A Survey on Event Extraction and Related Techniques in NLP

www.surveyx.cn

Abstract

This survey paper explores the interconnected fields of Event Extraction, Large Language Models (LLMs), Data Augmentation, Few-shot Learning, Natural Language Processing (NLP), Information Retrieval, and Text Mining, emphasizing their collective role in advancing artificial intelligence and computational linguistics. Event Extraction focuses on identifying and categorizing occurrences in text, while LLMs enhance language understanding through extensive training on vast datasets. Data Augmentation addresses data scarcity by generating synthetic variations, and Few-shot Learning enables models to generalize from limited examples. The integration of NLP with Information Retrieval and Text Mining facilitates the extraction of valuable insights from unstructured data. This survey systematically examines the methodologies, challenges, and applications of event extraction, highlighting the role of LLMs in enhancing accuracy and adaptability. It discusses traditional, rule-based, machine learning, and deep learning approaches, along with document-level event extraction innovations. The paper also explores the architectural and training advancements in LLMs, their applications in event extraction, and the challenges they face. Additionally, it delves into the significance of data augmentation and few-shot learning, evaluating their impact on model performance. The integration of NLP with related fields, such as information retrieval and text mining, is analyzed for its contribution to advancing AI capabilities. The survey concludes by suggesting future research directions, including the development of large-scale diffusion models, enhancements in multilingual event extraction, and the application of advanced meta-prompting techniques, to further elevate the effectiveness of NLP and its related disciplines.

1 Introduction

1.1 Interconnected Fields in AI and Computational Linguistics

The integration of artificial intelligence (AI) and computational linguistics is crucial for tackling complex challenges across diverse domains. Large language models (LLMs) exemplify this synergy, particularly in automating tasks such as information extraction from real estate contracts, thereby streamlining processes that traditionally demanded significant human effort [1]. In finance, the application of AI techniques, especially LLMs, is vital for enhancing technological capabilities and predicting financial risks, highlighting their transformative potential [2].

Natural language processing (NLP) methods further illustrate the collaboration between AI and computational linguistics, particularly in policymaking, where they enrich the analysis of textual data, bridging critical knowledge gaps in political communication and decision-making [3]. This collaboration is essential for the automated extraction of socio-political events from news articles, necessitating improved methodologies and interdisciplinary efforts [4].

Moreover, AI tools demonstrate the role of these fields across various industries, advancing technological applications [5]. In NLP, text structuralization—transforming unstructured text into structured

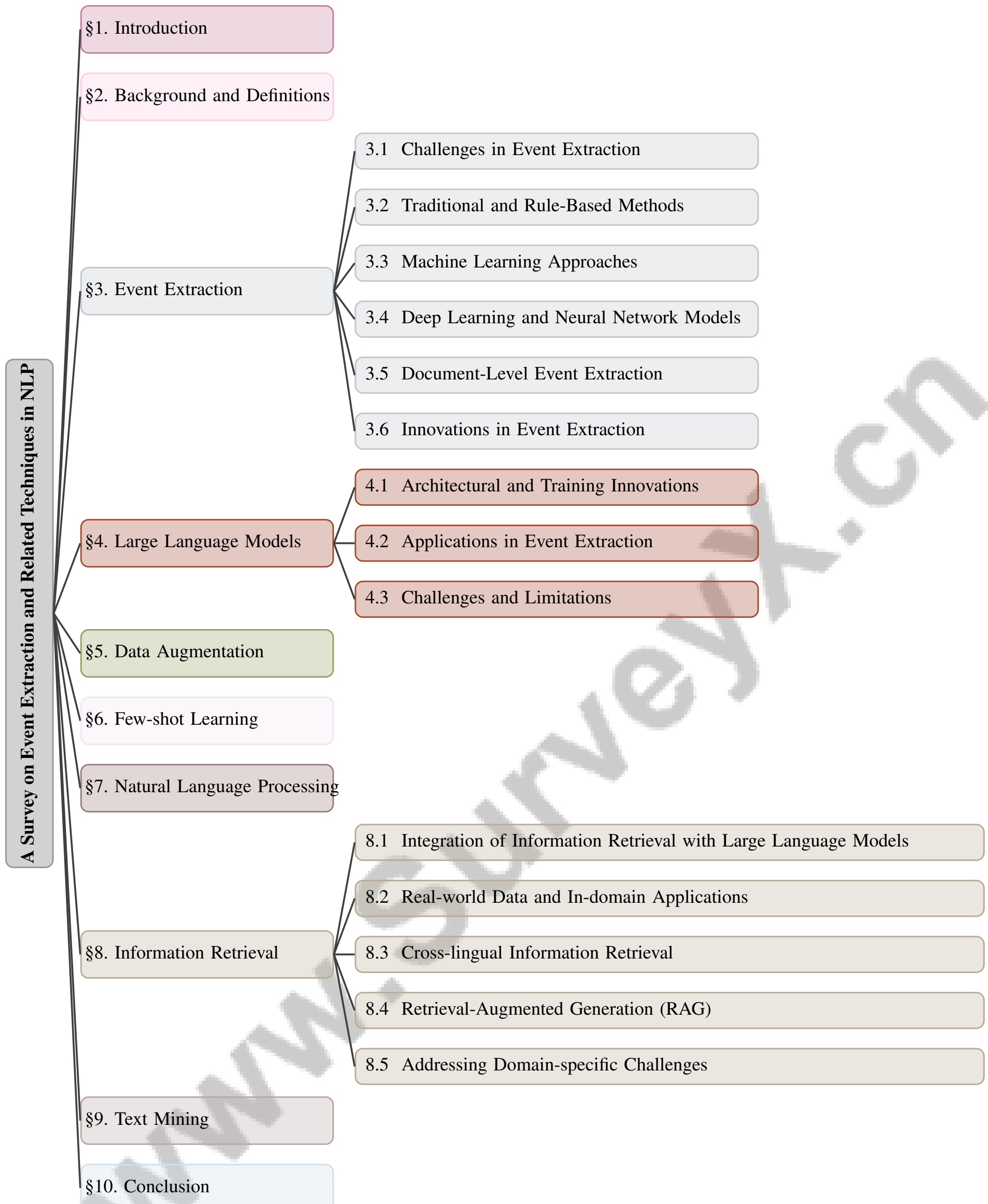


Figure 1: chapter structure

data—is vital for enhancing AI model performance [6]. This is particularly evident in named entity recognition (NER), where traditional methods often rely on extensive annotated corpora, underscoring the need for innovative approaches [7].

The collaboration between AI and computational linguistics is instrumental in developing event extraction methods, contributing to a deeper understanding and effective processing of complex data across various sectors [8].

1.2 Structure of the Survey

This survey is meticulously organized to explore event extraction and its related techniques within natural language processing (NLP) and artificial intelligence (AI). The introductory section lays the groundwork for understanding the interconnectedness of AI and computational linguistics, emphasiz-

AI-generated, for reference only.

ing their significance in addressing complex challenges. Following this, a detailed background section elucidates core concepts such as event extraction, LLMs, data augmentation, few-shot learning, NLP, information retrieval, and text mining, highlighting their interrelations.

Subsequent sections delve into the intricacies of event extraction, discussing its processes, challenges, and applications. The survey reviews both traditional and rule-based methods alongside contemporary machine learning and deep learning approaches, offering a comprehensive perspective. Innovations in document-level event extraction are also addressed, shedding light on recent advancements and ongoing challenges.

The critical role of LLMs in understanding and generating human-like language is examined, particularly their application in event extraction—a key NLP task focused on identifying and classifying events from unstructured text. This analysis considers significant challenges, such as data scarcity and imbalance, and explores novel methodologies for utilizing LLMs as expert annotators, alongside the development of high-quality datasets for 5W1H extraction, thus enhancing the performance and reliability of event extraction systems across domains like news, biomedical research, and cybersecurity [9, 10, 8]. Data augmentation techniques are also discussed for their role in improving training datasets and model robustness.

Few-shot learning is highlighted for its importance in enabling model generalization from limited examples, with discussions on techniques, evaluation, and benchmarking. The survey provides a thorough overview of NLP, detailing its primary objectives, essential tasks—including event extraction, argument extraction, and role labeling—and its integration with related fields such as automated literature review generation across various application domains, including biomedical research and historical analysis [8, 11].

Finally, the role of information retrieval in NLP is critically assessed, particularly its applications and synergy with event extraction and text mining. This exploration underscores how text mining techniques enhance event extraction's effectiveness across diverse domains, including news, biomedical fields, and cybersecurity. The survey concludes by summarizing key insights, emphasizing the interconnectedness of the fields, and proposing future research directions [12, 8].The following sections are organized as shown in Figure 1.

2 Background and Definitions

2.1 Background and Definitions

Event Extraction (EE) is a critical component of Information Extraction, aimed at identifying and categorizing events within text by detecting events and their participants and classifying them into predefined categories. This task is challenged by semantic complexities and linguistic ambiguities, further compounded by the scarcity of annotated datasets, especially in emerging or low-resource domains [8]. Addressing these challenges requires innovative methodologies [8].

Large Language Models (LLMs) have revolutionized Natural Language Processing (NLP) by significantly improving the understanding and generation of human-like language through comprehensive training on vast datasets. Despite their success in tasks like text classification and topic modeling, LLMs encounter difficulties in domain-specific applications requiring structured information extraction, highlighting the inefficiencies and high costs of relying on manually annotated datasets [13].

Data Augmentation techniques enhance the robustness and performance of NLP models by creating synthetic variations of training data, crucial for mitigating the scarcity of annotated datasets. This improves model accuracy and generalizability, particularly in fine-grained tasks such as Named Entity Recognition (NER), by addressing challenges related to insufficient data and imbalanced entity pair distributions [8, 13].

Few-shot Learning offers significant progress by enabling models to generalize from limited examples, which is particularly beneficial for extracting structured information with minimal human annotation. This is vital in complex tasks like event extraction from literary or narrative content, where unique structures and emotive elements pose substantial challenges [8].

Information Retrieval (IR) focuses on extracting relevant information from extensive datasets. Its integration with LLMs enhances retrieval accuracy and efficiency, as demonstrated by the development of entity-centric and event-centric knowledge graphs through information extraction techniques [13].

Text Mining, which involves extracting valuable information and patterns from textual data, complements event extraction by converting unstructured text into actionable insights. This is essential for applications across various domains [8]. The synergy among EE, LLMs, data augmentation, few-shot learning, NLP, IR, and text mining advances AI capabilities, enhancing their real-world applicability and underscoring the interconnectedness and importance of these fields within artificial intelligence and computational linguistics [8].

3 Event Extraction

Category	Feature	Method
Challenges in Event Extraction	Data Enhancement Techniques	InPars[14]
Traditional and Rule-Based Methods	Contextual Techniques	CPE[15]
Machine Learning Approaches	Few-Shot and Zero-Shot Capabilities	N/A[7]
	Data Synthesis and Ontology	VF-Event[16]
	Integrated Learning Strategies	CSM[17]
Deep Learning and Neural Network Models	Semantic Integration Techniques	EOIEO[18], TSAR[19]
	Objective-Driven Enhancements	sepBERT[20]
	Sequence Processing Frameworks	E2E-EE[21]
Document-Level Event Extraction	Robustness Enhancement	AEM[22]
	Event Interaction and Integration	SEBERTNets[23], DEED[24], JMEE[25], TER-MCEE[26]
	Granularity Fusion	MGR[27]
Innovations in Event Extraction	Probabilistic and Contextual Techniques	COFFEE[28], ProCE[29]
	Cross-Language and Resource Integration	IE-PRIME[30], VKI[31]
	Prompt and Generation Methods	EMRE2llm[13], PGLEE[32]

Table 1: This table presents a comprehensive summary of various methods employed in event extraction, categorized into challenges, traditional methods, machine learning approaches, deep learning models, document-level extraction, and innovative techniques. Each category is further detailed with specific features and corresponding methodologies, providing a structured overview of the current state and advancements in event extraction research. The table includes references to seminal works, highlighting the diversity and evolution of techniques addressing key challenges in the field.

In event extraction, overcoming inherent challenges is crucial for improving methodologies and outcomes. Key complexities include the necessity for high-quality annotated datasets, modeling long-distance dependencies, and integrating contextual information, all of which require a detailed examination of the specific challenges encountered in event extraction processes. Figure 2 illustrates the hierarchical structure of event extraction methodologies, challenges, and innovations. This figure categorizes the main challenges in event extraction, detailing traditional and rule-based methods, machine learning approaches, deep learning models, document-level extraction, and recent innovations. Each category is further subdivided into specific techniques, limitations, and advancements, thereby providing a comprehensive overview of the current landscape and future directions in event extraction research. Table 1 provides a detailed summary of the methodologies and techniques utilized in event extraction, categorizing them into distinct approaches such as traditional methods, machine learning, deep learning, and recent innovations. Additionally, Table 3 offers a comprehensive comparison of various methodologies employed in event extraction, focusing on their flexibility, dependency modeling capabilities, and adaptability across domains. Such a visual representation not only enhances our understanding of the complexities involved but also underscores the importance of addressing these challenges to advance the field.

3.1 Challenges in Event Extraction

Event extraction faces significant challenges that hinder optimization and accuracy. A notable issue is the scarcity of high-quality, human-annotated datasets, especially in specialized domains like military event extraction, where overlapping events and co-reference arguments complicate the process [33]. The labor-intensive creation of domain-specific training data further restricts model efficacy [14]. Traditional methods are resource-intensive, requiring extensive annotations and task decomposition [21].

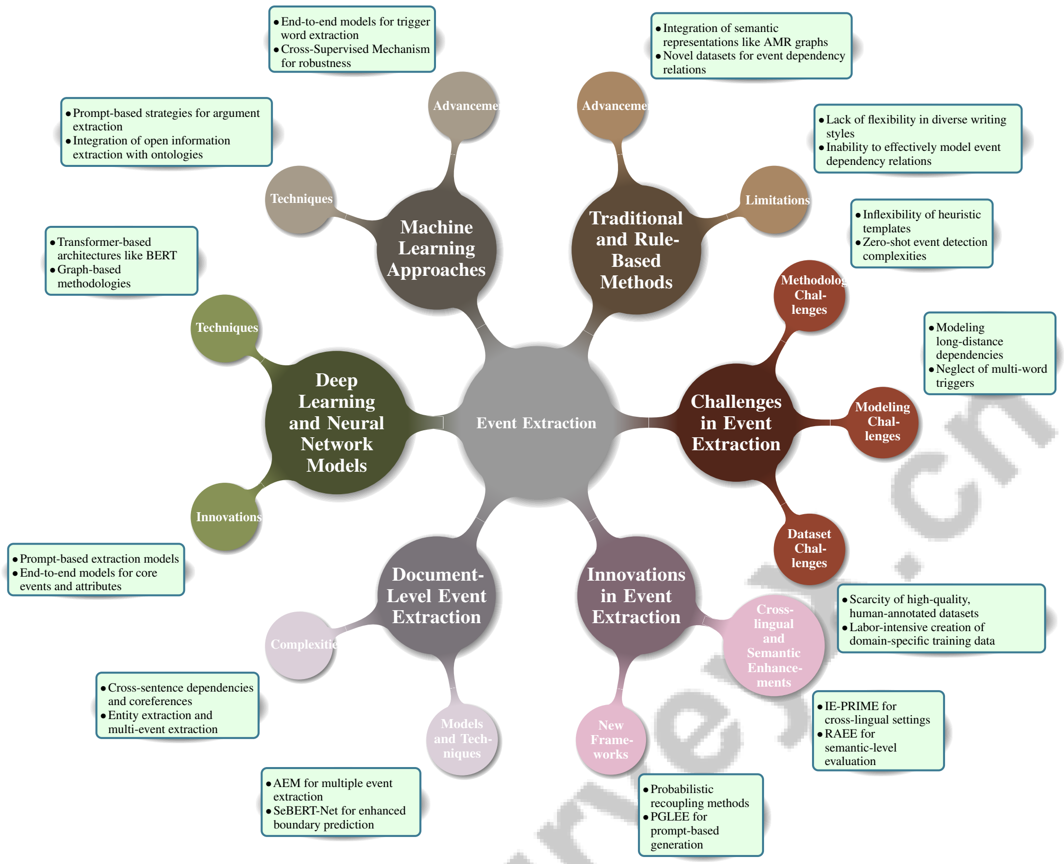


Figure 2: This figure illustrates the hierarchical structure of event extraction methodologies, challenges, and innovations. It categorizes the main challenges in event extraction, traditional and rule-based methods, machine learning approaches, deep learning models, document-level extraction, and recent innovations. Each category is further subdivided into specific techniques, limitations, and advancements, providing a comprehensive overview of the current landscape and future directions in event extraction research.

Modeling long-distance dependencies between triggers and arguments while excluding irrelevant context is a major challenge [19]. Current methods often fail to leverage deeper linguistic insights about verbs, critical for nuanced event understanding [31]. The focus on single-word triggers, neglecting multi-word triggers, limits model effectiveness [20].

Heuristic templates and extensive manual annotations limit adaptability to diverse text structures, leading to inaccurate event identification and categorization [28]. Existing methods struggle to model relationships among multiple occurrences of the same triggers and arguments, resulting in ambiguous interpretations [29].

Zero-shot event detection introduces complexities, as models must identify unseen event types based on definitions, necessitating benchmarks for this challenge [34]. Inconsistent task settings and evaluation methods across benchmarks undermine performance assessments [35].

Sequential prediction inefficiencies, particularly in longer texts, degrade performance due to error propagation from candidate spans [36]. Existing benchmarks often lack annotated historical datasets, limiting model evaluation in historical event extraction [37]. Accurate evaluation of event extraction results is complicated by the need to focus on the semantic correctness of predicted triggers and arguments, inadequately addressed by current benchmarks [38].

These challenges necessitate innovative approaches to enhance event extraction accuracy and comprehensiveness. Figure 3 illustrates the primary challenges faced in event extraction, categorized

AI-generated, for reference only.

into data scarcity, modeling challenges, and evaluation complexities, highlighting the need for such innovative approaches.

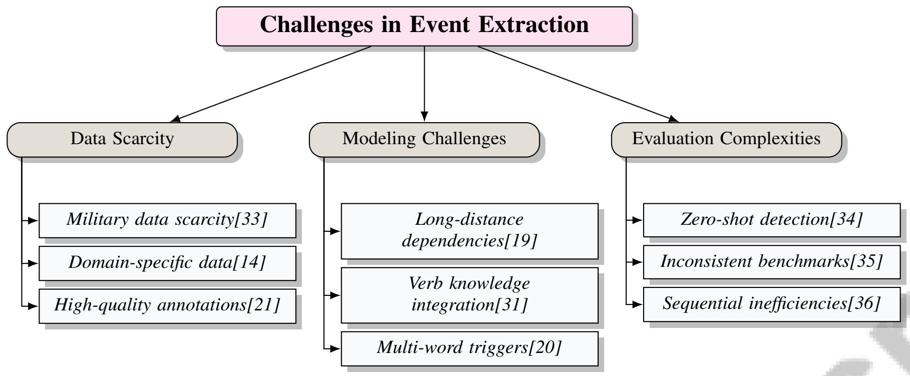


Figure 3: This figure illustrates the primary challenges faced in event extraction, categorized into data scarcity, modeling challenges, and evaluation complexities, highlighting the need for innovative approaches to enhance accuracy and comprehensiveness.

3.2 Traditional and Rule-Based Methods

Traditional and rule-based methods in event extraction rely on predefined templates and heuristic rules to identify and categorize events within text. These approaches, while effective in controlled environments, often lack flexibility to accommodate diverse writing styles and domain-specific language variations [15]. The rigidity of rule-based systems limits their applicability across different domains, struggling with nuanced expressions and complex sentence structures.

A significant limitation is their inability to effectively model event dependency relations, crucial for understanding event context and significance. Recent advancements, such as novel datasets distinguishing required and optional arguments in event dependency relations, aim to address these shortcomings [39]. This underscores the need for sophisticated datasets to enhance rule-based methods in capturing intricate event relationships.

Integrating semantic representations, like Abstract Meaning Representation (AMR) graphs, into traditional frameworks has improved event argument extraction. The TSAR model exemplifies this by encoding text through dual perspectives and using AMR graphs for a richer semantic understanding, overcoming some limitations of purely rule-based systems [19]. This innovative use of semantic structures signifies an evolution in traditional methods, offering pathways toward more adaptive and context-aware event extraction processes.

3.3 Machine Learning Approaches

Method Name	Learning Techniques	Domain Adaptability	Integration Approaches
E2E-EE[21]	Seq2seq Model	Low-resource Environments	Trigger Word Extraction
CSM[17]	Heterogeneous Information Networks	Low-resource Environments	Cross-supervised Mechanism
PGLEE[32]	Prompt-based Model	Different Domains	Prompt-based Generation
N/A[7]	Large Language Models	Various Domains	Prompting And Parsing
VF-Event[16]	Few-shot Learning	Low-resource Settings	Visual Imagination Integration
EOIEO[18]	Open Information Extraction	Different Text Structures	Two-phase Process
VKI[31]	Adapter Modules	Different Domains	Trigger Identification

Table 2: Comparison of machine learning methods for event extraction, detailing their learning techniques, domain adaptability, and integration approaches. The table highlights the diversity of strategies, ranging from sequence-to-sequence models to large language models, and their application in various domains and environments.

Machine learning approaches have advanced event extraction by enabling models to learn complex patterns from data, reducing reliance on manual feature engineering and enhancing adaptability across domains. Integrating trigger word extraction and event attribute identification into a single end-to-end model exemplifies this advancement, improving extraction efficiency [21]. Table 2 provides

AI-generated, for reference only.

a comprehensive comparison of various machine learning approaches utilized in event extraction, illustrating their distinct learning techniques, adaptability across domains, and integration strategies.

The Cross-Supervised Mechanism (CSM) illustrates machine learning's power by using co-occurrence patterns from heterogeneous information networks to alternately supervise trigger and entity extraction, enhancing robustness and accuracy [17]. Prompt-based strategies, like PAIE, optimize event argument extraction using prompt tuning [36]. The Prompt-based Graph Model for Liberal Event Extraction (PGLEE) constructs heterogeneous event graphs to generate candidate triggers and arguments, improving interaction and accuracy [32].

Large language models (LLMs) have propelled the field by demonstrating effectiveness in zero- and few-shot scenarios, as seen in the llmNER library for Named Entity Recognition (NER) tasks [7]. In low-resource environments, supervised methods like SFT and SIT outperform in-context learning (ICL), highlighting structured learning approaches' importance [40].

Few-shot learning techniques using K-shot sampling strategies are crucial in training models under low-resource and class-transfer settings, broadening event extraction models' applicability [35]. Innovative approaches, like the VF-Event model synthesizing images from text for event detection, address domain-specific challenges in few-shot learning contexts [16].

Integrating open information extraction (OIE) with ontologies provides a semantic layer enhancing interpretability and accuracy of extracted events [18]. Verb Knowledge Injection (VKI) enriches pretrained language models by incorporating verb knowledge from lexical resources, improving event extraction performance [31].

Diverse neural architectures, including BiLSTM, BiGRU, ConvBiLSTM, BERT, RoBERTa, and DistilBERT, underscore the variety of machine learning techniques in event extraction, each contributing unique strengths [41]. These advancements expand event extraction capabilities, facilitating precise and comprehensive methodologies across diverse applications.

3.4 Deep Learning and Neural Network Models

Deep learning and neural network models have significantly advanced event extraction by providing sophisticated frameworks for handling complex linguistic structures in text data. Transformer-based architectures, like BERT, offer robust mechanisms for capturing intricate semantic relationships through sequence labeling tasks, enhancing event detection precision [20]. The sepBERT model employs token-level and sentence-level objectives, improving overall event extraction performance.

Integrating Abstract Meaning Representation (AMR) into neural network models refines argument extraction processes. The TSAR model incorporates a two-stream encoding mechanism and an AMR-guided interaction module, enhancing event argument extraction across sentences by providing deeper semantic understanding [19]. This emphasizes semantic frameworks' importance in improving event extraction accuracy.

Graph-based methodologies have gained prominence, with models like TSAR leveraging graph structures for richer semantic interactions. Ontology-based event modeling in open information extraction (OIE) frameworks allows for extracting relationships and entities from text, enriching contextual understanding [18].

Recent innovations in prompt-based extraction models, such as PAIE, introduce flexible role prompts and a bipartite matching loss to optimize span assignments, distinguishing them from traditional single-role extraction methods [36]. This flexibility enhances adaptability to various event extraction scenarios, particularly in low-resource environments.

End-to-end models streamline event extraction by simultaneously outputting core events and attributes. The seq2seq model used in clinical event extraction exemplifies this, employing special tokens to format outputs and improve efficiency [21]. This highlights seq2seq architectures' potential in automating complex extraction tasks.

Advancements in deep learning and neural network models push event extraction boundaries, offering promising avenues for future research and application. These architectures enhance performance and versatility, enabling effective identification and analysis of event-related information across diverse fields, including news, biomedical research, and cybersecurity, addressing challenges like data imbalance and scarcity [9, 42, 8, 43].

3.5 Document-Level Event Extraction

Document-Level Event Extraction (DEE) extends beyond sentence-level methods, requiring identification and extraction of events from entire documents where arguments are distributed across sentences. This complexity, compounded by multiple interrelated events within a document, necessitates techniques capturing cross-sentence dependencies and coreferences [24]. DEE involves sub-tasks like entity extraction, event type judgment, and multi-event extraction, forming a comprehensive information retrieval problem [26].

Key DEE advancements include models like AEM, extracting multiple events from longer texts without needing a predetermined event count, essential for processing documents with undefined event numbers and nature [22]. Methods identifying event-specific role fillers across sentences enhance event representation accuracy within broader document contexts [27].

The SeBERT-Net model integrates global and sequence semantic information, enhancing event entity boundary prediction accuracy by combining sequence-level insights with document-level semantics [23]. This integration captures nuanced relationships and dependencies characterizing document-level events. Joint modeling techniques, as demonstrated by the JMEE framework on the ACE 2005 dataset, effectively extract multiple events and roles, showcasing joint approaches' potential in handling DEE intricacies [25].

Challenges remain, particularly regarding training data availability in languages other than English, like Spanish and Portuguese, limiting model performance in multilingual contexts [44]. Developing DEE-specific benchmarks is crucial for evaluating model performance across diverse scenarios and driving further innovation [45]. These benchmarks provide a standardized framework for assessing models' abilities to manage document-level event extraction complexities, paving the way for continued refinement and advancement.

3.6 Innovations in Event Extraction

Recent innovations in event extraction have advanced methodologies' precision and scope through novel techniques and frameworks. Probabilistic recoupling and regularization methods effectively address event extraction entanglement issues, enhancing model performance and accuracy [29]. These methods provide a robust framework for disentangling complex event interactions, crucial for accurate extraction.

The PGLEE framework integrates a prompt-based generation model with heterogeneous event graphs, allowing simultaneous event extraction and schema induction without predefined templates [32]. This approach enhances adaptability and efficiency, particularly where traditional methods fall short.

The CMNEE benchmark introduces a two-stage, multi-turn annotation strategy, including co-reference argument annotations, often missing in existing benchmarks [33]. This comprehensive approach enables more nuanced and accurate event extraction, particularly in document-level contexts where argument co-reference is critical.

In cross-lingual settings, the IE-PRIME framework improves trigger and argument detection and classification, especially in low-resource and zero-shot scenarios [30]. This highlights cross-lingual methodologies' importance in expanding event extraction technologies' applicability to diverse linguistic contexts.

The EusIE dataset for Basque introduces typological features exploring linguistic characteristics' effects on transfer quality, offering broader linguistic features than previous benchmarks [46]. This dataset is pivotal in understanding typological differences' impact on event extraction, providing insights into models' adaptability across languages.

The RAEE framework introduces a semantic-level evaluation using large language models (LLMs) to assess event extraction results, offering a more nuanced understanding of model performance than traditional benchmarks [38]. This semantic evaluation framework enhances interpretability and reliability of event extraction outcomes.

The COFFEE framework demonstrates superior performance in event extraction tasks without relying on oracle information, indicating effective extraction can be achieved solely based on context [28]. This context-driven approach simplifies extraction while maintaining high accuracy.

In historical NLP, framing event extraction as an extractive QA task within a multilingual dataset addresses low-resource setting challenges [37]. This facilitates historical event extraction, expanding event extraction scope to underrepresented temporal contexts.

Data augmentation techniques, like Flexible Threshold Loss (FTL) and Multimodal Relation Data Augmentation (MRDA), consistently improve model performance by addressing data scarcity and imbalanced distributions, especially in multilingual tasks [13]. These techniques underscore synthetic data's value in bolstering event extraction models' robustness and accuracy.

Integrating linguistic resources into NLP models, demonstrated by the Verb Knowledge Injection method, significantly enhances language models' performance in event extraction tasks across multiple languages [31]. This highlights leveraging linguistic knowledge's importance to improve model capabilities.

Advancements in event extraction research, highlighted by recent studies, underscore its evolving landscape and potential for innovative applications across fields like news media, biomedical research, history, and cybersecurity. These innovations encompass traditional document-level extraction and introduce cross-document methodologies, integrating information from multiple sources to enhance event detection comprehensiveness and accuracy. By providing a detailed overview of state-of-the-art techniques, common challenges, and future research directions, this work paves the way for further exploration and practical implementation of event extraction technologies in diverse domains [12, 8, 45]. The continued evolution of these methodologies promises to enhance event extraction technologies' precision and applicability.

Feature	Traditional and Rule-Based Methods	Machine Learning Approaches	Deep Learning and Neural Network Models
Flexibility	Limited	High	High
Dependency Modeling	Ineffective	Enhanced	Sophisticated
Cross-Domain Adaptability	Low	Improved	Versatile

Table 3: This table provides a comparative analysis of three major methodological approaches in event extraction: traditional and rule-based methods, machine learning approaches, and deep learning and neural network models. It highlights key attributes such as flexibility, dependency modeling, and cross-domain adaptability, illustrating the strengths and limitations of each approach. This comparison is critical for understanding the evolution and current state of event extraction technologies.

4 Large Language Models

Large language models (LLMs) have emerged as pivotal tools in natural language processing, significantly impacting various applications, including event extraction. This section explores the architectural and training innovations that have elevated LLMs within computational linguistics, providing insights into their enhanced capabilities and broader implications.

4.1 Architectural and Training Innovations

Recent advancements in LLM architecture and training have markedly improved language comprehension and generation. The CLEVE framework exemplifies this by integrating a text encoder for semantics with a graph encoder for structural modeling, enhancing both supervised and unsupervised event extraction [47]. Architectural innovations such as adaptive integration of LLM predictions with classical machine learning outputs and multi-granularity readers, leveraging BERT for enhanced contextual understanding, illustrate progress in accuracy and semantic correlation [48].

LLMs also advance anomaly detection through supervised fine-tuning (SFT) and in-context learning (ICL), learning from labeled data and prompts to detect anomalies across datasets [49]. Evaluating model resilience by introducing dataset noise further underscores the importance of robustness [50]. The adoption of transformer models like BERT and RoBERTa has been crucial for evaluating data augmentation techniques, thereby enhancing model robustness [51]. The PAIE framework introduces a novel prompt tuning method for joint event argument extraction, accommodating multiple arguments per role without heuristic thresholds [36].

These architectural and training innovations significantly enhance LLM performance and adaptability, facilitating complex natural language processing tasks. Automated data augmentation, improved text structuralization, and tailored instructions address scalability and consistency challenges, leading to superior NLP outcomes [52, 53, 6, 54].

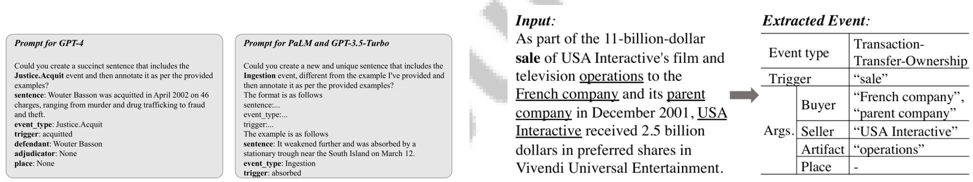
4.2 Applications in Event Extraction

LLMs have transformed event extraction methodologies, offering sophisticated processing of complex linguistic structures. In task-oriented dialogue systems, they enhance argument relationship modeling through iterative question-answer interactions, improving event extraction accuracy [55]. The COFFEE framework exemplifies this application by utilizing a generator for event extraction and a context-refining selector [28].

Fine-tuned LLMs outperform models like ChatGPT in extracting detailed 5W1H information, showcasing comprehensive event understanding [10]. Generative models in Monte Carlo-based approaches manage language and semantic complexities, enhancing event extraction [56]. In low-resource scenarios, LLM adaptability is demonstrated through few-shot learning with structured prompts, improving performance in specialized domains [57]. Multilingual contexts benefit from models like XLM-RoBERTa, which extract event relationships across languages without parallel training data.

The EMRE2llm framework highlights LLM utility in multimodal tasks, combining small model guidance with LLM knowledge for flexible task adaptation [13]. Experiments with models like TabEAE achieve state-of-the-art performance in event argument extraction by leveraging event co-occurrences [58]. LLMs also generate synthetic training data for information retrieval, offering valuable training resources [14]. Incorporating explicit verb knowledge into language models enhances reasoning about events [31].

These studies underscore LLMs' transformative potential in enhancing event extraction, improving precision and adaptability across domains. Innovative approaches like schema-aware augmented retrieval and document-level trigger exploration advance automated event extraction systems, optimizing knowledge extraction for situational awareness and informed decision-making [59, 60, 12].



(a) Prompt for GPT-4 and PaLM/GPT-3.5-Turbo[9]

(b) USA Interactive's Film and Television Operations Sale to French Company and Parent Company in December 2001[61]

Figure 4: Examples of Applications in Event Extraction

As illustrated in Figure 4, LLMs like GPT-4 and PaLM/GPT-3.5-Turbo demonstrate significant potential in event identification and categorization. The first example shows how these models generate concise sentences for specific event types, such as 'Justice.Acquit,' highlighting their capability to understand structured information. The second example details a real-world transaction, emphasizing LLMs' efficacy in extracting and organizing event-specific information [9, 61].

4.3 Challenges and Limitations

LLMs face several challenges affecting their effectiveness in natural language processing tasks. Development and maintenance costs pose scalability issues [62]. The reliance on extensive labeled datasets limits adaptability to new event types due to their costly production [8]. Integrating LLMs with existing systems often requires unavailable external features, hindering fully end-to-end solutions [63].

Bias in training data significantly impacts detection accuracy and reliability [49]. Over-correction in sentence structures by LLMs can degrade performance, especially in noisy text [50]. Limited diversity in initial datasets and biases in paraphrasing hinder benchmarks from fully capturing natural language complexities [51]. Maintaining and updating document indices during training poses computational

AI-generated, for reference only.

challenges, affecting efficiency [64]. Variability in LLM performance across languages, linked to typological features, indicates a need for refined cross-linguistic models [46].

These challenges highlight the need for ongoing innovation in LLM methodologies, particularly for automating complex tasks like literature review generation and data augmentation. Leveraging diverse NLP techniques, including retrieval-augmented generation (RAG) with LLMs, can enhance effectiveness across tasks. For instance, LLM GPT-3.5-turbo has excelled in generating literature reviews, achieving the highest ROUGE-1 score. Exploring data augmentation strategies with LLMs reveals transformative opportunities and challenges, emphasizing the importance of high-quality data and innovative training paradigms to meet NLP application demands [65, 11].

5 Data Augmentation

5.1 Introduction to Data Augmentation

Data augmentation is pivotal in natural language processing (NLP) for enhancing model robustness and generalization by synthetically generating variations of training data. This technique addresses challenges posed by language's discrete nature, which complicates traditional augmentation strategies prevalent in computer vision. Innovations like back-translation and generative modeling have improved data diversity, especially in low-resource settings and few-shot tasks. Despite its potential, data augmentation in NLP remains underexplored, with further research needed to integrate these techniques with large-scale pretrained models effectively [66, 67]. It mitigates data scarcity and class imbalance issues, reducing the manual effort in dataset preparation.

Incorporating large language models (LLMs) into data augmentation has introduced novel methods for generating augmented datasets. The Self-LLMDA framework, for example, automates task-specific augmentation instructions, enhancing model training efficiency [54]. Techniques like label-flipping, which alters labels in augmented data, improve learning in few-shot scenarios by providing varied examples.

LLMs facilitate training data generation for relation extraction tasks through structured prompts, enabling domain-specific annotations and expanding model applicability in specialized contexts. Data augmentation is particularly beneficial in low-resource environments, generating synthetic training data and reducing reliance on extensive manual annotations. Techniques such as label-conditioned word replacement and generative models significantly enhance model performance, even when traditional methods fall short [66, 67, 68, 69]. For instance, generating numerous positive training examples from a limited set of annotated pairs can expand datasets and improve accuracy.

As demand for diverse training data grows with large-scale neural networks and low-resource domain exploration, data augmentation becomes essential in NLP. It enhances model performance and adaptability by creating varied examples without extensive data collection. Recent advancements, particularly LLM use, have improved augmented data quality, facilitating effective training across diverse NLP tasks. However, data augmentation in NLP remains an area ripe for exploration, with opportunities for future research to address challenges and refine methodologies [66, 70, 67, 54]. The ongoing evolution of this field, especially through LLM integration, holds promise for enhancing NLP systems' capabilities in managing diverse datasets.

5.2 Challenges in Data Augmentation

Data augmentation in NLP faces challenges affecting synthetic datasets' efficacy. Traditional methods often yield limited gains or degrade performance in complex tasks [67], exacerbated by initial datasets' lack of diversity, which introduces biases and restricts variability critical for robust training [51].

Manual augmentation instruction creation is labor-intensive, leading to inconsistent augmented dataset quality [54], hindering scalability and real-world complexity capture necessary for effective training [71]. Multilingual language models' reliance limits language support, impacting performance consistency across linguistic contexts [72].

Defending against adversarial examples is crucial, as augmented data must maintain semantic integrity and syntactic correctness [71]. The scarcity of high-quality annotated data further constrains advanced techniques' application, especially in non-English languages, where larger model fine-tuning is rare [51].

These challenges underscore the need for ongoing research and innovation in augmentation methodologies to enhance robustness and applicability across NLP tasks, particularly with increasing reliance on large-scale neural networks and low-resource domain exploration. Existing methods often yield marginal improvements or degrade performance, highlighting the necessity for more effective strategies leveraging generative modeling and pretrained models [66, 70, 67, 52].

5.3 Innovative Techniques and Frameworks

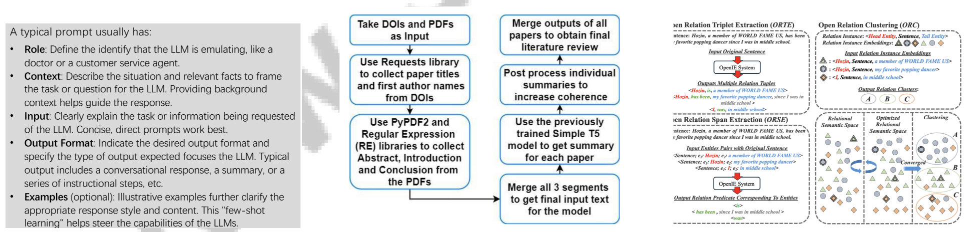
Recent advancements in data augmentation have introduced innovative techniques and frameworks leveraging large language models (LLMs) to enhance training dataset creation, labeling, and transformation. The Mask-then-Fill method exemplifies this by masking sentence segments and infilling them with variable-length text spans generated by a fine-tuned model, enriching training data diversity [73]. This showcases LLMs' potential in generating contextually relevant and varied data augmentations.

The denoised structure-to-text augmentation framework (DAEE) employs a knowledge-based generation model with reinforcement learning to select effective samples, addressing noise and enhancing robustness [74]. This framework highlights the importance of integrating knowledge-based techniques to elevate augmentation quality.

The Self-LLMDA framework introduces a dual approach, generating diverse augmentation instructions and using a scoring model to select the most relevant for specific tasks, enhancing data quality by tailoring strategies to unique NLP tasks [54].

FlipDA, focusing on label-flipping, contrasts with traditional methods aiming to preserve labels. By altering labels, FlipDA introduces variability, challenging models to adapt to diverse scenarios and enhancing learning efficiency [67].

Feng et al. categorize data augmentation techniques based on methodological approaches, providing a comprehensive overview of existing research and emphasizing the need for unified frameworks addressing multimodal challenges and leveraging self-supervised learning [66]. This categorization serves as a foundation for future research aimed at developing more cohesive and adaptable strategies.



(a) A Typical Prompt for Language Models[75]

(b) Flowchart for Literature Review Generation[11]

(c) Open Relation Extraction and Clustering[76]

Figure 5: Examples of Innovative Techniques and Frameworks

As illustrated in Figure 5, data augmentation is crucial in AI, enhancing model performance by expanding training datasets. The figure showcases innovative techniques and frameworks exemplifying diverse applications. The first subfigure, "A Typical Prompt for Language Models," summarizes effective prompting for language models, emphasizing context, input/output formats, and prompt examples, laying the groundwork for structuring data to enhance learning. The second subfigure, "Flowchart for Literature Review Generation," presents a systematic approach to automating literature reviews, illustrating how data augmentation can streamline academic research tasks. Lastly, the "Open Relation Extraction and Clustering" subfigure depicts a sophisticated method for extracting and clustering relational data using the OpenIE system. Collectively, these examples underscore data augmentation techniques' transformative potential in enhancing AI systems' capabilities across domains [75, 11, 76].

AI-generated, for reference only.

6 Few-shot Learning

6.1 Concept and Challenges of Few-shot Learning

Few-shot learning (FSL) is a crucial paradigm that enables machine learning models to generalize from minimal training examples, proving invaluable in domains where large labeled datasets are impractical to obtain, such as those requiring domain-specific insights [64]. Models like llmNER leverage FSL to reduce dependency on extensive annotated corpora, thus enhancing generalization in low-resource settings where conventional models may struggle [7, 16].

However, FSL presents challenges primarily due to small sample sizes, which can lead to sample bias and hinder generalization to new event types [35]. Large language models (LLMs) often lack the task-specific capabilities needed in few-shot scenarios, as they require extensive prior knowledge for effective performance [40]. Additionally, selecting effective prompts and hyperparameters for LLMs in few-shot contexts is critical for realistic evaluations [36].

Class imbalance in text classification further complicates FSL, as models struggle with predicting minority classes due to limited training data. Innovative strategies are necessary to balance class distribution and improve prediction accuracy [67]. Moreover, existing few-shot information extraction methods typically focus on positive examples, highlighting the need for approaches incorporating both positive and negative instances to enhance model robustness [49].

Bias in model outputs remains a concern, affecting result reliability, especially when human annotators evaluate outputs based on faithfulness and coherence [77]. Current data augmentation methods often fail to produce samples with diverse sentence structures, limiting their effectiveness in few-shot text classification [13]. Despite these challenges, FSL continues to advance, particularly in low-data environments, where novel methodologies have shown improved performance and robustness in event extraction tasks. Future research should refine these methodologies, integrate LLMs with databases, and address data privacy concerns to broaden the applicability and effectiveness of FSL across diverse domains.

6.2 Techniques for Enhancing Few-shot Learning

Enhancing few-shot learning involves diverse strategies that employ various learning paradigms to improve model performance with limited data. Retrieval-augmented language models, like Atlas, demonstrate the effectiveness of retrieving relevant documents to aid in generating responses, allowing models to learn from fewer examples [64]. This approach underscores the potential of combining retrieval mechanisms with language models to enhance learning efficiency and accuracy in few-shot contexts.

The LLM-AD framework utilizes both supervised fine-tuning and in-context learning to classify log entries as normal or anomalous, highlighting the importance of employing multiple learning strategies to boost model adaptability [49]. This dual approach fosters a more robust classification process, particularly in scenarios where traditional models face challenges due to data scarcity.

A unified framework by Ma et al. consolidates various prototype-based methods, establishing a new baseline that significantly outperforms existing techniques [35]. This benchmark emphasizes the necessity for innovative methodologies that integrate prototype-based learning to enhance few-shot learning outcomes.

Future research should develop inclusive models that leverage both supervised and unsupervised learning paradigms, alongside improvements in language adaptation techniques [40]. Such advancements are essential for enhancing model adaptability and performance across diverse linguistic and contextual scenarios.

The literature discusses various innovative strategies aimed at enhancing few-shot learning capabilities across different applications. Integrating data augmentation with pre-trained language models has improved machine translation accuracy, even with minimal training examples. Retrieval-augmented models like Atlas have excelled in performing knowledge-intensive tasks with limited data, outperforming larger models with fewer parameters. Additionally, generative language models for data augmentation in specialized text classification tasks demonstrate that carefully selected seed examples can yield substantial performance gains. Collectively, these approaches highlight the unique strengths

AI-generated, for reference only.

of each method and suggest promising avenues for future research and practical applications in few-shot learning [78, 79, 80].

6.3 Evaluation and Benchmarking

Benchmark	Size	Domain	Task Format	Metric
CLIN-IE[81]	18,164	Clinical Nlp	Information Extraction	Accuracy, F1-score
DLEB[60]	8,000	Event Extraction	Template Filling	Micro-F1
InDEE-2019[82]	56,000	Disaster Management	Event Extraction	Micro-F1, Macro-F1
AMU-EURANOVA[44]	1,000	Event Information Extraction	Named Entity Recognition	macro-F1
DiscourseEE[83]	7,464	Health	Event Argument Extraction	RM$_F1, EM_F1$
CLES[45]	20,059	Event Extraction	Cross-Document Event Extraction	F1
ChatGPT-EE[84]	1,000	Event Extraction	Event Detection	F1
NED[85]	193	Narrative Understanding	New Event Detection	Precision, Recall

Table 4: This table presents a selection of representative benchmarks used for evaluating few-shot learning models across various domains. It includes details on dataset size, task format, and evaluation metrics, highlighting the diversity and specificity of tasks such as clinical NLP, event extraction, and narrative understanding. These benchmarks provide a comprehensive framework for assessing model performance in low-resource settings.

Evaluating and benchmarking few-shot learning models are crucial for assessing their performance across diverse tasks and datasets, especially where models must generalize from minimal training data. The complexity of few-shot learning, particularly in dynamic fields like machine translation and relation extraction, necessitates comprehensive evaluation methodologies that accurately gauge model performance across various contexts [78, 79, 86]. Table 4 illustrates the diverse benchmarks employed in the evaluation of few-shot learning models, emphasizing their application across different domains and task formats.

In text classification, few-shot learning models have been evaluated using benchmark datasets such as AG News, SST-2, IMDB, and Yahoo! Answers, emphasizing the importance of dataset diversity in minimizing bias and ensuring comprehensive model assessment [87]. Model performance is typically assessed using precision, recall, and F1 scores, providing a balanced view of their capabilities in handling limited labeled samples per class.

Micro-F1 scores in relation extraction tasks, particularly on datasets like TACRED, TACREV, and Re-TACRED, underscore the necessity of balancing precision and recall for accurate evaluation of few-shot learning models. This approach ensures that models maintain high precision while effectively recalling relevant instances, a vital capability for tasks involving intricate relationships between entities [43, 88].

Few-shot learning approaches in named entity recognition (NER) have been extensively evaluated using benchmark datasets including CoNLL 2003, OntoNotes 5.0, MIT Movie, and FEW-NERD. Evaluations across various shot settings—5-shot, 10-shot, 15-shot, and 20-shot—demonstrate the effectiveness of different few-shot learning techniques, such as the Mining Undefined Classes from Other-class (MUCO) model, which enhances classifier performance by leveraging semantics from other-class words, and innovative data augmentation strategies that improve model robustness in low-resource conditions [87, 69]. These evaluations highlight the adaptability of few-shot learning models to different data constraints and the critical need for comprehensive benchmarking to capture the nuances of model performance.

The unified baseline proposed by Ma et al. marks a significant advancement, outperforming existing methods with an average F1 gain of 2.7

Challenges related to effective prompt selection in true few-shot contexts have been highlighted by findings indicating that models often perform worse than random selection, emphasizing the critical role of prompt engineering in few-shot learning. Addressing sampling bias and outliers in few-shot event detection tasks requires innovative methodologies, which have shown significant performance improvements. For instance, recent research has demonstrated the effectiveness of modeling relationships between training tasks using cross-task prototypes and enforcing prediction consistency among classifiers, leading to enhanced robustness against outliers. This approach mitigates the limitations of traditional supervised learning, which often necessitates extensive labeled

AI-generated, for reference only.

data, while improving the model's ability to detect novel event types with minimal examples, as evidenced by consistent enhancements across multiple few-shot learning datasets [35, 89, 90, 8, 91].

Future research should explore the integration of few-shot learning techniques in n-ary relation extraction, leveraging the capabilities of large language models (LLMs) alongside traditional relation extraction methods. This exploration could lead to advanced models that enhance precision in argument identification by effectively combining in-context learning and schema-constrained data generation. Additionally, refining these models to address current limitations—such as the lack of task-specific capabilities in LLMs and insufficient prior knowledge in traditional methods—may significantly improve performance in low-resource scenarios and facilitate better domain adaptation for relation extraction tasks [92, 79, 86, 57]. Dynamic learning of proxy nodes and methods to capture inter-sentence relations in document-level multi-event extraction represent promising areas for further investigation.

The evaluation and benchmarking of few-shot learning models are crucial for advancing the field, as they provide a structured framework for understanding the capabilities of different models and informing future research directions. This process is particularly significant in contexts such as machine translation, where models must adapt to evolving vocabulary with minimal examples, and in relation extraction tasks, where integrating traditional methods with large language models can enhance performance. By systematically assessing various fine-tuning and data augmentation strategies, researchers can identify effective methodologies that improve model accuracy and adaptability, ultimately leading to more robust and versatile few-shot learning approaches [93, 79, 78, 92, 86].

7 Natural Language Processing

7.1 Goals and Importance of NLP

Natural Language Processing (NLP) is integral to artificial intelligence (AI), enabling machines to comprehend and generate human language, thus facilitating natural human-machine interactions and enhancing user experiences across applications. It plays a crucial role in human-computer interaction, automating content analysis, and extracting insights from unstructured data [3]. NLP aims to bridge human language and machine understanding, enhancing tasks like sentiment analysis, machine translation, and information retrieval, which are vital for applications ranging from chatbots to data analytics [1]. The integration of NLP with large language models (LLMs) automates complex processes such as information extraction and summarization, reducing the need for human intervention [2].

NLP facilitates multilingual content translation, breaking language barriers and supporting global communication [4, 46]. Its significance extends to healthcare, finance, and social media, where it analyzes textual data to derive actionable insights. For instance, in healthcare, NLP extracts information from clinical notes to support decision-making, while in finance, it analyzes market trends to predict risks [21, 2].

The goals of NLP revolve around developing intelligent systems for language processing, enhancing AI technologies, and driving innovation across sectors. As NLP evolves, its integration with fields like information extraction, data augmentation, and LLMs will likely catalyze sophisticated applications, improving efficiencies and accessibility while addressing the challenges of language data's discrete nature and the need for extensive training datasets [11, 66, 5, 88, 70].

7.2 Key NLP Tasks

NLP encompasses tasks essential for enabling machines to process and understand human language, including sentiment analysis, named entity recognition (NER), machine translation, and information retrieval. In event extraction, tasks such as event detection, argument extraction, and role labeling enhance methodologies across domains like news and cybersecurity, improving information retrieval and understanding [60, 12, 8, 26, 43].

Sentiment analysis determines text sentiment—positive, negative, or neutral—useful in social media monitoring and customer feedback analysis, where public sentiment informs decisions. In event extraction, it provides an emotional tone layer to events [3]. NER identifies and classifies entities like names and locations, vital for ensuring relevant entities are accurately contextualized within event narratives [7].

Machine translation, crucial for cross-lingual applications, ensures accurate event extraction across languages, enhancing global applicability [46]. Information retrieval involves obtaining relevant data from large datasets, facilitating the extraction of pertinent events from unstructured data, thus enhancing extraction models' accuracy [14].

Foundational tasks like event detection, argument extraction, and role labeling are crucial for advancing NLP methodologies, enabling accurate and contextually rich event extraction across domains [43, 12, 8].

7.3 Integration with Related Fields

NLP is foundational within AI, integrating with fields like information retrieval, text mining, and information extraction. Advances in automating literature reviews using NLP and retrieval-augmented generation with LLMs tackle the challenges of vast research literature. NLP's role in transforming unstructured text into structured knowledge enhances its relevance across domains, including knowledge graph construction and recommendation systems [11, 8, 5, 6, 88]. This integration enhances NLP systems' capabilities, allowing sophisticated analysis and extraction of insights from unstructured data.

The synergy between NLP and information retrieval (IR) is evident in tasks requiring data extraction from large datasets. IR techniques filter and retrieve pertinent data, which NLP systems process to extract complex events, crucial for applications like document-level event extraction [14]. The integration with LLMs enhances retrieval accuracy, enabling precise extraction of event-centric information [13].

Text mining complements NLP by extracting useful patterns from textual data. Techniques like entity linking and sentiment analysis provide layers of understanding vital for comprehensive event extraction, transforming unstructured text into structured knowledge for decision-making in finance and healthcare [2].

Developing event-centric knowledge graphs exemplifies NLP's integration with text mining and IR, structuring information from unstructured texts to enhance understanding of relationships between events and entities, providing a comprehensive data landscape view [13]. This holistic approach is essential for applications requiring nuanced understanding, such as socio-political event extraction [4].

The integration with fields like IR and text mining significantly enhances AI systems, enabling automation of complex tasks like literature review generation and information extraction. Recent research shows how NLP techniques, combined with retrieval-augmented generation using LLMs, synthesize literature reviews from extensive datasets. Frameworks like Code4UIE illustrate LLMs' potential in transforming textual information into structured knowledge for various applications. Advances in data augmentation methods within NLP address challenges in low-resource domains, improving model performance across tasks, expanding AI systems' functional scope [11, 66, 5, 88, 70]. By leveraging each domain's strengths, this integration facilitates advanced methodologies for processing and understanding human language, driving innovation across sectors.

8 Information Retrieval

The evolution of information retrieval (IR) has been propelled by technological advancements and increasing data complexity. This section delves into the foundational aspects of IR, focusing on its integration with large language models (LLMs) to enhance data interaction and extraction. The synergy between IR and LLMs is explored, highlighting its transformative impact on natural language processing (NLP).

8.1 Integration of Information Retrieval with Large Language Models

Integrating IR with LLMs represents a significant leap in NLP, enhancing models' ability to process extensive textual data. This synergy improves information extraction accuracy and efficiency, particularly in retrieval-augmented generation, where LLMs leverage retrieval capabilities to access relevant data, thereby enhancing contextual understanding and response generation [94]. The inclusion of IR techniques within LLM frameworks facilitates effective data-model interaction, crucial for zero-shot event extraction and other complex analyses, demonstrating superior performance over traditional methods [56].

This collaboration addresses challenges in natural language understanding and generation. By utilizing LLM capabilities like in-context learning and retrieval-augmented generation, it enhances information extraction accuracy from diverse sources and automates tasks such as literature reviews and conversational information seeking. Practical applications, exemplified by frameworks like Code4UIE and toolkits such as RETA-LLM, showcase improved performance across various NLP tasks through task-specific schemas and domain-specific data [14, 11, 94, 88, 15].

8.2 Real-world Data and In-domain Applications

IR's application in real-world and in-domain scenarios enhances the accessibility and relevance of data-driven insights across sectors. In NLP, IR techniques are vital for extracting pertinent information from large datasets, aiding tasks like event extraction and sentiment analysis. The integration of IR with LLMs amplifies these capabilities, enabling precise retrieval of contextually relevant information essential for decision-making in complex domains [14].

In finance, IR processes vast textual data to analyze market trends and predict financial risks, informing strategic decisions and risk management [2]. In healthcare, IR organizes information from clinical notes and research papers, supporting medical decisions and improving patient outcomes [21].

In socio-political contexts, IR analyzes political communication by retrieving relevant data from news and social media, facilitating automated socio-political event extraction and offering insights into political dynamics and public sentiment [4]. Moreover, integrating IR with LLMs in cross-lingual settings expands applicability to multilingual contexts, enabling information retrieval across languages without parallel training data [46].

The diverse applications of IR underscore its critical role in enhancing data-driven insights. By integrating IR systems with LLMs, applications improve decision-making and stimulate innovation across fields, including automated literature reviews and information extraction, by providing accurate, contextually relevant insights [95, 94, 11, 88].

8.3 Cross-lingual Information Retrieval

Cross-lingual information retrieval (CLIR) addresses the challenge of retrieving information across languages, often without parallel corpora. Accurate translation and alignment of queries and documents are essential, requiring sophisticated machine translation and language models capable of handling diverse linguistic structures [46]. The EusIE dataset for Basque illustrates CLIR complexities by examining typological features' effects on transfer quality [46].

Techniques like multilingual embeddings and cross-lingual semantic representations align semantic spaces across languages, facilitating accurate retrieval. Integrating LLMs with CLIR systems enhances their capabilities, improving handling of linguistic diversity and relevant information extraction across language barriers [46].

Despite advancements, CLIR systems face challenges like language-specific biases and scarcity of high-quality training data for low-resource languages. Developing robust evaluation benchmarks capturing cross-lingual retrieval nuances is crucial for advancing the field and ensuring accurate, contextually relevant results [46].

Research in CLIR focuses on enhancing translation accuracy, achieving semantic alignment, and ensuring model adaptability across linguistic environments. Recent advancements include zero-shot cross-lingual information extraction, leveraging pretrained multilingual encoders, data projection, and self-training methods to boost model performance. Integrating LLMs in retrieval tasks shows promise in generating synthetic training data, outperforming traditional baselines, and facilitating better zero-shot transfer capabilities, underscoring the importance of tailored strategies for optimizing retrieval effectiveness [96, 11, 14].

AI-generated, for reference only.

8.4 Retrieval-Augmented Generation (RAG)

Retrieval-Augmented Generation (RAG) combines information retrieval with generative models to enhance content quality and relevance. This methodology leverages both retrieval systems and LLMs to produce contextually accurate outputs. By integrating retrieval mechanisms, RAG models access external knowledge bases to retrieve pertinent information, guiding the generation process. This integration is beneficial for tasks requiring deep contextual understanding, allowing models to draw from a broader information range than their pre-trained parameters [94].

RAG applications include question answering, summarization, and dialogue systems. In question answering, RAG models retrieve relevant documents or snippets, providing necessary background for precise answers to complex queries not explicitly encoded in the model's training data [94]. In summarization, RAG incorporates contextually relevant information from external sources, ensuring comprehensive outputs.

RAG enhances dialogue systems by generating coherent responses enriched with factual information from trusted sources, crucial in customer service contexts where accuracy is vital. By leveraging advanced pre-trained language models like GPT-3 and innovative in-context learning techniques, RAG effectively extracts and generates relevant information tailored to user inquiries, addressing data scarcity challenges in deep learning-based NLP and enhancing conversational systems' performance [53, 15].

RAG represents a significant advancement in NLP, offering a framework for improving generated content's accuracy and relevance. As retrieval and generation technologies advance, RAG is poised to become a crucial component in creating advanced AI systems, particularly for automating complex tasks like literature reviews and information extraction. This evolution is driven by the demand for efficient data processing, with recent research demonstrating RAG's effectiveness in leveraging LLMs to enhance performance in NLP applications, including knowledge graph construction and question-answering [14, 88, 79, 11].

8.5 Addressing Domain-specific Challenges

Information retrieval (IR) encounters domain-specific challenges affecting its effectiveness across fields. A major challenge is the scarcity of high-quality, annotated datasets tailored to specific domains, limiting model performance in specialized contexts [14]. This scarcity is pronounced in finance and healthcare, where data complexity and specificity demand models that comprehend nuanced language and terminology.

In finance, the dynamic nature of market data and real-time analysis needs challenge IR systems. Models must synthesize information from diverse sources for accurate predictions and insights, requiring robust data processing and domain-specific knowledge [2]. Similarly, in healthcare, extracting and interpreting information from clinical notes and research papers demands deep understanding of medical terminology and data integration from disparate sources [21].

Integrating IR with LLMs offers potential solutions by enhancing models' domain-specific language processing. However, reliance on large training data volumes and biases in existing datasets can affect LLM performance, necessitating more targeted datasets [46]. Adaptive models capable of learning from limited examples are critical in domains where data is scarce or costly to annotate [7].

Addressing domain-specific IR challenges requires developing specialized datasets, integrating domain knowledge into model architectures, and adopting adaptive learning techniques. These strategies will significantly improve IR systems' performance, enabling precise, contextually relevant insights across applications like automated literature reviews, event extraction, and structured knowledge generation, leveraging advanced techniques like retrieval-augmented generation with LLMs and in-context learning [11, 90, 97, 88, 43].

9 Text Mining

9.1 Techniques in Text Mining

Text mining utilizes various techniques to derive insights from unstructured text, converting it into actionable data. Central to this is Natural Language Processing (NLP), which employs algorithms for

AI-generated, for reference only.

language understanding through tokenization, part-of-speech tagging, and parsing [8]. Sentiment analysis assesses emotional tones in text, crucial for social media and customer feedback analysis, using machine learning on annotated datasets to classify sentiments as positive, negative, or neutral [3]. Named Entity Recognition (NER) identifies and categorizes entities like names and locations, essential for event extraction [7]. Clustering and topic modeling, such as Latent Dirichlet Allocation (LDA) and k-means, reveal thematic structures in large corpora [8]. Information extraction (IE) extracts structured data from text, crucial for knowledge graph construction and enhancing retrieval systems [13]. Advanced models like transformers and neural networks improve text analysis by learning complex patterns from extensive data [8]. These evolving techniques, driven by AI advancements, support applications in knowledge graph construction, biomedical research, and content generation, addressing diverse schemas and text structures [11, 90, 8, 88, 43].

9.2 Entity Linking in Text Mining

Entity linking aligns named entities in text with structured knowledge bases, transforming unstructured data into structured insights. This process disambiguates similar-named entities and captures context, essential for applications like information retrieval and event extraction [8]. It involves entity recognition, candidate generation, and disambiguation using NER techniques [7]. Advances in LLMs and machine learning enhance entity recognition and disambiguation, capturing complex semantic relationships for precise linking [13]. Semantic representations, like embeddings and graph-based models, improve disambiguation by leveraging contextual information [8]. In multilingual contexts, cross-lingual embeddings and multilingual knowledge bases ensure accurate entity associations across languages [46]. Entity linking enhances text mining by improving structured information extraction, with LLMs automating literature reviews and frameworks like Code4UIE transforming text into structured knowledge [43, 88, 53, 11].

9.3 Event Extraction as a Complementary Process

Event extraction complements text mining by structuring insights from unstructured data. It identifies and categorizes occurrences in text, facilitating event-centric knowledge graph construction and enriching data relationships [8, 13]. Event extraction clarifies semantic understanding and linguistic ambiguity, identifying triggers and arguments to provide context, crucial in socio-political analysis [4]. In multilingual contexts, it enables consistent event identification across languages using cross-lingual embeddings, expanding text mining applicability globally [46]. Integrating advanced text mining frameworks enhances insights' depth and accuracy, supporting applications like automated literature reviews and socio-political analyses, leveraging sophisticated NLP and LLMs for structured knowledge extraction [11, 90, 88, 98, 43].

9.4 Applications of Text Mining

Text mining is pivotal across domains, extracting insights from unstructured data. In healthcare, it analyzes clinical notes and EHRs, supporting decision-making and improving patient outcomes by structuring clinical data for disease prediction [21]. In finance, it processes market reports and social media feeds, informing investment strategies and risk assessments by identifying trends [2]. In socio-political analysis, it extracts events from news and social media, analyzing discourse and sentiment to provide insights into political dynamics [4]. In customer relationship management, it analyzes feedback and interactions, identifying themes to enhance service delivery [3]. Text mining aids information retrieval, managing large datasets for tasks like classification and summarization [14]. Driven by NLP and machine learning advancements, applications include automated literature review generation and retrieval-augmented generation with LLMs, addressing research article volume challenges. LLMs augment data for text classification, enhancing model performance. Event extraction, crucial in NLP, supports tasks across news, biomedical research, and cybersecurity, illustrating text mining's expanding capabilities [53, 8, 11]. The ability to extract and analyze unstructured text drives innovation and efficiency across domains, highlighting text mining's transformative impact.

10 Conclusion

10.1 Future Directions

The trajectory of natural language processing (NLP) and its associated disciplines presents numerous opportunities to enhance event extraction techniques. A primary focus for upcoming research is the advancement of large-scale diffusion language models tailored for few-shot learning, which promise to enhance model adaptability in resource-constrained environments. Moreover, integrating multimodal diffusion models that combine visual and textual elements could significantly enrich the capabilities of event extraction systems.

Broadening the CMNEE framework to include a diverse array of event types and pioneering new annotation strategies will be crucial in strengthening the robustness and variety of event extraction tasks. The joint modeling of triggers and arguments, along with the use of document-level information, is particularly advantageous for cross-lingual applications. Additionally, automating ontology updates and adapting current methods for niche domains can further enhance the flexibility and robustness of event extraction systems.

Future research should also focus on incorporating co-reference resolution as a supplementary task and leveraging entity information to improve argument boundary identification. Expanding evaluation frameworks to encompass a wider range of scenarios and criteria will allow for a more nuanced assessment of model performance.

In the realm of information retrieval, refining retrieval mechanisms, enhancing fine-tuning strategies, and integrating Atlas with other model architectures are essential for boosting retrieval accuracy and efficiency. Extending analyses to cover a broader spectrum of tasks and languages, particularly in terms of typology and cross-lingual transfer capabilities, will advance multilingual event extraction.

The ProCE framework offers potential for extending to other information extraction tasks, such as relation extraction, thereby improving generalization capabilities. Expanding the COFFEE framework to support document-level extraction and developing zero-shot methods for novel events are vital for enhancing adaptability in event extraction systems.

Future endeavors should also aim to expand benchmarks to incorporate a variety of event types and languages, while refining strategies to handle noisy historical texts. Domain-specific augmentation strategies must address the challenges posed by legal jargon and protected terms in data generation. Additionally, scaling synthetic datasets, refining question generation prompts, and incorporating negative examples into training processes will bolster event extraction effectiveness.

Optimizing resource creation methodologies and extending approaches to specialized domains with existing high-quality lexica are crucial. Enhancements to the event presence prediction task and exploration of other auxiliary objectives could further improve model accuracy. Progress in liberal event extraction and the development of complex event schemas also represent promising avenues for future research.

In the medical field, enhancing models' capabilities to recognize rare core words, distinguish similar contextual structures, and improve the pre-training process for better adaptation to medical terminology is imperative. Investigating advanced meta-prompting techniques, ensemble augmentation methods, and the application of Self-LLMDA across a wider range of LLM architectures should be prioritized.

These future directions map out a path for continued innovation and efficacy in NLP and its related fields, paving the way for the creation of more advanced and capable AI systems.

References

[1] Shiqi Yang, Yu Zhao, and Haoxiang Gao. Using large language models in real estate transactions: A few-shot learning approach. *OSF Preprints*, 2024.

[2] Yupeng Cao, Zhi Chen, Qingyun Pei, Fabrizio Dimino, Lorenzo Ausiello, Prashant Kumar, K. P. Subbalakshmi, and Papa Momar Ndiaye. Risklabs: Predicting financial risk using large language model based on multi-sources data, 2024.

[3] Zhijing Jin and Rada Mihalcea. Natural language processing for policymaking, 2023.

[4] Ali Hürriyetoğlu, Vanni Zavarella, Hristo Tanev, Erdem Yörük, Ali Safaya, and Osman Mutlu. Automated extraction of socio-political events from news (aespen): Workshop and shared task report, 2020.

[5] Hitesh Mohapatra and Soumya Ranjan Mishra. Exploring ai tool's versatile responses: An in-depth analysis across different industries and its performance evaluation, 2023.

[6] Xuanfan Ni, Piji Li, and Huayang Li. Unified text structuralization with instruction-tuned language models, 2023.

[7] Fabián Villena, Luis Miranda, and Claudio Aracena. llmner: (zero|few)-shot named entity recognition, exploiting the power of large language models, 2024.

[8] Viet Dac Lai. Event extraction: A survey, 2022.

[9] Ruirui Chen, Chengwei Qin, Weifeng Jiang, and Dongkyu Choi. Is a large language model a good annotator for event extraction? In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 17772–17780, 2024.

[10] Yang Cao, Yangsong Lan, Feiyan Zhai, and Piji Li. 5w1h extraction with large language models, 2024.

[11] Nurshat Fateh Ali, Md. Mahdi Mohtasim, Shakil Mosharrof, and T. Gopi Krishna. Automated literature review using nlp techniques and llm-based retrieval-augmented generation, 2024.

[12] Jiangwei Liu, Liangyu Min, and Xiaohong Huang. An overview of event extraction and its applications, 2021.

[13] Wentao He, Hanjie Ma, Shaohua Li, Hui Dong, Haixiang Zhang, and Jie Feng. Using augmented small multimodal models to guide large language models for multimodal relation extraction. *Applied Sciences*, 13(22):12208, 2023.

[14] Luiz Bonifacio, Hugo Abonizio, Marzieh Fadaee, and Rodrigo Nogueira. Inpars: Data augmentation for information retrieval using large language models. *arXiv preprint arXiv:2202.05144*, 2022.

[15] Patrizio Bellan, Mauro Dragoni, and Chiara Ghidini. Leveraging pre-trained language models for conversational information seeking from text, 2022.

[16] Farhad Moghimifar, Fatemeh Shiri, Van Nguyen, Reza Haffari, and Yuan-Fang Li. Few-shot domain-adaptive visually-fused event detection from text, 2023.

[17] Yue Wang, Zhuo Xu, Lu Bai, Yao Wan, Lixin Cui, Qian Zhao, Edwin R. Hancock, and Philip S. Yu. Cross-supervised joint-event-extraction with heterogeneous information networks, 2020.

[18] Sihem Sahnoun. Event extraction based on open information extraction and ontology, 2019.

[19] Runxin Xu, Peiyi Wang, Tianyu Liu, Shuang Zeng, Baobao Chang, and Zhifang Sui. A two-stream amr-enhanced model for document-level event argument extraction, 2022.

[20] Parul Awasthy, Tahira Naseem, Jian Ni, Taesun Moon, and Radu Florian. Event presence prediction helps trigger detection across languages, 2020.

[21] Wei Feng, Ruochen Huang, Yun Yu, Huiting Sun, and Yun Liu. End-to-end clinical event extraction from chinese electronic health record, 2022.

[22] Rui Wang, Deyu Zhou, and Yulan He. Open event extraction from online text using a generative adversarial network, 2019.

[23] Congqing He, Xiangyu Zhu, Yuquan Le, Yuzhong Liu, and Jianhong Yin. Sebertnets: Sequence enhanced bert networks for event entity extraction tasks oriented to the finance field, 2024.

[24] Kung-Hsiang Huang and Nanyun Peng. Document-level event extraction with efficient end-to-end learning of cross-event dependencies, 2021.

[25] Xiao Liu, Zhunchen Luo, and Heyan Huang. Jointly multiple events extraction via attention-based graph information aggregation, 2018.

[26] Qizhi Wan, Changxuan Wan, Keli Xiao, Hui Xiong, Dexi Liu, and Xiping Liu. Token-event-role structure-based multi-channel document-level event extraction, 2023.

[27] Xinya Du and Claire Cardie. Document-level event role filler extraction using multi-granularity contextualized encoding, 2020.

[28] Meiru Zhang, Yixuan Su, Zaiqiao Meng, Zihao Fu, and Nigel Collier. Coffee: A contrastive oracle-free framework for event extraction, 2024.

[29] Xingyu Bai, Taiqiang Wu, Han Guo, Zhe Zhao, Xuefeng Yang, Jiayi Li, Weijie Liu, Qi Ju, Weigang Guo, and Yujiu Yang. Recouple event field via probabilistic bias for event extraction, 2023.

[30] Steven Fincke, Shantanu Agarwal, Scott Miller, and Elizabeth Boschee. Language model priming for cross-lingual event extraction, 2021.

[31] Olga Majewska, Ivan Vulić, Goran Glavaš, Edoardo M. Ponti, and Anna Korhonen. Verb knowledge injection for multilingual event processing, 2020.

[32] Haochen Li and Di Geng. Prompt-based graph model for joint liberal event extraction and event schema induction, 2024.

[33] Mengna Zhu, Zijie Xu, Kaisheng Zeng, Kaiming Xiao, Mao Wang, Wenjun Ke, and Hongbin Huang. Cmnee: A large-scale document-level event extraction dataset based on open-source chinese military news, 2024.

[34] Zefan Cai, Po-Nien Kung, Ashima Suvarna, Mingyu Derek Ma, Hritik Bansal, Baobao Chang, P. Jeffrey Brantingham, Wei Wang, and Nanyun Peng. Improving event definition following for zero-shot event detection, 2024.

[35] Yubo Ma, Zehao Wang, Yixin Cao, and Aixin Sun. Few-shot event detection: An empirical study and a unified view. *arXiv preprint arXiv:2305.01901*, 2023.

[36] Yubo Ma, Zehao Wang, Yixin Cao, Mukai Li, Meiqi Chen, Kun Wang, and Jing Shao. Prompt for extraction? paie: Prompting argument interaction for event argument extraction, 2022.

[37] Nadav Borenstein, Natalia da Silva Perez, and Isabelle Augenstein. Multilingual event extraction from historical newspaper adverts, 2023.

[38] Yi-Fan Lu, Xian-Ling Mao, Tian Lan, Chen Xu, and Heyan Huang. Beyond exact match: Semantically reassessing event extraction by large language models, 2024.

[39] Ruiqi Li, Patrik Haslum, and Leyang Cui. Eder: A dataset for exploring dependency relations between events, 2023.

[40] Evgeniia Razumovskaia, Ivan Vulić, and Anna Korhonen. Analyzing and adapting large language models for few-shot multilingual nlu: Are we there yet?, 2024.

[41] Reshma Sheik, KP Siva Sundara, and S Jaya Nirmala. Neural data augmentation for legal overruling task: Small deep learning models vs. large language models. *Neural Processing Letters*, 56(2):121, 2024.

[42] Qian Li, Jianxin Li, Jiawei Sheng, Shiyao Cui, Jia Wu, Yiming Hei, Hao Peng, Shu Guo, Lihong Wang, Amin Beheshti, and Philip S. Yu. A survey on deep learning event extraction: Approaches and applications, 2022.

[43] Yaduo Liu, Longhui Zhang, Shujuan Yin, Xiaofeng Zhao, and Feiliang Ren. An effective system for multi-format information extraction, 2021.

[44] Léo Bouscarrat, Antoine Bonnefoy, Cécile Capponi, and Carlos Ramisch. Amu-euranova at case 2021 task 1: Assessing the stability of multilingual bert, 2021.

[45] Qiang Gao, Zixiang Meng, Bobo Li, Jun Zhou, Fei Li, Chong Teng, and Donghong Ji. Harvesting events from multiple sources: Towards a cross-document event extraction paradigm, 2024.

[46] Mikel Zubillaga, Oscar Sainz, Ainara Estarrona, Oier Lopez de Lacalle, and Eneko Agirre. Event extraction in basque: Typologically motivated cross-lingual transfer-learning analysis, 2024.

[47] Ziqi Wang, Xiaozhi Wang, Xu Han, Yankai Lin, Lei Hou, Zhiyuan Liu, Peng Li, Juanzi Li, and Jie Zhou. Cleve: Contrastive pre-training for event extraction, 2021.

[48] Yuhang Wu, Yingfei Wang, Chu Wang, and Zeyu Zheng. Large language model enhanced machine learning estimators for classification, 2024.

[49] Hongwei Jin, George Papadimitriou, Krishnan Raghavan, Pawel Zuk, Prasanna Balaprakash, Cong Wang, Anirban Mandal, and Ewa Deelman. Large language models for anomaly detection in computational workflows: from supervised fine-tuning to in-context learning, 2024.

[50] Xiaowu Zhang, Xiaotian Zhang, Cheng Yang, Hang Yan, and Xipeng Qiu. Does correction remain a problem for large language models?, 2023.

[51] Jordan J. Bird, Anikó Ekárt, and Diego R. Faria. Chatbot interaction with artificial intelligence: Human data augmentation with t5 and language transformer ensemble for text classification, 2020.

[52] Bosheng Ding, Chengwei Qin, Ruochen Zhao, Tianze Luo, Xinze Li, Guizhen Chen, Wenhan Xia, Junjie Hu, Anh Tuan Luu, and Shafiq Joty. Data augmentation using large language models: Data perspectives, learning paradigms and challenges, 2024.

[53] Huanhuan Zhao, Haihua Chen, Thomas A Ruggles, Yunhe Feng, Debjani Singh, and Hong-Jun Yoon. Improving text classification with large language model-based data augmentation. *Electronics*, 13(13):2535, 2024.

[54] Yichuan Li, Kaize Ding, Jianling Wang, and Kyumin Lee. Empowering large language models for textual data augmentation. *arXiv preprint arXiv:2404.17642*, 2024.

[55] Qian Li, Hao Peng, Jianxin Li, Jia Wu, Yuanxing Ning, Lihong Wang, Philip S. Yu, and Zheng Wang. Reinforcement learning-based dialogue guided event extraction to exploit argument relations, 2021.

[56] Erica Cai and Brendan O'Connor. A monte carlo language model pipeline for zero-shot sociopolitical event extraction, 2024.

[57] Vanni Zavarella, Juan Carlos Gamero-Salinas, and Sergio Consoli. A few-shot approach for relation extraction domain adaptation using large language models, 2024.

[58] Yuxin He, Jingyue Hu, and Buzhou Tang. Revisiting event argument extraction: Can eae models learn better when being aware of event co-occurrences?, 2023.

[59] Fatemeh Shiri, Van Nguyen, Farhad Moghimifar, John Yoo, Gholamreza Haffari, and Yuan-Fang Li. Decompose, enrich, and extract! schema-aware event extraction using llms, 2024.

[60] Shaden Shaar, Wayne Chen, Maitreyi Chatterjee, Barry Wang, Wenting Zhao, and Claire Cardie. Are triggers needed for document-level event extraction?, 2024.

[61] Xinya Du and Claire Cardie. Event extraction by answering (almost) natural questions, 2021.

[62] Xiang Wei, Yufeng Chen, Ning Cheng, Xingyu Cui, Jinan Xu, and Wenjuan Han. Collabkg: A learnable human-machine-cooperative information extraction toolkit for (event) knowledge graph construction, 2023.

[63] Trung Minh Nguyen and Thien Huu Nguyen. One for all: Neural joint modeling of entities and events, 2018.

[64] Gautier Izacard, Patrick Lewis, Maria Lomeli, Lucas Hosseini, Fabio Petroni, Timo Schick, Jane Dwivedi-Yu, Armand Joulin, Sebastian Riedel, and Edouard Grave. Atlas: Few-shot learning with retrieval augmented language models. *Journal of Machine Learning Research*, 24(251):1–43, 2023.

[65] Bosheng Ding, Chengwei Qin, Ruochen Zhao, Tianze Luo, Xinze Li, Guizhen Chen, Wenhan Xia, Junjie Hu, Luu Anh Tuan, and Shafiq Joty. Data augmentation using llms: Data perspectives, learning paradigms and challenges. In *Findings of the Association for Computational Linguistics ACL 2024*, pages 1679–1705, 2024.

[66] Steven Y Feng, Varun Gangal, Jason Wei, Sarath Chandar, Soroush Vosoughi, Teruko Mitamura, and Eduard Hovy. A survey of data augmentation approaches for nlp. *arXiv preprint arXiv:2105.03075*, 2021.

[67] A: Effective and robust data aug.

[68] Mosleh Mahamud, Zed Lee, and Isak Samsten. Distributional data augmentation methods for low resource language, 2023.

[69] Jian Liu, Yufeng Chen, and Jinan Xu. Low-resource ner by data augmentation with prompting. In *IJCAI*, pages 4252–4258, 2022.

[70] Steven Y. Feng, Varun Gangal, Jason Wei, Sarath Chandar, Soroush Vosoughi, Teruko Mitamura, and Eduard Hovy. A survey of data augmentation approaches for nlp, 2021.

[71] Yue Zhou, Chenlu Guo, Xu Wang, Yi Chang, and Yuan Wu. A survey on data augmentation in large model era. *arXiv preprint arXiv:2401.15422*, 2024.

[72] Chris Jenkins, Shantanu Agarwal, Joel Barry, Steven Fincke, and Elizabeth Boschee. Massively multi-lingual event understanding: Extraction, visualization, and search, 2023.

[73] Jun Gao, Changlong Yu, Wei Wang, Huan Zhao, and Ruifeng Xu. Mask-then-fill: A flexible and effective data augmentation framework for event extraction, 2023.

[74] bo wang, Heyan Huang, Xiaochi Wei, Ge Shi, Xiao Liu, Chong Feng, Tong Zhou, Shuaiqiang Wang, and Dawei Yin. Boosting event extraction with denoised structure-to-text augmentation, 2023.

[75] Yuanfeng Song, Yuanqin He, Xuefang Zhao, Hanlin Gu, Di Jiang, Haijun Yang, Lixin Fan, and Qiang Yang. A communication theory perspective on prompting engineering methods for large language models, 2023.

[76] Pai Liu, Wenyang Gao, Wenjie Dong, Lin Ai, Ziwei Gong, Songfang Huang, Zongsheng Li, Ehsan Hoque, Julia Hirschberg, and Yue Zhang. A survey on open information extraction from rule-based model to large language model, 2024.

[77] Ali Hürriyetoğlu. Tense, aspect and mood based event extraction for situation analysis and crisis management, 2020.

[78] Few-shot learning through contex.

[79] Gautier Izacard, Patrick Lewis, Maria Lomeli, Lucas Hosseini, Fabio Petroni, Timo Schick, Jane Dwivedi-Yu, Armand Joulin, Sebastian Riedel, and Edouard Grave. Few-shot learning with retrieval augmented language models. *arXiv preprint arXiv:2208.03299*, 1(2):4, 2022.

[80] Aleksandra Edwards, Asahi Ushio, Jose Camacho-Collados, Hélène de Ribaupierre, and Alun Preece. Guiding generative language models for data augmentation in few-shot text classification. *arXiv preprint arXiv:2111.09064*, 2021.

[81] Monica Agrawal, Stefan Hegselmann, Hunter Lang, Yoon Kim, and David Sontag. Large language models are few-shot clinical information extractors. *arXiv preprint arXiv:2205.12689*, 2022.

[82] Ayush Maheshwari, Hrishikesh Patel, Nandan Rathod, Ritesh Kumar, Ganesh Ramakrishnan, and Pushpak Bhattacharyya. Tale of tails using rule augmented sequence labeling for event extraction, 2020.

[83] Omar Sharif, Joseph Gatto, Madhusudan Basak, and Sarah M. Preum. Explicit, implicit, and scattered: Revisiting event extraction to capture complex arguments, 2024.

[84] Jun Gao, Huan Zhao, Changlong Yu, and Ruifeng Xu. Exploring the feasibility of chatgpt for event extraction, 2023.

[85] Seyed Mahed Mousavi, Shohei Tanaka, Gabriel Roccabruna, Koichiro Yoshino, Satoshi Nakamura, and Giuseppe Riccardi. Whats new? identifying the unfolding of new events in narratives, 2023.

[86] Ye Liu, Kai Zhang, Aoran Gan, Linan Yue, Feng Hu, Qi Liu, and Enhong Chen. Empowering few-shot relation extraction with the integration of traditional re methods and large language models, 2024.

[87] Meihan Tong, Shuai Wang, Bin Xu, Yixin Cao, Minghui Liu, Lei Hou, and Juanzi Li. Learning from miscellaneous other-class words for few-shot named entity recognition, 2021.

[88] Yucan Guo, Zixuan Li, Xiaolong Jin, Yantao Liu, Yutao Zeng, Wenxuan Liu, Xiang Li, Pan Yang, Long Bai, Jiafeng Guo, et al. Retrieval-augmented code generation for universal information extraction. In *CCF International Conference on Natural Language Processing and Chinese Computing*, pages 30–42. Springer, 2024.

[89] Rui Feng, Jie Yuan, and Chao Zhang. Probing and fine-tuning reading comprehension models for few-shot event extraction, 2020.

[90] Zhuo Xu, Yue Wang, Lu Bai, and Lixin Cui. Writing style aware document-level event extraction, 2022.

[91] Viet Lai, Franck Dernoncourt, and Thien Huu Nguyen. Learning prototype representations across few-shot tasks for event detection. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP 2021)*, 2021.

[92] Xin Xu, Yuqi Zhu, Xiaohan Wang, and Ningyu Zhang. How to unleash the power of large language models for few-shot relation extraction? *arXiv preprint arXiv:2305.01555*, 2023.

[93] Ethan Perez, Douwe Kiela, and Kyunghyun Cho. True few-shot learning with language models. *Advances in neural information processing systems*, 34:11054–11070, 2021.

[94] Jiongnan Liu, Jiajie Jin, Zihan Wang, Jiehan Cheng, Zhicheng Dou, and Ji-Rong Wen. Reta-llm: A retrieval-augmented large language model toolkit, 2023.

[95] Yubo Ma, Yixin Cao, YongChing Hong, and Aixin Sun. Large language model is not a good few-shot information extractor, but a good reranker for hard samples! *arXiv preprint arXiv:2303.08559*, 2023.

[96] Mahsa Yarmohammadi, Shijie Wu, Marc Marone, Haoran Xu, Seth Ebner, Guanghui Qin, Yunmo Chen, Jialiang Guo, Craig Harman, Kenton Murray, Aaron Steven White, Mark Dredze, and Benjamin Van Durme. Everything is all it takes: A multipronged strategy for zero-shot cross-lingual information extraction, 2021.

[97] Ying Mo, Jiahao Liu, Jian Yang, Qifan Wang, Shun Zhang, Jingang Wang, and Zhoujun Li. C-icl: Contrastive in-context learning for information extraction, 2024.

[98] Andrew Halterman. Geolocating political events in text, 2019.

Disclaimer:

SurveyX is an AI-powered system designed to automate the generation of surveys. While it aims to produce high-quality, coherent, and comprehensive surveys with accurate citations, the final output is derived from the AI's synthesis of pre-processed materials, which may contain limitations or inaccuracies. As such, the generated content should not be used for academic publication or formal submissions and must be independently reviewed and verified. The developers of SurveyX do not assume responsibility for any errors or consequences arising from the use of the generated surveys.

AI-generated, for reference only.