers. Existing studies predominantly focus on extracting relations among various speakers or individuals mentioned in conversations, leveraging contextualized representations to capture intricate relationships within dialogues. However, due to the properties of low information density and high personal pronoun frequency in dialogue, more research efforts are necessary to effectively capture relational clues within dialogues, particularly in capturing sparse semantics among multiple speakers.

4.3 Domain-Specific RE

In real-world scenarios, RE approaches are typically applied to different specific domains. However, general-purpose RE models, when directly applied to domain-specific data, can yield unsatisfactory results due to the shift in word distribution from general domain data to domain-specific data. Therefore, it is necessary to explore how to endow RE models with the ability to adapt to domain-specific corpora. Although RE studies have been thriving for a few decades, few researchers have reviewed domain-specific fields so far. In this section, we discuss recent RE methods tailored to different specific domains, including biomedical [31, 53, 88], financial [162], legal [5], and scientific domains.

4.3.1 RE in the Biomedical Field. In the biomedical field, the RE models aim at automatically extracting relations between biomedical entities (proteins, genes, diseases, etc.) from a rich source of biomedical texts [126, 169]. BioBERT [83] is a representative model of PLMs that inject biomedical information. Most works on biomedical RE focus on one type of relation, which can be categorized into several types according to biomedical relation types. These types include drug-drug interaction RE [7], disease-protein RE [183], chemical-protein RE [173], and protein-protein interaction RE [3]. For example, Asada [7] utilized heterogeneous domain information for drug-drug interaction RE, combining drug description and molecular structure information. Weber et al. [173] defined the Humboldt contribution task as an RE problem, where the chemical-protein relations are modeled with PLMs by incorporating entity descriptions. Ahmed et al. [3] designed a tree LSTM model with structured attention architecture for identifying protein-protein interaction relationships. Zhao et al. [233] explored modeling the global dependency relation of sentences by self-attention mechanism and graph convolutional networks. Haq et al. [61] introduced accuracy-optimized and a speed-optimized architecture. The systems understand different aspects of clinical documents, thereby enhancing the accuracy of extracting entity pairs and clinical relations, including extracting and correlating dates to generate a timeline of a patient's data, as well as parsing and comprehending trial results for analysis.

Recently, with the success of PLMs, several Transformer-based approaches have been widely explored for biomedical RE. Wei et al. [174] first explored implementing the BERT model for clinical RE tasks, where the unstructured clinical data is typically documented by specific professionals. Thillaisundaram and Togia [153] proposed extracting biomedical triplets with an extended BERT model, which encoded gene-disease pairs and their textual context to predict the "function change" relation. Yadav et al. [195] proposed a multi-task learning framework for RE in biomedical and clinical domains, modeling the RE task with three subtasks to better utilize the shared representation. Kanjirangat and Rinaldi [79] proposed a distantly supervised biomedical RE method using the SDP for selecting representative samples. Moreover, Sarrouti et al. [132] did an empirical study on encoder-only and encoder-decoder transformers over ten biomedical RE datasets. These comparisons also included the four major biomedical subtasks, i.e., chemical-protein RE, disease-protein RE, drug-drug RE, and protein-protein RE. They further explored multi-task fine-tuning to examine correlations among these subtasks.

In addition, similar to the idea discussed in the knowledge-enhanced methods Section 4.1.2, in the biomedical field, KGs play a significant role in enriching manually annotated information [35, 106]. They offer substantial potential for leveraging external knowledge sources to enhance our understanding of entity relations and extract biomedical relationships. Besides, there is an increasing demand to extract n-ary relations [74] from multiple documents, where $n > 2$. It is essential to extract relations between more than a pair of entities in the biomedical field. For example, detecting the relationship between a drug, a cancer patient, and a specific gene mutation is crucial for determining whether a drug is relevant for treating cancer patients with a certain mutation in a given gene. Lee et al. [84] proposed cross-sentence N-ary RE by utilizing entity linking and discourse relations, respectively. Tiktinsky et al. [155] proposed an N-ary drug combination RE dataset to assist professionals in identifying beneficial drug combinations. They also proposed a baseline model to predict if a subset of drugs used together in combination therapy is effective.

4.3.2 RE in the Finance Field. In the financial domain, the RE systems focus on identifying specific relations within financial texts, such as automatically extracting and linking **key performance indicators** (**KPIs**) from financial documents [65]. For example, Deußer et al. [37] explored extracting KPIs from financial documents where a word-level weighting scheme models the inherently fuzzy borders of the entity pairs and the corresponding relations. Wu et al. [178] focused on Chinese financial entity recognition and RE and proposed a mixed pattern with POS tagging to generate the quadruples (entity1, entity2, relation, text) from the unstructured finical text. Jabbari et al. [69] presented a domain-specific ontology for financial entities and relations in French news and created a corpus to build a KB of financial relations. Sharma et al. [138] released the first financial RE dataset and demonstrated that the RE models trained on the general domain might be ineffective in understanding financial relations in texts due to the discrepancies in the set of relations. The research value of financial RE is to make full use of financial information and help investors make better investment decisions.

4.3.3 RE in the Legal Field. In the legal domain, RE systems aim at extracting the legal relationships between entities contained in judicial documents, such as the relationship between a person and a company. To automatically identify entities and relations in legal documents, Andrew [6] explored combining statistical and rule-based techniques without labeled data. Hendrycks et al. [62] created the first legal dataset for contract reviews. Previous studies focused on modeling implicit relations within legal documents, such as criminal relations in judgment documents [26] and clause relations in contracts [186]. Thomas and Sangeetha [154] proposed a semi-supervised pattern-based learning method to extract relational facts from the judicial text. This work combines bootstrapping and OBIE techniques to expedite the extraction of judicial facts. Wang et al. [172] focused on cross-domain contract element extraction and proposed a Bi-FLEET model, which incorporates a clause-element relation encoder with a bi-directional feedback scheme. A multi-task framework is applied to capture interactions between contract element extraction and clause classification. Xu et al. [187] proposed a ConReader framework for a better contract understanding, which explores the long-range context relation, term-definition relation, and similar clause relation in the contract clause extraction.

We evaluate our model on the task of **question answering** using … .

Section: Dataset
SQuAD is a **machine comprehension** dataset on a large set of **Wikipedia** articles, … .
Two metrics are used to evaluate models: **Exact Match (EM)** and a softer metric, **F1 score** … .

Section: Model Details
… Each paragraph and question are tokenized by a regular-expression-based word tokenizer (**PTB Tokenizer**) and fed into the model.

Section: Results
The results of our model and computing approaches on the hidden test are summarized in Table [reference]. **BiDAF (ensemble)** achieves an **EM** score of 73.3 and an **F1** score of 81.1, outperforming all previous approaches.

Fig. 4. An example in [70] for the document-level N-ary relation (dataset: SQuaD, task: Machine comprehension, method: BiDAF (ensemble), metric: EM/F1).

4.3.4 RE in the Scientific Field. In the scientific domain, in order to minimize the time invested in the scientific literature search, researchers have proposed methods to automatically extract the relations of scientific articles automatically. Augenstein et al. [8] proposed the SemEval task for extracting keyphrases and the corresponding relations between them in the scientific texts. Luan [96] proposed a semi-supervised learning framework for scientific RE. Luan et al. [98] developed a unified framework SciIE for extracting entities, relations, and coreferences in scientific documents. Hou et al. [66] constructed a scientific leaderboard for extracting four items from NLP articles, including task, dataset, metric, and score. This benefits the community in keeping track of interesting scientific results. Eberts and Ulges [45] proposed a transformer-based joint RE model based on SciERC. Jain et al. [70] created the SciREX dataset for the document-level N-ary RE from scientific articles. As shown in Figure 4, the key challenge is to detect the target triplets residing in multiple modalities, including paragraphs and tables of the document. Kruiper et al. [81] introduced the semi-open RE task to comprehend the most significant relationships governing the central concepts in the document. Different from previous works solely considering the content of the article, CitaionIE further leverages the citation graph of referential links, showing the article's place in the broader literature. Magnusson and Friedman [102] built a SciClaim KG with entities, relations, and attributes. SciClaim contains both coarse-grained and fine-grained entity spans and relations from scientific claims.

4.3.5 Prospects on Domain-Specific RE. Overall, we notice that: (1) extensive studies have focused on biomedical RE. However, there is still a great demand for publicly available data resources and effective approaches in other specific domains. (2) Some domain-specific PLMs have been proposed to address the lack of high-quality, large-scale labeled domain data. The corresponding PLMs injecting domain-specific information include BioBERT [83], SciBERT [11], FinBERT [194], and Legal-BERT [20]. These tasks are challenging due to the specialized vocabulary and the complexity of the relationships involved. Therefore, the continued advancement of RE techniques specifically tailored to these domains is essential for various domain-specific applications. For example, Roy and Pan [127] incorporated entity-level KG into pre-trained BERT for clinical RE, which integrates medical knowledge by several techniques.

Summary. Despite the remarkable progress made by previous works, there is still substantial room for improving the RE performance in specific domains. (1) It is essential to further develop benchmark datasets and methods to identify and extract more practical and specific relations in

Table 5. Two Typical Input and Output Examples of Generation Methods for RE

Work	Input Example	Output Example
TANL [113]	Tolkien's epic novel The Lord of the Rings was published in 1954-1955.	[Tolkien | person]'s epic novel [The Lord of the Rings | book | author = Tolkien] was published in 1954-1955
ChatIE [175]	Given sentence: "Japan then laid siege to the Syrian penalty area and had a goal disallowed for offside in the 16th minute." The known entity types are: ['LOC', 'ISC', 'ORG', 'PER']. Please answer: What types of entities are included in this sentence?	LOC, MISC

different application domains. Current domain-specific datasets are either too narrow, containing only a small number of semantic relations, or too broad, containing an unbounded number of generic relations extracted from large and generic corpora [81]. (2) PLMs have made a significant contribution to RE, and it would be a promising direction to further tailor domain-specific PLMs by injecting domain knowledge into the general PLMs to understand specialized vocabulary and tackle the complexity of the involved relations.

5 RE with PLMs

Recently, PLMs have proved to be powerful in improving the performance of RE [188], as demonstrated in Section 3.2.4, where PLMs are used for context encoding. As illustrated in Table 7, the first part [28] of Table 7 shows the results of state-of-the-art RE approaches with fine-tuned PLMs on two benchmark datasets in the general domain (i.e., NYT and WebNLG). The second part [57] of Table 7 compares the few-shot performance between fine-tuned PLMs on two biomedical RE datasets, i.e., ChemProt and DDI. Specifically, three PLMs (i.e., PubMedBERT-base [54], BioBERT-large [83], and RoBERTa-large [95]) are fine-tuned on 100 labeled training samples. From Table 7, we can observe that PLMs with more parameters usually outperform those with fewer parameters. Some well-designed BERT-based models produce competitive results compared to models with larger PLMs (i.e., BART and RoBERTa).

Although PLMs have contributed significantly, supervised fine-tuning still suffers from a lack of sufficient supervised RE data in practice. In addition, there is a significant gap between the training objectives of the pre-training and fine-tuning processes in PLMs, which may hinder the adaptation of the knowledge in PLMs, especially for FSRE. To overcome this limitation, prompt-tuning techniques [144] have been proposed to bridge the gap between pre-training and fine-tuning processes by converting downstream RE tasks into a language model format. This approach aims at leveraging the capabilities of the PLM to perform a specific task by adapting it to the target task through training on a smaller, task-specific dataset. The key idea is to reformulate the tasks by appending an instruction phase that can be directly solved by PLMs. Therefore, prompt learning casts the RE task as the text generation problem. This approach appends the templates to input sentences, introducing additional information into templates to aid the generation process. The prompts/templates appropriately define the relationship and order for the entity spans and labels. For example, in the first example in Table 5, Paolini et al. [113] enclosed each entity and possibly some relations with special tokens [ ]. The sequence of |-separated tags represents the entity type and a list of relations in the format "X = Y", where X is the relation type, and Y is the tail entity of the relation. Besides, some recent developments [60, 77, 134] in the field of RE include the use of prompt-based approaches to prompt a PLM by converting the extraction of relation to predict the missing words. As shown in Table 6, the mined prompts [77] are constructed from Wikipedia through both middle words and dependency paths. The manual prompts are created

Table 6. Examples of Prompts for RE in [77]

ID	Relations	Manual Prompts	Mined Prompts
P140	religion	x is affiliated with the y religion	x who converted to y
P159	headquarters location	The headquarter of x is in y	x is based in y
P20	place of death	x died in y	x died at his home in y
P264	record label	x is represented by music label y	x recorded for y
P279	subclass of	x is a subclass of y	x is a type of y
P39	position held	x has the position of y	x is elected y

Table 7. An Overview Performance Comparison of RE Methods with PLMs in General and Specific Domains

		NYT			WebNLG		
Method	**# PLM Param.**	**Prec.**	**Rec.**	**F1**	**Prec.**	**Rec.**	**F1**
RE methods with PLMs (In general domain)							
CasRel [176]	$\text{BERT}_{(110M)}$	89.7	89.5	89.6	93.4	90.1	91.7
TPLinker [171]	$\text{BERT}_{(110M)}$	91.3	92.5	91.9	91.8	92.0	91.9
CGT [207]	$\text{UniLM}_{(110M)}$	94.7	84.2	89.1	92.9	75.6	83.4
PRGC [244]	$\text{BERT}_{(110M)}$	93.3	91.9	92.6	94.0	92.1	93.0
REBEL [17]	$\text{BART}_{(406M)}$	–	–	93.4	–	–	–
R-BPtrNet [27]	$\text{RoBERTa}_{(335M)}$	94.0	92.9	93.5	94.3	93.3	93.8
MTG [29]	$\text{T5-large}_{(770M)}$	95.6	93.1	94.3	94.8	95.1	94.9
		ChemProt			**DDI**		
Method	**# PLM Param.**	**Prec.**	**Rec.**	**F1**	**Prec.**	**Rec.**	**F1**
RE methods with PLMs (In specific domain)							
PubMedBERT-Base	100 M	17.9	62.0	27.7	19.9	79.1	31.8
BioBERT-Large	345 M	19.0	60.6	28.7	17.3	75.4	28.2
RoBERTa-Large	354 M	22.0	69.7	33.4	25.5	77.9	38.4

by experts according to the relation semantics, which is more complicated syntactically. However, manually defining the appropriate mapping phrase is time-consuming and non-intuitive [77] since it requires task-specific knowledge and manual identification words that the PLM can sufficiently understand.

To avoid the labor-intensive process of constructing prompts, recent works [48, 133, 141] pay attention to automatically generating and searching prompts. For example, Shin et al. [141] designed AUTOPROMPT to automatically create prompts by a gradient-guided search. It shows that **masked language models (MLMs)** can be effectively used as relation extractors without additional fine-tuning. Moreover, some studies [85, 91, 119] propose continuous prompts while fixing all PLM parameters, and experiments show that such soft prompts work well on few-shot RE datasets. Drawing inspiration from prompting, Li and Liang [91] proposed the prefix-tuning, which attends the subsequence tokens to prompt the PLMs. Qin and Eisner [119] and Lester et al. [85] proposed to model prompts as continuous vectors optimized by a mixture of prompts.

What's more, recent advances [75, 149, 189] in LLMs, such as GPT-3 [16], ChatGPT, and GPT-4[6] [1], have demonstrated their exceptional performance across various NLP tasks. While PLMs primarily strive for high performance in predefined NLP tasks, LLMs exhibit emergent capabilities extending beyond task-specific learning. GPT-3 represents a significant milestone in

[6]The corresponding model versions of GPT-3, ChatGPT, and GPT-4 are text-DaVinci-003, GPT-3.5-turbo, and GPT-4-turbo, respectively.

the evolution from PLMs to LLMs [239]. With the continuous growth in model parameters and training corpus size, LLMs exhibit emergent abilities that enable them to engage in **in-context learning** (**ICL**), where the models can reason from a small number of demonstration examples within the input context [43]. For example, in the second example in Table 5, LLMs can effectively perform RE given specific prompts [175]. Besides, some LLM-based methods [2, 75, 175] also provide several example demonstrations in the input, fully taking advantage of LLMs' larger number of parameters and longer input context lengths. Jiang et al. [75] tested the capabilities of the leading LLMs to perform RE in a zero-shot manner, which includes the GPT Family [16], i.e., text-davinci-003, gpt-3.5-turbo, gpt-3.5-turbo-instruct and gpt-4-turbo [1], and the LLaMA family [156], i.e., LLaMA-2-7B, LLaMA-2-70B, Vicuna-1.5-7B, Vicuna-1.3-33B, and WizardLM-70B [182]. Agrawal et al. [2] showed that LLMs perform well at zero- and few-shot clinical RE despite not being trained specifically in clinical texts. The construction of demonstrations facilitates the LLMs' comprehension and easy answer extraction. Wei et al. ([175]) explored the helpfulness of ChatGPT in the RE task and proposed a two-stage framework (ChatIE). This framework transforms the zero-shot IE task into a multi-turn QA problem by prompting ChatGPT and improves the experimental results.

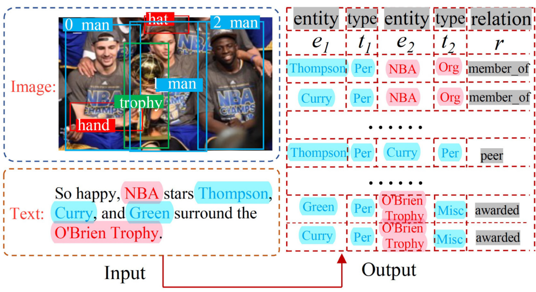
Summary. RE tasks have benefited significantly from both PLMs and LLMs. PLMs have shown remarkable performance in improving RE by leveraging pre-training on large corpora, especially for fine-tuning PLMs for specific RE tasks. However, existing methods relying solely on PLMs often face challenges when dealing with newly emerging relations due to the need for extensive data annotation, which can be time-consuming and labor-intensive. LLMs showcase impressive capabilities in generation and have inspired exploration of alternative approaches for obtaining auto-labeled documents with new relations. They excel in scenarios with limited annotations, where their memorization and reasoning capabilities contribute significantly to relation extraction tasks. However, the inference latency and financial cost associated with calling LLMs' APIs are higher compared to fine-tuning PLMs.

6 Future Directions

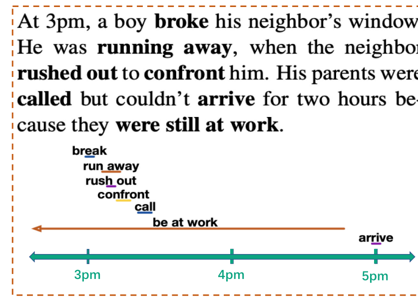
RE studies have made significant progress in recent years regarding new neural RE model designs and subtasks. However, challenges and limitations remain that need to be addressed, including the need for more diverse data in practical scenarios, handling complex and unevenly distributed relations, and incorporating additional new relation types.

6.1 Multi-Modal RE

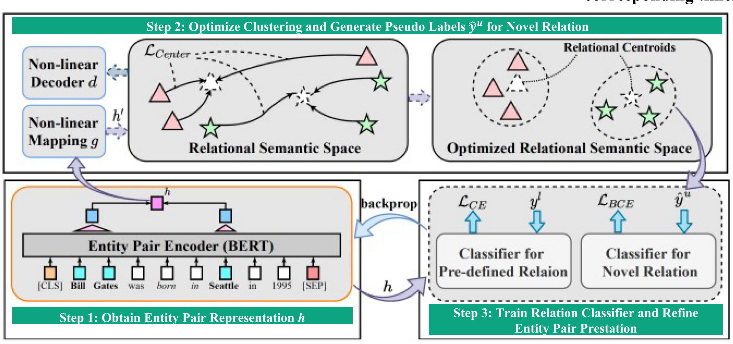
Along with text, images, and videos have become popular ways to convey information on the internet. This highlights the importance of extracting relations from multi-modal input rather than textual data alone. Multi-modal RE takes advantage of the large visual-text corpus by focusing on extracting relations from these media forms. Zheng et al. [243] proposed the multi-modal RE dataset MNRE containing visual evidence collected from social media posts. Subsequently, Zheng et al. [242] introduced the multi-modal RE model to capture the knowledge from related information in the texts and images. However, many interesting problems remain to be explored [167]. As shown in Figure 5(a) [216], the multi-modal RE task takes images and texts as input, then recognizes the entities and corresponding relations from the multi-modal data. This task is expected to align the entity-entity relations in the text with the object-object relations in the images. Since multi-modal data is often closely related, the visual content can supplement missing semantics of the textual content and improve the performance of RE methods. Thus, it is crucial to develop well-constructed multi-modal RE methods combining visual and textual information to extract relations more accurately.



(a) Multimodal relation extraction. It requires extracting different characteristic information from vision.



(b) Temporal relation extraction. The narrative and corresponding timeline are shown.



(c) Open relation extraction. There are three steps iteratively performed: (1) encode both the labeled and unlabeled instances into entity pair representations. (2) transfer the pair of entity representations into the relation-oriented representations by gathering towards their relational centroids. (3) optimize the entity pair representations and classifier by minimizing a joint objective function to reduce the clustering bias on predefined classes.

Fig. 5. Illustration of the future directions mentioned in Section 6.

6.2 Cross-Lingual RE

Existing state-of-the-art RE systems are primarily available for English because they heavily rely on annotated corpora and PLMs. These methods perform RE on a sentence in a source language by first translating it into English, then performing RE on the translated sentence, and finally projecting the identified phrase back to the source language. However, these methods assume that parallel bilingual corpora can be obtained by existing machine translation systems. It is challenging to mitigate the noisy data problem caused by machine translation systems and align the sentences and extracted triples between different languages. One future research direction is to explore the *cross-lingual projection methods* for language-independent RE [47]. Therefore, some works have been proposed to improve cross-lingual transfer for RE, including utilizing universal dependency structure parses [146] and mBERT [122]. Recent progress [12, 135] demonstrates that multi-lingual training can improve performance across all languages in RE since the relation information from other languages might help encode the information in a given language. These methods learn language-agnostic sentence representations in complex and multi-lingual common spaces. As shown in Table 8 [135], the distribution for each language is quite different. One of the main challenges in cross-lingual RE is dealing with language differences. Languages vary in their grammatical structure, vocabulary, and syntax, which makes it difficult to identify relationships between entities across languages. Another challenge is the ambiguity of words and their translations across languages. To tackle these challenges, it is crucial to investigate diverse approaches for aligning relation semantics between resource-rich languages and those with more limited data availability.

6.3 Temporal RE

Temporal RE aims at identifying relations between entities subject to temporal constraints, enhancing the applicability of RE systems in complex reasoning. As shown in Figure 5(b) [160],

Table 8. The Number of Sentences and Relations for Each Language Shown in the Multi-Lingual RE Dataset [135]

	EN	KO	IT	FR	DE	PT	NL	PL	ES	AR	RU	SV	FA	UK	**Total**
Sentences	748 k	20 k	76 k	62 k	53 k	45 k	40 k	17 k	12 k	9 k	7 k	5 k	3 k	1 k	**1.1 M**
Relations	36	28	22	22	22	22	22	22	22	9	8	22	8	7	**36**

a timeline illustrates the fine-grained (real-valued) temporal relations implicated in the text, mapping the temporal relations and event durations to real-valued scales. There are two mainstream approaches dedicated to temporal RE: dealing with relations between events and time expressions [151] and extracting relations between entities at a given time spot through temporal reasoning [197]. Although previous studies have attempted to address this issue by generating patterns for time-variant relations, many challenges remain, including (i) the complex dependencies between entities, relations, and conditions; (ii) the difficulty of handling conditions in various forms in free text; and (iii) the lack of well-annotated data. Therefore, a general framework is needed to formalize the conditional dimensions.

6.4 Evolutionary RE

In recent years, most RE paradigms have been designed on pre-defined relation sets. However, as our world experiences continuous expansion of new relations, it is infeasible for RE systems to handle all emerging relation types. Therefore, there is a demand for RE systems that can generalize to new relations beyond pre-defined schemes. Several works have been proposed to handle new relations, which mainly fall into two groups: ((1) **open RE**. As illustrated in Figure 5(c) [235], open information extraction approaches [67] extract related phrases as representations of relations and entities from the text. Another type is the clustering-based unsupervised relation discovery method [235], which discovers unseen relation types using clustering optimization; (2) **lifelong RE**. This group of methods [237] aims at continuously training an RE model to learn new relations while avoiding forgetting the accurate classification of old ones. Evolutionary RE is a promising research area, giving RE models the ability to generalize beyond the training data and learn from new data. However, many unresolved challenges remain. For open RE, where phrases of the same relation can have various forms, the key challenge is to canonicalize relation phrases to reduce ambiguity and redundancy. For lifelong RE, more efforts are needed to prevent RE models from overfitting the experience memory. It is worth exploring more effective methods leveraging LLMs (as discussed in Section 5) to tackle the challenges in evolutionary RE effectively.

6.5 Explainable RE

Despite significant advancements in RE over the past decade, the opacity of DL-based RE models has led to an increasing demand for explainability. The core challenge in achieving explainability lies in the intrinsic complexity of RE models [124], which often function as black boxes. Another obstacle is that the features extracted by RE models may not be directly interpretable by humans. This disconnect complicates efforts to comprehend the underlying rationale behind the model's decisions [128], obscuring the reliability and accuracy of the extracted relations. Such opacity hampers users' ability to trust the models. To address these challenges, future research needs to focus on developing methods that provide accurate, real-time explanations of model predictions, particularly shedding light on how these models arrive at their conclusions. By enhancing explainability, the RE models could advance in their capabilities and become more trustworthy, enabling broader adoption in critical domains where transparency and reliability are essential.

7 Conclusion

This survey provided an up-to-date and comprehensive review of recent advances in RE. We first designed a novel taxonomy to systematically summarize the model architectures used in existing DNN-based RE approaches, fully combing recent research trends in categories, and illustrating the differences and connections between RE subtasks. Then, we analyzed several important yet challenging RE problems and their corresponding solutions. Specifically, we discussed the performance of RE on current solutions in diverse, challenging settings (i.e., the low-resource setting and the cross-sentence setting) and specific domains (i.e., biomedical, finance, legal, and scientific fields). Considering the new frontiers in RE studies, we also presented in-depth analyses that revealed the issues of RE with PLMs and LLMs. Finally, we pointed out several promising future directions and prospects. We hope this survey provides insightful perspectives and inspires the widespread implementation of real-life RE systems.

References

[1] Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, and Shyamal Anadkat. 2023. Gpt-4 technical report. arXiv:2303.08774. Retrieved from https://arxiv.org/abs/2303.08774

[2] Monica Agrawal, Stefan Hegselmann, Hunter Lang, Yoon Kim, and David Sontag. 2022. Large language models are few-shot clinical information extractors. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*. 1998–2022.

[3] Mahtab Ahmed, Jumayel Islam, Muhammad Rifayat Samee, and Robert E. Mercer. 2019. Identifying protein-protein interaction using tree LSTM and structured attention. In *Proceedings of the 2019 IEEE 13th International Conference on Semantic Computing (ICSC)*. IEEE, 224–231.

[4] Alon Albalak, Varun Embar, Yi-Lin Tuan, Lise Getoor, and William Yang Wang. 2022. D-REX: Dialogue relation extraction with explanations. In *Proceedings of the 4th Workshop on NLP for Conversational AI*. 34–46.

[5] Judith Jeyafreeda Andrew. 2018. Automatic extraction of entities and relation from legal documents. In *Proceedings of the 7th Named Entities Workshop*. 1–8.

[6] Judith Jeyafreeda Andrew. 2018. Automatic extraction of entities and relation from legal documents. In *Proceedings of the NEWS@ACL*.

[7] Masaki Asada. 2022. Integrating heterogeneous domain information into relation extraction: A case study on drug-drug interaction extraction. *arXiv preprint arXiv:2212.10714* (2022).

[8] Isabelle Augenstein, Mrinal Das, Sebastian Riedel, Lakshmi Vikraman, and Andrew McCallum. 2017. SemEval 2017 Task 10: ScienceIE - extracting keyphrases and relations from scientific publications. In *Proceedings of the International Workshop on Semantic Evaluation*.

[9] Sachin Pawar, Girish K. Palshikar, and Pushpak Bhattacharyya. 2017. Relation extraction: A survey. *arXiv preprint arXiv:1712.05191* (2017).

[10] Elisa Bassignana and Barbara Plank. 2022. What do you mean by relation extraction? a survey on datasets and study on scientific relation classification. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*.

[11] Iz Beltagy, Kyle Lo, and Arman Cohan. 2019. SciBERT: A pretrained language model for scientific text. 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)*. 3615–3620.

[12] Abhyuday Bhartiya, Kartikeya Badola, and Mausam. 2021. DiS-ReX: A multilingual dataset for distantly supervised relation extraction. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*.

[13] Junyi Bian, Li Huang, Xiaodi Huang, Hong Zhou, and Shanfeng Zhu. 2021. Grantrel: Grant information extraction via joint entity and relation extraction. In *Proceedings of the ACL-IJCNLP*. 2674–2685.

[14] Christian Bizer, Jens Lehmann, Georgi Kobilarov, Sören Auer, Christian Becker, Richard Cyganiak, and Sebastian Hellmann. 2009. Dbpedia-a crystallization point for the web of data. *Journal of Web Semantics* 7, 3 (2009), 154–165.

[15] Kurt Bollacker, Colin Evans, Praveen Paritosh, Tim Sturge, and Jamie Taylor. 2008. Freebase: A collaboratively created graph database for structuring human knowledge. In *Proceedings of the 2008 ACM SIGMOD International Conference on Management of Data* (2008). 1247–1250.

[16] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D. Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, and Amanda Askell. 2020. Language models are few-shot learners. *Advances in Neural Information Processing Systems* 33 (2020), 1877–1901.

[17] Pere-Lluís Huguet Cabot and Roberto Navigli. 2021. REBEL: Relation extraction by end-to-end language generation. In *Findings of the Association for Computational Linguistics: EMNLP 2021*. 2370–2381.

[18] Deng Cai and Wai Lam. 2019. Graph transformer for graph-to-sequence learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*.

[19] Rui Cai, Xiaodong Zhang, and Houfeng Wang. 2016. Bidirectional recurrent convolutional neural network for relation classification. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. 756–765.

[20] Ilias Chalkidis, Manos Fergadiotis, Prodromos Malakasiotis, Nikolaos Aletras, and Ion Androutsopoulos. 2020. Legal-bert: The muppets straight out of law school. In *Proceedings of the EMNLP*.

[21] Yee Seng Chan and Dan Roth. 2011. Exploiting syntactico-semantic structures for relation extraction. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*. 551–560.

[22] Hui Chen, Pengfei Hong, Wei Han, Navonil Majumder, and Soujanya Poria. 2023. Dialogue relation extraction with document-level heterogeneous graph attention networks. *Cognitive Computation* 15, 2 (2023), 793–802.

[23] Xiang Chen, Ningyu Zhang, Lei Li, Shumin Deng, Chuanqi Tan, Changliang Xu, Fei Huang, Luo Si, and Huajun Chen. 2022. Hybrid transformer with multi-level fusion for multimodal knowledge graph completion. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 904–915.

[24] Xiang Chen, Ningyu Zhang, Xin Xie, Shumin Deng, Yunzhi Yao, Chuanqi Tan, Fei Huang, Luo Si, and Huajun Chen. 2022. Knowprompt: Knowledge-aware prompt-tuning with synergistic optimization for relation extraction. In *Proceedings of the ACM Web Conference 2022*. 2778–2788.

[25] Yunmo Chen, William Gantt, Weiwei Gu, Tongfei Chen, Aaron Steven White, and Benjamin Van Durme. 2022. Iterative document-level information extraction via imitation learning. In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, 1858–1874.

[26] Yanguang Chen, Yuanyuan Sun, Zhihao Yang, and Hongfei Lin. 2020. Joint entity and relation extraction for legal documents with legal feature enhancement. In *Proceedings of the International Conference on Computational Linguistics*.

[27] Yubo Chen, Yunqi Zhang, Changran Hu, and Yongfeng Huang. 2021. Jointly extracting explicit and implicit relational triples with reasoning pattern enhanced binary pointer network. In *Proceedings of the ACL*. 5694–5703.

[28] Yubo Chen, Yunqi Zhang, and Yongfeng Huang. 2022. Learning reasoning patterns for relational triple extraction with mutual generation of text and graph. In *Proceedings of the ACL*. 1638–1647.

[29] Yubo Chen, Yunqi Zhang, and Yongfeng Huang. 2022. Learning reasoning patterns for relational triple extraction with mutual generation of text and graph. In *Findings of the Association for Computational Linguistics: ACL 2022*.

[30] Zi-Yuan Chen, Chih-Hung Chang, Yi-Pei Chen, Jijnasa Nayak, and Lun-Wei Ku. 2019. UHop: An unrestricted-hop relation extraction framework for knowledge-based question answering. 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)*. 345–356.

[31] Sung-Pil Choi. 2018. Extraction of protein-protein interactions (PPIs) from the literature by deep convolutional neural networks with various feature embeddings. *Journal of Information Science* 44, 1 (2018), 60–73.

[32] Fenia Christopoulou, Makoto Miwa, and Sophia Ananiadou. 2019. Connecting the Dots: Document-level neural relation extraction with edge-oriented graphs. 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)*. 4925–4936.

[33] Fenia Christopoulou, Makoto Miwa, and Sophia Ananiadou. 2019. Connecting the Dots: Document-level neural relation extraction with edge-oriented graphs. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*.

[34] Qin Dai, Benjamin Heinzerling, and Kentaro Inui. 2022. Cross-stitching text and knowledge graph encoders for distantly supervised relation extraction. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*.

[35] Qin Dai, Naoya Inoue, Paul Reisert, Ryo Takahashi, and Kentaro Inui. 2019. Distantly supervised biomedical knowledge acquisition via knowledge graph based attention. In *Proceedings of the Workshop on Extracting Structured Knowledge from Scientific Publications*. 1–10.

[36] Meiji Cui, Li Li, Zhihong Wang, and Mingyu You. 2017. A survey on relation extraction. In *Knowledge Graph and Semantic Computing. Language, Knowledge, and Intelligence: Second China Conference (CCKS'17)*. Springer, 50–58.

[37] Tobias Deußer, Syed Musharraf Ali, Lars Hillebrand, Desiana Nurchalifah, Basil Jacob, Christian Bauckhage, and Rafet Sifa. 2022. KPI-EDGAR: A novel dataset and accompanying metric for relation extraction from financial documents. In *21st IEEE International Conference on Machine Learning and Applications (ICMLA'22)*. IEEE, 1654–1659.

[38] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. 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), 4171–4186.
[39] Ning Ding, Xiaobin Wang, Yao Fu, Guangwei Xu, Rui Wang, Pengjun Xie, Ying Shen, Fei Huang, Hai-Tao Zheng, and Rui Zhang. 2020. Prototypical representation learning for relation extraction. In *Proceedings of the International Conference on Learning Representations*.
[40] Kalpit Dixit and Yaser Al-Onaizan. 2019. Span-level model for relation extraction. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*.
[41] George R. Doddington, Alexis Mitchell, Mark Przybocki, Lance Ramshaw, Stephanie Strassel, and Ralph Weischedel. 2004. The automatic content extraction (ACE) program–tasks, data, and evaluation. In *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC'04)*.
[42] Manqing Dong, Chunguang Pan, and Zhipeng Luo. 2021. MapRE: An effective semantic mapping approach for low-resource relation extraction. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. 2694–2704.
[43] Qingxiu Dong, Lei Li, Damai Dai, Ce Zheng, Zhiyong Wu, Baobao Chang, Xu Sun, Jingjing Xu, and Zhifang Sui. 2022. A survey for in-context learning. *arXiv preprint arXiv:2301.00234* (2022).
[44] Jinhua Du, Jingguang Han, Andy Way, and Dadong Wan. 2018. Multi-level structured self-attentions for distantly supervised relation extraction. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 2216–2225.
[45] Markus Eberts and Adrian Ulges. 2019. Span-based joint entity and relation extraction with transformer pre-training. In *(ECAI'20)*. IOS Press, 2006–2013.
[46] Markus Eberts and Adrian Ulges. 2020. Span-based joint entity and relation extraction with transformer pre-training. In *Proceedings of the ECAI 2020*. IOS, 2006–2013.
[47] Manaal Faruqui and Shankar Kumar. 2015. Multilingual open relation extraction using cross-lingual projection. In *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. 1351–1356.
[48] Tianyu Gao, Adam Fisch, and Danqi Chen. 2021. Making pre-trained language models better few-shot learners. 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)*. 3816–3830.
[49] Tianyu Gao, Xu Han, Zhiyuan Liu, and Maosong Sun. 2019. Hybrid attention-based prototypical networks for noisy few-shot relation classification. In *Proceedings of the AAAI Conference on Artificial Intelligence*. 6407–6414.
[50] Tianyu Gao, Xu Han, Hao Zhu, Zhiyuan Liu, Peng Li, Maosong Sun, and Jie Zhou. 2019. FewRel 2.0: Towards more challenging few-shot relation classification. In *Proceedings of the EMNLP-IJCNLP*.
[51] Claire Gardent, Anastasia Shimorina, Shashi Narayan, and Laura Perez-Beltrachini. 2017. Creating training corpora for nlg micro-planning. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (ACL)*.
[52] Yanjie Gou, Yinjie Lei, Lingqiao Liu, Pingping Zhang, and Xi Peng. 2020. A dynamic parameter enhanced network for distant supervised relation extraction. *Knowledge-Based Systems* 197 (2020), 105912.
[53] Jinghang Gu, Longhua Qian, and Guodong Zhou. 2016. Chemical-induced disease relation extraction with various linguistic features. 2016 (2016).
[54] Yu Gu, Robert Tinn, Hao Cheng, Michael Lucas, Naoto Usuyama, Xiaodong Liu, Tristan Naumann, Jianfeng Gao, and Hoifung Poon. 2021. Domain-specific language model pretraining for biomedical natural language processing. *ACM Transactions on Computing for Healthcare* 3, 1 (2021), 1–23.
[55] Qipeng Guo, Yuqing Yang, Hang Yan, Xipeng Qiu, and Zheng Zhang. 2022. DORE: Document ordered relation extraction based on generative framework. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*.
[56] Zhijiang Guo, Yan Zhang, and Wei Lu. 2019. Attention guided graph convolutional networks for relation extraction. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. 241–251.
[57] Bernal Jiménez Gutiérrez, Nikolas McNeal, Clay Washington, You Chen, Lang Li, Huan Sun, and Yu Su. 2022. Thinking about GPT-3 in-context learning for biomedical ie? think again. In *Findings of the Association for Computational Linguistics (EMNLP'22)*.
[58] Jiale Han, Bo Cheng, and Wei Lu. 2021. Exploring task difficulty for few-shot relation extraction. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. 2605–2616.
[59] Xu Han, Tianyu Gao, Yankai Lin, Hao Peng, Yaoliang Yang, Chaojun Xiao, Zhiyuan Liu, Peng Li, Maosong Sun, and Jie Zhou. 2020. More data, more relations, more context and more openness: A review and outlook for relation extraction. In *Proceedings of the AACL*.
[60] Xu Han, Weilin Zhao, Ning Ding, Zhiyuan Liu, and Maosong Sun. 2021. Ptr: Prompt tuning with rules for text classification. *AI Open* 3 (2022), 182–192.

[61] H. Ul Haq, V. Kocaman, D. Talby, A. P. B. Veyseh, F. Dernoncourt, T. H. Nguyen, and V. D. Lai. 2022. Deeper clinical document understanding using relation extraction. In *Proceedings of the Workshop on Scientific Document Understanding*. CEUR.

[62] Dan Hendrycks, Collin Burns, Anya Chen, and Spencer Ball. 2021. Cuad: An expert-annotated nlp dataset for legal contract review. *arXiv preprint arXiv:2103.06268* (2021).

[63] María Herrero-Zazo, Isabel Segura-Bedmar, Paloma Martínez, and Thierry Declerck. 2013. The DDI corpus: An annotated corpus with pharmacological substances and drug–drug interactions. *Journal of Biomedical Informatics* 46, 5 (2013), 914–920.

[64] Daniel Hewlett, Alexandre Lacoste, Llion Jones, Illia Polosukhin, Andrew Fandrianto, Jay Han, Matthew Kelcey, and David Berthelot. 2016. WikiReading: A novel large-scale language understanding task over wikipedia. In *Proceedings of the ACL*.

[65] Lars Hillebrand, Tobias Deußer, Tim Dilmaghani, Bernd Kliem, Rüdiger Loitz, Christian Bauckhage, and Rafet Sifa. 2022. Kpi-bert: A joint named entity recognition and relation extraction model for financial reports. In *Proceedings of the 2022 26th International Conference on Pattern Recognition (ICPR)*. IEEE, 606–612.

[66] Yufang Hou, Charles Jochim, Martin Gleize, Francesca Bonin, and Debasis Ganguly. 2019. Identification of tasks, datasets, evaluation metrics, and numeric scores for scientific leaderboards construction. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. 5203–5213.

[67] Xuming Hu, Lijie Wen, Yusong Xu, Chenwei Zhang, and S. Yu Philip. 2020. SelfORE: Self-supervised relational feature learning for open relation extraction. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 3673–3682.

[68] Zhiheng Huang, Wei Xu, and Kai Yu. 2015. Bidirectional LSTM-CRF models for sequence tagging. *arXiv preprint arXiv:1508.01991* (2015).

[69] Ali Akbar Jabbari, Olivier Sauvage, Hamada Zeine, and Hamza Chergui. 2020. A french corpus and annotation schema for named entity recognition and relation extraction of financial news. In *Proceedings of the International Conference on Language Resources and Evaluation.*

[70] Sarthak Jain, Madeleine van Zuylen, Hannaneh Hajishirzi, and Iz Beltagy. 2020. SciREX: A challenge dataset for document-level information extraction. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. 7506–7516.

[71] Linmei Hu, Luhao Zhang, Chuan Shi, Liqiang Nie, Weili Guan, and Cheng Yang. 2019. Improving distantly-supervised relation extraction with joint label embedding. 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)*, 3821–3829.

[72] Bin Ji, Jie Yu, Shasha Li, Jun Ma, Qingbo Wu, Yusong Tan, and Huijun Liu. 2020. Span-based joint entity and relation extraction with attention-based span-specific and contextual semantic representations. In *Proceedings of the 28th International Conference on Computational Linguistics*, 88–99.

[73] Bin Ji, Jie Yu, Shasha Li, Jun Ma, Q. Wu, Yusong Tan, and Huijun Liu. 2020. Span-based Joint Entity and Relation Extraction with Attention-based Span-specific and Contextual Semantic Representations. In *Proceedings of the International Conference on Computational Linguistics.*

[74] Robin Jia, Cliff Wong, and Hoifung Poon. 2019. Document-level N-ary relation extraction with multiscale representation learning. In *Proceedings of the ACL.*

[75] Pengcheng Jiang, Jiacheng Lin, Zifeng Wang, Jimeng Sun, and Jiawei Han. 2024. GenRES: Rethinking evaluation for generative relation extraction in the era of large language models. In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, 2820–2837.

[76] Xiaobo Jiang, Kun He, Jiajun He, and Guangyu Yan. 2021. A new entity extraction method based on machine reading comprehension. *arXiv preprint arXiv:2108.06444* (2021).

[77] Zhengbao Jiang, Frank F Xu, Jun Araki, and Graham Neubig. 2020. How can we know what language models know? *Transactions of the Association for Computational Linguistics* 8 (2020), 423–438.

[78] Mandar Joshi, Danqi Chen, Yinhan Liu, Daniel S. Weld, Luke Zettlemoyer, and Omer Levy. 2020. SpanBERT: Improving pre-training by representing and predicting spans. *Transactions of the Association for Computational Linguistics* 8 (2020), 64–77.

[79] Vani Kanjirangat and Fabio Rinaldi. 2021. Enhancing biomedical relation extraction with transformer models using shortest dependency path features and triplet information. *Journal of Biomedical Informatics* 122 (2021), 103893.

[80] Nora Kassner, Philipp Dufter, and Hinrich Schütze. 2021. Multilingual LAMA: Investigating knowledge in multilingual pretrained language models. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*. 3250–3258.

[81] Ruben Kruiper, Julian Vincent, Jessica Chen-Burger, Marc Desmulliez, and Ioannis Konstas. 2020. In layman's terms: Semi-open relation extraction from scientific texts. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 1489–1500.

[82] Bongseok Lee and Yong Suk Choi. 2021. Graph based network with contextualized representations of turns in dialogue. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing.*

[83] Jinhyuk Lee, Wonjin Yoon, Sungdong Kim, Donghyeon Kim, Sunkyu Kim, Chan Ho So, and Jaewoo Kang. 2020. BioBERT: A pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics* 36, 4 (2020), 1234–1240.

[84] Sanghak Lee, Seungmin Seo, Byungkook Oh, Kyong-Ho Lee, Donghoon Shin, and Yeonsoo Lee. 2020. Cross-sentence N-ary relation extraction using entity link and discourse relation. In *Proceedings of the 29th ACM International Conference on Information and Knowledge Management.* 705–714.

[85] Brian Lester, Rami Al-Rfou, and Noah Constant. 2021. The power of scale for parameter-efficient prompt tuning. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing.* 3045–3059.

[86] Omer Levy, Minjoon Seo, Eunsol Choi, and Luke Zettlemoyer. 2017. Zero-shot relation extraction via reading comprehension. In *Proceedings of the 21st Conference on Computational Natural Language Learning (CoNLL 2017).* 333–342.

[87] Bo Li, Wei Ye, Zhonghao Sheng, Rui Xie, Xiangyu Xi, and Shikun Zhang. 2020. Graph enhanced dual attention network for document-level relation extraction. In *Proceedings of the International Conference on Computational Linguistics.*

[88] Fei Li, Meishan Zhang, Guohong Fu, and Donghong Ji. 2017. A neural joint model for entity and relation extraction from biomedical text. *BMC Bioinformatics* 18, 1 (2017), 1–11.

[89] Xianming Li, Xiaotian Luo, Chenghao Dong, Daichuan Yang, Beidi Luan, and Zhen He. 2021. TDEER: An efficient translating decoding schema for joint extraction of entities and relations. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing.* 8055–8064.

[90] Xiaoya Li, Fan Yin, Zijun Sun, Xiayu Li, Arianna Yuan, Duo Chai, Mingxin Zhou, and Jiwei Li. 2019. Entity-relation extraction as multi-turn question answering. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics.* 1340–1350.

[91] Xiang Lisa Li and Percy Liang. 2021. Prefix-tuning: Optimizing continuous prompts for generation. 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).* 4582–4597.

[92] Yang Li, Guodong Long, Tao Shen, Tianyi Zhou, Lina Yao, Huan Huo, and Jing Jiang. 2020. Self-attention enhanced selective gate with entity-aware embedding for distantly supervised relation extraction. In *Proceedings of the AAAI Conference on Artificial Intelligence.* 8269–8276.

[93] Weijie Liu, Peng Zhou, Zhe Zhao, Zhiruo Wang, Qi Ju, Haotang Deng, and Ping Wang. 2020. K-bert: Enabling language representation with knowledge graph. In *Proceedings of the AAAI Conference on Artificial Intelligence.* 2901–2908.

[94] Yang Liu, Jinpeng Hu, Xiang Wan, and Tsung-Hui Chang. 2022. Learn from relation information: Towards prototype representation rectification for few-shot relation extraction. In *Findings of the Association for Computational Linguistics: NAACL 2022.* 1822–1831.

[95] Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692* (2019).

[96] Yi Luan. 2018. Information extraction from scientific literature for method recommendation. *arXiv preprint arXiv:1901.00401* (2018).

[97] Yi Luan, Luheng He, Mari Ostendorf, and Hannaneh Hajishirzi. 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.* 3219–3232.

[98] Yi Luan, Luheng He, Mari Ostendorf, and Hannaneh Hajishirzi. 2018. Multi-task identification of entities, relations, and coreference for scientific knowledge graph construction. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing.*

[99] Yi Luan, Dave Wadden, Luheng He, Amy Shah, Mari Ostendorf, and Hannaneh Hajishirzi. 2019. A general framework for information extraction using dynamic span graphs. 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).* 3036–3046.

[100] Kangqi Luo, Fengli Lin, Xusheng Luo, and Kenny Zhu. 2018. Knowledge base question answering via encoding of complex query graphs. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing* (2018). 2185–2194.

[101] Xuezhe Ma and Eduard Hovy. 2016. End-to-end sequence labeling via bi-directional lstm-cnns-crf. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 1064–1074.
[102] Ian H. Magnusson and Scott E. Friedman. 2021. Extracting fine-grained knowledge graphs of scientific claims: Dataset and transformer-based results. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*.
[103] Angrosh Mandya, Danushka Bollegala, and Frans Coenen. 2020. Graph convolution over multiple dependency subgraphs for relation extraction.. In *Proceedings of the COLING*. International Committee on Computational Linguistics, 6424–6435.
[104] Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781* (2013).
[105] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S. Corrado, and Jeff Dean. 2013. Distributed representations of words and phrases and their compositionality. *Advances in Neural Information Processing Systems* 26 (2013).
[106] Nikola Milosevic and Wolfgang Thielemann. 2022. Relationship extraction for knowledge graph creation from biomedical literature. *arXiv preprint arXiv:2201.01647* (2022).
[107] Mike Mintz, Steven Bills, Rion Snow, and Dan Jurafsky. 2009. Distant supervision for relation extraction without labeled data. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP*. 1003–1011.
[108] Makoto Miwa and Mohit Bansal. 2016. End-to-end relation extraction using LSTMs on sequences and tree structures. In *Proceedings of the ACL*.
[109] Guoshun Nan, Zhijiang Guo, Ivan Sekulić, and Wei Lu. 2020. Reasoning with latent structure refinement for document-level relation extraction. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. 1546–1557.
[110] Guoshun Nan, Zhijiang Guo, Ivan Sekulic, and Wei Lu. 2020. Reasoning with latent structure refinement for document-level relation extraction. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*.
[111] Tapas Nayak, Navonil Majumder, Pawan Goyal, and Soujanya Poria. 2021. Deep neural approaches to relation triplets extraction: A comprehensive survey. *Cognitive Computation* 13, 5 (2021), 1215–1232.
[112] Tapas Nayak and Hwee Tou Ng. 2020. Effective modeling of encoder-decoder architecture for joint entity and relation extraction. In *Proceedings of the AAAI Conference on Artificial Intelligence*. 8528–8535.
[113] Giovanni Paolini, Ben Athiwaratkun, Jason Krone, Jie Ma, Alessandro Achille, RISHITA ANUBHAI, Cicero Nogueira dos Santos, Bing Xiang, and Stefano Soatto. 2021. Structured prediction as translation between augmented natural languages. In *Proceedings of the International Conference on Learning Representations*.
[114] Sachin Pawar, Girish Keshav Palshikar, and Pushpak Bhattacharyya. 2017. Relation extraction : A survey. *arXiv preprint arXiv:1712.05191* (2017).
[115] Yifan Peng, Shankai Yan, and Zhiyong Lu. 2019. Transfer learning in biomedical natural language processing: An evaluation of BERT and ELMo on ten benchmarking datasets. In *Proceedings of the 18th BioNLP Workshop and Shared Task*.
[116] Jeffrey Pennington, Richard Socher, and Christopher D. Manning. 2014. Glove: Global vectors for word representation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 1532–1543.
[117] Matthew E. Peters, Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark, Kenton Lee, and Luke Zettlemoyer. 2018. Deep contextualized word representations. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics*. 2227–2237.
[118] Van-Thuy Phi, Joan Santoso, Van-Hien Tran, Hiroyuki Shindo, Masashi Shimbo, and Yuji Matsumoto. 2019. Distant supervision for relation extraction via piecewise attention and bag-level contextual inference. *IEEE Access* 7 (2019), 103570–103582.
[119] Guanghui Qin and Jason Eisner. 2021. Learning how to ask: Querying LMs with mixtures of soft prompts. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. 5203–5212.
[120] Liang Qiu, Yuan Liang, Yizhou Zhao, Pan Lu, Baolin Peng, Zhou Yu, Ying Nian Wu, and Song-chun Zhu. 2021. SocAoG: Incremental graph parsing for social relation inference in dialogues. 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)*. 658–670.
[121] Meng Qu, Tianyu Gao, Louis-Pascal Xhonneux, and Jian Tang. 2020. Few-shot relation extraction via bayesian meta-learning on relation graphs. In *International Conference on Machine Learning*. PMLR, 7867–7876.
[122] Vipul Rathore, Kartikeya Badola, and Parag Singla. 2022. PARE: A simple and strong baseline for monolingual and multilingual distantly supervised relation extraction. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. 340–354.

[123] Xiang Ren, Zeqiu Wu, Wenqi He, Meng Qu, Clare R. Voss, Heng Ji, Tarek F. Abdelzaher, and Jiawei Han. 2017. Cotype: Joint extraction of typed entities and relations with knowledge bases. In *Proceedings of the 26th International Conference on World Wide Web*. 1015–1024.

[124] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. 2016. " Why should i trust you?" Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 1135–1144.

[125] Sebastian Riedel, Limin Yao, and Andrew McCallum. 2010. Modeling relations and their mentions without labeled text. In *Proceedings of the Joint European Conference on Machine Learning and Knowledge Discovery in Databases*.

[126] Arpita Roy and Shimei Pan. 2021. Incorporating medical knowledge in BERT for clinical relation extraction. In *Proceedings of the EMNLP*.

[127] Arpita Roy and Shimei Pan. 2021. Incorporating medical knowledge in BERT for clinical relation extraction. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*.

[128] Waddah Saeed and Christian Omlin. 2023. Explainable AI (XAI): A systematic meta-survey of current challenges and future opportunities. *Knowledge-Based Systems* 263 (2023), 110273.

[129] Sunil Kumar Sahu, Fenia Christopoulou, Makoto Miwa, and Sophia Ananiadou. 2019. Inter-sentence relation extraction with document-level graph convolutional neural network. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. 4309–4316.

[130] Oscar Sainz, Oier Lopez de Lacalle, Gorka Labaka, Ander Barrena, and Eneko Agirre. 2021. Label verbalization and entailment for effective zero and few-shot relation extraction. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. 1199–1212.

[131] TYSS Santosh, Prantika Chakraborty, Sudakshina Dutta, Debarshi Kumar Sanyal, and Partha Pratim Das. 2021. Joint entity and relation extraction from scientific documents: Role of linguistic information and entity types. *EEKE@ JCDL* 21 (2021), 15–19.

[132] Mourad Sarrouti, Carson Tao, and Yoann Mamy Randriamihaja. 2022. Comparing encoder-only and encoder-decoder transformers for relation extraction from biomedical texts: An empirical study on ten benchmark datasets. In *Proceedings of the BIONLP*.

[133] Timo Schick, Helmut Schmid, and Hinrich Schütze. 2020. Automatically identifying words that can serve as labels for few-shot text classification. In *Proceedings of the 28th International Conference on Computational Linguistics*. 5569–5578.

[134] Timo Schick and Hinrich Schütze. 2021. Exploiting cloze-questions for few-shot text classification and natural language inference. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*. 255–269.

[135] Alessandro Seganti, Klaudia Firląg, Helena Skowronska, Michał Satława, and Piotr Andruszkiewicz. 2021. Multilingual entity and relation extraction dataset and model. In *Proceedings of the ACL*.

[136] Yuming Shang, He-Yan Huang, Xian-Ling Mao, Xin Sun, and Wei Wei. 2020. Are noisy sentences useless for distant supervised relation extraction?. In *Proceedings of the AAAI Conference on Artificial Intelligence*. 8799–8806.

[137] Yu-Ming Shang, Heyan Huang, Xin Sun, Wei Wei, and Xian-Ling Mao. 2023. Learning relation ties with a force-directed graph in distant supervised relation extraction. *ACM Transactions on Information Systems* 41, 1 (2023), 1–23.

[138] Soumya Sharma, Tapas Nayak, Arusarka Bose, Ajay Kumar Meena, Koustuv Dasgupta, Niloy Ganguly, and Pawan Goyal. 2022. FinRED: A dataset for relation extraction in financial domain. In *Companion Proceedings of the Web Conference 2022*. 595–597.

[139] Yatian Shen and Xuan-Jing Huang. 2016. Attention-based convolutional neural network for semantic relation extraction. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*. 2526–2536.

[140] Yongliang Shen, Xinyin Ma, Yechun Tang, and Weiming Lu. 2021. A trigger-sense memory flow framework for joint entity and relation extraction. In *Proceedings of the Web Conference 2021*. 1704–1715.

[141] Taylor Shin, Yasaman Razeghi, Robert L. Logan I. V., Eric Wallace, and Sameer Singh. 2020. AutoPrompt: Eliciting knowledge from language models with automatically generated prompts. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 4222–4235.

[142] Livio Baldini Soares, Nicholas Fitzgerald, Jeffrey Ling, and Tom Kwiatkowski. 2019. Matching the blanks: Distributional similarity for relation learning. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. 2895–2905.

[143] Richard Socher, Brody Huval, Christopher D. Manning, and Andrew Y. Ng. 2012. Semantic compositionality through recursive matrix-vector spaces. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*. 1201–1211.

[144] Junyoung Son, Jinsung Kim, Jung Hoon Lim, and Heu-Jeoung Lim. 2022. GRASP: Guiding model with relational semantics using prompt for dialogue relation extraction. In *Proceedings of the International Conference on Computational Linguistics*.

[145] Wenhui Song, Xiang Zhou, Ping Yang, Yuanyuan Sun, Liang Yang, and Hongfei Lin. 2021. Joint entity and relation extraction for legal texts. In *Proceedings of the 20th Chinese National Conference on Computational Linguistics*. 589–599.

[146] Ananya Subburathinam, Di Lu, Heng Ji, Jonathan May, Shih-Fu Chang, Avirup Sil, and Clare Voss. 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)*. 313–325.

[147] Dianbo Sui, Xiangrong Zeng, Yubo Chen, Kang Liu, and Jun Zhao. 2023. Joint entity and relation extraction with set prediction networks. *IEEE Transactions on Neural Networks and Learning Systems* (2023).

[148] Kai Sun, Richong Zhang, Samuel Mensah, Yongyi Mao, and Xudong Liu. 2020. Recurrent interaction network for jointly extracting entities and classifying relations. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 3722–3732.

[149] Qi Sun, Kun Huang, Xiaocui Yang, Rong Tong, Kun Zhang, and Soujanya Poria. 2024. Consistency guided knowledge retrieval and denoising in LLMs for zero-shot document-level relation triplet extraction. In *Proceedings of the ACM on Web Conference*, 4407–4416.

[150] Bruno Taillé, Vincent Guigue, Geoffrey Scoutheeten, and Patrick Gallinari. 2020. Let's stop incorrect comparisons in end-to-end relation extraction!. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 3689–3701.

[151] Xingwei Tan, Gabriele Pergola, and Yulan He. 2021. Extracting event temporal relations via hyperbolic geometry. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. 8065–8077.

[152] Wei Tang, Benfeng Xu, Yuyue Zhao, Zhendong Mao, Yifeng Liu, Yong Liao, and Haiyong Xie. 2022. UniRel: Unified representation and interaction for joint relational triple extraction. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*. 7087–7099.

[153] Ashok Thillaisundaram and Theodosia Togia. 2019. Biomedical relation extraction with pre-trained language representations and minimal task-specific architecture. In *Proceedings of the 5th Workshop on BioNLP Open Shared Tasks*. 84–89.

[154] Anu Thomas and S. Sangeetha. 2021. Semi-supervised, knowledge-integrated pattern learning approach for fact extraction from judicial text. *Expert Systems* 38, 3 (2021), e12656.

[155] Aryeh Tiktinsky, Vijay Viswanathan, Danna Niezni, Dana Meron Azagury, Yosi Shamay, Hillel Taub-Tabib, Tom Hope, and Yoav Goldberg. 2022. A dataset for N-ary relation extraction of drug combinations. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. 3190–3203.

[156] Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, and Faisal Azhar. 2023. Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971* (2023).

[157] Hieu Minh Tran, Trung Minh Nguyen, and Thien Huu Nguyen. 2020. The dots have their values: Exploiting the node-edge connections in graph-based neural models for document-level relation extraction. In *Proceedings of the Findings*.

[158] Joseph Turian, Lev Ratinov, and Yoshua Bengio. 2010. Word representations: A simple and general method for semi-supervised learning. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*. 384–394.

[159] Shikhar Vashishth, Rishabh Joshi, Sai Suman Prayaga, Chiranjib Bhattacharyya, and Partha Talukdar. 2018. RESIDE: Improving distantly-supervised neural relation extraction using side information. In *Proceedings of the EMNLP*. 1257–1266.

[160] Siddharth Vashishtha, Benjamin Van Durme, and Aaron Steven White. 2019. Fine-grained temporal relation extraction. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. 2906–2919.

[161] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. *Advances in Neural Information Processing Systems* 30 (2017).

[162] Mihaela Vela and Thierry Declerck. 2009. Concept and relation extraction in the finance domain. In *Proceedings of the 8th International Conference on Computational Semantics*. 346–350.

[163] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. 2017. Graph attention networks. *arXiv preprint arXiv:1710.10903* (2017).

[164] Denny Vrandečić and Markus Krötzsch. 2014. Wikidata: A free collaborative knowledgebase. *Communications of the ACM* 57, 10 (2014), 78–85.

[165] David Wadden, Ulme Wennberg, Yi Luan, and Hannaneh Hajishirzi. 2019. Entity, relation, and event extraction with contextualized span 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)*. 5784–5789.

[166] Mengru Wang, Jianming Zheng, Fei Cai, Taihua Shao, and Honghui Chen. 2022. DRK: Discriminative rule-based knowledge for relieving prediction confusions in few-shot relation extraction. In *Proceedings of the International Conference on Computational Linguistics*.

[167] Xinyu Wang, Jiong Cai, Yong Jiang, Pengjun Xie, Kewei Tu, and Wei Lu. 2022. Named entity and relation extraction with multi-modal retrieval. In *Findings of the Association for Computational Linguistics: EMNLP*, 5925–5936.

[168] Ye Wang, Xin-Xin Liu, Wen zhong Hu, and T. Zhang. 2022. A unified positive-unlabeled learning framework for document-level relation extraction with different levels of labeling. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*.

[169] Yuxia Wang, Karin Verspoor, and Timothy Baldwin. 2020. Learning from unlabelled data for clinical semantic textual similarity. In *Proceedings of the 3rd Clinical Natural Language Processing Workshop*. 227–233.

[170] Yucheng Wang, Bowen Yu, Yueyang Zhang, Tingwen Liu, Hongsong Zhu, and Limin Sun. 2020. TPLinker: Single-stage joint extraction of entities and relations through token pair linking. In *Proceedings of the 28th International Conference on Computational Linguistics*. 1572–1582.

[171] Yucheng Wang, Bowen Yu, Yueyang Zhang, Tingwen Liu, Hongsong Zhu, and Limin Sun. 2020. TPLinker: Single-stage joint extraction of entities and relations through token pair linking. In *Proceedings of the International Conference on Computational Linguistics*.

[172] Zihan Wang, Hongye Song, Zhaochun Ren, Pengjie Ren, Zhumin Chen, Xiaozhong Liu, Hongsong Li, and M. de Rijke. 2021. Cross-domain contract element extraction with a bi-directional feedback clause-element relation network. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval.*

[173] Leon Weber, Mario Sänger, Samuele Garda, Fabio Barth, Christoph Alt, and Ulf Leser. 2021. Humboldt@ drugprot: Chemical-protein relation extraction with pretrained transformers and entity descriptions. In *Proceedings of the BioCreative VII Challenge Evaluation Workshop, Online*. 22–25.

[174] Qiang Wei, Zongcheng Ji, Yuqi Si, Jingcheng Du, Jingqi Wang, Firat Tiryaki, Stephen Tze-Inn Wu, Cui Tao, Kirk Roberts, and Hua Xu. 2019. Relation extraction from clinical narratives using pre-trained language models. In *Proceedings of the AMIA Annual Symposium Proceedings.* 1236–1245.

[175] Xiang Wei, Xingyu Cui, Ning Cheng, Xiaobin Wang, Xin Zhang, Shen Huang, Pengjun Xie, Jinan Xu, Yufeng Chen, and Meishan Zhang. 2023. Zero-shot information extraction via chatting with ChatGPT. *arXiv preprint arXiv:2302.10205* (2023).

[176] Zhepei Wei, Jianlin Su, Yue Wang, Yuan Tian, and Yi Chang. 2019. A novel cascade binary tagging framework for relational triple extraction. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*.

[177] Max Welling and Thomas N. Kipf. 2016. Semi-supervised classification with graph convolutional networks. In *Proceedings of the International Conference on Learning Representations (ICLR 2017)*.

[178] Haoyu Wu, Qing Lei, Xinyue Zhang, and Zhengqian Luo. 2020. Creating a large-scale financial news corpus for relation extraction. In *Proceedings of the 2020 3rd International Conference on Artificial Intelligence and Big Data (ICAIBD)*. 259–263.

[179] Shanchan Wu, Kai Fan, and Qiong Zhang. 2019. Improving distantly supervised relation extraction with neural noise converter and conditional optimal selector. In *Proceedings of the AAAI Conference on Artificial Intelligence*. 7273–7280.

[180] Shanchan Wu and Yifan He. 2019. Enriching pre-trained language model with entity information for relation classification. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*. 2361–2364.

[181] Benfeng Xu, Quan Wang, Yajuan Lyu, Yong Zhu, and Zhendong Mao. 2021. Entity structure within and throughout: Modeling mention dependencies for document-level relation extraction. In *Proceedings of the AAAI Conference on Artificial Intelligence*.

[182] Can Xu, Qingfeng Sun, Kai Zheng, Xiubo Geng, Pu Zhao, Jiazhan Feng, Chongyang Tao, and Daxin Jiang. 2023. Wizardlm: Empowering large language models to follow complex instructions. *arXiv preprint arXiv:2304.12244* (2023).

[183] Jia Xu, Pengwei Yang, Shang Xue, Bhuvan Sharma, Marta Sanchez-Martin, Fang Wang, Kirk A. Beaty, Elinor Dehan, and Baiju Parikh. 2019. Translating cancer genomics into precision medicine with artificial intelligence: Applications, challenges and future perspectives. *Human Genetics* 138, 2 (2019), 109–124.

[184] Kun Xu, Yansong Feng, Songfang Huang, and Dongyan Zhao. 2015. Semantic relation classification via convolutional neural networks with simple negative sampling. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*. 536–540.

[185] Wang Xu, Kehai Chen, and Tiejun Zhao. 2021. Document-level relation extraction with reconstruction. In *Proceedings of the AAAI Conference on Artificial Intelligence*. 14167–14175.

[186] Weiwen Xu, Yang Deng, Wenqiang Lei, Wenlong Zhao, Tat-Seng Chua, and Wai Lam. 2022. ConReader: Exploring implicit relations in contracts for contract clause extraction. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, 2581–2594.

[187] Weiwen Xu, Yang Deng, Wenqiang Lei, Wenlong Zhao, Tat-Seng Chua, and Wai Lam. 2022. ConReader: Exploring implicit relations in contracts for contract clause extraction. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*.

[188] Xingtian Xu, Xiangnan Chen, Ningyu Zhang, Xin Xie, Xi Chen, and Huajun Chen. 2022. Towards realistic low-resource relation extraction: A benchmark with empirical baseline study. In *Findings of the Association for Computational Linguistics: EMNLP*, 413–427.

[189] Xin Xu, Yuqi Zhu, Xiaohan Wang, and Ningyu Zhang. 2023. How to unleash the power of large language models for few-shot relation extraction?. In *Proceedings of the 4th Workshop on Simple and Efficient Natural Language Processing (SustaiNLP)*. 190–200.

[190] Yan Xu, Lili Mou, Ge Li, Yunchuan Chen, Hao Peng, and Zhi Jin. 2015. Classifying relations via long short term memory networks along shortest dependency paths. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*. 1785–1794.

[191] Fuzhao Xue, Aixin Sun, Hao Zhang, and Eng Siong Chng. 2020. An embarrassingly simple model for dialogue relation extraction. In *Proceedings of the ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 6707–6711.

[192] Fuzhao Xue, Aixin Sun, Hao Zhang, and Eng Siong Chng. 2020. GDPNet: Refining latent multi-view graph for relation extraction. In *Proceedings of the AAAI Conference on Artificial Intelligence*.

[193] Fuzhao Xue, Aixin Sun, Hao Zhang, and Eng Siong Chng. 2021. Gdpnet: Refining latent multi-view graph for relation extraction. In *Proceedings of the AAAI*. 14194–14202.

[194] Zhuang Liu, Degen Huang, Kaiyu Huang, Zhuang Li, and Jun Zhao. 2021. Finbert: A pre-trained financial language representation model for financial text mining. In *Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence*, 4513–4519.

[195] Shweta Yadav, Srivastsa Ramesh, Sriparna Saha, and Asif Ekbal. 2020. Relation extraction from biomedical and clinical text: Unified multitask learning framework. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 19, 2 (2020), 1105–1116.

[196] Yadollah Yaghoobzadeh, Heike Adel, and Hinrich Schütze. 2017. Noise mitigation for neural entity typing and relation extraction. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers*. 1183–1194.

[197] Jianhao Yan, Lin He, Ruqin Huang, Jian Li, and Ying Liu. 2019. Relation extraction with temporal reasoning based on memory augmented distant supervision. 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)*. 1019–1030.

[198] Zhaohui Yan, Zixia Jia, and Kewei Tu. 2022. An empirical study of pipeline vs. joint approaches to entity and relation extraction. In *Proceedings of the ACL*.

[199] Zhiheng Yan, Chong Zhang, Jinlan Fu, Qi Zhang, and Zhongyu Wei. 2021. A partition filter network for joint entity and relation extraction. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. 185–197.

[200] Kaijia Yang, Nantao Zheng, Xinyu Dai, Liang He, Shujian Huang, and Jiajun Chen. 2020. Enhance prototypical network with text descriptions for few-shot relation classification. In *Proceedings of the 29th ACM International Conference on Information and Knowledge Management*. 2273–2276.

[201] Shan Yang, Yongfei Zhang, Guanglin Niu, Qinghua Zhao, and Shiliang Pu. 2021. Entity concept-enhanced few-shot relation extraction. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*. 987–991.

[202] Zuoxi Yang. 2020. Biomedical information retrieval incorporating knowledge graph for explainable precision medicine. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. 2486–2486.

[203] Shaowei Yao, Tianming Wang, and Xiaojun Wan. 2020. Heterogeneous graph transformer for graph-to-sequence learning. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*.

[204] Yuan Yao, Deming Ye, Peng Li, Xu Han, Yankai Lin, Zhenghao Liu, Zhiyuan Liu, Lixin Huang, Jie Zhou, and Maosong Sun. 2019. DocRED: A large-scale document-level relation extraction dataset. In *Proceedings of the ACL*.

[205] Yuan Yao, Deming Ye, Peng Li, Xu Han, Yankai Lin, Zhenghao Liu, Zhiyuan Liu, Lixin Huang, Jie Zhou, and Maosong Sun. 2019. DocRED: A large-scale document-level relation extraction dataset. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 764–777.

[206] Deming Ye, Yankai Lin, Peng Li, and Maosong Sun. 2022. Packed levitated marker for entity and relation extraction. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. 4904–4917.

[207] Hongbin Ye, Ningyu Zhang, Shumin Deng, Mosha Chen, Chuanqi Tan, Fei Huang, and Huajun Chen. 2021. Contrastive triple extraction with generative transformer. In *Proceedings of the AAAI Conference on Artificial Intelligence*. 14257–14265.

[208] Zhi-Xiu Ye and Zhen-Hua Ling. 2019. Distant supervision relation extraction with intra-bag and inter-bag attentions. 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)*. 2810–2819.

[209] Bowen Yu, Zhenyu Zhang, Tingwen Liu, Bin Wang, Sujian Li, and Quangang Li. 2019. Beyond word attention: Using segment attention in neural relation extraction. In *Proceedings of the IJCAI*. 5401–5407.

[210] Bowen Yu, Zhenyu Zhang, Xiaobo Shu, Tingwen Liu, Yubin Wang, Bin Wang, and Sujian Li. 2020. Joint extraction of entities and relations based on a novel decomposition strategy. In *(ECAI'20)*. IOS Press, 2282–2289.

[211] Dian Yu, Kai Sun, Claire Cardie, and Dong Yu. 2020. Dialogue-based relation extraction. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. 4927–4940.

[212] Erxin Yu, Wenjuan Han, Yuan Tian, and Yi Chang. 2020. Tohre: A top-down classification strategy with hierarchical bag representation for distantly supervised relation extraction. In *Proceedings of the 28th International Conference on Computational Linguistics*. 1665–1676.

[213] Haiyang Yu, Ningyu Zhang, Shumin Deng, Hongbin Ye, Wei Zhang, and Huajun Chen. 2020. Bridging text and knowledge with multi-prototype embedding for few-shot relational triple extraction. In *Proceedings of the 28th International Conference on Computational Linguistics*. 6399–6410.

[214] Tianshu Yu, Min Yang, and Xiaoyan Zhao. 2022. Dependency-aware prototype learning for few-shot relation classification. In *Proceedings of the 29th International Conference on Computational Linguistics*. 2339–2345.

[215] Changsen Yuan, Heyan Huang, Chong Feng, Xiao Liu, and Xiaochi Wei. 2019. Distant supervision for relation extraction with linear attenuation simulation and non-iid relevance embedding. In *Proceedings of the AAAI Conference on Artificial Intelligence*. 7418–7425.

[216] Li Yuan, Yi Cai, Jin Wang, and Qing Li. 2023. Joint multimodal entity-relation extraction based on edge-enhanced graph alignment network and word-pair relation tagging. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 11051–11059.

[217] Yujin Yuan, Liyuan Liu, Siliang Tang, Zhongfei Zhang, Yueting Zhuang, Shiliang Pu, Fei Wu, and Xiang Ren. 2019. Cross-relation cross-bag attention for distantly-supervised relation extraction. In *Proceedings of the AAAI Conference on Artificial Intelligence*. 419–426.

[218] Yue Yuan, Xiaofei Zhou, Shirui Pan, Qiannan Zhu, Zeliang Song, and Li Guo. 2020. A relation-specific attention network for joint entity and relation extraction. In *Proceedings of the International Joint Conference on Artificial Intelligence*.

[219] Y. Yuan, X. Zhou, S. Pan, Q. Zhu, Z. Song, and L. Guo. 2021. A relation-specific attention network for joint entity and relation extraction. In *Proceedings of the International Joint Conference on Artificial Intelligence*. International Joint Conference on Artificial Intelligence.

[220] Dmitry Zelenko, Chinatsu Aone, and Anthony Richardella. 2002. Kernel methods for relation extraction. In *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing*.

[221] Daojian Zeng, Kang Liu, Yubo Chen, and Jun Zhao. 2015. Distant supervision for relation extraction via piecewise convolutional neural networks. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*. 1753–1762.

[222] Daojian Zeng, Kang Liu, Siwei Lai, Guangyou Zhou, and Jun Zhao. 2014. Relation classification via convolutional deep neural network. In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*. 2335–2344.

[223] Daojian Zeng, Haoran Zhang, and Qianying Liu. 2020. Copymtl: Copy mechanism for joint extraction of entities and relations with multi-task learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*. 9507–9514.

[224] Shuang Zeng, Yuting Wu, and Baobao Chang. 2021. SIRE: Separate intra-and inter-sentential reasoning for document-level relation extraction. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*. 524–534.

[225] Shuang Zeng, Runxin Xu, Baobao Chang, and Lei Li. 2020. Double graph based reasoning for document-level relation extraction. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 1630–1640.

[226] Xiangrong Zeng, Shizhu He, Daojian Zeng, Kang Liu, Shengping Liu, and Jun Zhao. 2019. Learning the extraction order of multiple relational facts in a sentence with reinforcement learning. 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)*. 367–377.

[227] Xiangrong Zeng, Daojian Zeng, Shizhu He, Kang Liu, and Jun Zhao. 2018. Extracting relational facts by an end-to-end neural model with copy mechanism. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. 506–514.

[228] Min Zhang, Jie Zhang, Jian Su, and Guodong Zhou. 2006. A composite kernel to extract relations between entities with both flat and structured features. In *Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics*. 825–832.
[229] Rui Zhang, Bayu Distiawan Trisedya, Miao Li, Yong Jiang, and Jianzhong Qi. 2022. A benchmark and comprehensive survey on knowledge graph entity alignment via representation learning. *The VLDB Journal* 31, 5 (2022), 1143–1168.
[230] Ranran Haoran Zhang, Qianying Liu, Aysa Xuemo Fan, Heng Ji, Daojian Zeng, Fei Cheng, Daisuke Kawahara, and Sadao Kurohashi. 2020. Minimize exposure bias of seq2seq models in joint entity and relation extraction. In *Findings of the Association for Computational Linguistics: EMNLP 2020*. 236–246.
[231] Yuhao Zhang, Peng Qi, and Christopher D. Manning. 2018. Graph convolution over pruned dependency trees improves relation extraction. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. 2205–2215.
[232] Yuhao Zhang, Victor Zhong, Danqi Chen, Gabor Angeli, and Christopher D. Manning. 2017. Position-aware attention and supervised data improve slot filling. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*.
[233] Di Zhao, Jian Wang, Hongfei Lin, Xin Wang, Zhihao Yang, and Yijia Zhang. 2021. Biomedical cross-sentence relation extraction via multihead attention and graph convolutional networks. *Applied Soft Computing* 104 (2021), 107230.
[234] Hongyan Zhao, Ru Li, Xiaoli Li, and Hongye Tan. 2020. CFSRE: Context-aware based on frame-semantics for distantly supervised relation extraction. *Knowledge-Based Systems* 210 (2020), 106480.
[235] Jun Zhao, Tao Gui, Qi Zhang, and Yaqian Zhou. 2021. A relation-oriented clustering method for open relation extraction. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. 9707–9718.
[236] Kang Zhao, Hua Xu, Yue Cheng, Xiaoteng Li, and Kai Gao. 2021. Representation iterative fusion based on heterogeneous graph neural network for joint entity and relation extraction. *Knowledge-Based Systems* 219 (2021), 106888.
[237] Kang Zhao, Hua Xu, Jiangong Yang, and Kai Gao. 2022. Consistent representation learning for continual relation extraction. In *Findings of the Association for Computational Linguistics: ACL 2022*. 3402–3411.
[238] Tianyang Zhao, Zhao Yan, Yunbo Cao, and Zhoujun Li. 2021. Asking effective and diverse questions: A machine reading comprehension based framework for joint entity-relation extraction. In *Proceedings of the 29th International Conference on International Joint Conferences on Artificial Intelligence*. 3948–3954.
[239] Wayne Xin Zhao, Kun Zhou, Junyi Li, Tianyi Tang, Xiaolei Wang, Yupeng Hou, Yingqian Min, Beichen Zhang, Junjie Zhang, and Zican Dong. 2023. A survey of large language models. *arXiv preprint arXiv:2303.18223* (2023).
[240] Xiaoyan Zhao, Min Yang, Qiang Qu, and Ruifeng Xu. 2024. Few-shot relation extraction with automatically generated prompts. *IEEE Transactions on Neural Networks and Learning Systems* (2024).
[241] Xiaoyan Zhao, Min Yang, Qiang Qu, Ruifeng Xu, and Jieke Li. 2023. Exploring privileged features for relation extraction with contrastive student-teacher learning. *IEEE Transactions on Knowledge and Data Engineering* 35, 8 (2023), 7953–7965.
[242] Changmeng Zheng, Junhao Feng, Ze Fu, Yi Cai, Qing Li, and Tao Wang. 2021. Multimodal relation extraction with efficient graph alignment. In *Proceedings of the MM*.
[243] Changmeng Zheng, Zhiwei Wu, Junhao Feng, Ze Fu, and Yi Cai. 2021. MNRE: A challenge multimodal dataset for neural relation extraction with visual evidence in social media posts. In *Proceedings of the ICME*.
[244] Hengyi Zheng, Rui Wen, Xi Chen, Yifan Yang, Yunyan Zhang, Ziheng Zhang, Ningyu Zhang, Bin Qin, Xu Ming, and Yefeng Zheng. 2021. PRGC: Potential relation and global correspondence based joint relational triple extraction. 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)*. 6225–6235.
[245] Suncong Zheng, Feng Wang, Hongyun Bao, Yuexing Hao, Peng Zhou, and Bo Xu. 2017. Joint extraction of entities and relations based on a novel tagging scheme. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 1227–1236.
[246] Zexuan Zhong and Danqi Chen. 2021. A frustratingly easy approach for entity and relation extraction. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. 50–61.
[247] GuoDong Zhou, Jian Su, Jie Zhang, and Min Zhang. 2005. Exploring various knowledge in relation extraction. In *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL–05)*. 427–434.
[248] Peng Zhou, Suncong Zheng, Jiaming Xu, Zhenyu Qi, Hongyun Bao, and Bo Xu. 2017. Joint extraction of multiple relations and entities by using a hybrid neural network. In *Proceedings of the Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data*. Springer, 135–146.

[249] Wenxuan Zhou, Kevin Huang, Tengyu Ma, and Jing Huang. 2021. Document-level relation extraction with adaptive thresholding and localized context pooling. In *Proceedings of the AAAI Conference on Artificial Intelligence*. 14612–14620.
[250] Tong Zhu, Haitao Wang, Junjie Yu, Xiabing Zhou, Wenliang Chen, Wei Zhang, and Min Zhang. 2020. Towards accurate and consistent evaluation: A dataset for distantly-supervised relation extraction. In *Proceedings of the 28th International Conference on Computational Linguistics*. 6436–6447.

Received 9 May 2023; revised 28 May 2024; accepted 2 June 2024