Information Extraction in Low-Resource Scenarios: Survey and Perspective

Shumin Deng
National University of Singapore
NUS-NCS Joint Lab
Singapore
shumin@nus.edu.sg

Yubo Ma
S-Lab, Nanyang Technological University
Singapore
yubo001@e.ntu.edu.sg

Ningyu Zhang[†]
Zhejiang University
China
zhangningyu@zju.edu.cn

Yixin Cao
Fudan University
China
yxcao@fudan.edu.cn

Bryan Hooi[†]
National University of Singapore
NUS-NCS Joint Lab
Singapore
dcsbhk@nus.edu.sg

Abstract—**Information Extraction (IE) seeks to derive structured information from unstructured texts, often encountering obstacles in low-resource scenarios due to data scarcity and unseen classes. This paper presents a review of neural approaches to low-resource IE from *traditional* and *LLM-based* perspectives, systematically organizing them into a fine-grained taxonomy. Our empirical studies compare LLM-based methods with prior leading models, revealing that: (1) well-tuned LMs perform relatively best; (2) tuning open-resource LLMs and in-context learning with GPT family are generally effective; (3) LLMs struggle to tackle complex tasks with intricate schema. Furthermore, we compare traditional methods and discuss LLM-based approaches, spotlighting promising applications and delineating future research directions. This survey aims to foster understanding of this field, inspire new ideas, and encourage widespread applications in both academia and industry.**

I. INTRODUCTION

Information Extraction (IE) [1], [2] offers essential support for a diverse range of domain-specific [3], [4] and knowledge-intensive [5], [6] tasks, thus appealing to AI community. Common IE tasks, including Named Entity Recognition (NER) [7], Relation Extraction (RE) [8], and Event Extraction (EE) [9], aim to derive structured data from unstructured texts. While deep learning has boosted IE performance, it demands massive amounts of labeled data, which is difficult to acquire due to task and domain variability in practice. This issue motivates research in **low-resource IE** (data-efficient IE), which seeks to enhance learning efficiency under real world contexts with sparse data.

Low-resource IE has been widely investigated due to its potential for making models data-efficient and adaptable to various scenarios. Among numerous approaches proposed in recent years, Large Language Models (LLMs) [10], [11] have demonstrated promising performance in low-resource scenarios. Hence, a timely review on low-resource IE is beneficial in offering insights to both researchers and industry practitioners. Though several surveys on task-specific low-resource IE (*e.g.*,

NER [12], [13], RE [14], [15], EE [16]) and low-resource NLP [17], [18] have been published, a comprehensive survey for low-resource IE encompassing both traditional models and modern LLMs is still absent. Therefore, in this paper, we provide a literature review on low-resource IE, hoping to systematically analyse the methodologies, compare different technical solutions and inspire diverse new ideas.

In this paper, we begin by introducing the basics of low-resource learning and IE in §II. We then categorize low-resource IE approaches[1] into **Traditional** (§III-A) and **LLM-based Methods** (§III-B). In view of *which aspect is chiefly enhanced when training*, we further categorize **traditional low-resource IE approaches** into three paradigms:

Exploiting Higher-Resource Data (§III-A1): Increasing the size and diversity of the original sparse data with auxiliary resources which are generated endogenously or imported exogenously, aiming to obtain enriched samples and more precise semantic representations.

Developing Stronger Data-Efficient Models (§III-A2): Developing more robust models to better cope with maldistribution of samples and new unseen classes, aiming to improve model learning abilities so as to reduce dependence on samples.

Optimizing Data & Models Together (§III-A3): Jointly utilizing representative samples and robust models, aiming at adapting to low-resource scenarios promptly, and searching more suitable strategies for learning with sparse data.

We also divide **LLM-based low-resource IE approaches** based on *whether the LLM is tuned*, into two paradigms:

Direct Inference Without Tuning (§III-B1): Prompting LLMs to generate answers with given instructions, where prompts can either be in textual or code format.

Model Specialization With Tuning (§III-B2): Fine-tuning LLMs, either via prompt-tuning or standard fine-tuning. The fine-tuned LLMs are often tailored for specific IE tasks.

[†]Corresponding Author.

[1]Due to page limits, complete papers are listed: https://github.com/zjunlp/Low-resource-KEPapers.

We then conduct empirical studies comparing traditional and LLM-based low-resource IE approaches in §IV. We explore the effectiveness of different techniques, especially those based on LLMs, to identify the most suitable approach for various low-resource IE tasks. The **key findings** are: (1) Well-tuned LMs continue to dominate in performance; (2) Tuning open-resource LLMs and using in-context learning (ICL) with GPTs presents promising results; (3) LLMs struggle to address the complex task with intricate schema. Besides, we summarize widely used benchmarks and promising applications in §V. Finally, we compare existing low-resource IE methods, discuss future directions in §VI, and make a conclusion in §VII.

II. PRELIMINARY ON LOW-RESOURCE IE

A. Information Extraction

Generally, IE [19], [20] can be regarded as structured prediction tasks [21], where a classifier is trained to approximate a target function $F(x) \rightarrow y$, where $x \in \mathcal{X}$ denotes the input data and $y \in \mathcal{Y}$ denotes the output class sequence. This includes the subtasks of Named Entity Recognition (NER), Relation Extraction (RE) and Event Extraction (EE). For instance, given a sentence "*Jack is married to the Iraqi microbiologist known as Dr. Germ.*":

Named Entity Recognition [7] should identify the types of entities in given texts, *e.g.*, '*Jack*', '*Dr. Germ*' $\Rightarrow$ PERSON;

Relation Extraction [8] should identify the relationship of the given entity pair ⟨*Jack, Dr. Germ*⟩ as husband_of;

Event Extraction [9] should identify the event type as Marry, where the word '*married*' triggers the event (subtask: ED, Event Detection); *Jack* and *Dr. Germ* are participants as husband and wife in the event respectively (subtask: EAE, Event Argument Extraction).

B. Low-Resource Scenarios

Most traditional IE models [22]–[24] assume that sufficient training data are indispensable to achieve satisfactory performance. However, in real-world applications, task-specific labeled samples tend to be unevenly distributed and new unseen classes may emerge over time, which result in low-resource scenarios [17]. Considering *maldistribution of samples* and *new unseen classes*, we systematically categorize low-resource scenarios into three aspects.

Long-tail Scenario [25] means that only a minority of the classes are data-rich, while the majority of the classes have extremely few labeled sample data. Formally, given classes $\mathcal{Y} = \mathcal{Y}_h \cup \mathcal{Y}_t$ with head classes ($\mathcal{Y}_h$) and long-tail classes ($\mathcal{Y}_t$), we denote the number of labeled samples for $\mathcal{Y}_h$ and $\mathcal{Y}_t$ as $|\mathcal{X}_h|$ and $|\mathcal{X}_t|$ respectively. $|\mathcal{X}_h|$ is much larger than $|\mathcal{X}_t|$, *i.e.*, $|\mathcal{X}_h| \gg |\mathcal{X}_t|$, while $|\mathcal{Y}_h| \ll |\mathcal{Y}_t|$.

Few-shot Scenario [26], [27] means that the classes for testing $\mathcal{Y}_{te}$ have only a small number of samples, where the small number can be *fixed* following the N-Way-K-Shot setting [28] (each of the N classes contains only K samples), or *unfixed* and relatively small comparing with the total data size. Besides, $\mathcal{Y}_{te}$ is nonexistent (unseen) in the training dataset with classes $\mathcal{Y}_{tr}$, *i.e.*, $\mathcal{Y}_{tr} \cap \mathcal{Y}_{te} = \emptyset$.

Zero-shot Scenario [29] means that the testing classes $\mathcal{Y}_{te}$ to predict have no samples during the training phase. The model can classify these new unseen $\mathcal{Y}_{te}$ by utilizing prior knowledge or relationships between seen and unseen classes, represented as class attributes or embeddings. In *standard zero-shot scenarios*, $\mathcal{Y}_{te}$ only contains unseen classes in the training set; in *generalized zero-shot scenarios* [30], $\mathcal{Y}_{te}$ contains both seen and unseen classes.

III. TAXONOMY OF TECHNICAL SOLUTIONS

In general, for this survey, we select influential low-resource IE papers mostly published within the last five years. The majority of these papers are sourced from prestigious conferences and journals, especially in NLP and machine learning domains, *e.g.*, ACL, EMNLP, NAACL, NeurIPS, ICLR, TACL, TKDE and so on. We generally categorize existing approaches into traditional and LLM-based methods based on the *model parameter size* with a threshold: 6B. For **Traditional Methods**, the backbone model parameter size is less than 6B, *e.g.*, BERT [174], BART [175], T5-large [176], GPT-2 [177], GPT-J[2], *etc*. For **LLM-based Methods**, the backbone model parameter size is larger than 6B, *e.g.*, FlanT5-11B [178], LLaMA [179], GLM [180], GPT-3 [181], ChatGPT[3], GPT-4 [182], *etc*. We refine representative methods in the two categories, and establish a fine-grained taxonomy as presented in Figure 1.

A. Traditional Methods

In view of *which aspect is chiefly enhanced when addressing low-resource IE problems*, we categorize some representative traditional methods into three general paradigms, as shown in Figure 1. In summary, traditional low-resource IE models tend to (1) exploit *higher*-resource data; (2) develop *stronger* data-efficient models; and (3) optimize data & models *together*.

1) Exploiting Higher-Resource Data: This paradigm mainly refers to data augmentation [183] or knowledge enhancement on the original small dataset, with endogenous or exogenous auxiliary resources. *The goal is to create more representative samples and improve semantic representations with higher-resource data.*

Weakly Supervised Augmentation involves synthesizing more training data through weak/distant supervision. This kind of methods usually utilize a knowledge base (KB) and some heuristic rules to automatically relabel training instances in corpus, potentially resulting in a noisy synthesized dataset.

[184] proposed *distant supervision* for RE, utilizing a large semantic KB — Freebase to label relations in an unlabeled corpus. Similarly, [31] used a dictionary for distantly supervised NER. [32] leveraged Freebase and FrameNet (an event KB) to automatically label training data for EE. Recent studies have proposed to improve quality of weakly supervised data with pretrained LMs (PLMs) [33], [34] and optimization strategies [35]–[39]. To mitigate selection bias, [57] have proposed a gradient imitation reinforcement learning framework for low-resource RE.

---

[2]https://huggingface.co/EleutherAI/gpt-j-6b.
[3]https://openai.com/blog/chatgpt.

Low-Resource IE

Traditional Methods (§III-A)

Exploiting *Higher*-Resource *Data* (§III-A1)

- **Weakly Supervised Augmentation**
  - *KB-based*: AutoNER [31]; [32]
  - *PLM-based*: LADA [33]; BOND [34]
  - *Strategy-enhanced*: [35]; DMBERT$_{ADV}$ [36]; REIF [37]; Jointprop [38]; Top-Neg [39]
- **Multi-Modal Augmentation**
  - *Attention Mechanism*: [40]; DRMM [41]
  - *Space Mapping*: WASE [42]; JMMT [43]; MEGA [44]
  - *Prefix Guided*: MKGformer [45]; HVPNeT [46]
- **Multi-Lingual Augmentation**
  - *Knowledge Transfer*: [47]; MAFT [48]; [49]
  - *Consistency Capture*: [50]; GMLATT [51]; PRAM [52]
  - *Input Augmentation*: IE-PRIME [53]; CLRED [54]
- **Auxiliary Knowledge Enhancement**
  - *Textual Knowledge*: [55]; ZSBERT [56]; GradLRE [57]; Mask-then-Fill [58]; MELM [59]; GPDA [60]; EnTDA [61]; ACLM [62]; S2ynRE [63]
  - *Structured Knowledge*: [64]; NERO [65]; KaNa [66]; PathLM [67]; OntoED [68]; kHPN [69]

Developing *Stronger* Data-Efficient *Models* (§III-A2)

- **Meta Learning**
  - *Metric-based*: HATT [70]; DMBPN [71]; MetaNER [72]; COL [73]; PA-CRF [74]; MCMN [75]; [76]; ProML [77]; HEProto [78]; MultiPLe [79]
  - *MAML-based*: MIML [80]; FewNER [81]; MetaEvent [82]
  - *Memory-based*: MCML [83]
  - *Model-based*: REGRAB [84]; AKE-BML [85]
- **Transfer Learning**
  - ZSEE [86]; [87]; MZET [88]; [89]; wRAN [90]; DualGraph [91]; MANNER [92]
- **Fine-Tuning PLM**
  - Matching the Blanks [93]; PLMEE [94]; [95]; CoFEE-MRC [96]; SDNet [97]; MCCL [98]; CLEVE [99]; GPTEDOT [100]; ClarET [101]; Fɪsʜ-Dɪᴘ [102]; ZED [103]; MSDP [104]

Optimizing *Data & Models* Together (§III-A3)

- **Multi-Task Learning**
  - *NEN+NER*: [105] *WSD+ED*: [106]; GraphTransfer [107]
  - *NER+RE*: GraphRel [108]; CopyMTL [109];
  - *Joint IE*: DyGIE++ [110]; Sᴘᴇᴇᴄʜ [111]; Mirror [112]
- **Task Reformulation**
  - Multi-turn QA [113]; BERT-MRC [114]; RCEE [115]; QAEE [116]; EE-GQA [117]; IE-QA [118]; Text2Event [119]; TANL [120]; Set [121]
- **Prompt-Tuning PLM**
  - *Vanilla Prompt-Tuning*: [122]; LightNER [123]; PAIE [124]; RelationPrompt [125]; EntLM [126]; [127]; PromptNER [128]; DSP [129]; SPEAE [130]
  - *Augmented Prompt-Tuning*: KnowPrompt [131]; RetroPrompt [132]; OntoPrompt [133]; DEGREE [134]; AMPERE [135]; UIE [136]; LasUIE [137]; USM [138]; RexUIE [139]; PTR [140]; RAP [141]; [15]

LLM-Based Methods (§III-B)

Direct Inference *Without Tuning* (§III-B1)

- **Instruction Prompting**
  - [142]; ChatIE [143]; [144]; SᴜᴍAsᴋ [145]; PromptNER [146]; [147]; AutoKG [148]
- **Code Prompting**
  - Code4Struct [149]; CodeIE [150]; CodeKGC [151]; ViStruct [152]; Code4UIE [153]; GoLLIE [154]
- **In-Context Learning**
  - In-context NER [155]; [156]; GPT-RE [157]; GL [158]; Reranker [159]; GPT-NER [160]; REPLM [161]; [162]; [163]; [164]; [165]; LOKE [166]; [167]; PGIM [168]

Model Specialization *With Tuning* (§III-B2)

- **Prompt-Tuning LLM**
  - DeepStruct [169]; QA4RE [170]; Oᴅɪᴇ [171]; UniversalNER [172]; InstructUIE [173]
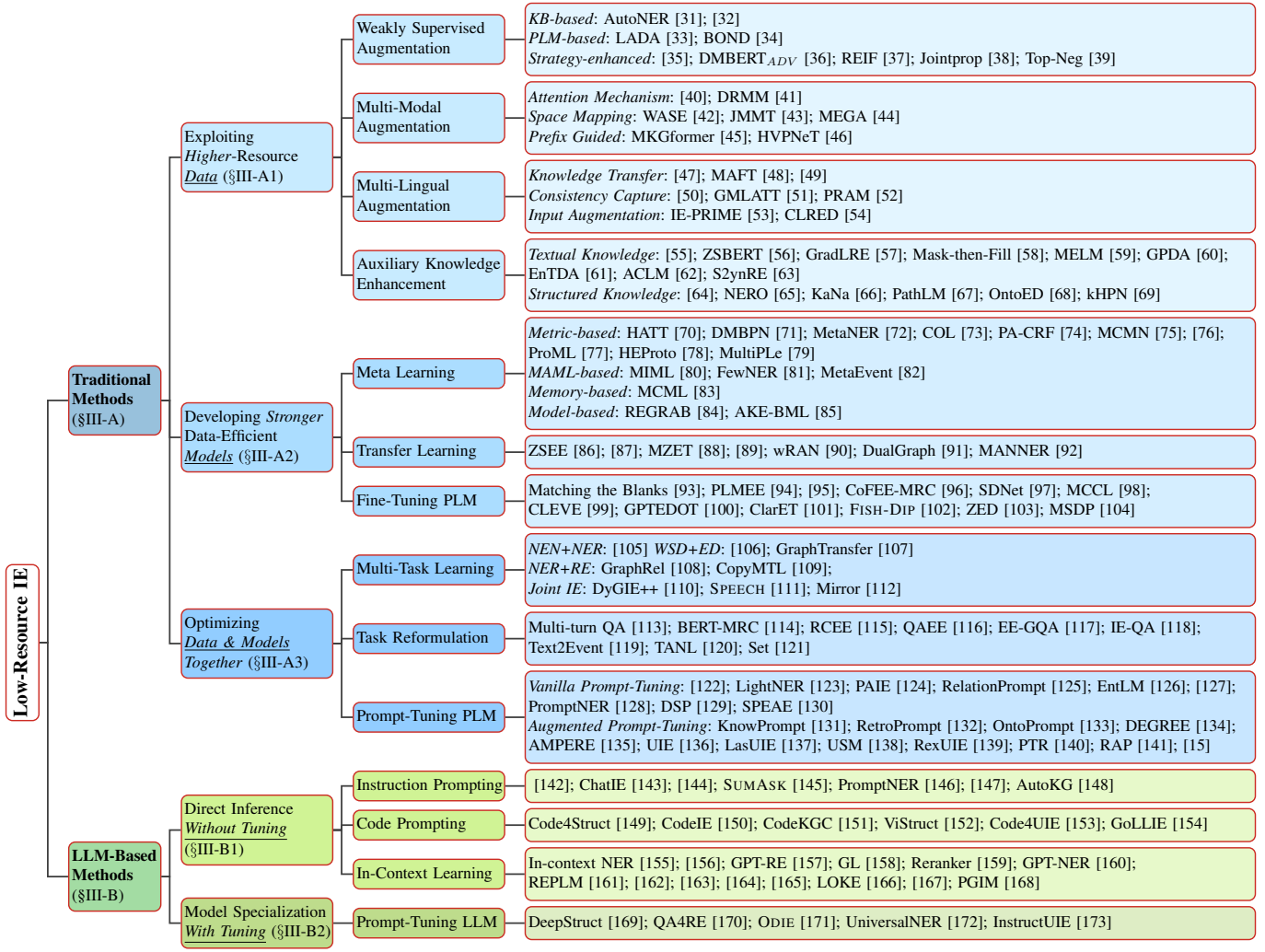
Fig. 1. Taxonomy of low-resource IE methods. (We only list part of representative approaches, referring to the complete paper list[1] for more details.)

Multi-Modal Augmentation supplements single modal with multi-modal samples to enhance semantics and facilitate disambiguation. Intuitively, the main challenge of such methods lies in effectively fusing data from different modalities.

To fuse multi-modal data in low-resource IE, [40], [41] utilized attention mechanisms, [42]–[44] imposed multi-modal embedding space mapping, and [45], [46] leveraged prefix-guided multi-modal fusion.

Multi-Lingual Augmentation involves incorporating multi-lingual samples to achieve diverse and robust sample representation. Intuitively, the main challenge of such methods is to obtain language representations cross linguistics.

To tackle the issue, transferring cross-lingual knowledge [47]–[49] and capturing the consistency cross languages [50]–[52], [185] demonstrated effectiveness. Other promising models are to import additional context [54] and task-specific knowledge [53] for cross-lingual IE.

Auxiliary Knowledge Enhancement employs external knowledge as remedies, intended to learn semantic representation of samples more precisely. Different from weakly supervised augmentation using KB, this paradigm adopts more diverse knowledge formats and knowledge enhancement methods. Given diverse formats of auxiliary knowledge, we divide them into two categories.

(i) Textual Knowledge

Textual knowledge can be *class-related knowledge* [186], [187], like class descriptions [56], [188], [189] and class-specific texts [55]. It can also be *synthesized data* [58]–[63], [190]–[194] via data augmentation.

(ii) Structured Knowledge

Structured knowledge for low-resource IE can take the form of KG triples [195], task-specific ontology [196], and rules [65]. For example, [64] studied zero-shot RE by leveraging KG embeddings and logic rules to connect seen and unseen relations; [66] proposed a KB-aware NER framework to utilize class-heterogeneous knowledge in KBs; [67]–[69] resolved low-resource EE problems by utilizing association knowledge among classes.

2) Developing Stronger Data-Efficient Models: The paradigm focuses on developing models to handle sample

maldistribution and unseen classes more effectively. *The stronger models aim to improve learning abilities, so as to maximize the utilization of small data and minimize the dependence of samples.*

Meta Learning [197] promptly assimilates emerging knowledge and deduces new classes by learning from few instances, with the ability of "learning to learn", which is naturally suitable for few-shot IE tasks.

[70]–[79] utilized *metric-based* methods, mostly equipped with prototypical networks [28]. [80]–[82], [198]–[200] leveraged *model-agnostic* methods based on MAML [201]. [83] proposed a *memory-based* method [202]. [84], [85] used *model-based* methods with Bayesian meta learning [203].

Transfer Learning [204] reduces the dependence on labeled target data by transferring learned class-invariant features, especially from high-resource to low-resource classes.

[86] leveraged class structures to transfer knowledge from existing to unseen classes. [87] proposed a weighted adversarial network to adapt features learned from high-resource to low-resource classes. [87], [91] utilized graph neural networks to facilitate knowledge transfer. [88], [92] used a memory module for information retrieval or similarity comparison from the source to target domain. [89] incorporated linguistic representations for cross-domain few-shot transfer.

Fine-Tuning PLM leverages PLMs [205] to utilize contextual representations and pre-trained parameters for fine-tuning. It adapts PLMs' powerful language understanding capabilities to specific low-resource IE tasks.

Fine-tuning PLMs for low-resource IE aims at learning task-specific entity representations [96], [97], [104]; relation representations [93], [98], [102]; and event representations [94], [95], [99]–[101], [103] with data-efficient learning.

3) Optimizing Data & Models Together: This paradigm refers to jointly optimizing representative samples and data-efficient models, enabling swift adaptation to low-resource scenarios. *The aim is to identify more suitable strategies for learning with sparse data.*

Multi-Task Learning signifies learning multiple related tasks simultaneously by exploiting both task-generic commonality and task-specific diversity [169], contributing to improved performance for task-specific models, which naturally enables boost the target low-resource IE task.

(i) IE & IE-Related Tasks

[105], [206], [207] proposed to jointly model NER and Named Entity Normalization (NEN), as these two tasks can benefit each other with enhanced entity mention features. [106], [107] proposed to transfer the knowledge learned on Word Sense Disambiguation (WSD) to ED (a subtask of EE), considering ED and WSD are two similar tasks in which they both involve identifying the classes (*i.e.*, event types or word senses) of some words in a given sentence.

(ii) Joint IE & Other Structured Prediction Tasks

Different IE and structured prediction [21] tasks can also benefit from each other considering the similar task structures and progressive task process. For example, [108], [109] tackled the joint NER and RE task, respectively leveraging the

relational graph and a copy mechanism; [110] incorporated global context in a general multi-task IE framework, *w.r.t.* the NER, RE and EE; [111], [112] jointly addressed IE and other structured prediction tasks, such as event-relation extraction, multi-span extraction, and n-ary tuple extraction.

Task Reformulation refers to formulating IE tasks into other formats which imports task-related knowledge, to capitalize on model architecture and data advantages. For example, IE can be reformulated as *machine reading comprehension (MRC)* or *text-to-structure generation* tasks. MRC-based IE identifies answer spans in context of questions, providing crucial knowledge for target tasks. Generative IE [208]–[210] employs generative LMs (GenLMs) for IE, which can reduce error propagation while increasing adaptability for IE.

(i) Reformulating IE as QA/MRC. [115] tackled low-resource EE by transferring event schema into natural questions. [116], [117], [211] studied how question generation strategies affect QA-based EE. [113], [114], [212] explored QA-based IE to efficiently encode of crucial information of classes.

(ii) Reformulating IE as Text-to-Structure Generation. [119] proposed a sequence-to-structure generation paradigm for EE. [120] framed IE as a translation task, efficiently extracting task-relevant information. [121] employed sequence-to-sequence IE with set learning to reduce structure order bias.

Prompt-Tuning PLM [213] involves inserting text pieces, *i.e.*, templates, into the input to convert a classification task into a masked language modeling problem. This enables IE approaches to benefit from LMs' pretrained knowledge, improving sample efficiency.

(i) Vanilla Prompt Tuning

Vanilla prompt tuning methods leverage the basic prompt learning framework, excelling in low-resource scenarios. [122], [123], [126]–[128], [214] respectively applied a template-based, template-free, dynamic-template-filling, demonstration-based, lightweight and contrastive prompt-based method for few-shot NER. [125], [129] leveraged the structured template and discriminative soft prompts for zero-shot RE. [124], [130], [215], [216] respectively utilized template-based, extractive, contextualized, and type-specific prompts for low-resource EE.

(ii) Augmented Prompt Tuning

Augmented prompt tuning methods enhance vanilla prompt learning by incorporating diverse knowledge, facilitating low-resource IE. [131] incorporated knowledge among labels to prompt-tuning. [134], [141] enhanced prompts with label semantics and class description. [15], [133], [135], [140] demonstrated the effectiveness of enhancing prompt-tuning with ontology, rules, semantic structures and reasoning rationales. [136], [138] introduced a unified text generation (UIE) and semantic matching (USM) framework for different IE tasks. [137] proposed a structure-aware GenLM to harness syntactic knowledge for UIE. [139] introduced UIE for any kind of schemas. [132], [217]–[219] utilized retrieval-augmented prompts to import task-specific knowledge.

B. LLM-Based Methods

Comparing with traditional PLMs, LLMs possess more powerful pretrained abilities, allowing for more complex prompt learning. In view of *whether the LLM is tuned (i.e., LLM initial parameters are modified)*, we categorize some representative LLM-based methods into two general paradigms, as shown in Figure 1.

1) Direct Inference without Tuning: This kind of methods typically involve ways to leverage LLMs without extensive additional training. They can boost low-resource IE by leveraging the inherent capabilities of LLMs to understand and process contexts, further obtaining valuable insights from scarce data, thereby reducing the requirement of fine-tuning.

Instruction Prompting involves giving the LLM explicit instructions (without demonstrations) to perform a specific task. For low-resource IE, instruction prompting can be effective as it allows the model to execute tasks using its pre-existing knowledge and language understanding.

As instruction prompting does not require demonstrations, it is naturally suitable for zero-shot [143], [147], [148], [220] and cross-domain [146] IE tasks. [142] investigated zero-shot ED task and found that ChatGPT is competitive in simple scenarios, but struggles in more complex and long-tail scenarios. [144] used global constraints with prompting for zero-shot EE, demonstrating adaptability to any other datasets. [145], [221] observed that ChatGPT is advantageous in zero-shot NER and RE tasks with instruction prompting.

Code Prompting involves presenting the LLM with snippets of code (or code-like instructions [154]) to guide its responses. This kind of methods can be particularly useful in low-resource IE tasks that involve structured output, as the code implies the schema of the specific task.

[149]–[151] tackled low-resource IE tasks with code prompting on Code-LLMs, such as CODEX [222], demonstrating the effectiveness of code-style prompts. [153] proposed a universal retrieval-augmented code generation framework. Moreover, code prompting can also be applied to multimodal IE tasks [152].

In-Context Learning (ICL) utilizes the ability of LLMs to learn from the context provided in prompts. The model uses few relevant examples (demonstrations) to "understand" the specific IE task and then applies this understanding to new data, particularly useful in low-resource scenarios.

Recent research leveraged ICL for low-resource NER [155], [160], [162], [167], [168], RE [156], [157], [161], [223], joint IE [158], [159], [163], [164] and OpenIE [165], [166]. *The key challenges of ICL for IE are: (1) input prompting fails to thoroughly express intricate IE tasks, and (2) aligning input and labels is not effective enough. To tackle these issues, synthesizing data with LLMs [224], [225] and inputting more task-specific prompts are promising.*

2) Model Specialization with Tuning: This kind of methods enhance low-resource IE by tailoring the model capabilities to specific tasks, and can be divided into prompt-tuning and fine-tuning. Generally, prompt-tuning is particularly valuable for its efficiency and minimal data requirements. Fine-tuning,

while more resource-intensive, offers deeper customization and potentially better performance.

Prompt-Tuning LLM involves keeping the LLM weights fixed and only tuning a small set of parameters associated with the prompts. In low-resource IE tasks, prompt-tuning allows for adapting the model to specific tasks or domains with minimal data, where prompts act as a guide.

[173] introduced a benchmark of diverse IE tasks with expert-written instructions, and proposed a unified IE framework, InstructUIE, with instruction tuning on FlanT5-11B [178]. [171] proposed InstructIE dataset and finetune LLaMA-7B [179] with instruction-following capability on IE. [170] hypothesized that instruction-tuning has been unable to elicit strong RE capabilities in LLMs by aligning RE with QA. [169] pretrained GLM-10B [180] on a collection of task-agnostic corpora to generate structures from text. [172] explored targeted distillation with mission-focused instruction tuning to train student models on NER task.

Fine-Tuning LLM involves adjusting the weights of the LLM on a smaller and task-specific dataset. This approach is more data-intensive than prompt-tuning but expectantly leads to performance gains.

Actually, this kind of methods are currently underdeveloped for low-resource IE tasks due to the limit of sufficient computing resources.

IV. EMPIRICAL STUDY ON TECHNICAL SOLUTIONS

A. Setup

Datasets. We adopt some widely used datasets for each subtask of low-resource IE. *NER*: we use CoNLL03 [227]; OntoNotes5.0 [228]; FewNERD [229]. *RE*: we adopt TACREV [230]; NYT [231]; FewRel [232]. *ED* of EE: we use ACE05 [233]; MAVEN [234]; FewEvent [71]. *EAE* (event argument extraction) of EE: we adopt ACE05 [233]; RAMS [235]; WikiEvents [236].

We follow [159] to construct few-shot datasets in this paper. In all experiments, we utilize micro F1 score following previous methods like [136] to evaluate performance. We release the code and data on GitHub[4].

Models. We conduct empirical study on LLM-based methods comparing with previous SOTA approaches. For prompt-/fine-tuning LLM, we adopt three typical models: *InstructUIE* [173] with prompt-tuning on FlanT5-11B [178]; *KnowLM*[5] which fine-tuned LLaMA [179] with LoRA [237]; and *fine-tuning ChatGPT*[6]. For ICL with LLM, we select three widely used LLMs: *Vicuna(1.5)*[7]; *ChatGPT*[3]; *GPT-4* [182].

B. Implementation Details

We test *InstructUIE* and *KnowLM* following their guidelines. In addition to the default parameters, for *InstructUIE*, we respectively set the maximum length for source, target and generation to 512, 50, 50, and set maximum instance

---

[4]https://github.com/mayubo2333/LLM_project.
[5]https://github.com/zjunlp/KnowLM.
[6]https://platform.openai.com/docs/guides/fine-tuning.
[7]https://github.com/lm-sys/FastChat.

Task	Dataset	Previous SOTA			Prompt-Tuning/Fine-Tuning LLM			In-Context Learning with LLM			
		full	0-shot all-way | 10-way	5-shot all-way | 10-way	InstructUIE 0-shot test all-way	KnowLM 0-shot test all-way	FT ChatGPT 0-shot test all-way	Vicuna 5-shot ICL all-way	Vicuna 1.5 5-shot ICL all-way	ChatGPT 5-shot ICL all-way	GPT-4 5-shot ICL all-way
NER	CoNLL03	94.60	▷74.99 | -	83.25 | -	**92.94**	92.65	75.26	34.00	44.40	77.80	**84.89**
	OntoNotes5.0	92.30	-	59.70 | -	**90.19**	80.40	55.34	22.38	26.38	59.39	**68.71**
	FewNERD	70.90	-	59.41 | 79.00	35.41	**71.47**	57.22	23.30	27.43	56.03	**63.08**
RE	NYT	93.50	-	-	90.47	**91.82**	50.42	8.34	13.31	21.75	**43.37**
	TACREV	85.80	▷59.40 | -	47.12 | -	3.48	35.44	**60.33**	16.67	19.80	46.16	**66.46**
	FewRel	-	-	- | 84.20	39.55	68.76	**80.00**	45.60	42.53	61.80	**78.80**
ED	ACE05	83.65	51.20 | 54.50	**55.61** | 64.80	**77.13**	51.99	49.50	13.81	20.63	38.75	53.94
	MAVEN	79.09	*59.90 | 36.86◇	**64.80** | 93.06†	-	33.19	38.53	7.22	9.49	25.29	33.39
	FewEvent	96.58	**58.14** | 68.37	‡**60.67** | 93.18	28.80	43.27	42.80	21.82	25.23	35.36	46.44
EAE	ACE05	73.50	▷31.20 | 31.40	45.90 | 42.70	**72.94**	35.62	23.60 / 47.35	4.75 / 24.89	9.05 / 30.96	25.17 / 52.79	39.82 / **58.54**
	RAMS	59.66	-	**54.08** | -	-	6.08	**49.08**	26.78	33.55	44.16	51.91
	WikiEvents	70.08	-	-	7.58	5.10	**51.28**	20.00	21.37	38.20	**45.46**

TABLE I

MICRO F1 RESULTS (MARKING THE **BEST** ALL-WAY <u>**0-SHOT**</u> & <u>**5-SHOT**</u>) OF PREVIOUS SOTA AND LLM-BASED METHODS. ▷: LLM-BASED; ∗: TYPE-SPECIFIC PROMPTING [216]; ◇: PROMPT-BASED META LEARNING [82]; †: 45-WAY-5-SHOT [82]; ‡: ALL-WAY-4-SHOT [226]; r_o/r_w FOR ACE05 DENOTES THE EAE RESULT W/O OR W/ GOLDEN TRIGGERS. DUE TO SPACE LIMITS, WE LIST THE METHODS ACHIEVING PREVIOUS SOTA IN APPENDIX A.

number for each task to 200; for *KnowLM*, we use the version of knowlm-13b-base-v1.0, and respectively set the maximum length for source and new tokens to 512 and 300. For *Vicuna*, we adopt Vicuna-13B and Vicuna-1.5-13B without fine-tuning, setting the maximum input length to 1800; the batch size to 1; and both frequency_penalty and presence_penalty to 0. About the number of demonstrations, we set it to 4 for NER and ED; 8 for EAE; and 16 for RE. We run each experiment on 4 NVIDIA V100 GPUs for InstructUIE and KnowLM, and a single one for Vicuna. To save memory, we leverage the Accelerate framework and fp16 inference. For *Fine-Tuning ChatGPT*, we fine-tune gpt-turbo-3.5-0613 on each specific dataset. We split 10% samples from training set to construct valid set and train over 5 epochs with 5-shot samples, then conduct instruction prompting on fine-tuned ChatGPT. For *ChatGPT* and *GPT-4*, we call official APIs (gpt-3.5-turbo-0301/gpt-4-0314) without fine-tuning, and set the maximum input length to 3600 for all tasks. For *ICL with LLMs (Vicuna(1.5), ChatGPT, GPT-4)*, We generate output with sampling temperature as 0 (*i.e.*, greedy decoding). We unify the maximum output length to 32 for RE; 96 for NER, ED and EAE.

C. Experimental Analysis

In this section, we intend to analyze the performance of various (especially LLM-based) models, and explore the optimal selection of low-resource IE approaches. We summarize main experimental results in Table I, and elaborate on our key findings as below:

Open-source or Proprietary LLMs? We observe that prompt-/fine-tuning InstructUIE/KnowLM (*open-source LLMs*) significantly outperform fine-tuned ChatGPT (*proprietary LLMs*) on most tasks, except for under-trained datasets, *e.g.*, TACREV, MAVEN, RAMS and WikiEvents. Nevertheless, ICL with Vicuna(1.5) (*open-source LLMs*) lag far behind ChatGPT and GPT-4 (*proprietary LLMs*). Thus we conclude that *for ICL, proprietary LLMs outperform open-source LLMs; when it comes to prompt-tuning or fine-tuning, open-source LLMs outperform their proprietary counterparts.* We speculate that the possible reasons are: (1) The fundamental abilities of proprietary LLMs are much better (larger parameters, more abundant training data and effective training strategies), so that they understand the task from in-context demonstrations more effectively; (2) Open-source LLMs are much more lightweight than proprietary LLMs, contributing to thoroughly training and adapting to a specific task. While for proprietary LLMs like ChatGPT, the limited flexibility of fine-tuning APIs makes it difficult to obtain a well-trained task-specific model.

Tuning or ICL? Regarding tuning InstructUIE and KnowLM, we find they achieve satisfying performance on well-trained tasks (*e.g.*, NER and RE) but show limited generalization abilities on under-trained tasks (*e.g.*, ED and EAE). For Chat-GPT, we observe that fine-tuning leads to generally consistent improvement on RE, ED and EAE tasks than ICL, especially on datasets which ICL with ChatGPT struggles, *e.g.*, NYT, TACREV, ACE05 and MAVEN. However, even fine-tuned ChatGPT only shows similar performance with GPT-4 and performs worse than well-tuned SOTA small PLMs and open-resource LLMs. The possible reason is that for ICL with not-fine-tuned LLMs, where the instructions of LLMs are under alignment with the specific IE task [170], it could be difficult to fully leverage the power of LLMs. As prompt-/fine-tuning enables adapting the LLM to specific tasks by adjusting its inputs or training it on specific data, we can infer that *tuning performs generally best in low-resource IE, and more thoroughly training contributes to better performance.*

Simple v.s. Complex Tasks. Across different IE tasks, we observe that LLMs are more proficient on NER and EAE (event class given) tasks than ED task. Given that ED is a more complex task (which requires instantiating abstract event classes on words in the context), we speculate that prompt-tuned LLMs are not well-acquainted with this task, and unable to fully understand the complex task through instructions and demonstrations. Furthermore, we find that the performance

Task	Dataset	Summary
Low-res NER	Few-NERD [229]	human-annotated; 8 super- and 66 sub- entity types; 188,238 sentences from Wikipedia with 4,601,160 words.
Low-res RE	FewRel [232]	human-annotated; 100 relations; 70,000 sentences derived from Wikipedia with 124,577 unique tokens in total.
	FewRel2.0 [238]	more challenging than FewRel, considering biomedical domain and NOTA setting.
	Entail-RE [69]	human-annotated; 80 relation types; 220k instances from Wikipedia; 19 relation entailment pairs.
Low-res EE	FewEvent [71]	human- & machine- annotated; 19 super- and 100 sub- event types; 70,852 instances from Wikipedia.
	Causal-EE [69]	human-annotated; 80 event types; 3.7k instances from Wikipedia; 112 event causality pairs.
	OntoEvent [68]	human- & machine- annotated; 13 super- and 100 sub- event types; 60,546 instances from Wikipedia; 3,804 event correlation pairs.

TABLE II
A SUMMARY OF SOME PUBLICLY-RELEASED LOW-RESOURCE IE BENCHMARKS.

gaps between small SOTA PLMs and LLMs are larger on the heterogeneous dataset (*e.g.*, NYT, MAVEN, WikiEvents), suggesting that LLMs have difficulty in understanding the task containing numerous labels. The possible reason is related to low relevance between fine-grained labels and limited prompting input. Thus we infer that *LLMs struggle on more complex and fine-grained tasks.*

V. BENCHMARKS AND APPLICATIONS

A. Benchmarks

In this section, we briefly introduce some public benchmarks specially established for low-resource IE, as shown in Table II. Besides, there are also some low-resource IE methods [115] directly sampling part of data from public IE benchmarks in general scenarios.

Low-resource NER. Few-NERD [229] is designed for *few-shot NER*, where each fine-grained entity type has sufficient examples for few-shot learning, including Few-NERD (INTRA) and Few-NERD (INTER). All entities in different sets of Few-NERD (INTRA) belong to different types, in contrast with Few-NERD (INTER).

Low-resource RE. FewRel [232], designed for *few-shot RE*, is a large-scale supervised dataset. FewRel2.0 [238] is derived from FewRel, and used for more challenging tasks: (1) adapt to a new domain with only a handful of instances, and (2) detect none-of-the-above (NOTA) relations. Entail-RE [69], designed for *low-resource RE*, is augmented with relation entailment annotations. Wiki-ZSL [56]

is proposed for *zero-shot RE*. More low-resource RE studies are shown in FewREBench [239].

Low-resource EE. FewEvent [71] designed for *few-shot ED*, is derived from widely-used EE datasets and data augmentation. Causal-EE [69] and OntoEvent [68], designed for *low-resource ED*, is augmented with event-relations. FewDocAE [240] is used for *few-shot document-level EAE*. More EE studies are shown in TEXTEE [241].

B. Applications

We propose some promising applications of low-resource IE, and summarize them into two aspects.

(1) *Domain-specific Applications* [3]. With the continuous emergence of knowledge from various domains, such as health care, natural science, and finance, it's challenging to annotate large-scale structured knowledge. Especially for specific domains involving data privacy, where sufficient labeled data is not available, low-resource IE is naturally crucial.

(2) *Knowledge-intensive Tasks* [5]. Low-resource IE has the potential to enhance knowledge-intensive tasks such as KG-based QA, fact verification, text-to-structure generation, and commonsense reasoning. Intuitively, those tasks require background KG for reasoning; however, collecting high-quality KGs is cumbersome and time-consuming. Low-resource IE can efficiently collect diverse knowledge across resources and simultaneously enhance performance.

VI. COMPARISON, DISCUSSION AND OUTLOOK

Comparison of Traditional Methods. From the methodology perspective, (1) *Exploiting higher-resource data*, *essentially to enrich the data features in low-resource scenarios*, serves as a general solution, which can be applied to widespread backbone models. (2) *Developing stronger data-efficient models*, *essentially to constrain the hypothesis space more precisely*, is a fundamental advancement to enable models equipped with specific low-resource learning abilities, typically through leveraging or transferring "modeledge" (*e.g.*, implicit knowledge captured in model weights) in PLMs. (3) *Optimizing data & models together*, *essentially to pursue the best strategy for training on representative data and searching the optimal hypothesis space*, represents a synthesis of the aforementioned two paradigms.

From the task perspective, low-resource NER and RE are two typical tasks that have been well investigated. Since their schema of NER and RE are relatively simple, they can normally obtain satisfactory performance, especially with the utilization of traditional PLMs and recent LLMs. However, low-resource EE still suffers from poor generalization due to the complicated task schema. Thus, it remains challenging but promising to resolve the EE issues of event interpretation and representation in low-resource scenarios.

Discussion on LLM-Based Methods. Comparing to traditional PLMs, LLMs equipped with more powerful pretrained abilities enable more complex prompt learning, and can be given more complicated instructions. Similar to traditional PLMs, LLMs with limited input capacity also struggle to

tackle IE tasks with the intricate schema. Through the empirical study in §IV-C, we deduce that *tuning open-resource LLMs and ICL with GPT family is promising in general, and the optimal LLM-based technical solution for low-resource IE can be task-dependent.*

Furthermore, [159] demonstrated that LLMs are generally not effective few-shot information extractors, but excel at reranking challenging samples. Besides, LLMs have potential to serve as data creators [242] for low-resource IE. In addition, as LLMs allow for multi-agent interaction [243], [244], tool use [221], [245], *etc*, it's promising to achieve low-resource IE through collaboration between different models, *e.g.*, large and small LM cooperation.

Future Directions. Despite many models were proposed as surveyed, potential directions remain:

- *Utilizing more informative knowledge* [246]–[249]. Advanced sample selection techniques, like active learning, sample reranking [159], [250] and data synthesizing [225], [242] with LLMs, are promising to efficiently identify and utilize informative samples, reducing large-scale annotation requirements.
- *Focusing on practical low-resource settings and applications.* Current low-resource IE models are mostly research-oriented which is unrealistic in real world settings. It is essential to explore practical low-resource IE and applications, especially involving out-of-distribution (OOD) data, establishing benchmarks and evaluation metrics (*e.g.*, low-computation/memory cost).
- *Equipping models with adaptable inference.* We assume that low-resource IE should be robust to domain shifts, rather than being restricted to a specific domain. Thus, exploring lifelong low-resource IE and OpenIE [165], [166], [251] then schema induction [252] are promising.
- *Exploring robust, faithful and interpretable IE.* To help LMs exactly understand IE targets and rationales [15] can also be effective, especially with utilization of LLMs. Additionally, incorporating human feedback in training can further enhance accuracy, fairness, and reduce risks of biased or discriminatory IE.

VII. Conclusion

In this paper, we present a literature review on low-resource IE methodologies from *traditional* and *LLM-based* perspectives, systematically categorizing **traditional** methods into three paradigms: (1) exploiting *higher*-resource data, (2) developing *stronger* data-efficient models, and (3) optimizing data and models *together*; and **LLM-based** methods into two paradigms: (1) direct inference without tuning, and (2) model specialization with tuning. We also summarize widely used benchmarks and suggest some applications. Furthermore, we compare traditional paradigms, discuss low-resource IE with LLMs, and provide some insights into future work. Our survey aims to assist researchers in understanding this field and inspire innovative algorithms, while guiding industry practitioners in selecting appropriate technical solutions for real-world applications.

References

[1] J. R. Cowie and W. G. Lehnert, "Information extraction," *Commun. ACM*, vol. 39, no. 1, pp. 80–91, 1996. [Online]. Available: https://doi.org/10.1145/234173.234209

[2] R. Grishman, "Information extraction," *IEEE Intell. Syst.*, vol. 30, no. 5, pp. 8–15, 2015. [Online]. Available: https://doi.org/10.1109/MIS.2015.68

[3] Y. Wang, L. Wang, M. Rastegar-Mojarad, S. Moon, F. Shen, N. Afzal, S. Liu, Y. Zeng, S. Mehrabi, S. Sohn, and H. Liu, "Clinical information extraction applications: A literature review," *J. Biomed. Informatics*, vol. 77, pp. 34–49, 2018. [Online]. Available: https://doi.org/10.1016/j.jbi.2017.11.011

[4] S. Deng, C. Wang, Z. Li, N. Zhang, Z. Dai, H. Chen, F. Xiong, M. Yan, Q. Chen, M. Chen, J. Chen, J. Z. Pan, B. Hooi, and H. Chen, "Construction and applications of billion-scale pre-trained multimodal business knowledge graph," in *ICDE*. IEEE, 2023, pp. 2988–3002. [Online]. Available: https://doi.org/10.1109/ICDE55515.2023.00229

[5] S. Ji, S. Pan, E. Cambria, P. Marttinen, and P. S. Yu, "A survey on knowledge graphs: Representation, acquisition, and applications," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 33, no. 2, pp. 494–514, 2022. [Online]. Available: https://doi.org/10.1109/TNNLS.2021.3070843

[6] S. Deng, N. Zhang, W. Zhang, J. Chen, J. Z. Pan, and H. Chen, "Knowledge-driven stock trend prediction and explanation via temporal convolutional network," in *WWW (Companion Volume)*. ACM, 2019, pp. 678–685. [Online]. Available: https://doi.org/10.1145/3308560.3317701

[7] J. Li, A. Sun, J. Han, and C. Li, "A survey on deep learning for named entity recognition," *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 1, pp. 50–70, 2022. [Online]. Available: https://doi.org/10.1109/TKDE.2020.2981314

[8] H. Wang, K. Qin, R. Y. Zakari, G. Lu, and J. Yin, "Deep neural network-based relation extraction: an overview," *Neural Comput. Appl.*, vol. 34, no. 6, pp. 4781–4801, 2022. [Online]. Available: https://doi.org/10.1007/s00521-021-06667-3

[9] Q. Li, J. Li, J. Sheng, S. Cui, J. Wu, Y. Hei, H. Peng, S. Guo, L. Wang, A. Beheshti, and P. S. Yu, "A survey on deep learning event extraction: Approaches and applications," *IEEE Trans. Neural Networks Learn. Syst.*, pp. 1–21, 2022. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9927311

[10] W. X. Zhao, K. Zhou, J. Li, T. Tang, X. Wang, Y. Hou, Y. Min, B. Zhang, J. Zhang, Z. Dong, Y. Du, C. Yang, Y. Chen, Z. Chen, J. Jiang, R. Ren, Y. Li, X. Tang, Z. Liu, P. Liu, J. Nie, and J. Wen, "A survey of large language models," *CoRR*, vol. abs/2303.18223, 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2303.18223

[11] Z. Zhang, H. Li, Z. Zhang, Y. Qin, X. Wang, and W. Zhu, "Large graph models: A perspective," *CoRR*, vol. abs/2308.14522, 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2308.14522

[12] J. Huang, C. Li, K. Subudhi, D. Jose, S. Balakrishnan, W. Chen, B. Peng, J. Gao, and J. Han, "Few-shot named entity recognition: An empirical baseline study," in *EMNLP (1)*. Association for Computational Linguistics, 2021, pp. 10408–10423. [Online]. Available: https://doi.org/10.18653/v1/2021.emnlp-main.813

[13] V. Moscato, M. Postiglione, and G. Sperlí, "Few-shot named entity recognition: Definition, taxonomy and research directions," *ACM Trans. Intell. Syst. Technol.*, vol. 14, no. 5, pp. 94:1–94:46, 2023. [Online]. Available: https://doi.org/10.1145/3609483

[14] O. Sabo, Y. Elazar, Y. Goldberg, and I. Dagan, "Revisiting few-shot relation classification: Evaluation data and classification schemes," *Trans. Assoc. Comput. Linguistics*, vol. 9, pp. 691–706, 2021. [Online]. Available: https://doi.org/10.1162/tacl_a_00392

[15] S. Wadhwa, S. Amir, and B. C. Wallace, "Revisiting relation extraction in the era of large language models," in *ACL (1)*. Association for Computational Linguistics, 2023, pp. 15 566–15 589. [Online]. Available: https://aclanthology.org/2023.acl-long.868/

[16] Y. Ma, Z. Wang, Y. Cao, and A. Sun, "Few-shot event detection: An empirical study and a unified view," in *ACL (1)*. Association for Computational Linguistics, 2023, pp. 11 211–11 236. [Online]. Available: https://aclanthology.org/2023.acl-long.628/

[17] M. A. Hedderich, L. Lange, H. Adel, J. Strötgen, and D. Klakow, "A survey on recent approaches for natural language processing in low-resource scenarios," in *NAACL-HLT*. Association for Computational Linguistics, 2021, pp. 2545–2568. [Online]. Available: https://doi.org/10.18653/v1/2021.naacl-main.201

[18] M. Treviso, J.-U. Lee, T. Ji, B. van Aken, Q. Cao, M. R. Ciosici, M. Hassid, K. Heafield, S. Hooker, C. Raffel, P. H. Martins, A. F. T. Martins, J. Z. Forde, P. Milder, E. Simpson, N. Slonim, J. Dodge, E. Strubell, N. Balasubramanian, L. Derczynski, I. Gurevych, and R. Schwartz, "Efficient methods for natural language processing: A survey," *Trans. Assoc. Comput. Linguistics*, vol. 11, pp. 826–860, 2023. [Online]. Available: https://aclanthology.org/2023.tacl-1.48

[19] N. Zhang, S. Deng, Z. Bi, H. Yu, J. Yang, M. Chen, F. Huang, W. Zhang, and H. Chen, "Openue: An open toolkit of universal extraction from text," in *EMNLP (Demos)*. Association for Computational Linguistics, 2020, pp. 1–8. [Online]. Available: https://doi.org/10.18653/v1/2020.emnlp-demos.1

[20] N. Zhang, X. Xu, L. Tao, H. Yu, H. Ye, S. Qiao, X. Xie, X. Chen, Z. Li, and L. Li, "Deepke: A deep learning based knowledge extraction toolkit for knowledge base population," in *EMNLP (Demos)*. Association for Computational Linguistics, 2022, pp. 98–108. [Online]. Available: https://doi.org/10.18653/v1/2022.emnlp-demos.10

[21] B. Taskar, V. Chatalbashev, D. Koller, and C. Guestrin, "Learning structured prediction models: a large margin approach," in *ICML*, ser. ACM International Conference Proceeding Series, vol. 119. ACM, 2005, pp. 896–903. [Online]. Available: https://doi.org/10.1145/1102351.1102464

[22] J. P. C. Chiu and E. Nichols, "Named entity recognition with bidirectional lstm-cnns," *Trans. Assoc. Comput. Linguistics*, vol. 4, pp. 357–370, 2016. [Online]. Available: https://transacl.org/ojs/index.php/tacl/article/view/792

[23] D. Zeng, K. Liu, Y. Chen, and J. Zhao, "Distant supervision for relation extraction via piecewise convolutional neural networks," in *EMNLP*. Association for Computational Linguistics, 2015, pp. 1753–1762. [Online]. Available: https://doi.org/10.18653/v1/d15-1203

[24] Y. Chen, L. Xu, K. Liu, D. Zeng, and J. Zhao, "Event extraction via dynamic multi-pooling convolutional neural networks," in *ACL (1)*. The Association for Computer Linguistics, 2015, pp. 167–176. [Online]. Available: https://doi.org/10.3115/v1/p15-1017

[25] Y. Zhang, B. Kang, B. Hooi, S. Yan, and J. Feng, "Deep long-tailed learning: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 9, pp. 10 795–10 816, 2023. [Online]. Available: https://doi.org/10.1109/TPAMI.2023.3268118

[26] Y. Wang, Q. Yao, J. T. Kwok, and L. M. Ni, "Generalizing from a few examples: A survey on few-shot learning," *ACM Comput. Surv.*, vol. 53, no. 3, pp. 63:1–63:34, 2021. [Online]. Available: https://doi.org/10.1145/3386252

[27] J. Lu, P. Gong, J. Ye, J. Zhang, and C. Zhang, "A survey on machine learning from few samples," *Pattern Recognit.*, vol. 139, p. 109480, 2023. [Online]. Available: https://doi.org/10.1016/j.patcog.2023.109480

[28] J. Snell, K. Swersky, and R. S. Zemel, "Prototypical networks for few-shot learning," in *NIPS*, 2017, pp. 4077–4087. [Online]. Available: https://proceedings.neurips.cc/paper/2017/hash/cb8da6767461f2812ae4290eac7cbc42-Abstract.html

[29] W. Wang, V. W. Zheng, H. Yu, and C. Miao, "A survey of zero-shot learning: Settings, methods, and applications," *ACM Trans. Intell. Syst. Technol.*, vol. 10, no. 2, pp. 13:1–13:37, 2019. [Online]. Available: https://doi.org/10.1145/3293318

[30] F. Pourpanah, M. Abdar, Y. Luo, X. Zhou, R. Wang, C. P. Lim, X. Wang, and Q. M. J. Wu, "A review of generalized zero-shot learning methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 4, pp. 4051–4070, 2023. [Online]. Available: https://doi.org/10.1109/TPAMI.2022.3191696

[31] J. Shang, L. Liu, X. Gu, X. Ren, T. Ren, and J. Han, "Learning named entity tagger using domain-specific dictionary," in *EMNLP*. Association for Computational Linguistics, 2018, pp. 2054–2064. [Online]. Available: https://doi.org/10.18653/v1/d18-1230

[32] Y. Chen, S. Liu, X. Zhang, K. Liu, and J. Zhao, "Automatically labeled data generation for large scale event extraction," in *ACL (1)*. Association for Computational Linguistics, 2017, pp. 409–419. [Online]. Available: https://doi.org/10.18653/v1/P17-1038

[33] J. Chen, Z. Wang, R. Tian, Z. Yang, and D. Yang, "Local additivity based data augmentation for semi-supervised NER," in *EMNLP (1)*. Association for Computational Linguistics, 2020, pp. 1241–1251. [Online]. Available: https://doi.org/10.18653/v1/2020.emnlp-main.95

[34] C. Liang, Y. Yu, H. Jiang, S. Er, R. Wang, T. Zhao, and C. Zhang, "BOND: bert-assisted open-domain named entity recognition with distant supervision," in *KDD*. ACM, 2020, pp. 1054–1064. [Online]. Available: https://doi.org/10.1145/3394486.3403149

[35] P. Qin, W. Xu, and W. Y. Wang, "Robust distant supervision relation extraction via deep reinforcement learning," in *ACL (1)*. Association for Computational Linguistics, 2018, pp. 2137–2147. [Online]. Available: https://aclanthology.org/P18-1199/

[36] X. Wang, X. Han, Z. Liu, M. Sun, and P. Li, "Adversarial training for weakly supervised event detection," in *NAACL-HLT (1)*. Association for Computational Linguistics, 2019, pp. 998–1008. [Online]. Available: https://doi.org/10.18653/v1/n19-1105

[37] Z. Wang, R. Wen, X. Chen, S. Huang, N. Zhang, and Y. Zheng, "Finding influential instances for distantly supervised relation extraction," in *COLING*. International Committee on Computational Linguistics, 2022, pp. 2639–2650. [Online]. Available: https://aclanthology.org/2022.coling-1.233

[38] Y. Zheng, A. Hao, and A. T. Luu, "Jointprop: Joint semi-supervised learning for entity and relation extraction with heterogeneous graph-based propagation," in *ACL (1)*. Association for Computational Linguistics, 2023, pp. 14 541–14 555. [Online]. Available: https://doi.org/10.18653/v1/2023.acl-long.813

[39] L. Xu, L. Bing, and W. Lu, "Better sampling of negatives for distantly supervised named entity recognition," in *ACL (Findings)*, 2023, pp. 4874–4882. [Online]. Available: https://aclanthology.org/2023.findings-acl.300

[40] D. Lu, L. Neves, V. Carvalho, N. Zhang, and H. Ji, "Visual attention model for name tagging in multimodal social media," in *ACL (1)*. Association for Computational Linguistics, 2018, pp. 1990–1999. [Online]. Available: https://aclanthology.org/P18-1185/

[41] M. Tong, S. Wang, Y. Cao, B. Xu, J. Li, L. Hou, and T. Chua, "Image enhanced event detection in news articles," in *AAAI*. AAAI Press, 2020, pp. 9040–9047. [Online]. Available: https://ojs.aaai.org/index.php/AAAI/article/view/6437

[42] M. Li, A. Zareian, Q. Zeng, S. Whitehead, D. Lu, H. Ji, and S. Chang, "Cross-media structured common space for multimedia event extraction," in *ACL*. Association for Computational Linguistics, 2020, pp. 2557–2568. [Online]. Available: https://doi.org/10.18653/v1/2020.acl-main.230

[43] B. Chen, X. Lin, C. Thomas, M. Li, S. Yoshida, L. Chum, H. Ji, and S. Chang, "Joint multimedia event extraction from video and article," in *EMNLP (Findings)*. Association for Computational Linguistics, 2021, pp. 74–88. [Online]. Available: https://doi.org/10.18653/v1/2021.findings-emnlp.8

[44] C. Zheng, J. Feng, Z. Fu, Y. Cai, Q. Li, and T. Wang, "Multimodal relation extraction with efficient graph alignment," in *ACM Multimedia*. ACM, 2021, pp. 5298–5306. [Online]. Available: https://doi.org/10.1145/3474085.3476968

[45] X. Chen, N. Zhang, L. Li, S. Deng, C. Tan, C. Xu, F. Huang, L. Si, and H. Chen, "Hybrid transformer with multi-level fusion for multimodal knowledge graph completion," in *SIGIR*. ACM, 2022, pp. 904–915. [Online]. Available: https://doi.org/10.1145/3477495.3531992

[46] X. Chen, N. Zhang, L. Li, Y. Yao, S. Deng, C. Tan, F. Huang, L. Si, and H. Chen, "Good visual guidance makes A better extractor: Hierarchical visual prefix for multimodal entity and relation extraction," in *NAACL (Findings)*, 2022. [Online]. Available: https://aclanthology.org/2022.findings-naacl.121

[47] X. Feng, X. Feng, B. Qin, Z. Feng, and T. Liu, "Improving low resource named entity recognition using cross-lingual knowledge transfer," in *IJCAI*. ijcai.org, 2018, pp. 4071–4077. [Online]. Available: https://doi.org/10.24963/ijcai.2018/566

[48] J. O. Alabi, D. I. Adelani, M. Mosbach, and D. Klakow, "Adapting pre-trained language models to african languages via multilingual adaptive fine-tuning," in *COLING*. International Committee on Computational Linguistics, 2022, pp. 4336–4349. [Online]. Available: https://aclanthology.org/2022.coling-1.382

[49] L. Guzman-Nateras, F. Dernoncourt, and T. Nguyen, "Hybrid knowledge transfer for improved cross-lingual event detection via hierarchical sample selection," in *ACL (1)*. Association for Computational Linguistics, 2023, pp. 5414–5427. [Online]. Available: https://doi.org/10.18653/v1/2023.acl-long.296

[50] Y. Lin, Z. Liu, and M. Sun, "Neural relation extraction with multi-lingual attention," in *ACL (1)*. Association for Computational Linguistics, 2017, pp. 34–43. [Online]. Available: https://doi.org/10.18653/v1/P17-1004

[51] J. Liu, Y. Chen, K. Liu, and J. Zhao, "Event detection via gated multilingual attention mechanism," in *AAAI*. AAAI Press, 2018, pp. 4865–4872. [Online]. Available: https://www.aaai.org/ocs/index.php/AAAI/AAAI18/paper/view/16371

[52] Y. Huang, W. Liu, X. Zhang, J. Lang, T. Gong, and C. Li, "PRAM: an end-to-end prototype-based representation alignment model for zero-resource cross-lingual named entity recognition," in *ACL (Findings)*. Association for Computational Linguistics, 2023, pp. 3220–3233. [Online]. Available: https://doi.org/10.18653/v1/2023.findings-acl.201

[53] S. Fincke, S. Agarwal, S. Miller, and E. Boschee, "Language model priming for cross-lingual event extraction," in *AAAI*. AAAI Press, 2022, pp. 10 627–10 635. [Online]. Available: https://doi.org/10.1609/aaai.v36i10.21307

[54] C. Nguyen, L. N. Van, and T. Nguyen, "Retrieving relevant context to align representations for cross-lingual event detection," in *ACL (Findings)*. Association for Computational Linguistics, 2023, pp. 2157–2170. [Online]. Available: https://doi.org/10.18653/v1/2023.findings-acl.135

[55] M. Tong, B. Xu, S. Wang, Y. Cao, L. Hou, J. Li, and J. Xie, "Improving event detection via open-domain trigger knowledge," in *ACL*. Association for Computational Linguistics, 2020, pp. 5887–5897. [Online]. Available: https://doi.org/10.18653/v1/2020.acl-main.522

[56] C. Chen and C. Li, "ZS-BERT: towards zero-shot relation extraction with attribute representation learning," in *NAACL-HLT*. Association for Computational Linguistics, 2021, pp. 3470–3479. [Online]. Available: https://doi.org/10.18653/v1/2021.naacl-main.272

[57] X. Hu, C. Zhang, Y. Yang, X. Li, L. Lin, L. Wen, and P. S. Yu, "Gradient imitation reinforcement learning for low resource relation extraction," in *EMNLP (1)*. Association for Computational Linguistics, 2021, pp. 2737–2746. [Online]. Available: https://doi.org/10.18653/v1/2021.emnlp-main.216

[58] J. Gao, C. Yu, W. Wang, H. Zhao, and R. Xu, "Mask-then-fill: A flexible and effective data augmentation framework for event extraction," in *EMNLP (Findings)*, 2022, pp. 4537–4544. [Online]. Available: https://aclanthology.org/2022.findings-emnlp.332

[59] R. Zhou, X. Li, R. He, L. Bing, E. Cambria, L. Si, and C. Miao, "MELM: data augmentation with masked entity language modeling for low-resource NER," in *ACL (1)*, 2022, pp. 2251–2262. [Online]. Available: https://doi.org/10.18653/v1/2022.acl-long.160

[60] J. Cai, S. Huang, Y. Jiang, Z. Tan, P. Xie, and K. Tu, "Improving low-resource named entity recognition with graph propagated data augmentation," in *ACL (Short)*. Association for Computational Linguistics, 2023, pp. 110—-118. [Online]. Available: https://aclanthology.org/2023.acl-short.11.pdf

[61] X. Hu, Y. Jiang, A. Liu, Z. Huang, P. Xie, F. Huang, L. Wen, and P. S. Yu, "Entity-to-text based data augmentation for various named entity recognition tasks," in *ACL (Findings)*. Association for Computational Linguistics, 2023, pp. 9072–9087. [Online]. Available: https://doi.org/10.18653/v1/2023.findings-acl.578

[62] S. Ghosh, U. Tyagi, M. Suri, S. Kumar, R. S, and D. Manocha, "ACLM: A selective-denoising based generative data augmentation approach for low-resource complex NER," in *ACL (1)*. Association for Computational Linguistics, 2023, pp. 104–125. [Online]. Available: https://aclanthology.org/2023.acl-long.8/

[63] B. Xu, Q. Wang, Y. Lyu, D. Dai, Y. Zhang, and Z. Mao, "S2ynre: Two-stage self-training with synthetic data for low-resource relation extraction," in *ACL (1)*. Association for Computational Linguistics, 2023, pp. 8186–8207. [Online]. Available: https://doi.org/10.18653/v1/2023.acl-long.455

[64] J. Li, R. Wang, N. Zhang, W. Zhang, F. Yang, and H. Chen, "Logic-guided semantic representation learning for zero-shot relation classification," in *COLING*. International Committee on Computational Linguistics, 2020, pp. 2967–2978. [Online]. Available: https://doi.org/10.18653/v1/2020.coling-main.265

[65] W. Zhou, H. Lin, B. Y. Lin, Z. Wang, J. Du, L. Neves, and X. Ren, "NERO: A neural rule grounding framework for label-efficient relation extraction," in *WWW*. ACM / IW3C2, 2020, pp. 2166–2176. [Online]. Available: https://doi.org/10.1145/3366423.3380282

[66] B. Nie, R. Ding, P. Xie, F. Huang, C. Qian, and L. Si, "Knowledge-aware named entity recognition with alleviating heterogeneity," in *AAAI*. AAAI Press, 2021, pp. 13 595–13 603. [Online]. Available: https://ojs.aaai.org/index.php/AAAI/article/view/17603

[67] M. Li, Q. Zeng, Y. Lin, K. Cho, H. Ji, J. May, N. Chambers, and C. R. Voss, "Connecting the dots: Event graph schema induction with path language modeling," in *EMNLP (1)*. Association for Computational Linguistics, 2020, pp. 684–695. [Online]. Available: https://doi.org/10.18653/v1/2020.emnlp-main.50

[68] S. Deng, N. Zhang, L. Li, C. Hui, H. Tou, M. Chen, F. Huang, and H. Chen, "Ontoed: Low-resource event detection with ontology embedding," in *ACL/IJCNLP (1)*. Association for Computational Linguistics, 2021, pp. 2828–2839. [Online]. Available: https://doi.org/10.18653/v1/2021.acl-long.220

[69] S. Deng, N. Zhang, H. Chen, C. Tan, F. Huang, C. Xu, and H. Chen, "Low-resource extraction with knowledge-aware pairwise prototype learning," *Knowl. Based Syst.*, vol. 235, p. 107584, 2022. [Online]. Available: https://doi.org/10.1016/j.knosys.2021.107584

[70] T. Gao, X. Han, Z. Liu, and M. Sun, "Hybrid attention-based prototypical networks for noisy few-shot relation classification," in *AAAI*. AAAI Press, 2019, pp. 6407–6414. [Online]. Available: https://doi.org/10.1609/aaai.v33i01.33016407

[71] S. Deng, N. Zhang, J. Kang, Y. Zhang, W. Zhang, and H. Chen, "Meta-learning with dynamic-memory-based prototypical network for few-shot event detection," in *WSDM*. ACM, 2020, pp. 151–159.

[72] J. Li, S. Shang, and L. Shao, "Metaner: Named entity recognition with meta-learning," in *WWW*, 2020, pp. 429–440. [Online]. Available: https://doi.org/10.1145/3366423.3380127

[73] N. Ding, X. Wang, Y. Fu, G. Xu, R. Wang, P. Xie, Y. Shen, F. Huang, H. Zheng, and R. Zhang, "Prototypical representation learning for relation extraction," in *ICLR*. OpenReview.net, 2021. [Online]. Available: https://openreview.net/forum?id=aCgLmfhIy_f

[74] X. Cong, S. Cui, B. Yu, T. Liu, Y. Wang, and B. Wang, "Few-shot event detection with prototypical amortized conditional random field," in *ACL/IJCNLP (Findings)*, ser. Findings of ACL, vol. ACL/IJCNLP 2021. Association for Computational Linguistics, 2021, pp. 28–40. [Online]. Available: https://doi.org/10.18653/v1/2021.findings-acl.3

[75] F. Liu, H. Lin, X. Han, B. Cao, and L. Sun, "Pre-training to match for unified low-shot relation extraction," in *ACL (1)*. Association for Computational Linguistics, 2022, pp. 5785–5795. [Online]. Available: https://doi.org/10.18653/v1/2022.acl-long.397

[76] J. Ma, M. Ballesteros, S. Doss, R. Anubhai, S. Mallya, Y. Al-Onaizan, and D. Roth, "Label semantics for few shot named entity recognition," in *ACL (Findings)*. Association for Computational Linguistics, 2022, pp. 1956–1971. [Online]. Available: https://doi.org/10.18653/v1/2022.findings-acl.155

[77] Y. Chen, Y. Zheng, and Z. Yang, "Prompt-based metric learning for few-shot NER," in *ACL (Findings)*. Association for Computational Linguistics, 2023, pp. 7199–7212. [Online]. Available: https://doi.org/10.18653/v1/2023.findings-acl.451

[78] W. Chen, L. Zhao, P. Luo, T. Xu, Y. Zheng, and E. Chen, "Heproto: A hierarchical enhancing protonet based on multi-task learning for few-shot named entity recognition," in *CIKM*. ACM, 2023, pp. 296–305. [Online]. Available: https://doi.org/10.1145/3583780.3614908

[79] S. Wang, J. Zheng, W. Chen, F. Cai, and X. Luo, "Multiple: Multilingual prompt learning for relieving semantic confusions in few-shot event detection," in *CIKM*. ACM, 2023, pp. 2676–2685. [Online]. Available: https://doi.org/10.1145/3583780.3614984

[80] B. Dong, Y. Yao, R. Xie, T. Gao, X. Han, Z. Liu, F. Lin, L. Lin, and M. Sun, "Meta-information guided meta-learning for few-shot relation classification," in *COLING*. International Committee on Computational Linguistics, 2020, pp. 1594–1605. [Online]. Available: https://doi.org/10.18653/v1/2020.coling-main.140

[81] J. Li, B. Chiu, S. Feng, and H. Wang, "Few-shot named entity recognition via meta-learning," *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 9, pp. 4245–4256, 2022. [Online]. Available: https://doi.org/10.1109/TKDE.2020.3038670

[82] Z. Yue, H. Zeng, M. Lan, H. Ji, and D. Wang, "Zero- and few-shot event detection via prompt-based meta learning," in *ACL (1)*. Association for Computational Linguistics, 2023, pp. 7928–7943. [Online]. Available: https://doi.org/10.18653/v1/2023.acl-long.440

[83] H. Wang, Z. Wang, G. P. C. Fung, and K. Wong, "MCML: A novel memory-based contrastive meta-learning method for few shot slot tagging," *CoRR*, vol. abs/2108.11635, 2021. [Online]. Available: https://arxiv.org/abs/2108.11635

[84] M. Qu, T. Gao, L. A. C. Xhonneux, and J. Tang, "Few-shot relation extraction via bayesian meta-learning on relation graphs," in *ICML*, ser. Proceedings of Machine Learning Research, vol. 119. PMLR, 2020, pp. 7867–7876. [Online]. Available: http://proceedings.mlr.press/v119/qu20a.html

[85] S. Shen, T. Wu, G. Qi, Y. Li, G. Haffari, and S. Bi, "Adaptive knowledge-enhanced bayesian meta-learning for few-shot event detection," in *ACL/IJCNLP (Findings)*, ser. Findings of ACL, vol. ACL/IJCNLP 2021. Association for Computational Linguistics, 2021, pp. 2417–2429. [Online]. Available: https://doi.org/10.18653/v1/2021.findings-acl.214

[86] L. Huang, H. Ji, K. Cho, I. Dagan, S. Riedel, and C. R. Voss, "Zero-shot transfer learning for event extraction," in *ACL (1)*. Association for Computational Linguistics, 2018, pp. 2160–2170. [Online]. Available: https://www.aclweb.org/anthology/P18-1201/

[87] N. Zhang, S. Deng, Z. Sun, G. Wang, X. Chen, W. Zhang, and H. Chen, "Long-tail relation extraction via knowledge graph embeddings and graph convolution networks," in *NAACL-HLT (1)*. Association for Computational Linguistics, 2019, pp. 3016–3025. [Online]. Available: https://doi.org/10.18653/v1/n19-1306

[88] T. Zhang, C. Xia, C. Lu, and P. S. Yu, "MZET: memory augmented zero-shot fine-grained named entity typing," in *COLING*. International Committee on Computational Linguistics, 2020, pp. 77–87. [Online]. Available: https://doi.org/10.18653/v1/2020.coling-main.7

[89] S. Gururaja, R. Dutt, T. Liao, and C. P. Rosé, "Linguistic representations for fewer-shot relation extraction across domains," in *ACL (1)*. Association for Computational Linguistics, 2023, pp. 7502–7514. [Online]. Available: https://doi.org/10.18653/v1/2023.acl-long.414

[90] N. Zhang, S. Deng, Z. Sun, J. Chen, W. Zhang, and H. Chen, "Relation adversarial network for low resource knowledge graph completion," in *WWW*. ACM, 2020, pp. 1–12. [Online]. Available: https://doi.org/10.1145/3366423.3380089

[91] J. Li, S. Feng, and B. Chiu, "Few-shot relation extraction with dual graph neural network interaction," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–13, 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10143375

[92] J. Fang, X. Wang, Z. Meng, P. Xie, F. Huang, and Y. Jiang, "MANNER: A variational memory-augmented model for cross domain few-shot named entity recognition," in *ACL (1)*. Association for Computational Linguistics, 2023, pp. 4261–4276. [Online]. Available: https://doi.org/10.18653/v1/2023.acl-long.234

[93] L. B. Soares, N. FitzGerald, J. Ling, and T. Kwiatkowski, "Matching the blanks: Distributional similarity for relation learning," in *ACL (1)*. Association for Computational Linguistics, 2019, pp. 2895–2905. [Online]. Available: https://doi.org/10.18653/v1/p19-1279

[94] S. Yang, D. Feng, L. Qiao, Z. Kan, and D. Li, "Exploring pre-trained language models for event extraction and generation," in *ACL (1)*. Association for Computational Linguistics, 2019, pp. 5284–5294. [Online]. Available: https://doi.org/10.18653/v1/p19-1522

[95] S. Wang, M. Yu, S. Chang, L. Sun, and L. Huang, "Query and extract: Refining event extraction as type-oriented binary decoding," in *ACL (Findings)*. Association for Computational Linguistics, 2022, pp. 169–182. [Online]. Available: https://doi.org/10.18653/v1/2022.findings-acl.16

[96] M. Xue, B. Yu, Z. Zhang, T. Liu, Y. Zhang, and B. Wang, "Coarse-to-fine pre-training for named entity recognition," in *EMNLP (1)*. Association for Computational Linguistics, 2020, pp. 6345–6354. [Online]. Available: https://doi.org/10.18653/v1/2020.emnlp-main.514

[97] J. Chen, Q. Liu, H. Lin, X. Han, and L. Sun, "Few-shot named entity recognition with self-describing networks," in *ACL (1)*. Association for Computational Linguistics, 2022, pp. 5711–5722. [Online]. Available: https://doi.org/10.18653/v1/2022.acl-long.392

[98] W. Zhou, S. Zhang, T. Naumann, M. Chen, and H. Poon, "Continual contrastive finetuning improves low-resource relation extraction," in *ACL (1)*. Association for Computational Linguistics, 2023, pp. 13 249–13 263. [Online]. Available: https://doi.org/10.18653/v1/2023.acl-long.739

[99] Z. Wang, X. Wang, X. Han, Y. Lin, L. Hou, Z. Liu, P. Li, J. Li, and J. Zhou, "CLEVE: contrastive pre-training for event extraction," in *ACL/IJCNLP (1)*. Association for Computational Linguistics, 2021, pp. 6283–6297. [Online]. Available: https://doi.org/10.18653/v1/2021.acl-long.491

[100] A. P. B. Veyseh, V. D. Lai, F. Dernoncourt, and T. H. Nguyen, "Unleash GPT-2 power for event detection," in *ACL/IJCNLP (1)*. Association for Computational Linguistics, 2021, pp. 6271–6282. [Online]. Available: https://doi.org/10.18653/v1/2021.acl-long.490

[101] Y. Zhou, T. Shen, X. Geng, G. Long, and D. Jiang, "Claret: Pre-training a correlation-aware context-to-event transformer for event-centric generation and classification," in *ACL (1)*. Association for Computational Linguistics, 2022, pp. 2559–2575. [Online]. Available: https://doi.org/10.18653/v1/2022.acl-long.183

[102] S. S. S. Das, H. Zhang, P. Shi, W. Yin, and R. Zhang, "Unified low-resource sequence labeling by sample-aware dynamic sparse finetuning," in *EMNLP*. Association for Computational Linguistics, 2023, pp. 6998–7010. [Online]. Available: https://aclanthology.org/2023.emnlp-main.433

[103] H. Zhang, W. Yao, and D. Yu, "Efficient zero-shot event extraction with context-definition alignment," in *EMNLP (Findings)*. Association for Computational Linguistics, 2022, pp. 7169–7179. [Online]. Available: https://doi.org/10.18653/v1/2022.findings-emnlp.531

[104] G. Dong, Z. Wang, J. Zhao, G. Zhao, D. Guo, D. Fu, T. Hui, C. Zeng, K. He, X. Li, L. Wang, X. Cui, and W. Xu, "A multi-task semantic decomposition framework with task-specific pre-training for few-shot NER," in *CIKM*. ACM, 2023, pp. 430–440. [Online]. Available: https://doi.org/10.1145/3583780.3614766

[105] B. Zhou, X. Cai, Y. Zhang, and X. Yuan, "An end-to-end progressive multi-task learning framework for medical named entity recognition and normalization," in *ACL/IJCNLP (1)*. Association for Computational Linguistics, 2021, pp. 6214–6224. [Online]. Available: https://doi.org/10.18653/v1/2021.acl-long.485

[106] W. Lu and T. H. Nguyen, "Similar but not the same - word sense disambiguation improves event detection via neural representation matching," in *EMNLP*. Association for Computational Linguistics, 2018, pp. 4822–4828. [Online]. Available: https://aclanthology.org/D18-1517/

[107] V. D. Lai, M. V. Nguyen, T. H. Nguyen, and F. Dernoncourt, "Graph learning regularization and transfer learning for few-shot event detection," in *SIGIR*. ACM, 2021, pp. 2172–2176. [Online]. Available: https://doi.org/10.1145/3404835.3463054

[108] T. Fu, P. Li, and W. Ma, "Graphrel: Modeling text as relational graphs for joint entity and relation extraction," in *ACL (1)*. Association for Computational Linguistics, 2019, pp. 1409–1418. [Online]. Available: https://doi.org/10.18653/v1/p19-1136

[109] D. Zeng, H. Zhang, and Q. Liu, "Copymtl: Copy mechanism for joint extraction of entities and relations with multi-task learning," in *AAAI*. AAAI Press, 2020, pp. 9507–9514. [Online]. Available: https://ojs.aaai.org/index.php/AAAI/article/view/6495

[110] D. Wadden, U. Wennberg, Y. Luan, and H. Hajishirzi, "Entity, relation, and event extraction with contextualized span representations," in *EMNLP/IJCNLP (1)*. Association for Computational Linguistics, 2019, pp. 5783–5788. [Online]. Available: https://doi.org/10.18653/v1/D19-1585

[111] S. Deng, S. Mao, N. Zhang, and B. Hooi, "Speech: Structured prediction with energy-based event-centric hyperspheres," in *ACL (1)*. Association for Computational Linguistics, 2023, pp. 351–363. [Online]. Available: https://aclanthology.org/2023.acl-long.21/

[112] T. Zhu, J. Ren, Z. Yu, M. Wu, G. Zhang, X. Qu, W. Chen, Z. Wang, B. Huai, and M. Zhang, "Mirror: A universal framework for various information extraction tasks," in *EMNLP*. Association for Computational Linguistics, 2023, pp. 8861–8876. [Online]. Available: https://aclanthology.org/2023.emnlp-main.548

[113] X. Li, F. Yin, Z. Sun, X. Li, A. Yuan, D. Chai, M. Zhou, and J. Li, "Entity-relation extraction as multi-turn question answering," in *ACL (1)*. Association for Computational Linguistics, 2019, pp. 1340–1350. [Online]. Available: https://doi.org/10.18653/v1/p19-1129

[114] X. Li, J. Feng, Y. Meng, Q. Han, F. Wu, and J. Li, "A unified MRC framework for named entity recognition," in *ACL*. Association for Computational Linguistics, 2020, pp. 5849–5859. [Online]. Available: https://doi.org/10.18653/v1/2020.acl-main.519

[115] J. Liu, Y. Chen, K. Liu, W. Bi, and X. Liu, "Event extraction as machine reading comprehension," in *EMNLP (1)*. Association for Computational Linguistics, 2020, pp. 1641–1651. [Online]. Available: https://doi.org/10.18653/v1/2020.emnlp-main.128

[116] X. Du and C. Cardie, "Event extraction by answering (almost) natural questions," in *EMNLP (1)*. Association for Computational Linguistics, 2020, pp. 671–683. [Online]. Available: https://doi.org/10.18653/v1/2020.emnlp-main.49

[117] D. Lu, S. Ran, J. R. Tetreault, and A. Jaimes, "Event extraction as question generation and answering," in *ACL (2)*. Association for Computational Linguistics, 2023, pp. 1666–1688. [Online]. Available: https://doi.org/10.18653/v1/2023.acl-short.143

[118] Y. Zhang, F. Meng, Y. Chen, J. Xu, and J. Zhou, "Complex question enhanced transfer learning for zero-shot joint information extraction," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, pp. 1–15, 2023. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10214665

[119] Y. Lu, H. Lin, J. Xu, X. Han, J. Tang, A. Li, L. Sun, M. Liao, and S. Chen, "Text2event: Controllable sequence-to-structure generation for end-to-end event extraction," in *ACL/IJCNLP (1)*. Association for Computational Linguistics, 2021, pp. 2795–2806. [Online]. Available: https://doi.org/10.18653/v1/2021.acl-long.217

[120] G. Paolini, B. Athiwaratkun, J. Krone, J. Ma, A. Achille, R. Anubhai, C. N. dos Santos, B. Xiang, and S. Soatto, "Structured prediction as translation between augmented natural languages," in *ICLR*. OpenReview.net, 2021. [Online]. Available: https://openreview.net/forum?id=US-TP-xnXI

[121] J. Li, Y. Zhang, B. Liang, K. Wong, and R. Xu, "Set learning for generative information extraction," in *EMNLP*. Association for Computational Linguistics, 2023, pp. 13 043–13 052. [Online]. Available: https://aclanthology.org/2023.emnlp-main.806

[122] L. Cui, Y. Wu, J. Liu, S. Yang, and Y. Zhang, "Template-based named entity recognition using BART," in *ACL/IJCNLP (Findings)*, ser. Findings of ACL, vol. ACL/IJCNLP 2021. Association for Computational Linguistics, 2021, pp. 1835–1845. [Online]. Available: https://doi.org/10.18653/v1/2021.findings-acl.161

[123] X. Chen, L. Li, S. Deng, C. Tan, C. Xu, F. Huang, L. Si, H. Chen, and N. Zhang, "Lightner: A lightweight tuning paradigm for low-resource NER via pluggable prompting," in *COLING*. International Committee on Computational Linguistics, 2022, pp. 2374–2387. [Online]. Available: https://aclanthology.org/2022.coling-1.209

[124] Y. Ma, Z. Wang, Y. Cao, M. Li, M. Chen, K. Wang, and J. Shao, "Prompt for extraction? PAIE: prompting argument interaction for event argument extraction," in *ACL (1)*. Association for Computational Linguistics, 2022, pp. 6759–6774. [Online]. Available: https://doi.org/10.18653/v1/2022.acl-long.466

[125] Y. K. Chia, L. Bing, S. Poria, and L. Si, "Relationprompt: Leveraging prompts to generate synthetic data for zero-shot relation triplet extraction," in *ACL (Findings)*. Association for Computational Linguistics, 2022, pp. 45–57. [Online]. Available: https://doi.org/10.18653/v1/2022.findings-acl.5

[126] R. Ma, X. Zhou, T. Gui, Y. Tan, L. Li, Q. Zhang, and X. Huang, "Template-free prompt tuning for few-shot NER," in *NAACL-HLT*. Association for Computational Linguistics, 2022, pp. 5721–5732. [Online]. Available: https://doi.org/10.18653/v1/2022.naacl-main.420

[127] D. Lee, A. Kadakia, K. Tan, M. Agarwal, X. Feng, T. Shibuya, R. Mitani, T. Sekiya, J. Pujara, and X. Ren, "Good examples make A faster learner: Simple demonstration-based learning for low-resource NER," in *ACL (1)*. Association for Computational Linguistics, 2022, pp. 2687–2700. [Online]. Available: https://doi.org/10.18653/v1/2022.acl-long.192

[128] Y. Shen, Z. Tan, S. Wu, W. Zhang, R. Zhang, Y. Xi, W. Lu, and Y. Zhuang, "Promptner: Prompt locating and typing for named entity recognition," in *ACL (1)*. Association for Computational Linguistics, 2023, pp. 12 492–12 507. [Online]. Available: https://aclanthology.org/2023.acl-long.698/

[129] B. Lv, X. Liu, S. Dai, N. Liu, F. Yang, P. Luo, and Y. Yu, "DSP: discriminative soft prompts for zero-shot entity and relation extraction," in *ACL (Findings)*. Association for Computational Linguistics, 2023, pp. 5491–5505. [Online]. Available: https://doi.org/10.18653/v1/2023.findings-acl.339

[130] C. V. Nguyen, H. Man, and T. H. Nguyen, "Contextualized soft prompts for extraction of event arguments," in *ACL (Findings)*. Association for Computational Linguistics, 2023, pp. 4352–4361. [Online]. Available: https://doi.org/10.18653/v1/2023.findings-acl.266

[131] X. Chen, N. Zhang, X. Xie, S. Deng, Y. Yao, C. Tan, F. Huang, L. Si, and H. Chen, "Knowprompt: Knowledge-aware prompt-tuning with synergistic optimization for relation extraction," in *WWW*. ACM, 2022, pp. 2778–2788. [Online]. Available: https://doi.org/10.1145/3485447.3511998

[132] X. Chen, L. Li, N. Zhang, X. Liang, S. Deng, C. Tan, F. Huang, L. Si, and H. Chen, "Decoupling knowledge from memorization: Retrieval-augmented prompt learning," in *Advances in Neural Information Processing Systems*, vol. 35. Curran Associates, Inc., 2022, pp. 23 908–23 922. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2022/file/97011c648eda678424f9292dadeae72e-Paper-Conference.pdf

[133] H. Ye, N. Zhang, S. Deng, X. Chen, H. Chen, F. Xiong, X. Chen, and H. Chen, "Ontology-enhanced prompt-tuning for few-shot learning," in *WWW*. ACM, 2022, pp. 778–787. [Online]. Available: https://doi.org/10.1145/3485447.3511921

[134] I. Hsu, K. Huang, E. Boschee, S. Miller, P. Natarajan, K. Chang, and N. Peng, "DEGREE: A data-efficient generation-based event extraction model," in *NAACL-HLT*. Association for Computational Linguistics, 2022, pp. 1890–1908. [Online]. Available: https://doi.org/10.18653/v1/2022.naacl-main.138

[135] I.-H. Hsu, Z. Xie, K.-H. Huang, P. Natarajan, and N. Peng, "Ampere: Amr-aware prefix for generation-based event argument extraction model," in *ACL (1)*. Association for Computational Linguistics, 2023, pp. 10 976–10 993. [Online]. Available: https://aclanthology.org/2023.acl-long.615/

[136] Y. Lu, Q. Liu, D. Dai, X. Xiao, H. Lin, X. Han, L. Sun, and H. Wu, "Unified structure generation for universal information extraction," in *ACL (1)*. Association for Computational Linguistics, 2022, pp. 5755–5772. [Online]. Available: https://doi.org/10.18653/v1/2022.acl-long.395

[137] H. Fei, S. Wu, J. Li, B. Li, F. Li, L. Qin, M. Zhang, M. Zhang, and T.-S. Chua, "Lasuie: Unifying information extraction with latent adaptive structure-aware generative language model," in *Advances in Neural Information Processing Systems*, vol. 35. Curran Associates, Inc., 2022, pp. 15 460–15 475. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2022/file/63943ee9fe347f3d95892cf87d9a42e6-Paper-Conference.pdf

[138] J. Lou, Y. Lu, D. Dai, W. Jia, H. Lin, X. Han, L. Sun, and H. Wu, "Universal information extraction as unified semantic matching," in *AAAI*. AAAI Press, 2023, pp. 13 318–13 326. [Online]. Available: https://doi.org/10.1609/aaai.v37i11.26563

[139] C. Liu, F. Zhao, Y. Kang, J. Zhang, X. Zhou, C. Sun, K. Kuang, and F. Wu, "Rexuie: A recursive method with explicit schema instructor for universal information extraction," in *EMNLP (Findings)*. Association for Computational Linguistics, 2023, pp. 15 342–15 359. [Online]. Available: https://aclanthology.org/2023.findings-emnlp.1024

[140] X. Han, W. Zhao, N. Ding, Z. Liu, and M. Sun, "PTR: prompt tuning with rules for text classification," *AI Open*, vol. 3, pp. 182–192, 2022. [Online]. Available: https://doi.org/10.1016/j.aiopen.2022.11.003

[141] Y. Yao, S. Mao, X. Chen, N. Zhang, S. Deng, and H. Chen, "Schema-aware reference as prompt improves data-efficient relational triple and event extraction," in *SIGIR*, ser. SIGIR'23. Association for Computing Machinery, 2023, pp. 911—921. [Online]. Available: https://dl.acm.org/doi/10.1145/3539618.3591763

[142] J. Gao, H. Zhao, C. Yu, and R. Xu, "Exploring the feasibility of chatgpt for event extraction," *CoRR*, vol. abs/2303.03836, 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2303.03836

[143] X. Wei, X. Cui, N. Cheng, X. Wang, X. Zhang, S. Huang, P. Xie, J. Xu, Y. Chen, M. Zhang, Y. Jiang, and W. Han, "Zero-shot information extraction via chatting with chatgpt," *CoRR*, vol. abs/2302.10205, 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2302.10205

[144] Z. Lin, H. Zhang, and Y. Song, "Global constraints with prompting for zero-shot event argument classification," in *EACL (Findings)*. Association for Computational Linguistics, 2023, pp. 2482–2493. [Online]. Available: https://aclanthology.org/2023.findings-eacl.191

[145] G. Li, P. Wang, and W. Ke, "Revisiting large language models as zero-shot relation extractors," in *EMNLP (Findings)*. Association for Computational Linguistics, 2023, pp. 6877–6892. [Online]. Available: https://aclanthology.org/2023.findings-emnlp.459

[146] D. Ashok and Z. C. Lipton, "Promptner: Prompting for named entity recognition," *CoRR*, vol. abs/2305.15444, 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2305.15444

[147] B. Li, G. Fang, Y. Yang, Q. Wang, W. Ye, W. Zhao, and S. Zhang, "Evaluating chatgpt's information extraction capabilities: An assessment of performance, explainability, calibration, and faithfulness," *CoRR*, vol. abs/2304.11633, 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2304.11633

[148] Y. Zhu, X. Wang, J. Chen, S. Qiao, Y. Ou, Y. Yao, S. Deng, H. Chen, and N. Zhang, "Llms for knowledge graph construction and reasoning: Recent capabilities and future opportunities," *CoRR*, vol. abs/2305.13168, 2023. [Online]. Available: https://arxiv.org/abs/2305.13168

[149] X. Wang, S. Li, and H. Ji, "Code4struct: Code generation for few-shot event structure prediction," in *ACL (1)*. Association for Computational Linguistics, 2023, pp. 3640–3663. [Online]. Available: https://aclanthology.org/2023.acl-long.202/

[150] P. Li, T. Sun, Q. Tang, H. Yan, Y. Wu, X. Huang, and X. Qiu, "Codeie: Large code generation models are better few-shot information extractors," in *ACL (1)*. Association for Computational Linguistics, 2023, pp. 15339–15353. [Online]. Available: https://aclanthology.org/2023.acl-long.855/

[151] Z. Bi, J. Chen, Y. Jiang, F. Xiong, W. Guo, H. Chen, and N. Zhang, "Codekgc: Code language model for generative knowledge graph construction," *CoRR*, vol. abs/2304.09048, 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2304.09048

[152] Y. Chen, X. Wang, M. Li, D. Hoiem, and H. Ji, "Vistruct: Visual structural knowledge extraction via curriculum guided code-vision representation," in *EMNLP*. Association for Computational Linguistics, 2023, pp. 13342–13357. [Online]. Available: https://aclanthology.org/2023.emnlp-main.824

[153] Y. Guo, Z. Li, X. Jin, Y. Liu, Y. Zeng, W. Liu, X. Li, P. Yang, L. Bai, J. Guo, and X. Cheng, "Retrieval-augmented code generation for universal information extraction," *CoRR*, vol. abs/2311.02962, 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2311.02962

[154] O. Sainz, I. García-Ferrero, R. Agerri, O. L. de Lacalle, G. Rigau, and E. Agirre, "Gollie: Annotation guidelines improve zero-shot information-extraction," *CoRR*, vol. abs/2310.03668, 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2310.03668

[155] J. Chen, Y. Lu, H. Lin, J. Lou, W. Jia, D. Dai, H. Wu, B. Cao, X. Han, and L. Sun, "Learning in-context learning for named entity recognition," in *ACL (1)*. Association for Computational Linguistics, 2023, pp. 13661–13675. [Online]. Available: https://doi.org/10.18653/v1/2023.acl-long.764

[156] X. Xu, Y. Zhu, X. Wang, and N. Zhang, "How to unleash the power of large language models for few-shot relation extraction?" in *Proceedings of The Fourth Workshop on Simple and Efficient Natural Language Processing (SustaiNLP)*. Association for Computational Linguistics, 2023, pp. 190–200. [Online]. Available: https://aclanthology.org/2023.sustainlp-1.13/

[157] Z. Wan, F. Cheng, Z. Mao, Q. Liu, H. Song, J. Li, and S. Kurohashi, "GPT-RE: in-context learning for relation extraction using large language models," in *EMNLP*. Association for Computational Linguistics, 2023, pp. 3534–3547. [Online]. Available: https://aclanthology.org/2023.emnlp-main.214

[158] C. Pang, Y. Cao, Q. Ding, and P. Luo, "Guideline learning for in-context information extraction," in *EMNLP*. Association for Computational Linguistics, 2023, pp. 15372–15389. [Online]. Available: https://aclanthology.org/2023.emnlp-main.950

[159] Y. Ma, Y. Cao, Y. Hong, and A. Sun, "Large language model is not a good few-shot information extractor, but a good reranker for hard samples!" in *EMNLP (Findings)*. Association for Computational Linguistics, 2023, pp. 10572–10601. [Online]. Available: https://aclanthology.org/2023.findings-emnlp.710

[160] S. Wang, X. Sun, X. Li, R. Ouyang, F. Wu, T. Zhang, J. Li, and G. Wang, "GPT-NER: named entity recognition via large language models," *CoRR*, vol. abs/2304.10428, 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2304.10428

[161] Y. Ozyurt, S. Feuerriegel, and C. Zhang, "In-context few-shot relation extraction via pre-trained language models," *arXiv preprint arXiv:2310.11085*, 2023. [Online]. Available: https://arxiv.org/abs/2310.11085

[162] T. Xie, Q. Li, Y. Zhang, Z. Liu, and H. Wang, "Self-improving for zero-shot named entity recognition with large language models," *CoRR*, vol. abs/2311.08921, 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2311.08921

[163] H. Sousa, N. Guimarães, A. Jorge, and R. Campos, "Gpt struct me: Probing gpt models on narrative entity extraction," *CoRR*, vol. abs/2311.08921, 2023. [Online]. Available: https://arxiv.org/abs/2311.14583

[164] R. Han, T. Peng, C. Yang, B. Wang, L. Liu, and X. Wan, "Is information extraction solved by chatgpt? an analysis of performance, evaluation criteria, robustness and errors," *CoRR*, vol. abs/2305.14450, 2023. [Online]. Available: https://arxiv.org/abs/2305.14450

[165] C. Ling, X. Zhao, X. Zhang, Y. Liu, W. Cheng, H. Wang, Z. Chen, T. Osaki, K. Matsuda, H. Chen, and L. Zhao, "Improving open information extraction with large language models: A study on demonstration uncertainty," *CoRR*, vol. abs/2309.03433, 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2309.03433

[166] J. P. McCusker, "LOKE: linked open knowledge extraction for automated knowledge graph construction," *CoRR*, vol. abs/2311.09366, 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2311.09366

[167] C. Cai, Q. Wang, B. Liang, B. Qin, M. Yang, K. Wong, and R. Xu, "In-context learning for few-shot multimodal named entity recognition," in *EMNLP (Findings)*. Association for Computational Linguistics, 2023, pp. 2969–2979. [Online]. Available: https://aclanthology.org/2023.findings-emnlp.196

[168] J. Li, H. Li, Z. Pan, D. Sun, J. Wang, W. Zhang, and G. Pan, "Prompting chatgpt in MNER: enhanced multimodal named entity recognition with auxiliary refined knowledge," in *EMNLP (Findings)*. Association for Computational Linguistics, 2023, pp. 2787–2802. [Online]. Available: https://aclanthology.org/2023.findings-emnlp.184

[169] C. Wang, X. Liu, Z. Chen, H. Hong, J. Tang, and D. Song, "Deepstruct: Pretraining of language models for structure prediction," in *ACL (Findings)*. Association for Computational Linguistics, 2022, pp. 803–823. [Online]. Available: https://doi.org/10.18653/v1/2022.findings-acl.67

[170] K. Zhang, B. J. Gutierrez, and Y. Su, "Aligning instruction tasks unlocks large language models as zero-shot relation extractors," in *ACL (Findings)*. Association for Computational Linguistics, 2023, pp. 794–812. [Online]. Available: https://doi.org/10.18653/v1/2023.findings-acl.50

[171] Y. Jiao, M. Zhong, S. Li, R. Zhao, S. Ouyang, H. Ji, and J. Han, "Instruct and extract: Instruction tuning for on-demand information extraction," in *EMNLP*. Association for Computational Linguistics, 2023, pp. 10030–10051. [Online]. Available: https://aclanthology.org/2023.emnlp-main.620

[172] W. Zhou, S. Zhang, Y. Gu, M. Chen, and H. Poon, "Universalner: Targeted distillation from large language models for open named entity recognition," in *ICLR*. OpenReview.net, 2024. [Online]. Available: https://openreview.net/forum?id=r65xfUb76p

[173] X. Wang, W. Zhou, C. Zu, H. Xia, T. Chen, Y. Zhang, R. Zheng, J. Ye, Q. Zhang, T. Gui, J. Kang, J. Yang, S. Li, and C. Du, "Instructuie: Multi-task instruction tuning for unified information extraction," *CoRR*, vol. abs/2304.08085, 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2304.08085

[174] J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: pre-training of deep bidirectional transformers for language understanding," in *NAACL-HLT (1)*. Association for Computational Linguistics, 2019, pp. 4171–4186. [Online]. Available: https://doi.org/10.18653/v1/n19-1423

[175] M. Lewis, Y. Liu, N. Goyal, M. Ghazvininejad, A. Mohamed, O. Levy, V. Stoyanov, and L. Zettlemoyer, "BART: denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension," in *ACL*. Association for Computational Linguistics, 2020, pp. 7871–7880. [Online]. Available: https://doi.org/10.18653/v1/2020.acl-main.703

[176] C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li, and P. J. Liu, "Exploring the limits of transfer learning with a unified text-to-text transformer," *J. Mach. Learn. Res.*, vol. 21, pp. 140:1–140:67, 2020. [Online]. Available: http://jmlr.org/papers/v21/20-074.html

[177] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever, "Language models are unsupervised multitask learners," 2019. [Online]. Available: https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf

[178] H. W. Chung, L. Hou, S. Longpre, B. Zoph, Y. Tay, W. Fedus, E. Li, X. Wang, M. Dehghani, S. Brahma, A. Webson, S. S. Gu, Z. Dai, M. Suzgun, X. Chen, A. Chowdhery, S. Narang, G. Mishra, A. Yu, V. Y. Zhao, Y. Huang, A. M. Dai, H. Yu, S. Petrov, E. H. Chi, J. Dean, J. Devlin, A. Roberts, D. Zhou, Q. V. Le, and J. Wei, "Scaling instruction-finetuned language models," *CoRR*, vol. abs/2210.11416, 2022. [Online]. Available: https://doi.org/10.48550/arXiv.2210.11416

[179] H. Touvron, T. Lavril, G. Izacard, X. Martinet, M. Lachaux, T. Lacroix, B. Rozière, N. Goyal, E. Hambro, F. Azhar, A. Rodriguez, A. Joulin, E. Grave, and G. Lample, "Llama: Open and efficient foundation language models," *CoRR*, vol. abs/2302.13971, 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2302.13971

[180] Z. Du, Y. Qian, X. Liu, M. Ding, J. Qiu, Z. Yang, and J. Tang, "GLM: general language model pretraining with autoregressive blank infilling," in *ACL (1)*. Association for Computational Linguistics, 2022, pp. 320–335. [Online]. Available: https://doi.org/10.18653/v1/2022.acl-long.26

[181] T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. M. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, and D. Amodei, "Language models are few-shot learners," in *NeurIPS*, 2020. [Online]. Available: https://proceedings.neurips.cc/paper/2020/hash/1457c0d6bfcb4967418bfb8ac142f64a-Abstract.html

[182] OpenAI, "GPT-4 technical report," 2023. [Online]. Available: https://cdn.openai.com/papers/gpt-4.pdf

[183] J. Chen, D. Tam, C. Raffel, M. Bansal, and D. Yang, "An empirical survey of data augmentation for limited data learning in NLP," *Trans. Assoc. Comput. Linguistics*, vol. 11, pp. 191–211, 2023. [Online]. Available: https://aclanthology.org/2023.tacl-1.12

[184] M. Mintz, S. Bills, R. Snow, and D. Jurafsky, "Distant supervision for relation extraction without labeled data," in *ACL/IJCNLP*. Association for Computational Linguistics, 2009, pp. 1003–1011. [Online]. Available: https://aclanthology.org/P09-1113/

[185] N. Taghizadeh and H. Faili, "Cross-lingual transfer learning for relation extraction using universal dependencies," *Comput. Speech Lang.*, vol. 71, p. 101265, 2022. [Online]. Available: https://doi.org/10.1016/j.csl.2021.101265

[186] Y. Lu, H. Lin, X. Han, and L. Sun, "Distilling discrimination and generalization knowledge for event detection via delta-representation learning," in *ACL (1)*. Association for Computational Linguistics, 2019, pp. 4366–4376. [Online]. Available: https://doi.org/10.18653/v1/p19-1429

[187] M. Dong, C. Pan, and Z. Luo, "Mapre: An effective semantic mapping approach for low-resource relation extraction," in *EMNLP (1)*. Association for Computational Linguistics, 2021, pp. 2694–2704. [Online]. Available: https://doi.org/10.18653/v1/2021.emnlp-main.212

[188] B. Zhou, D. Khashabi, C. Tsai, and D. Roth, "Zero-shot open entity typing as type-compatible grounding," in *EMNLP*. Association for Computational Linguistics, 2018, pp. 2065–2076. [Online]. Available: https://doi.org/10.18653/v1/d18-1231

[189] R. Obeidat, X. Z. Fern, H. Shahbazi, and P. Tadepalli, "Description-based zero-shot fine-grained entity typing," in *NAACL-HLT (1)*. Association for Computational Linguistics, 2019, pp. 807–814. [Online]. Available: https://doi.org/10.18653/v1/n19-1087

[190] J. Liu, Y. Chen, and J. Xu, "Low-resource NER by data augmentation with prompting," in *IJCAI*. ijcai.org, 2022, pp. 4252–4258. [Online]. Available: https://doi.org/10.24963/ijcai.2022/590

[191] X. Hu, A. Liu, Z. Tan, X. Zhang, C. Zhang, I. King, and P. S. Yu, "Gda: Generative data augmentation techniques for relation extraction tasks," in *ACL (Findings)*. Association for Computational Linguistics, 2023, pp. 10 221–10 234. [Online]. Available: https://aclanthology.org/2023.findings-acl.649/

[192] A. P. B. Veyseh, F. Dernoncourt, B. Min, and T. Nguyen, "Generating labeled data for relation extraction: A meta learning approach with joint GPT-2 training," in *ACL (Findings)*. Association for Computational Linguistics, 2023, pp. 11 466–11 478. [Online]. Available: https://doi.org/10.18653/v1/2023.findings-acl.727

[193] J. Zhao, W. Zhan, X. Zhao, Q. Zhang, T. Gui, Z. Wei, J. Wang, M. Peng, and M. Sun, "Re-matching: A fine-grained semantic matching method for zero-shot relation extraction," in *ACL (1)*. Association for Computational Linguistics, 2023, pp. 6680–6691. [Online]. Available: https://aclanthology.org/2023.acl-long.369/

[194] Y. Pei, P. Bhuse, Z. Liu, and E. Xing, "Segmix: A simple structure-aware data augmentation method," *CoRR*, vol. abs/2311.09505, 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2311.09505

[195] S. Liu, Y. Chen, S. He, K. Liu, and J. Zhao, "Leveraging framenet to improve automatic event detection," in *ACL (1)*. The Association for Computer Linguistics, 2016. [Online]. Available: https://doi.org/10.18653/v1/p16-1201

[196] H. V. Nguyen, F. Gelli, and S. Poria, "DOZEN: cross-domain zero shot named entity recognition with knowledge graph," in *SIGIR*. ACM, 2021, pp. 1642–1646. [Online]. Available: https://doi.org/10.1145/3404835.3463113

[197] T. M. Hospedales, A. Antoniou, P. Micaelli, and A. J. Storkey, "Meta-learning in neural networks: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 9, pp. 5149–5169, 2022. [Online]. Available: https://doi.org/10.1109/TPAMI.2021.3079209

[198] A. Obamuyide and A. Vlachos, "Model-agnostic meta-learning for relation classification with limited supervision," in *ACL (1)*. Association for Computational Linguistics, 2019, pp. 5873–5879. [Online]. Available: https://doi.org/10.18653/v1/p19-1589

[199] S. Deng, N. Zhang, Z. Sun, J. Chen, and H. Chen, "When low resource NLP meets unsupervised language model: Meta-pretraining then meta-learning for few-shot text classification (student abstract)," in *AAAI*. AAAI Press, 2020, pp. 13 773–13 774. [Online]. Available: https://doi.org/10.1609/aaai.v34i10.7158

[200] T. Ma, H. Jiang, Q. Wu, T. Zhao, and C. Lin, "Decomposed meta-learning for few-shot named entity recognition," in *ACL (Findings)*. Association for Computational Linguistics, 2022, pp. 1584–1596. [Online]. Available: https://doi.org/10.18653/v1/2022.findings-acl.124

[201] C. Finn, P. Abbeel, and S. Levine, "Model-agnostic meta-learning for fast adaptation of deep networks," in *ICML*, ser. Proceedings of Machine Learning Research, vol. 70. PMLR, 2017, pp. 1126–1135. [Online]. Available: http://proceedings.mlr.press/v70/finn17a.html

[202] A. Santoro, S. Bartunov, M. M. Botvinick, D. Wierstra, and T. P. Lillicrap, "Meta-learning with memory-augmented neural networks," in *ICML*, ser. JMLR Workshop and Conference Proceedings, vol. 48. JMLR.org, 2016, pp. 1842–1850. [Online]. Available: http://proceedings.mlr.press/v48/santoro16.html

[203] M. Patacchiola, J. Turner, E. J. Crowley, M. F. P. O'Boyle, and A. J. Storkey, "Bayesian meta-learning for the few-shot setting via deep kernels," in *NeurIPS*, 2020. [Online]. Available: https://proceedings.neurips.cc/paper/2020/hash/b9cfe8b6042cf759dc4c0cccb27a6737-Abstract.html

[204] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, 2010. [Online]. Available: https://doi.org/10.1109/TKDE.2009.191

[205] X. Han, Z. Zhang, N. Ding, Y. Gu, X. Liu, Y. Huo, J. Qiu, Y. Yao, A. Zhang, L. Zhang, W. Han, M. Huang, Q. Jin, Y. Lan, Y. Liu, Z. Liu, Z. Lu, X. Qiu, R. Song, J. Tang, J. Wen, J. Yuan, W. X. Zhao, and J. Zhu, "Pre-trained models: Past, present and future," *AI Open*, vol. 2, pp. 225–250, 2021. [Online]. Available: https://doi.org/10.1016/j.aiopen.2021.08.002

[206] S. Zhao, T. Liu, S. Zhao, and F. Wang, "A neural multi-task learning framework to jointly model medical named entity recognition and normalization," in *AAAI*. AAAI Press, 2019, pp. 817–824. [Online]. Available: https://doi.org/10.1609/aaai.v33i01.3301817

[207] B. Zhou, X. Cai, Y. Zhang, W. Guo, and X. Yuan, "MTAAL: multi-task adversarial active learning for medical named entity recognition and normalization," in *AAAI*. AAAI Press, 2021, pp. 14 586–14 593. [Online]. Available: https://ojs.aaai.org/index.php/AAAI/article/view/17714

[208] M. Josifoski, N. D. Cao, M. Peyrard, F. Petroni, and R. West, "Genie: Generative information extraction," in *NAACL-HLT*. Association for Computational Linguistics, 2022, pp. 4626–4643. [Online]. Available: https://doi.org/10.18653/v1/2022.naacl-main.342

[209] H. Ye, N. Zhang, H. Chen, and H. Chen, "Generative knowledge graph construction: A review," in *EMNLP*. Association for Computational Linguistics, 2022, pp. 1–17. [Online]. Available: https://aclanthology.org/2022.emnlp-main.1

[210] D. Xu, W. Chen, W. Peng, C. Zhang, T. Xu, X. Zhao, X. Wu, Y. Zheng, and E. Chen, "Large language models for generative information extraction: A survey," *CoRR*, vol. abs/2312.17617, 2023. [Online]. Available: https://arxiv.org/abs/2312.17617

[211] H. Ye, N. Zhang, Z. Bi, S. Deng, C. Tan, H. Chen, F. Huang, and H. Chen, "Learning to ask for data-efficient event argument extraction (student abstract)," in *AAAI*. AAAI Press, 2022, pp. 13 099–13 100. [Online]. Available: https://ojs.aaai.org/index.php/AAAI/article/view/21686

[212] S. Najafi and A. Fyshe, "Weakly-supervised questions for zero-shot relation extraction," in *EACL*. Association for Computational Linguistics, 2023, pp. 3067–3079. [Online]. Available: https://doi.org/10.18653/v1/2023.eacl-main.224

[213] P. Liu, W. Yuan, J. Fu, Z. Jiang, H. Hayashi, and G. Neubig, "Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing," *ACM Comput. Surv.*, vol. 55, no. 9, 2023. [Online]. Available: https://doi.org/10.1145/3560815

[214] Y. Huang, K. He, Y. Wang, X. Zhang, T. Gong, R. Mao, and C. Li, "COPNER: contrastive learning with prompt guiding for few-shot named entity recognition," in *COLING*. International Committee on Computational Linguistics, 2022, pp. 2515–2527. [Online]. Available: https://aclanthology.org/2022.coling-1.222

[215] X. Liu, H. Huang, G. Shi, and B. Wang, "Dynamic prefix-tuning for generative template-based event extraction," in *ACL (1)*. Association for Computational Linguistics, 2022, pp. 5216–5228. [Online]. Available: https://doi.org/10.18653/v1/2022.acl-long.358

[216] S. Wang, M. Yu, and L. Huang, "The art of prompting: Event detection based on type specific prompts," in *ACL (2)*. Association for Computational Linguistics, 2023, pp. 1286–1299. [Online]. Available: https://doi.org/10.18653/v1/2023.acl-short.111

[217] X. Chen, L. Li, N. Zhang, C. Tan, F. Huang, L. Si, and H. Chen, "Relation extraction as open-book examination: Retrieval-enhanced prompt tuning," in *SIGIR*. ACM, 2022, pp. 2443–2448. [Online]. Available: https://doi.org/10.1145/3477495.3531746

[218] X. Du and H. Ji, "Retrieval-augmented generative question answering for event argument extraction," in *EMNLP*. Association for Computational Linguistics, 2022, pp. 4649–4666. [Online]. Available: https://doi.org/10.18653/v1/2022.emnlp-main.307

[219] H. Huang, X. Liu, G. S. Han, and Q. Liu, "Event extraction with dynamic prefix tuning and relevance retrieval," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 10, pp. 9946–9958, 2023. [Online]. Available: https://doi.org/10.1109/TKDE.2023.3266495

[220] C. Yuan, Q. Xie, and S. Ananiadou, "Zero-shot temporal relation extraction with chatgpt," in *BioNLP@ACL*. Association for Computational Linguistics, 2023, pp. 92–102. [Online]. Available: https://doi.org/10.18653/v1/2023.bionlp-1.7

[221] T. Xie, Q. Li, J. Zhang, Y. Zhang, Z. Liu, and H. Wang, "Empirical study of zero-shot NER with chatgpt," in *EMNLP*. Association for Computational Linguistics, 2023, pp. 7935–7956. [Online]. Available: https://aclanthology.org/2023.emnlp-main.493

[222] M. Chen, J. Tworek, H. Jun, Q. Yuan, H. P. de Oliveira Pinto, J. Kaplan, H. Edwards, Y. Burda, N. Joseph, G. Brockman, A. Ray, R. Puri, G. Krueger, M. Petrov, H. Khlaaf, G. Sastry, P. Mishkin, B. Chan, S. Gray, N. Ryder, M. Pavlov, A. Power, L. Kaiser, M. Bavarian, C. Winter, P. Tillet, F. P. Such, D. Cummings, M. Plappert, F. Chantzis, E. Barnes, A. Herbert-Voss, W. H. Guss, A. Nichol, A. Paino, N. Tezak, J. Tang, I. Babuschkin, S. Balaji, S. Jain, W. Saunders, C. Hesse, A. N. Carr, J. Leike, J. Achiam, V. Misra, E. Morikawa, A. Radford, M. Knight, M. Brundage, M. Murati, K. Mayer, P. Welinder, B. McGrew, D. Amodei, S. McCandlish, I. Sutskever, and W. Zaremba, "Evaluating large language models trained on code," *CoRR*, vol. abs/2107.03374, 2021. [Online]. Available: https://arxiv.org/abs/2107.03374

[223] X. Ma, J. Li, and M. Zhang, "Chain of thought with explicit evidence reasoning for few-shot relation extraction," in *EMNLP (Findings)*. Association for Computational Linguistics, 2023, pp. 2334–2352. [Online]. Available: https://aclanthology.org/2023-emnlp.153

[224] M. Josifoski, M. Sakota, M. Peyrard, and R. West, "Exploiting asymmetry for synthetic training data generation: Synthie and the case of information extraction," in *EMNLP*. Association for Computational Linguistics, 2023, pp. 1555–1574. [Online]. Available: https://aclanthology.org/2023.emnlp-main.96

[225] R. Zhang, Y. Li, Y. Ma, M. Zhou, and L. Zou, "Llmaaa: Making large language models as active annotators," in *EMNLP (Findings)*. Association for Computational Linguistics, 2023, pp. 13 088–13 103. [Online]. Available: https://aclanthology.org/2023.findings-emnlp.872

[226] S. Wang, J. Zheng, F. Cai, C. Song, and X. Luo, "Msprompt: Multi-step prompt learning for debiasing few-shot event detection," *Information Processing and Management*, vol. 60, no. 6, p. 103509, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306457323002467

[227] E. F. T. K. Sang and F. D. Meulder, "Introduction to the conll-2003 shared task: Language-independent named entity recognition," in *CoNLL*. ACL, 2003, pp. 142–147. [Online]. Available: https://aclanthology.org/W03-0419/

[228] R. Weischedel, M. Palmer, M. Marcus, E. Hovy, S. Pradhan, L. Ramshaw, N. Xue, A. Taylor, J. Kaufman, M. Franchini *et al.*, "Ontonotes release 5.0 ldc2013t19," *Linguistic Data Consortium, Philadelphia, PA*, 2013. [Online]. Available: https://catalog.ldc.upenn.edu/LDC2013T19

[229] N. Ding, G. Xu, Y. Chen, X. Wang, X. Han, P. Xie, H. Zheng, and Z. Liu, "Few-nerd: A few-shot named entity recognition dataset," in *ACL/IJCNLP (1)*. Association for Computational Linguistics, 2021, pp. 3198–3213. [Online]. Available: https://doi.org/10.18653/v1/2021.acl-long.248

[230] C. Alt, A. Gabryszak, and L. Hennig, "TACRED revisited: A thorough evaluation of the TACRED relation extraction task," in *ACL*. Association for Computational Linguistics, 2020, pp. 1558–1569. [Online]. Available: https://doi.org/10.18653/v1/2020.acl-main.142

[231] S. Riedel, L. Yao, and A. McCallum, "Modeling relations and their mentions without labeled text," in *ECML/PKDD (3)*, ser. Lecture Notes in Computer Science, vol. 6323. Springer, 2010, pp. 148–163. [Online]. Available: https://doi.org/10.1007/978-3-642-15939-8_10

[232] X. Han, H. Zhu, P. Yu, Z. Wang, Y. Yao, Z. Liu, and M. Sun, "Fewrel: A large-scale supervised few-shot relation classification dataset with state-of-the-art evaluation," in *EMNLP*. Association for Computational Linguistics, 2018, pp. 4803–4809. [Online]. Available: https://doi.org/10.18653/v1/d18-1514

[233] G. R. Doddington, A. Mitchell, M. A. Przybocki, L. A. Ramshaw, S. M. Strassel, and R. M. Weischedel, "The automatic content extraction (ACE) program - tasks, data, and evaluation," in *LREC*. European Language Resources Association, 2004. [Online]. Available: http://www.lrec-conf.org/proceedings/lrec2004/summaries/5.htm

[234] X. Wang, Z. Wang, X. Han, W. Jiang, R. Han, Z. Liu, J. Li, P. Li, Y. Lin, and J. Zhou, "MAVEN: A massive general domain event detection dataset," in *EMNLP (1)*. Association for Computational Linguistics, 2020, pp. 1652–1671. [Online]. Available: https://doi.org/10.18653/v1/2020.emnlp-main.129

[235] S. Ebner, P. Xia, R. Culkin, K. Rawlins, and B. V. Durme, "Multi-sentence argument linking," in *ACL*. Association for Computational Linguistics, 2020, pp. 8057–8077. [Online]. Available: https://doi.org/10.18653/v1/2020.acl-main.718

[236] S. Li, H. Ji, and J. Han, "Document-level event argument extraction by conditional generation," in *NAACL-HLT*. Association for Computational Linguistics, 2021, pp. 894–908. [Online]. Available: https://doi.org/10.18653/v1/2021.naacl-main.69

[237] E. J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, and W. Chen, "Lora: Low-rank adaptation of large language models," in *ICLR*. OpenReview.net, 2022. [Online]. Available: https://openreview.net/forum?id=nZeVKeeFYf9

[238] T. Gao, X. Han, H. Zhu, Z. Liu, P. Li, M. Sun, and J. Zhou, "Fewrel 2.0: Towards more challenging few-shot relation classification," in *EMNLP/IJCNLP (1)*. Association for Computational Linguistics, 2019, pp. 6249–6254. [Online]. Available: https://doi.org/10.18653/v1/D19-1649

[239] X. Xu, X. Chen, N. Zhang, X. Xie, X. Chen, and H. Chen, "Towards realistic low-resource relation extraction: A benchmark with empirical baseline study," in *EMNLP (Findings)*, 2022, pp. 413–427. [Online]. Available: https://aclanthology.org/2022.findings-emnlp.29

[240] X. Yang, Y. Lu, and L. R. Petzold, "Few-shot document-level event argument extraction," in *ACL (1)*. Association for Computational Linguistics, 2023, pp. 8029–8046. [Online]. Available: https://doi.org/10.18653/v1/2023.acl-long.446

[241] K. Huang, I. Hsu, T. Parekh, Z. Xie, Z. Zhang, P. Natarajan, K. Chang, N. Peng, and H. Ji, "A reevaluation of event extraction: Past, present, and future challenges," *CoRR*, vol. abs/2311.09562, 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2311.09562

[242] D. Lee, J. Pujara, M. Sewak, R. White, and S. K. Jauhar, "Making large language models better data creators," in *EMNLP*. Association for Computational Linguistics, 2023, pp. 15 349–15 360. [Online]. Available: https://aclanthology.org/2023.emnlp-main.948

[243] S. Abdelnabi, A. Gomaa, S. Sivaprasad, L. Schönherr, and M. Fritz, "Llm-deliberation: Evaluating llms with interactive multi-agent negotiation games," *CoRR*, vol. abs/2309.17234, 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2309.17234

[244] J. Zhang, X. Xu, and S. Deng, "Exploring collaboration mechanisms for LLM agents: A social psychology view," *CoRR*, vol. abs/2310.02124, 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2310.02124

[245] S. Qiao, H. Gui, H. Chen, and N. Zhang, "Making language models better tool learners with execution feedback," *CoRR*, vol. abs/2305.13068, 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2305.13068

[246] Z. Feng, W. Ma, W. Yu, L. Huang, H. Wang, Q. Chen, W. Peng, X. Feng, B. Qin, and T. Liu, "Trends in integration of knowledge and large language models: A survey and taxonomy of methods, benchmarks, and applications," *CoRR*, vol. abs/2311.05876, 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2311.05876

[247] S. Pan, L. Luo, Y. Wang, C. Chen, J. Wang, and X. Wu, "Unifying large language models and knowledge graphs: A roadmap," *CoRR*, vol. abs/2306.08302, 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2306.08302

[248] J. Z. Pan, S. Razniewski, J. Kalo, S. Singhania, J. Chen, S. Dietze, H. Jabeen, J. Omeliyanenko, W. Zhang, M. Lissandrini, R. Biswas, G. de Melo, A. Bonifati, E. Vakaj, M. Dragoni, and D. Graux, "Large language models and knowledge graphs: Opportunities and challenges," *Transactions on Graph Data and Knowledge*, vol. 1, no. 1, pp. 2:1–2:38, 2023. [Online]. Available: https://drops.dagstuhl.de/entities/document/10.4230/TGDK.1.1.2

[249] H. Chen, "Large knowledge model: Perspectives and challenges," 2023. [Online]. Available: https://arxiv.org/abs/2312.02706

[250] A. Amalvy, V. Labatut, and R. Dufour, "Learning to rank context for named entity recognition using a synthetic dataset," in *EMNLP*. Association for Computational Linguistics, 2023, pp. 10 372–10 382. [Online]. Available: https://aclanthology.org/2023.emnlp-main.642

[251] J. Qi, C. Zhang, X. Wang, K. Zeng, J. Yu, J. Liu, L. Hou, J. Li, and X. Bin, "Preserving knowledge invariance: Rethinking robustness evaluation of open information extraction," in *EMNLP*. Association for Computational Linguistics, 2023, pp. 5876–5890. [Online]. Available: https://aclanthology.org/2023.emnlp-main.360

[252] T. Zhang, I. Tham, Z. Hou, J. Ren, L. Zhou, H. Xu, L. Zhang, L. J. Martin, R. Dror, S. Li, H. Ji, M. Palmer, S. W. Brown, R. Suchocki, and C. Callison-Burch, "Human-in-the-loop schema induction," in *ACL (demo)*. Association for Computational Linguistics, 2023, pp. 1–10. [Online]. Available: https://doi.org/10.18653/v1/2023.acl-demo.1

[253] X. Wang, Y. Jiang, N. Bach, T. Wang, Z. Huang, F. Huang, and K. Tu, "Automated concatenation of embeddings for structured prediction," in *ACL/IJCNLP (1)*. Association for Computational Linguistics, 2021, pp. 2643–2660. [Online]. Available: https://doi.org/10.18653/v1/2021.acl-long.206

[254] Y. Fei and X. Xu, "GFMRC: A machine reading comprehension model for named entity recognition," *Pattern Recognit. Lett.*, vol. 172, pp. 97–105, 2023. [Online]. Available: https://doi.org/10.1016/j.patrec.2023.06.011

[255] D. Ye, Y. Lin, P. Li, and M. Sun, "Packed levitated marker for entity and relation extraction," in *ACL (1)*. Association for Computational Linguistics, 2022, pp. 4904–4917. [Online]. Available: https://doi.org/10.18653/v1/2022.acl-long.337

[256] Z. Yang, Y. Liu, and C. Ouyang, "Causal intervention-based few-shot named entity recognition," in *EMNLP (Findings)*. Association for Computational Linguistics, 2023, pp. 15 635–15 646. [Online]. Available: https://aclanthology.org/2023.findings-emnlp.1046

[257] Y. Zhan, Z. Li, X. Zhao, C. Zhang, and T. Chen, "A simple overlapping relation extraction method based on dropout," in *IJCNN*. IEEE, 2022, pp. 1–8. [Online]. Available: https://doi.org/10.1109/IJCNN55064.2022.9892310

[258] J. Zheng and Z. Chen, "Sentence-level relation extraction via contrastive learning with descriptive relation prompts," *CoRR*, vol. abs/2304.04935, 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2304.04935

[259] X. Ma, X. Qin, J. Liu, and W. Ran, "Interaction information guided prototype representation rectification for few-shot relation extraction," *Electronics*, vol. 12, no. 13, 2023. [Online]. Available: https://www.mdpi.com/2079-9292/12/13/2912

[260] S. Zhang, T. Ji, W. Ji, and X. Wang, "Zero-shot event detection based on ordered contrastive learning and prompt-based prediction," in *NAACL-HLT (Findings)*. Association for Computational Linguistics, 2022, pp. 2572–2580. [Online]. Available: https://doi.org/10.18653/v1/2022.findings-naacl.196

[261] W. Liu, S. Cheng, D. Zeng, and H. Qu, "Enhancing document-level event argument extraction with contextual clues and role relevance," in *ACL (Findings)*. Association for Computational Linguistics, 2023, pp. 12 908–12 922. [Online]. Available: https://doi.org/10.18653/v1/2023.findings-acl.817

APPENDICES

APPENDIX

For **Previous SOTA** results in Table I, we list the corresponding methods in Table III.

Task	Dataset	Previous SOTA				
		full	all-way-0-shot	10-way-0-shot	all-way-5-shot	10-way-5-shot
NER	CoNLL03	94.60 [253]	74.99 [162]▷	-	83.25 [138]	-
	OntoNotes5.0	92.30 [254]	-	-	59.70 [159]*	-
	FewNERD	70.90 [255]	-	-	59.41 [159]*	79.00 [256]
RE	NYT	93.50 [257]	-	-	-	-
	TACREV	85.80 [258]	59.40 [170]▷	-	47.12 [159]°	-
	FewRel	-	-	84.20 [129]	-	96.51 [259]
ED	ACE05	83.65 [260]	51.20 [138]	54.50 [216]	55.61 [134]	64.80 [134]
	MAVEN	79.09 [111]	59.90 [216]*	36.86 [82]°	64.80 [16]	93.06 [82]†
	FewEvent	96.58 [260]	58.14 [260]	68.37 [82]	60.67 [226]‡	93.18 [82]
EAE	ACE05	73.50 [134]	31.20 [144]▷	31.40 [134]	45.90 [159]*	42.70 [134]
	RAMS	59.66 [261]	-	-	54.08 [159]*	-
	WikiEvents	70.08 [261]	-	-	-	-

TABLE III

THE METHODS ACHIEVING THE PREVIOUS SOTA MICRO F1 PERFORMANCE, BASED ON RESULTS IN TABLE I. ▷: LLM-BASED; *: TYPE-SPECIFIC PROMPTING; ◇: PROMPT-BASED META LEARNING; †: 45-WAY-5-SHOT; ‡: ALL-WAY-4-SHOT.