A Survey of Large Language Models Event Extraction and Semantic Role Labeling in Natural Language Processing

Abstract

Large Language Models (LLMs) have become integral to advancing natural language processing (NLP) by enhancing event extraction and semantic role labeling. This survey explores the intersection of LLMs with these tasks, emphasizing their role in transforming unstructured text into structured data for applications across diverse domains. Key findings highlight the incorporation of knowledge graphs (KGs) into LLMs, which significantly improves their ability to generate factually accurate content and enhances performance in knowledge-grounded tasks. The integration of contextual information further bolsters event extraction accuracy, underscoring the potential of LLMs as zero-shot text classifiers, particularly beneficial for users with limited text classification expertise. The survey also discusses the challenges of data scarcity and annotation, particularly in specialized domains, and the innovative approaches like reinforcement learning frameworks that address these constraints. The synergy between LLMs and event extraction methods offers promising avenues for future research, aiming to refine the models' capabilities in handling complex language tasks. The ongoing evolution in this field is expected to drive significant advancements, offering new opportunities for innovation across various applications.

1 Introduction

1.1 Significance of Human Language Processing

Human language processing is essential in artificial intelligence (AI) and linguistics, driving technological progress and deepening our understanding of human communication. Neuro-symbolic AI techniques enhance analogical reasoning in Large Language Models (LLMs), allowing them to address complex analogies that necessitate knowledge beyond mere text [1]. Multi-format information extraction systems exemplify advanced models' effectiveness in deriving meaningful insights from unstructured text, achieving high performance across diverse tasks [2]. Furthermore, evaluating and calibrating probabilistic models in natural language processing (NLP) underscores the significance of understanding uncertainty, enabling users to gauge the reliability of model outputs [3]. These advancements illustrate the integral role of human language processing in enhancing AI technology and linguistic comprehension.

LLMs, a cornerstone of NLP, exhibit remarkable capabilities in generating coherent and contextually relevant text, although challenges persist in ensuring accurate recall and application of factual knowledge crucial for knowledge-grounded tasks [4]. The construction of knowledge graphs (KGs) is vital for organizing extensive data into structured formats, facilitating applications across various domains. The significance of human language processing also extends to real-time data analysis in dynamic environments, where rapid and accurate processing of large datasets is essential [5]. Additionally, privacy and regulatory compliance in training machine learning models on sensitive data remain critical considerations influencing NLP system development [6]. Addressing these challenges

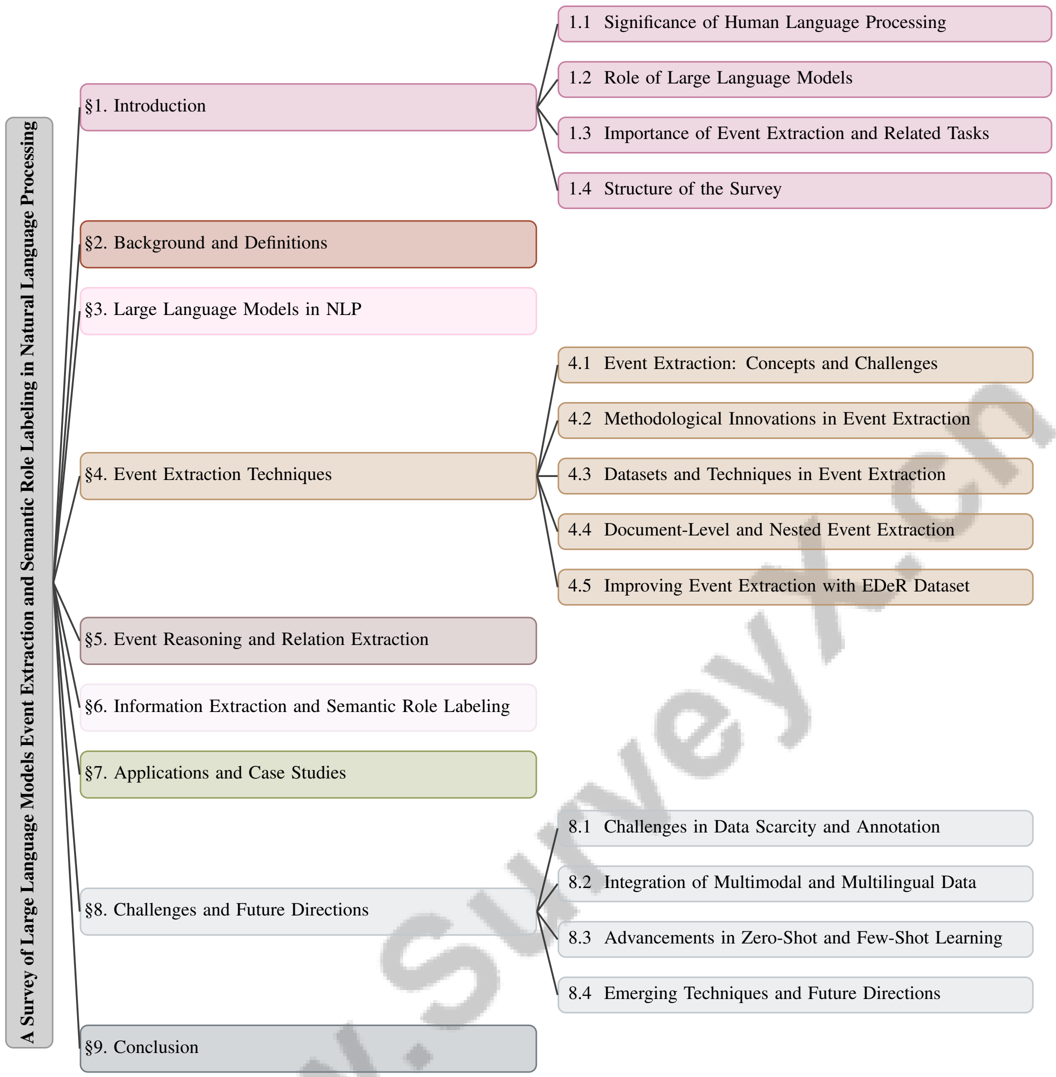


Figure 1: chapter structure

will be pivotal in advancing AI systems' capabilities to effectively understand and utilize human language.

1.2 Role of Large Language Models

Large Language Models (LLMs) significantly advance natural language processing (NLP) by enhancing human language generation and interpretation. Equipped with extensive parameters and trained on vast datasets, LLMs excel in various language tasks, including text generation, interpretation, and classification, establishing a new paradigm for automating knowledge acquisition and representation across multiple domains [7]. Their ability to manage complex analogies in unstructured text highlights their potential, with neuro-symbolic AI techniques suggested to further enhance these capabilities [1].

The performance of LLMs in causal discovery tasks is influenced by pre-training corpora, emphasizing the importance of understanding training data's impact on model capabilities [8]. In ontology learning, LLMs facilitate knowledge organization and structuring [9]. However, existing methods for extracting structured data from unstructured text are limited, necessitating approaches that leverage LLMs to overcome these challenges [10]. Deductive reasoning in natural language is crucial for LLM development, with accurate reasoning capabilities becoming increasingly vital [11]. Despite advancements, exploring reasoning biases and improving methodologies remains an active research

AI-generated, for reference only.

area. As LLMs evolve, their contributions to text generation and interpretation are anticipated to expand, offering new opportunities for innovation across various fields.

1.3 Importance of Event Extraction and Related Tasks

Event extraction (EE) is a vital aspect of natural language processing (NLP) that converts unstructured text into structured data, enabling applications like information retrieval, document summarization, and knowledge graph construction. This transformation is essential for downstream tasks such as news aggregation and event knowledge graph development, which aid in comprehending and organizing complex narratives [12]. EE encompasses subtasks including event trigger detection, entity mention detection, and argument role prediction, all contributing to comprehensive event information extraction.

The complexity of EE is heightened by the need to accommodate diverse event types and schemas present in online text, a challenge existing datasets often fail to address due to their focus on fixed event types [13]. Processing large volumes of textual data poses additional challenges, as traditional pattern-based approaches may lack accuracy, necessitating innovative methods [14]. Open domain event extraction involves extracting unconstrained event types from clusters of news reports, presenting further challenges in adaptability and scalability [15].

Traditional EE methods have primarily targeted single modalities, while contemporary journalism disseminates news through multimedia, requiring cross-media approaches for effective event capture and interpretation [16]. The dynamic nature of data patterns and network conditions poses significant obstacles, as existing methods often struggle to maintain optimal performance in changing environments [5]. Extracting events from specialized domains, such as biomedical texts, highlights EE's significance in building pathways and enriching databases [17].

EE is also critical in fields requiring precise relationship modeling, such as biomedicine, where extracting intricate relationships between entities is paramount [18]. In socio-political contexts, automated event information extraction supports decision-making processes in social sciences and policy-making. Despite advancements in deep learning, the need for extensive labeled data remains a bottleneck, with LLMs exhibiting unstable performance in tasks like named entity recognition [19]. The rise of financial crime emphasizes the necessity for specialized datasets to train and evaluate models addressing these issues [20].

The challenges of extracting structured information from unstructured data are pronounced in EE tasks [21]. In finance and economics, EE faces issues of annotation scarcity and class imbalance, which traditional supervised learning methods struggle to address [22]. As research progresses, tackling these challenges is essential for enhancing EE accuracy and efficiency across various applications. Integrating outcome supervision to improve event extraction accuracy is increasingly recognized as necessary to address limitations faced by current LLMs, such as instruction following and hallucination [23].

1.4 Structure of the Survey

This survey is systematically structured to comprehensively examine the intersection of large language models (LLMs), event extraction, and semantic role labeling within the broader field of natural language processing (NLP). It begins with an introduction that emphasizes the significance of human language processing in AI and linguistics, highlighting LLMs' pivotal role in text generation and interpretation. The importance of event extraction and related tasks is outlined, setting the stage for detailed exploration of these topics.

Following the introduction, the background and definitions section provides foundational knowledge by defining key concepts such as LLMs, event extraction, event reasoning, event relation extraction, information extraction, and semantic role labeling. This section elucidates their interconnections and relevance in NLP, establishing a framework for understanding subsequent discussions.

The survey then delves into LLM development and capabilities in NLP, exploring their applications in text generation and interpretation, and their influence on advancing event extraction and reasoning tasks. A review of event extraction techniques follows, discussing methodologies, challenges, and advancements in identifying and categorizing events within text. Innovative approaches such as the

ISI-CLEAR system, which enables cross-lingual, zero-shot event extraction across 100 languages, are also examined [24].

Subsequent sections address event reasoning and relation extraction, focusing on frameworks and models that facilitate understanding the implications and contexts of events, as well as techniques for identifying relationships between events. The survey also explores information extraction and semantic role labeling, discussing methods for retrieving data from text and assigning roles to words or phrases to clarify their relationships to main actions or events.

Applications and case studies illustrate the practical implementation of these techniques across various domains, such as news, biomedical, and social media, highlighting their impact and effectiveness. The survey concludes by addressing current challenges in integrating LLMs with event extraction and semantic role labeling in NLP. It discusses potential advancements aimed at overcoming these limitations, such as improved prompt engineering strategies, dynamic schema-aware retrieval techniques, and structured heuristic methods to enhance model performance. Additionally, it emphasizes the importance of adapting methodologies for document-level event extraction and the varying effectiveness of trigger inputs, suggesting that these innovations could significantly enhance the efficiency and accuracy of event extraction processes across diverse NLP applications [25, 26, 27, 28].The following sections are organized as shown in Figure 1.

2 Background and Definitions

2.1 Interconnections and Relevance in NLP

The synergy among large language models (LLMs), event extraction, event reasoning, event relation extraction, information extraction, and semantic role labeling is crucial for advancing natural language processing (NLP). These interconnected tasks enable nuanced text interpretation, underpinning sophisticated NLP systems. LLMs facilitate the extraction of structured information from unstructured text through techniques like extractive summarization and event extraction, ensuring factual accuracy and grammatical integrity. Frameworks such as EYEGLAXS leverage LLMs like LLAMA27B and ChatGLM2-6B for efficient document summarization, setting new performance standards. Additionally, LLMs automate event extraction by decomposing tasks into detection and argument extraction, employing schema-aware retrieval to mitigate hallucinations. Research highlights the importance of prompt engineering in optimizing LLM performance, addressing model complexity and resource demands, thereby enhancing LLM capabilities across diverse textual sources [26, 29, 30]. Challenges remain in open information extraction (OIE), especially regarding multi-slot extraction in relation extraction and managing overlapping event roles.

Event extraction, a vital NLP component, involves identifying and categorizing events within text, often hindered by data scarcity and noise [31]. Datasets like Multilingual Event Extraction (MEE) address these challenges, facilitating event mention recognition and argument identification across languages [32].

Relation extraction (RE) infers semantic relationships between entities but traditionally requires extensive labeled data, which is costly and time-consuming [33]. The complexity of extracting relations and events from biomedical texts, such as identifying interactions among multiple entities in lengthy sentences, illustrates these challenges [34]. This complexity is pronounced in domains like Finance and Economics, where a lack of annotated datasets hampers effective supervised approaches [35].

Integrating these tasks within NLP systems is essential for developing models capable of complex language tasks, such as extracting entities, relations, and coreferences from text documents. Current information extraction (IE) approaches are often task-specialized, resulting in dedicated architectures and isolated models that impede knowledge sharing and rapid development across different IE tasks [36].

In the biomedical domain, accurate event trigger identification is crucial for extracting relevant events. Existing models struggle to generalize across datasets, exhibiting high error rates due to reliance on handcrafted features or limited context [37]. Despite these challenges, integrating these concepts is vital for advancing NLP, enabling models to perform complex tasks like term typing, taxonomy discovery, and extracting non-taxonomic relations [7]. Comprehensive evaluations at the corpus level, rather than focusing solely on distinct subtasks, are necessary to address high-recall needs in social sciences [38]. The labor-intensive nature, ambiguity, and scalability challenges in ontology learning further necessitate domain-specific knowledge and innovative approaches [9].

In recent years, Large Language Models (LLMs) have significantly transformed the landscape of Natural Language Processing (NLP). Their capabilities extend beyond mere text generation, encompassing a range of functionalities that enhance model efficiency and promote cross-lingual transfer. To better understand the intricate nature of these capabilities, we can refer to Figure 2, which illustrates the hierarchical categorization of LLMs' capabilities in NLP. This figure delineates four primary categories: text generation, model efficiency, cross-lingual transfer, and addressing biases. Each of these categories is further subdivided into key techniques, applications, optimization strategies, scalability enhancements, cross-lingual capabilities, cross-domain adaptability, bias detection, and robustness enhancements. Such a comprehensive framework underscores the multifaceted impact of LLMs in advancing NLP technologies, providing a structured overview that facilitates deeper analysis and understanding of their contributions.

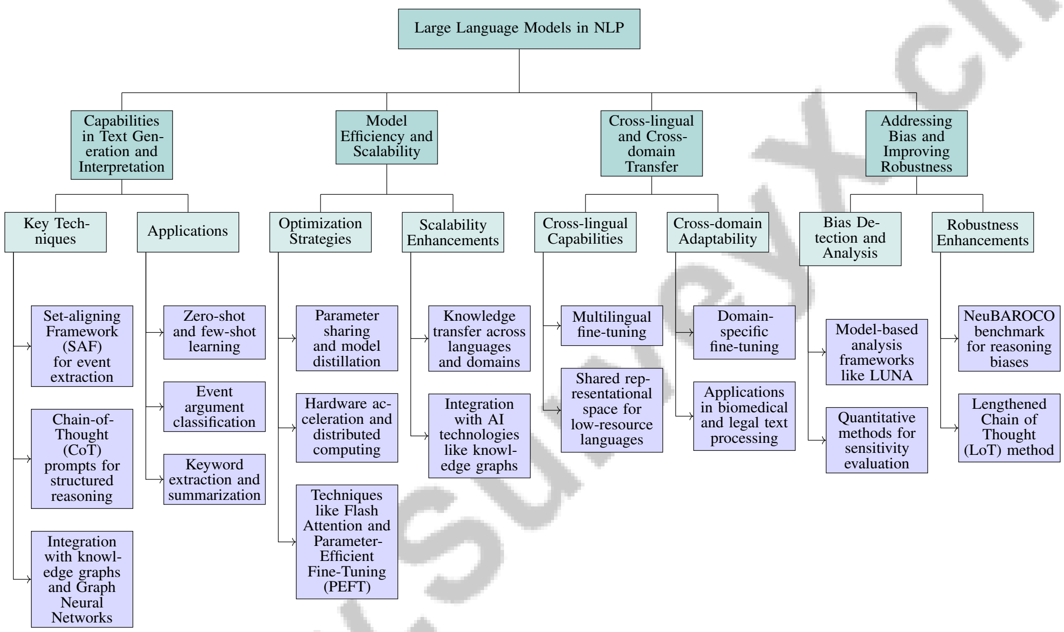


Figure 2: This figure illustrates the hierarchical categorization of Large Language Models' capabilities in NLP, focusing on text generation, model efficiency, cross-lingual transfer, and addressing biases. Each primary category is further divided into key techniques, applications, optimization strategies, scalability enhancements, cross-lingual capabilities, cross-domain adaptability, bias detection, and robustness enhancements, highlighting the multifaceted impact of LLMs in advancing NLP technologies.

3 Large Language Models in NLP

3.1 Capabilities in Text Generation and Interpretation

Large Language Models (LLMs) have revolutionized natural language processing (NLP) by enhancing text generation and interpretation through sophisticated architectures and extensive datasets. The Set-aligning Framework (SAF) exemplifies this by reformulating event extraction as conditional set generation, improving performance [39]. LLMs excel in zero-shot and few-shot learning, notably in event extraction, as demonstrated by ChatGPT's capabilities compared to traditional models [40]. The efficacy of Chain-of-Thought (CoT) prompts, which rely on reasoning steps, underscores the significance of structured reasoning in achieving superior outcomes [41].

Integrating LLMs with knowledge graphs and Graph Neural Networks enhances document-level event extraction by embedding structured knowledge, thereby improving accuracy [42]. LLMs also excel in zero-shot event argument classification, utilizing global constraints from domain knowledge

AI-generated, for reference only.

[43]. In specialized domains, LLMs are applied to tasks like keyword extraction and summarization, as seen in the clustering-based method for detecting news events from the GDELT dataset [44]. The DEGREE approach innovates event extraction by generating natural language outputs, enhancing data efficiency through event type definitions and keywords [45].

LLMs significantly impact question-answering models in event extraction by generating high-quality questions, emphasizing question quality's role in improving model effectiveness [46]. These advancements illustrate LLMs' transformative role in text generation and interpretation, continually pushing NLP boundaries.

As illustrated in Figure 3, the hierarchical categorization of capabilities in text generation and interpretation focuses on key areas such as event extraction, reasoning and question-answering, and knowledge integration. This figure highlights the significant frameworks and models contributing to advancements in each area, providing a visual representation of the interconnectedness of these capabilities.

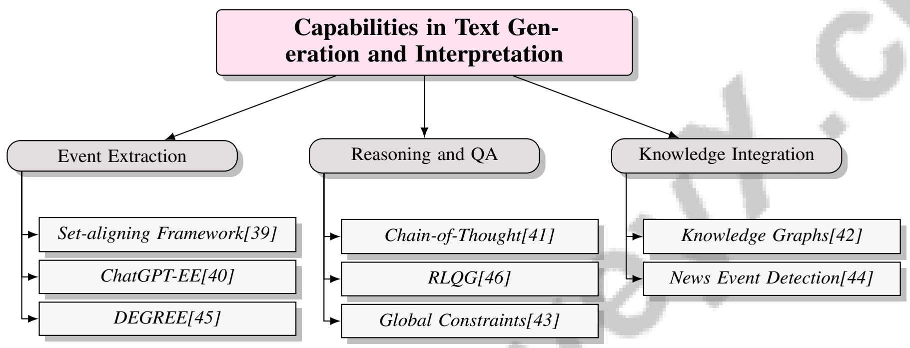


Figure 3: This figure illustrates the hierarchical categorization of capabilities in text generation and interpretation, focusing on event extraction, reasoning and question-answering, and knowledge integration. It highlights the significant frameworks and models contributing to advancements in each area.

3.2 Model Efficiency and Scalability

Efficiency and scalability are critical for Large Language Models (LLMs) in NLP tasks. Characterized by extensive parameters and computational demands, LLMs require optimization strategies to enhance efficiency while maintaining performance. Recent advancements include translating expert intuition into quantifiable features, improving model accuracy through human domain knowledge. Frameworks like EYEGLAXS address computational challenges via parameter-efficient fine-tuning and advanced attention mechanisms, facilitating applications in risk assessment and information extraction [29, 2, 47, 20, 36].

Efforts to enhance model efficiency focus on optimizing architecture and training processes. Techniques like parameter sharing and model distillation reduce size and complexity while preserving performance. Advances in hardware acceleration and distributed computing have improved LLM scalability, enabling efficient processing of larger datasets and sophisticated analyses in real-time. EYEGLAXS employs techniques like Flash Attention and Parameter-Efficient Fine-Tuning (PEFT) to overcome computational challenges, achieving effective extractive summarization of lengthy texts and setting new benchmarks on datasets like PubMed and ArXiv [29, 30].

Scalability is further enhanced through models capable of transferring knowledge across languages and domains, reducing the need for extensive retraining. This capability is vital for deploying LLMs in diverse environments, addressing linguistic and contextual variations. LLMs' ability to adapt swiftly to new tasks with minimal training underscores their versatility in fields like predictive analytics, integrating expert knowledge into structured features for improved decision-making [47, 29, 48]. Integrating LLMs with other AI technologies, such as knowledge graphs and neural networks, enhances efficiency by enabling structured and context-aware information processing [47, 26, 49].

Ongoing research aimed at improving LLMs' efficiency and scalability is crucial for their advancement and integration into real-world applications. As LLMs enhance efficiency and scalability, their influence on NLP and related fields is expected to grow, creating new avenues for innovation across domains like extractive text summarization, data augmentation, and keyword extraction [50, 29, 48, 30].

3.3 Cross-lingual and Cross-domain Transfer

LLMs exhibit remarkable capabilities in transferring knowledge across languages and domains, enhancing their effectiveness in various NLP tasks. This multilingual proficiency broadens applicability across cultural contexts and addresses usability challenges for diverse language groups. Ongoing research optimizes LLMs for applications like extractive summarization and ontology learning, improving performance in processing complex textual data [51, 29, 7, 30]. Cross-lingual and cross-domain transfer is facilitated by extensive pre-training on multilingual and multi-domain datasets, allowing LLMs to generalize linguistic patterns and semantic relationships beyond their original contexts.

The ability to perform cross-lingual transfer is advantageous in scenarios with scarce labeled data. By utilizing a shared representational space, LLMs apply knowledge from high-resource languages to low-resource counterparts, facilitating tasks like translation and information retrieval across diverse linguistic settings. Techniques such as multilingual fine-tuning enhance transferability, enabling models to adapt effectively to specific language pairs or tasks [51, 29, 52, 2].

In addition to cross-lingual transfer, LLMs demonstrate substantial cross-domain transfer capabilities, adapting to new subject areas with minimal additional training. This adaptability is essential for applications requiring domain-specific knowledge integration, such as biomedical text processing and legal document analysis. By leveraging pre-trained LLMs and employing domain-specific fine-tuning, researchers can integrate general language understanding with specialized domain needs, enhancing tasks like keyword extraction and predictive analytics [47, 30].

The cross-lingual and cross-domain transfer abilities of LLMs significantly expand the reach and applicability of NLP technologies. As research progresses in developing and optimizing transfer mechanisms, LLMs' capacity to enhance communication and comprehension across diverse linguistic and disciplinary landscapes is expected to grow, presenting possibilities for innovation across domains [47, 29].

3.4 Addressing Bias and Improving Robustness

The development of LLMs has advanced NLP, yet these models face limitations related to bias and robustness. Addressing these issues is essential for ensuring equitable and reliable applications across various domains. Frameworks like LUNA utilize model-based analysis for evaluating LLM quality through abstract modeling and semantics binding, aiding in identifying and addressing biases [53]. Introducing quantitative methods to evaluate the sensitivity of explanations to training randomness marks a significant advancement in understanding biases in LLMs [54]. This approach enhances transparency and fairness in AI applications.

A dynamic method for identifying and analyzing biases in LLMs enhances transparency and fairness in AI systems [55]. Additionally, the NeuBAROCO benchmark introduces a systematic evaluation of reasoning biases in LLMs, providing a robust framework for understanding biases in logical reasoning tasks [11]. Techniques such as the Lengthened Chain of Thought (LoT) method enhance logical reasoning and problem-solving capabilities in LLMs, contributing to more robust performance [41].

Addressing bias and enhancing robustness in LLMs involves methodological innovations and evaluative frameworks aimed at improving transparency, fairness, and reliability. As research progresses in integrating LLMs across applications, combining innovative methodologies—such as extractive summarization frameworks, predictive analytics enhancements, and schema-aware event extraction techniques—will foster more equitable and robust implementations of LLMs, improving accuracy and efficiency [47, 26, 29, 30].

4 Event Extraction Techniques

Event extraction (EE) is a vital task in natural language processing (NLP), focusing on identifying and categorizing events from unstructured text. Table 1 offers a detailed categorization of event extraction

Category	Feature	Method
Event Extraction: Concepts and Challenges	Graph-Based Approaches Model Enhancement Strategies Data Handling and Selection Neural Network Techniques	CE[56] TLEEF[22], EDF[44] RUSS[14] E2E-CCM[57]
Methodological Innovations in Event Extraction	Relation and Structure Recognition Generative and Adversarial Techniques Active and Reinforcement Learning Hybrid and Integration Approaches	OneEE[58] AEM[59] MBLP[60], RLQG[46] Galois[10]
Datasets and Techniques in Event Extraction	Span Identification Methods Contextual Information Extraction Graph-Based Techniques	PESE[61] QGA-EE[62] PGLEE[63]
Document-Level and Nested Event Extraction	Sequence Construction Techniques Graph-Based Modeling	MGR[64] JSEEGraph[65]
Improving Event Extraction with EDeR Dataset	Model Performance Improvement	EG[66], COFFEE[67], MTF[68], GCN-CST[69], DAEE[70], HRE[71], IEE[12], PTPCG[72], JETRE[73]

Table 1: This table provides a comprehensive overview of various methodologies and innovations in event extraction, categorizing them into key areas such as event extraction concepts and challenges, methodological advancements, datasets and techniques, document-level and nested event extraction, and improvements using the EDeR dataset. Each category highlights specific features and methods, along with relevant references, demonstrating the diversity and complexity of approaches in enhancing event extraction capabilities in natural language processing.

methodologies, illustrating the diverse techniques and innovations that address the complexities and challenges in this field. Additionally, Table 4 presents a comprehensive comparison of various event extraction methodologies, illustrating their core techniques, application domains, and unique features. This section reviews methodologies developed to enhance EE capabilities while addressing challenges such as event definition nuances and extraction complexities.

4.1 Event Extraction: Concepts and Challenges

EE is crucial for applications like information retrieval, document summarization, and knowledge graph construction. A key challenge lies in accurately identifying and classifying event triggers and arguments, often complicated by linguistic variability and contextual factors [74, 59]. Traditional methods inaccurately presume each word corresponds to a single event, leading to errors, particularly in longer texts. Extracting relational triplets from sentences with overlapping entities is another challenge where existing methods often fail [57]. Issues with nested events further complicate the extraction of triggers and arguments [58]. However, integrating Large Language Models (LLMs) and innovations like the Cluster Stability Assessment Index (CSAI) offer promising solutions for improving extraction accuracy [44].

In specialized domains like biomedicine and financial crime, the scarcity of tailored resources and datasets limits the effectiveness of named entity recognition (NER) methods [20]. The manual preparation of extraction pipelines adds to these challenges, highlighting the need for automated methodologies [10]. Models like the Memory-Based Loss Prediction (MBLP) and batch-based selection strategies enhance sample selection for EE, addressing traditional method limitations [60]. Event graph alignment loss based on optimal transport further improves understanding of complex events by capturing argument structures [56].

Automating explicit event descriptor extraction from news articles is crucial for analytical tasks, yet zero-shot event extraction presents challenges in accurately representing event types [14]. The Transfer Learning-based Event Extraction Framework (TLEEF) exemplifies efforts to enhance EE through transfer learning techniques [22]. The Prompt-based Graph Model for Liberal Event Extraction (PGLEE) offers a streamlined approach to extract events and discover schemas without relying on external knowledge bases [63]. Innovations like Oracle-Free Event Extraction (OFEE) and the integration of sequential information from entire sentences show promise in overcoming existing obstacles [67].

As illustrated in Figure 4, the key challenges and solutions in event extraction are multifaceted, reflecting the complexities faced in specialized domains and the innovative methods proposed to enhance event extraction techniques.

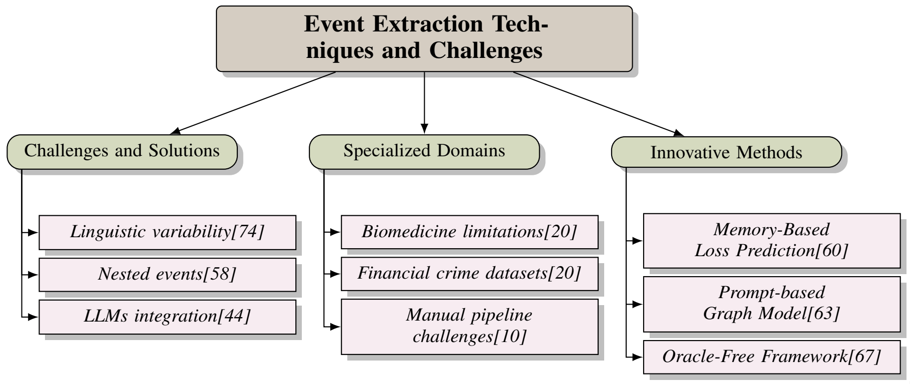


Figure 4: This figure illustrates the key challenges and solutions in event extraction, the complexities faced in specialized domains, and innovative methods proposed to enhance event extraction techniques.

Method Name	Methodological Approaches	Integration Techniques	Application Domains
AEM[59]	Generative Adversarial Networks	-	News Articles
MBLP[60]	Active Learning	-	Military, Legal
Galois[10]	Db-first Architecture	Hybrid Querying	-
RLQG[46]	Reinforcement Learning	-	Real-world Tasks
OneEE[58]	Tagging Scheme	Adaptive Event Fusion	Structured Prediction Tasks
PESE[61]	Pointer Networks	Bert Model	Natural Language Processing
PGLEE[63]	Prompt-based Model	Graph Representation Learning	Various Domains

Table 2: Overview of recent methodological innovations in event extraction, detailing various approaches, integration techniques, and application domains. This table highlights the diversity and specialization of methods such as generative adversarial networks, active learning, and prompt-based models across domains like news articles, military, and natural language processing.

4.2 Methodological Innovations in Event Extraction

Recent advancements in EE have introduced innovative methodologies to tackle challenges such as data scarcity and the integration of multimodal information. Table 2 presents a comprehensive summary of the latest methodological advancements in event extraction, illustrating the diverse strategies and their respective application areas. The Adversarial-neural Event Model (AEM) uses Generative Adversarial Networks (GANs) to learn event representations from documents, enhancing robustness and accuracy [59]. Active learning strategies, including batch-based selection and inter-exter sample loss ranking, optimize sample selection for improved training efficiency [60]. The Galois method integrates LLMs into traditional database systems, enabling hybrid querying of structured and unstructured data [10].

The RLQG method employs reinforcement learning to generate high-quality, context-dependent questions for QA-based EE, emphasizing the role of question generation in improving model accuracy [46]. The OneEE framework reformulates EE as a word-word relation recognition task, allowing parallel processing of event triggers and arguments, streamlining the extraction process [58]. The PESE method employs a pointer network-based encoder-decoder architecture for end-to-end event tuple generation, addressing limitations of previous models that separated trigger and argument extraction [61]. Furthermore, PGLEE's prompt-based approach combined with graph representation learning enhances EE and schema discovery without external dependencies [63].

These innovations signify substantial progress in NLP, addressing persistent challenges and enhancing the ability to analyze complex event data across diverse domains, including news, biomedicine, and cybersecurity [75, 28].

4.3 Datasets and Techniques in Event Extraction

The advancement of EE methodologies is significantly supported by diverse datasets and innovative techniques that serve as benchmarks for model performance. The ACE 2005 dataset is a cornerstone resource, providing a detailed collection of event types and argument roles across multiple languages,

Benchmark	Size	Domain	Task Format	Metric
EusIE[76]	1,500	Event Extraction	Entity And Event Extraction, Argument Extraction	F1-score
EDeR[77]	11,852	Event Extraction	Argument Classification	Accuracy, ROC$_AUC$
LogicBench[78]	12,908	Logical Reasoning	Binary Question-Answering (bqa) And Multiple-Choice Questions-Answering (mcqa)	Accuracy
SpeechEE[79]	40,850	Cybersecurity	Event Extraction	F1-score
DivED[80]	30,000	Event Detection	Event Extraction	F1-score; Recall
Vrittanta-EN[81]	174,559	Event Extraction	Event Classification	F1-score
CNC[82]	3,950	Commodity News	Event Extraction	Cohen's Kappa
BEES[83]	218,198	Biographical Event Extraction	Annotation OF Biographical Events	Inter-Annotator Agreement, F-measure

Table 3: The table presents a comprehensive overview of representative benchmarks used in event extraction research. It details various datasets, highlighting their size, domain, task format, and evaluation metrics, reflecting the diversity and scope of resources available for advancing event extraction methodologies.

facilitating a deeper understanding of event structures and semantics [84, 85, 86, 87]. This dataset is instrumental in evaluating models, offering standard splits for training, development, and testing.

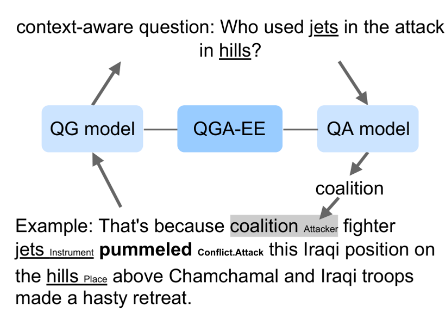
The TAC-KBP 2017 event extraction dataset comprises 8,026 training samples and 683 test samples, with predefined event schemas, supporting evaluations of EE in complex scenarios [63]. The GNBusiness dataset enhances research resources by providing a collection of 55,618 news reports organized into clusters, valuable for advancing methodologies in event detection and argument extraction [13, 28, 44]. Additionally, the New York Times and WebNLG datasets have been utilized to evaluate models for extracting relational triplets from overlapping entities.

Innovative techniques have propelled advancements in EE. The PESE model exemplifies an end-to-end approach that generates structured event frames from sentences, enhancing robustness and accuracy [61]. A comprehensive dataset containing 106,875 images supports zero-shot settings in Multimedia Event Extraction, showcasing the integration of multimodal information [28, 56, 88, 16]. Furthermore, experiments with approximately 15,000 news articles from GDELT highlight the potential of clustering algorithms and embedding methods in enhancing EE capabilities.

Table 3 provides a detailed overview of representative benchmarks that are instrumental in evaluating and advancing methodologies in event extraction. The ongoing development of diverse datasets and innovative techniques is crucial for advancing EE, encompassing tasks such as event detection and argument extraction across various domains. Recent developments, including cross-document event extraction methods and comprehensive benchmark datasets, emphasize the importance of these elements in pushing research boundaries [75, 28, 89]. The synergy between rich datasets and innovative methodologies continues to drive progress in NLP, facilitating the development of sophisticated EE systems.

Trigger	retirement
Event type	Personnel:End-Position
Person-Arg	Welch
Entity-Arg	GE
Position-Arg	-
Time-Arg	-
Place-Arg	-

(a) Trigger Retirement Event Type Personnel:End-Position Welch GE Position-Arg - Time-Arg - Place-Arg - [28]



(b) Context-Aware Question Answering with a Coalition Model [62]

Figure 5: Examples of Datasets and Techniques in Event Extraction

As shown in Figure 5, advancements in EE have led to the development of various datasets and techniques that enhance the accuracy and efficiency of information extraction. The first example, "Trigger Retirement Event Type Personnel:End-Position Welch GE Position-Arg - Time-Arg - Place-Arg," illustrates a structured representation of event data, categorizing attributes such as event type and associated arguments. The second example, "Context-Aware Question Answering with a Coalition Model," depicts a sophisticated framework integrating a Question Generation (QG) model, a Question Generation and Answering Event Extraction (QGA-EE) model, and a Question Answering (QA) model. This interaction highlights the dynamic process of leveraging context to improve answer relevance and accuracy. Together, these examples underscore the diversity of methodologies in EE, showcasing both structured data representation and advanced model integration to enhance event extraction from textual data [28, 62].

4.4 Document-Level and Nested Event Extraction

Document-level and nested event extraction are vital components of NLP, focusing on identifying and categorizing events across multi-sentence texts or within overlapping and nested structures. These tasks are essential for applications requiring deep comprehension of intricate narratives, such as information retrieval and summarization. Tools like Text Annotation Graphs (TAG) facilitate the annotation of complex relationships, enhancing extraction tasks in domains like biomedicine. Multi-format information extraction systems evaluate EE at both sentence and document levels, enabling nuanced understanding of narrative events, while research on identifying new events aids in summarizing and reasoning about unfolding events [90, 91, 28, 2].

Nested Event Extraction (NEE) addresses the challenge of extracting complex structures where events may contain other events recursively. The JSEEGraph model exemplifies an innovative approach to NEE by representing interrelations between entities, events, and relations as a unified graph, enhancing extraction accuracy [65]. Document-level extraction requires integrating contextualized information from sentence and paragraph-level representations, crucial for capturing broader event contexts and improving model performance [64].

Advancements in document-level and nested event extraction continue to propel progress in NLP, enabling sophisticated and accurate extraction of complex structures. Emerging methodologies, such as TAG for annotating relationships, the EYEGLAXS framework for efficient summarization using LLMs, and approaches incorporating writing style awareness, are anticipated to significantly enhance NLP capabilities in managing complex textual data across various domains [90, 29, 28, 92, 93].

4.5 Improving Event Extraction with EDeR Dataset

The EDeR dataset has become a pivotal resource for enhancing event extraction methodologies, offering a framework for modeling intricate event dependencies. By distinguishing between required and optional arguments, EDeR improves the accuracy of event extraction models, aligning with methodologies that focus on relational modeling [66]. The dataset's annotations facilitate the development of models that leverage relational information, significantly advancing argument role classification tasks.

Incorporating the EDeR dataset into event extraction frameworks has demonstrated substantial improvements, especially in low-resource environments. The Mask-then-Fill framework exemplifies this by balancing data diversity and distributional similarity, enhancing performance in such settings [68]. This aligns with findings emphasizing the importance of data quality in boosting model performance [70]. Additionally, integrating EDeR into models like EventGraph has achieved state-of-the-art results in argument role classification, showcasing its potential to enhance model effectiveness [66].

The EDeR dataset's focus on event dependencies complements advanced modeling techniques, such as graph convolutional networks, which have shown high F1 scores in extracting events from commodity news [69]. This synergy between detailed datasets and cutting-edge methodologies is crucial for improving event extraction outcomes. Furthermore, the dataset supports addressing challenges in document-level extraction, such as scattering-arguments and multi-events [71]. The PTPCG model, for instance, demonstrates significant improvements in efficiency and resource consumption for document-level extraction, achieving competitive results with fewer parameters and less training time compared to state-of-the-art models [72].

Future research could explore applying this approach to n-ary relation extraction and refining model performance across varied event types [12]. The COFFEE framework illustrates that effective event extraction can be achieved without relying on oracle information, outperforming traditional methods dependent on templates [67]. Moreover, the joint model proposed by [73] leverages structured inference to assign event and relation labels concurrently, representing an innovative methodology that could benefit from the EDeR dataset.

The EDeR dataset stands as a critical asset in EE, supporting the development of models capable of handling complex event structures and dependencies. Its contributions are particularly significant in low-resource settings, where there is a pressing demand for high-quality training data to enhance information extraction tasks. By leveraging advanced data generation methods, including those utilizing LLMs, this research addresses challenges posed by limited labeled data, ultimately improving the efficacy of event extraction across various domains, including economics and journalism [92, 94, 35, 95]. Insights from the EDeR dataset are likely to drive further innovations in event extraction methodologies, contributing to more accurate and efficient systems.

Feature	Adversarial-neural Event Model (AEM)	Active Learning Strategies	Galois Method
Core Technique	Gans	Batch Selection	Hybrid Querying
Application Domain	General Nlp	Training Efficiency	Database Systems
Unique Feature	Robustness Enhancement	Sample Optimization	Llm Integration

Table 4: This table provides a comparative analysis of three methodologies: the Adversarial-neural Event Model (AEM), Active Learning Strategies, and the Galois Method, highlighting their core techniques, application domains, and unique features. The comparison underscores the diverse approaches and innovations employed to enhance event extraction capabilities and address challenges in natural language processing and database systems.

5 Event Reasoning and Relation Extraction

Event reasoning plays a crucial role in natural language processing (NLP), focusing on understanding complex interactions within textual data. This section delves into innovative frameworks and models that advance event reasoning, highlighting their contributions to the field.

5.1 Frameworks and Models for Event Reasoning

Advanced frameworks for event reasoning significantly enhance NLP systems by improving the extraction and interpretation of complex event structures. One approach frames event extraction as an extractive question-answering (QA) task, which employs structured queries to identify event mentions and arguments, thereby addressing cross-lingual learning challenges in contexts with scarce annotated datasets [96, 26, 97, 28, 89].

Frameworks such as EventRL utilize reinforcement learning with outcome supervision to refine event extraction model outputs [23]. Memory mechanisms within these frameworks improve modeling of inter-dependencies between events, enhancing precision in document-level extraction [71].

Incorporating structured knowledge from knowledge graphs into unstructured data enriches event structure representation, offering a nuanced understanding of event dynamics [98]. Innovations like TabEAE address conventional system challenges where error propagation affects relation classification.

Exploratory studies on ChatGPT for event extraction reveal its potential in simpler scenarios but underscore the need for specialized frameworks for complex structures [40]. Graph-based structures, as utilized in frameworks like PGLEE, advance semantic understanding of event types [63].

The evolution of frameworks and models for event reasoning is vital for improving NLP systems' accuracy and adaptability. Incorporating extensive datasets, such as the SciEvents dataset, is expected to enhance complex event structure extraction, facilitating insights across fields like finance and biomedicine [99, 28].

5.2 Challenges in Event Reasoning

Event reasoning faces several challenges that hinder effective extraction and interpretation of complex structures. The variability and complexity of event data, including explicit and implicit information, pose significant hurdles. Scattered and implicit details in informal contexts complicate extraction, requiring methodologies to bridge these gaps [100].

Managing data sensitivity and privacy is another challenge, especially with heterogeneous data sources. Frameworks like DDP must adapt to varying sensitivity levels to enhance privacy without compromising performance [6]. This adaptability is crucial for robust event reasoning systems compliant with privacy regulations.

Integrating structured and unstructured data also presents challenges. Combining knowledge graphs with unstructured text data necessitates advanced methodologies to capture complex event relationships, enhancing model accuracy by managing real-world data complexities through dynamic schema-aware retrieval and end-to-end learning of cross-event dependencies [26, 101, 75, 102].

Addressing these challenges is essential for advancing event reasoning, enabling sophisticated models capable of extracting and interpreting complex structures across various domains. Overcoming issues like hallucination in large language models (LLMs) and implementing schema-aware event extraction are crucial for enhancing accuracy and reliability [75, 26]. Ongoing research and innovative approaches are expected to drive progress in overcoming these obstacles.

5.3 Techniques for Relation Extraction

Relation extraction, a fundamental NLP task, involves identifying and characterizing relationships between entities or events within text. Recent advancements have introduced various methodologies to enhance relation extraction models' accuracy and efficiency. The CKGC method treats relation and event classification as a sequence classification problem, offering a novel perspective [103].

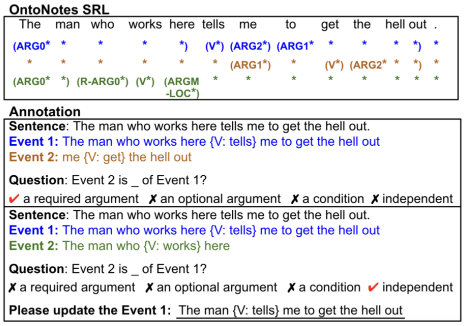
Integrating temporal, epistemic, and volitional dimensions in sentence analysis provides a comprehensive understanding of events, enhancing characterization granularity [104]. Cross-media approaches, evaluated by [16], highlight models' potential to identify event mentions and argument roles across text and images, using precision, recall, and F1-score metrics for robust evaluation.

Joint learning frameworks, such as those developed by [73], employ neural structured prediction models to extract events and temporal relations, ensuring global consistency in relation extraction. Benchmark assessments reveal the performance of state-of-the-art models, including fine-tuned deep learning models and LLMs like GPT-2, GPT-3.5, GPT-4, and Llama-2, emphasizing the importance of advanced models for improving entity recognition and relation extraction [19].

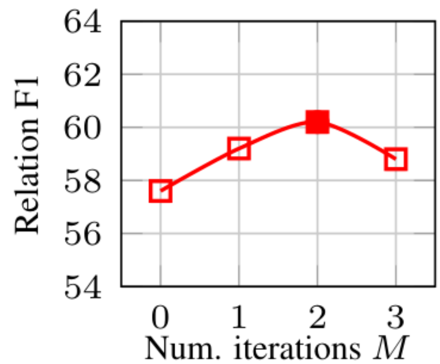
Evaluation of model performance through precision and F1 scores provides insights into the accuracy of predicted versus gold-standard annotations, ensuring models capture intended relationships within text [101]. In specialized domains like high-level corruption detection, low-billion parameter models show significant improvements over baseline methods [20].

The ongoing advancement of innovative techniques in relation extraction is crucial for NLP progression, enabling accurate inference of semantic relationships between entities in text. Recent approaches, such as treating relation extraction as a sequence-to-sequence task and leveraging large language models like GPT-3 and Flan-T5, have demonstrated significant performance improvements, achieving state-of-the-art results in few-shot settings. Furthermore, integrating multi-format information extraction systems and exploring event extraction methodologies highlight the diverse applications and challenges within this critical area of NLP research [57, 33, 28, 2]. These advancements are expected to refine model accuracy and efficiency, facilitating sophisticated extraction of complex event relationships across various domains.

As shown in Figure 6, various techniques have been developed to enhance the accuracy and efficiency of extracting relationships from textual data. Two notable examples are illustrated: the OntoNotes Semantic Role Labeling (SRL) framework, depicted in subfigure (a), which provides a structured representation of sentences for identifying and annotating event types, and a graph in subfigure (b) that explores the impact of the number of iterations (M) on the relation F1 score, highlighting the importance of optimizing parameters to improve model accuracy. Together, these examples



(a) OntoNotes SRL[77]



(b) The graph shows the relationship between the number of iterations (M) and the relation F1.[105]

Figure 6: Examples of Techniques for Relation Extraction

underscore the diverse methodologies employed in relation extraction, contributing to the broader goal of enhancing machine understanding of complex linguistic structures [77, 105].

6 Information Extraction and Semantic Role Labeling

6.1 Integration of Information Extraction Techniques

Integrating information extraction (IE) techniques into NLP systems is crucial for enhancing data extraction from text across domains like news, biomedical research, and cybersecurity. Methodologies in event extraction, such as event detection, argument extraction, and role labeling, are foundational for identifying and categorizing events and entities [28, 75, 92, 93, 106]. The EDeR dataset, with its structured labeling of event pairs as required, optional, or non-arguments, exemplifies advancements in distinguishing argument types and refining event extraction tasks [77].

Joint extraction methodologies have improved IE integration by capturing relationships among events, arguments, and entities through joint inference, as demonstrated by [106], enhancing coherence and reducing error propagation in extracted data. The MONTEE system advances modality-aware event extraction by tagging events based on modality and negation, enriching semantic understanding [107]. The Multi-Attribute Relation Extraction (MARE) approach further highlights the benefits of operating without structural constraints, capturing nuanced relationships in dynamic environments [108].

Innovative methodologies like EYEGLAXS for extractive summarization and the Text Annotation Graphs (TAG) tool are enhancing NLP sophistication, as evidenced by EYEGLAXS's performance on datasets like PubMed and ArXiv [90, 29, 30, 2]. These advancements set new benchmarks for information extraction, keyword categorization, and document summarization.

6.2 Semantic Role Labeling in Multimodal Contexts

Semantic Role Labeling (SRL) in multimodal contexts enhances the interpretation of semantic roles by integrating textual and non-textual data, crucial for applications requiring analysis across modalities like text, images, and videos. The Video MultiMedia Event Extraction (VM2E2) framework exemplifies the effectiveness of combining modalities to improve event extraction and narrative comprehension [81, 2, 92, 88, 91].

Models capable of jointly processing multimodal data enrich semantic understanding by utilizing contextualized representations and interdependencies among entities and events [96, 64, 2, 109]. Cross-modal attention mechanisms enhance coherence in role labeling, integrating contextual information from diverse text spans to improve event role identification [96, 64, 92, 110].

The integration of SRL with multimodal data has significant implications for video understanding and multimedia content analysis, enhancing tasks like event extraction by distinguishing actual from potential events. This capability improves systems in Question Answering, Knowledge Graph construction, and Fact-checking, as well as recognizing writing styles' influence on argument role extraction [107, 92].

Exploration of Self-Regulated Learning (SRL) in multimodal contexts is advancing NLP innovation, with workshops like CASE at RANLP 2023 emphasizing technical and social science integration in multimodal event information collection. Frameworks like EYEGLAXS showcase the potential of Large Language Models (LLMs) in enhancing extractive summarization, addressing factual accuracy and resource efficiency challenges [31, 29]. Continued research is poised to yield advanced methodologies, further enhancing SRL capabilities in multimodal contexts.

6.3 Challenges and Innovations in Semantic Role Labeling

Semantic role labeling (SRL) faces challenges and opportunities for innovation in NLP. A key challenge is the reliance on verb knowledge quality, which can be compromised by translation noise and linguistic structures, necessitating robust multilingual frameworks [111]. Confirmation bias in learning paradigms limits model generalization, particularly in low-resource scenarios, highlighting the need for innovative approaches to enhance SRL system adaptability [112].

Innovations focus on improving event extraction accuracy and efficiency while reducing hallucination in model outputs. Schema-aware extraction methods demonstrate significant advancements in SRL systems' precision and reliability [26]. Future research will likely develop automatic prompts for joint argument extraction and refine trigger identification, with prompt-based knowledge elicitation techniques offering promising directions for improving SRL efficiency [113].

Findings suggest that simpler models can provide informative explanations, offering a complementary approach to complex LLMs in SRL tasks [54]. Addressing document-level event argument extraction challenges and leveraging innovations like contextual clues and role relevance is essential for advancing NLP. Recent advancements, including the SCPRG model, showcase improvements in event extraction performance, emphasizing the need for comprehensive approaches that enhance interpretability and accuracy [110, 92, 9, 28]. As research evolves, integrating novel methodologies and comprehensive datasets is expected to drive progress, enabling sophisticated SRL systems capable of tackling complex linguistic phenomena.

7 Applications and Case Studies

7.1 Applications in Specific Domains

NLP techniques, notably event extraction and semantic role labeling, have transformative impacts across sectors such as biomedical, news, and social media, each with unique challenges. In the biomedical field, NLP facilitates complex entity relationship extraction, enhancing the development of comprehensive knowledge bases essential for healthcare decision-making [34]. In news media, NLP automates event extraction and contextual aggregation, crucial for managing digital information influx. Systems like Giveme5W1H apply syntactic and domain-specific rules to identify key events, improving information accessibility and narrative generation [92, 94, 35, 2]. Writing style analysis further refines event identification, enhancing news content aggregation quality.

Social media's informal, rapidly evolving content presents distinct challenges. Advanced text annotation tools address these complexities, enabling NLP to monitor and analyze content for public sentiment and emerging trends, crucial for marketing and crisis management [90, 31]. These applications highlight computational methodologies' potential in extracting complex information. Continued refinement, particularly with Large Language Models (LLMs), is expected to enhance effectiveness across sectors, driving innovation and informed decision-making [47, 92, 36, 2].

7.2 Event Entity Extraction in the Finance Sector

In finance, event entity extraction is crucial for converting unstructured data into structured formats, enhancing analysis and decision-making efficiency. The Doc2EDAG framework addresses challenges in extracting scattered event arguments from financial documents through a document-level model, with knowledge graphs improving relationship identification accuracy, optimizing extraction for financial insights [114, 42]. A major obstacle is the limited availability of annotated datasets, hindering effective supervised learning model development. Innovations like automatic training data generation and distant supervision address this, exemplified by large-scale datasets like the CrudeOilNews corpus, enhancing event extraction performance [115, 35].

Transfer learning and domain adaptation improve financial event extraction system accuracy. Advancements like Sequence Enhanced BERT Networks (SEBERTNets) and knowledge graph integration have significantly improved performance, with hybrid models achieving F1-scores up to 0.934 [116]. These methods enhance the management of financial text complexities and adaptability to market conditions. Integrating event extraction with sentiment analysis and semantic role labeling improves financial event contextual interpretation, capturing interdependencies among events and entities for accurate predictions and comprehensive market dynamics understanding [28, 75, 106].

7.3 Low-Resource Information Extraction with GIRL

The Gradient Imitation Reinforcement Learning (GIRL) framework addresses low-resource information extraction challenges by employing feedback mechanisms that enhance model generalization and pseudo-label quality, outperforming existing methods in data-scarce scenarios [112]. GIRL leverages reinforcement learning to iteratively refine predictions, ensuring robust extraction without extensive labeled datasets. Its real-time feedback during learning allows models to adapt dynamically to low-resource environments, enhancing performance across tasks like Named Entity Recognition (NER), Relation Extraction (RE), and Event Extraction (EE) [117, 112, 36].

GIRL improves pseudo-label error identification and correction, reducing confirmation bias seen in traditional self-training methods, leading to more reliable outputs, particularly in low-resource settings where pseudo-label quality significantly impacts performance [47, 112, 118]. As reinforcement learning techniques evolve, frameworks like GIRL are expected to enhance extraction efficiency and accuracy in environments lacking extensive annotations, broadening NLP technology applicability across diverse fields [40, 81].

7.4 Specialized Training in IT Tasks with OWL Model

The OWL model advances specialized IT operations training, leveraging large language models (LLMs) through techniques like Homogeneous Markov Context Extension (HMCE) and a mixture-of-adapter strategy for parameter-efficient tuning across IT tasks. Its performance on the Owl-Bench and established benchmarks demonstrates enhanced training efficiency and accuracy, addressing the gap in domain-specific LLM development for IT applications [119, 29]. OWL effectively integrates domain-specific knowledge into training, enhancing material relevance by extracting and structuring knowledge from diverse sources [9, 7].

Integrating OWL with existing IT training platforms enhances personalized learning experiences, allowing training module customization based on learner needs, mirroring predictive analytics techniques to transform human knowledge into quantifiable features for tailored pathways [47]. OWL facilitates multimodal data integration, such as text, code snippets, and diagrams, enhancing learning by providing a comprehensive understanding of IT concepts [96, 81]. As LLM capabilities progress, OWL is expected to enhance IT training program effectiveness and efficiency, fostering a skilled workforce capable of managing large data volumes in the evolving IT landscape [9, 7].

7.5 Comparative Performance of ChatGPT

ChatGPT, a prominent large language model, has been evaluated across various NLP tasks, revealing strengths and limitations. In event extraction, ChatGPT excels in zero-shot settings, identifying and categorizing events effectively with minimal task-specific training, as shown in benchmark evaluations [40]. However, ChatGPT faces challenges in tasks requiring intricate reasoning and domain-specific knowledge, where specialized models leveraging domain-specific training outperform it due to its reliance on general language patterns [63].

In low-resource settings, while the GIRL framework enhances model generalization and pseudo-label quality, ChatGPT's performance is limited by its dependence on pre-trained knowledge, constraining adaptability in data-scarce scenarios [112]. This contrast underscores the significance of reinforce-

AI-generated, for reference only.

ment learning techniques in addressing data scarcity, suggesting areas for potential methodological advancements in ChatGPT's capabilities.

8 Challenges and Future Directions

8.1 Challenges in Data Scarcity and Annotation

Event extraction systems face significant challenges due to data scarcity and annotation difficulties, particularly in complex NLP tasks. The reliance on human annotators for identifying event triggers and arguments restricts the volume and quality of training datasets, which is further exacerbated by the absence of oracle information, complicating the extraction of multiple events within a single context [67]. Inadequate data quality leads to subpar model performance, especially in low-resource environments where benchmarks are often insufficient [46].

The infrequency and variability of certain events, particularly in specialized domains, impair extraction accuracy. Furthermore, LLMs introduce issues such as hallucinations and non-deterministic outputs, affecting result reliability [20]. Incomplete corpora and biases from sources like social media further compound inaccuracies in entity recognition and event extraction [19].

Extracting implicit information, especially in 'why' and 'how' questions, presents challenges due to their nuanced nature in news articles [94]. Current benchmarks inadequately assess reasoning biases in LLMs, complicating performance evaluation in reasoning tasks [11].

Addressing these challenges requires standardized data sharing and model validation protocols, alongside exploring federated learning and explainable AI to enhance event extraction systems' robustness and generalizability. Active learning strategies, despite their promise, must overcome challenges in managing noisy predictions and ensuring robustness against errors in sample selection [60]. Future research should integrate domain-specific knowledge and machine reading comprehension frameworks to improve trigger identification and event extraction [17].

8.2 Integration of Multimodal and Multilingual Data

Integrating multimodal and multilingual data in NLP is crucial for developing models capable of processing diverse sources and languages. Techniques like joint neural frameworks and multi-format information extraction enhance systems across applications, such as named entity recognition, relation extraction, and event extraction. Incorporating global features and multi-task learning mechanisms further improves adaptability to new languages and complex tasks like extractive summarization [29, 109, 2].

Multimodal integration, combining textual, visual, and auditory information, creates a holistic understanding of content, particularly when single-modality data is insufficient. The synergy of visual data with textual information enhances contextual understanding necessary for accurately interpreting semantic roles in NLP. This approach aids in visualizing complex relationships and summarizing lengthy texts, with LLMs and visualization techniques improving model performance [93, 29, 48].

In multilingual integration, developing models that process data across languages is challenging. The Massively Multilingual Event Understanding framework exemplifies efforts to enhance model quality across languages, emphasizing zero-shot ontology expansion and improved multilingual capabilities [24]. Challenges remain in aligning translations and annotations, which can introduce noise and affect performance [84].

The Two-Stream Attention-Refinement (TSAR) model highlights complexities in document-level event extraction, where ambiguous argument boundaries and complex document structures pose challenges [120]. Future research should focus on enhancing LLMs' robustness against biases, exploring efficient training techniques, and expanding applicability to diverse workflows [121].

8.3 Advancements in Zero-Shot and Few-Shot Learning

Advancements in zero-shot and few-shot learning in NLP are driven by sophisticated models and innovative methodologies, enabling models to generalize effectively from limited examples. Techniques like data augmentation with LLMs and innovative event extraction frameworks enhance

AI-generated, for reference only.

model performance across domains, addressing data scarcity while integrating expert knowledge into predictive analytics [47, 122, 48].

Zero-shot learning, allowing models to predict unseen classes without specific training examples, is enhanced by LLMs pre-trained on diverse datasets. This is exemplified by LLMs in zero-shot event extraction, where models like ChatGPT effectively identify events without extensive task-specific training [40]. LLMs support knowledge transfer across languages and domains, facilitating wide-ranging applications.

Few-shot learning, involving limited examples, progresses through techniques enhancing adaptability and efficiency. Prompt-based approaches, such as the Prompt-based Graph Model for Liberal Event Extraction (PGLEE), demonstrate few-shot learning's potential to improve event extraction by leveraging minimal data [63]. The Gradient Imitation Reinforcement Learning (GIRL) framework refines few-shot learning by providing feedback that enhances generalization and pseudo-label quality [112].

These advancements enhance NLP models' versatility across tasks and domains, particularly valuable in low-resource settings. Zero-shot methods allow classification of known and unknown categories without retraining, ideal for small businesses with limited resources. Few-shot approaches leverage minimal examples to fine-tune models, demonstrating strong performance in complex tasks like event extraction. Collectively, these techniques reduce reliance on large annotated datasets while improving model generalization across applications [122, 123, 124]. Continued research and development of sophisticated methodologies are expected to drive further advancements, enabling effective learning from minimal data across applications.

8.4 Emerging Techniques and Future Directions

Event extraction and semantic role labeling are advancing with emerging techniques and future research directions poised to enhance NLP system capabilities. Developing gold-standard datasets for additional languages is a promising area, facilitating the refinement of event extraction methods and improving cross-linguistic applicability [21]. Expanding event coverage through zero-shot learning and domain adaptation to related fields, such as finance and economics, is another critical direction [22].

Future research should explore more complex natural language inferences, extending evaluations to encompass a broader range of reasoning tasks, addressing models' limitations in handling intricate reasoning scenarios [11]. Enhancing model capabilities for liberal event extraction and exploring complex event schema induction techniques are essential for advancing the field, allowing for more flexible event representation [63].

Integrating structured constraints between events and relations, alongside incorporating multiple datasets, is pivotal for improving event extraction systems' robustness and accuracy. This approach addresses the triplet overlap issue and leads to higher recall rates in entity and relation tasks. Continuous refinement of active learning methods to mitigate risks associated with sample selection remains a key focus area [60].

Research will also investigate optimizing query plans and enhancing LLM outputs' accuracy, improving system flexibility that integrates these models. Developing schema-less querying techniques will enhance NLP systems' adaptability in handling diverse data formats [10]. Exploring AI approaches for analyzing event progression and evaluating the Cluster Stability Assessment Index (CSAI) across datasets will be central to future studies [44].

Incorporating event detection mechanisms to improve the extraction process and extending frameworks to support additional languages and datasets are crucial for advancing the field. Identifying and addressing hidden pitfalls in evaluations will ensure continuous methodological improvement, enabling more accurate extraction of complex event structures across domains [101]. As research evolves, innovative approaches and refined methodologies are expected to drive significant advancements in the field.

AI-generated, for reference only.

9 Conclusion

This survey illustrates the transformative potential of integrating Large Language Models (LLMs) with event extraction and semantic role labeling, underscoring their pivotal role in advancing natural language processing (NLP). The inclusion of knowledge graphs (KGs) within LLMs is shown to significantly enhance their capability to produce factually accurate content, thereby improving performance on knowledge-grounded tasks. This integration not only strengthens the factual precision of generated text but also enriches the models' capacity for intricate reasoning and decision-making, aligning with the benefits of incorporating expert intuition into predictive analytics.

Furthermore, the survey highlights the importance of contextual information and the potential of knowledge bases in enhancing event extraction accuracy. LLMs demonstrate a remarkable ability to function as effective zero-shot text classifiers, providing substantial benefits, particularly for users with limited expertise in text classification, thus extending their applicability across diverse fields. This capability is complemented by LLMs' proficiency in ontology learning tasks, offering valuable insights for future research focused on refining and expanding LLM applications in NLP.

The integration of LLMs with event extraction and semantic role labeling represents a significant advancement in NLP, enabling more nuanced and precise interpretations of complex textual data. The potential for future research and development in this area is vast, presenting opportunities to explore advanced methodologies that enhance LLM capabilities in tackling complex language tasks. As the field progresses, the synergy between LLMs and event extraction techniques is expected to drive substantial advancements, opening new avenues for innovation and application across various domains.

References

[1] Thilini Wijesiriwardene, Amit Sheth, Valerie L. Shalin, and Amitava Das. Why do we need neuro-symbolic ai to model pragmatic analogies?, 2023.

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

[3] Khanh Nguyen and Brendan O'Connor. Posterior calibration and exploratory analysis for natural language processing models, 2015.

[4] Linyao Yang, Hongyang Chen, Zhao Li, Xiao Ding, and Xindong Wu. Give us the facts: Enhancing large language models with knowledge graphs for fact-aware language modeling, 2024.

[5] Jingqi Kang, Tongtong Wu, Jinming Zhao, Guitao Wang, Guilin Qi, Yuan-Fang Li, and Gholamreza Haffari. Towards event extraction from speech with contextual clues, 2024.

[6] Runxin Xu, Tianyu Liu, Lei Li, and Baobao Chang. Document-level event extraction via heterogeneous graph-based interaction model with a tracker, 2021.

[7] Hamed Babaei Giglou, Jennifer D'Souza, and Sören Auer. Llms4ol: Large language models for ontology learning, 2023.

[8] Tao Feng, Lizhen Qu, Niket Tandon, Zhuang Li, Xiaoxi Kang, and Gholamreza Haffari. From pre-training corpora to large language models: What factors influence llm performance in causal discovery tasks?, 2024.

[9] Rick Du, Huilong An, Keyu Wang, and Weidong Liu. A short review for ontology learning: Stride to large language models trend, 2024.

[10] Mohammed Saeed, Nicola De Cao, and Paolo Papotti. Querying large language models with sql, 2023.

[11] Kentaro Ozeki, Risako Ando, Takanobu Morishita, Hirohiko Abe, Koji Mineshima, and Mitsuhiro Okada. Exploring reasoning biases in large language models through syllogism: Insights from the neubaroco dataset, 2024.

[12] Yunmo Chen, Tongfei Chen, Seth Ebner, Aaron Steven White, and Benjamin Van Durme. Reading the manual: Event extraction as definition comprehension, 2020.

[13] Haolin Deng, Yanan Zhang, Yangfan Zhang, Wangyang Ying, Changlong Yu, Jun Gao, Wei Wang, Xiaoling Bai, Nan Yang, Jin Ma, Xiang Chen, and Tianhua Zhou. Title2event: Benchmarking open event extraction with a large-scale chinese title dataset, 2022.

[14] Sneha Mehta, Huzefa Rangwala, and Naren Ramakrishnan. Improving zero-shot event extraction via sentence simplification, 2022.

[15] Xiao Liu, Heyan Huang, and Yue Zhang. Open domain event extraction using neural latent variable models, 2019.

[16] Manling Li, Alireza Zareian, Qi Zeng, Spencer Whitehead, Di Lu, Heng Ji, and Shih-Fu Chang. Cross-media structured common space for multimedia event extraction, 2020.

[17] Pengchao Wu, Xuefeng Li, Jinghang Gu, Longhua Qian, and Guodong Zhou. Pipelined biomedical event extraction rivaling joint learning, 2024.

[18] Haohan Yuan, Siu Cheung Hui, and Haopeng Zhang. A structure-aware generative model for biomedical event extraction, 2024.

[19] Yiming Li, Deepthi Viswaroopan, William He, Jianfu Li, Xu Zuo, Hua Xu, and Cui Tao. Improving entity recognition using ensembles of deep learning and fine-tuned large language models: A case study on adverse event extraction from multiple sources, 2024.

[20] Panagiotis Koletsis, Panagiotis-Konstantinos Gemos, Christos Chronis, Iraklis Varlamis, Vasilis Efthymiou, and Georgios Th. Papadopoulos. Entity extraction from high-level corruption schemes via large language models, 2024.

[21] Ali Balali, Masoud Asadpour, and Seyed Hossein Jafari. Cofee: A comprehensive ontology for event extraction from text, 2021.

[22] Meisin Lee, Lay-Ki Soon, and Eu-Gene Siew. Crude oil-related events extraction and processing: A transfer learning approach, 2022.

[23] Jun Gao, Huan Zhao, Wei Wang, Changlong Yu, and Ruifeng Xu. Eventrl: Enhancing event extraction with outcome supervision for large language models, 2024.

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

[25] Zhuoyuan Liu and Yilin Luo. Document-level event extraction with definition-driven icl, 2024.

[26] 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.

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

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

[29] Léo Hemamou and Mehdi Debiane. Scaling up summarization: Leveraging large language models for long text extractive summarization, 2024.

[30] Sandeep Chataut, Tuyen Do, Bichar Dip Shrestha Gurung, Shiva Aryal, Anup Khanal, Carol Lushbough, and Etienne Gnimpieba. Comparative study of domain driven terms extraction using large language models, 2024.

[31] Ali Hürriyetoğlu, Hristo Tanev, Osman Mutlu, Surendrabikram Thapa, Fiona Anting Tan, and Erdem Yörük. Challenges and applications of automated extraction of socio-political events from text (case 2023): Workshop and shared task report, 2023.

[32] Amir Pouran Ben Veyseh, Javid Ebrahimi, Franck Dernoncourt, and Thien Huu Nguyen. Mee: A novel multilingual event extraction dataset, 2022.

[33] Somin Wadhwa, Silvio Amir, and Byron C. Wallace. Revisiting relation extraction in the era of large language models, 2024.

[34] Elaheh ShafieiBavani, Antonio Jimeno Yepes, Xu Zhong, and David Martinez Iraola. Global locality in biomedical relation and event extraction, 2020.

[35] Meisin Lee, Lay-Ki Soon, Eu-Gene Siew, and Ly Fie Sugianto. Crudeoilnews: An annotated crude oil news corpus for event extraction, 2022.

[36] Yaojie Lu, Qing Liu, Dai Dai, Xinyan Xiao, Hongyu Lin, Xianpei Han, Le Sun, and Hua Wu. Unified structure generation for universal information extraction. *arXiv preprint arXiv:2203.12277*, 2022.

[37] Patchigolla V S S Rahul, Sunil Kumar Sahu, and Ashish Anand. Biomedical event trigger identification using bidirectional recurrent neural network based models, 2017.

[38] Andrew Halterman, Katherine A. Keith, Sheikh Muhammad Sarwar, and Brendan O'Connor. Corpus-level evaluation for event qa: The indiapoliceevents corpus covering the 2002 gujarat violence, 2021.

[39] Xingwei Tan, Yuxiang Zhou, Gabriele Pergola, and Yulan He. Set-aligning framework for auto-regressive event temporal graph generation, 2024.

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

[41] Mingyu Jin, Qinkai Yu, Dong Shu, Haiyan Zhao, Wenyue Hua, Yanda Meng, Yongfeng Zhang, and Mengnan Du. The impact of reasoning step length on large language models, 2024.

[42] Kaihao Guo, Tianpei Jiang, and Haipeng Zhang. Knowledge graph enhanced event extraction in financial documents, 2021.

[43] Zizheng Lin, Hongming Zhang, and Yangqiu Song. Global constraints with prompting for zero-shot event argument classification, 2023.

[44] Adane Nega Tarekegn. Large language model enhanced clustering for news event detection, 2024.

[45] I-Hung Hsu, Kuan-Hao Huang, Elizabeth Boschee, Scott Miller, Prem Natarajan, Kai-Wei Chang, and Nanyun Peng. Degree: A data-efficient generation-based event extraction model, 2022.

[46] Zijin Hong and Jian Liu. Towards better question generation in qa-based event extraction, 2024.

[47] Phoebe Jing, Yijing Gao, Yuanhang Zhang, and Xianlong Zeng. Translating expert intuition into quantifiable features: Encode investigator domain knowledge via llm for enhanced predictive analytics, 2024.

[48] 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.

[49] Yuqi Zhu, Xiaohan Wang, Jing Chen, Shuofei Qiao, Yixin Ou, Yunzhi Yao, Shumin Deng, Huajun Chen, and Ningyu Zhang. Llms for knowledge graph construction and reasoning: Recent capabilities and future opportunities, 2024.

[50] Atin Sakkeer Hussain and Anu Thomas. Large language models for judicial entity extraction: A comparative study, 2024.

[51] Kaiyu Huang, Fengran Mo, Xinyu Zhang, Hongliang Li, You Li, Yuanchi Zhang, Weijian Yi, Yulong Mao, Jinchen Liu, Yuzhuang Xu, Jinan Xu, Jian-Yun Nie, and Yang Liu. A survey on large language models with multilingualism: Recent advances and new frontiers, 2025.

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

[53] Da Song, Xuan Xie, Jiayang Song, Derui Zhu, Yuheng Huang, Felix Juefei-Xu, and Lei Ma. Luna: A model-based universal analysis framework for large language models, 2024.

[54] Jeremie Bogaert and Francois-Xavier Standaert. A question on the explainability of large language models and the word-level univariate first-order plausibility assumption, 2024.

[55] Abel Salinas, Louis Penafiel, Robert McCormack, and Fred Morstatter. "im not racist but...": Discovering bias in the internal knowledge of large language models, 2023.

[56] Manling Li, Ruochen Xu, Shuohang Wang, Luowei Zhou, Xudong Lin, Chenguang Zhu, Michael Zeng, Heng Ji, and Shih-Fu Chang. Clip-event: Connecting text and images with event structures, 2022.

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

[58] Hu Cao, Jingye Li, Fangfang Su, Fei Li, Hao Fei, Shengqiong Wu, Bobo Li, Liang Zhao, and Donghong Ji. Oneee: A one-stage framework for fast overlapping and nested event extraction, 2022.

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

[60] Shirong Shen, Zhen Li, and Guilin Qi. Active learning for event extraction with memory-based loss prediction model, 2023.

[61] Alapan Kuila and Sudeshan Sarkar. Pese: Event structure extraction using pointer network based encoder-decoder architecture, 2022.

[62] Di Lu, Shihao Ran, Joel Tetreault, and Alejandro Jaimes. Event extraction as question generation and answering, 2023.

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

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

[65] Huiling You, Samia Touileb, and Lilja Øvrelid. Jseegraph: Joint structured event extraction as graph parsing, 2023.

[66] Huiling You, David Samuel, Samia Touileb, and Lilja Øvrelid. Eventgraph: Event extraction as semantic graph parsing, 2022.

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

[68] 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.

[69] Meisin Lee, Lay-Ki Soon, and Eu-Gene Siew. Effective use of graph convolution network and contextual sub-tree forcommodity news event extraction, 2021.

[70] 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.

[71] Shiyao Cui, Xin Cong, Bowen Yu, Tingwen Liu, Yucheng Wang, and Jinqiao Shi. Document-level event extraction via human-like reading process, 2022.

[72] Tong Zhu, Xiaoye Qu, Wenliang Chen, Zhefeng Wang, Baoxing Huai, Nicholas Jing Yuan, and Min Zhang. Efficient document-level event extraction via pseudo-trigger-aware pruned complete graph, 2022.

[73] Rujun Han, Qiang Ning, and Nanyun Peng. Joint event and temporal relation extraction with shared representations and structured prediction, 2020.

[74] 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.

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

[76] 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.

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

[78] Mihir Parmar, Nisarg Patel, Neeraj Varshney, Mutsumi Nakamura, Man Luo, Santosh Mashetty, Arindam Mitra, and Chitta Baral. Logicbench: Towards systematic evaluation of logical reasoning ability of large language models, 2024.

[79] Bin Wang, Meishan Zhang, Hao Fei, Yu Zhao, Bobo Li, Shengqiong Wu, Wei Ji, and Min Zhang. Speechee: A novel benchmark for speech event extraction, 2024.

AI-generated, for reference only.

[80] 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.

[81] Chaitanya Kirti, Ayon Chattopadhyay, Ashish Anand, and Prithwijit Guha. Enhancing event extraction from short stories through contextualized prompts, 2024.

[82] Meisin Lee, Lay-Ki Soon, Eu-Gene Siew, and Ly Fie Sugianto. An annotated commodity news corpus for event extraction, 2021.

[83] Marco Antonio Stranisci, Enrico Mensa, Ousmane Diakite, Daniele Radicioni, and Rossana Damiano. Guidelines and a corpus for extracting biographical events, 2022.

[84] Luís Filipe Cunha, Ricardo Campos, and Alípio Jorge. Event extraction for portuguese: A qa-driven approach using ace-2005, 2024.

[85] Sen Yang, Dawei Feng, Linbo Qiao, Zhigang Kan, and Dongsheng Li. Exploring pre-trained language models for event extraction and generation. In *Proceedings of the 57th annual meeting of the association for computational linguistics*, pages 5284–5294, 2019.

[86] Luís Filipe Cunha, Purificação Silvano, Ricardo Campos, and Alípio Jorge. Ace-2005-pt: Corpus for event extraction in portuguese, 2024.

[87] Yee Seng Chan, Joshua Fasching, Haoling Qiu, and Bonan Min. Rapid customization for event extraction, 2018.

[88] Brian Chen, Xudong Lin, Christopher Thomas, Manling Li, Shoya Yoshida, Lovish Chum, Heng Ji, and Shih-Fu Chang. Joint multimedia event extraction from video and article, 2021.

[89] 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.

[90] Angus G. Forbes, Kristine Lee, Gus Hahn-Powell, Marco A. Valenzuela-Escárcega, and Mihai Surdeanu. Text annotation graphs: Annotating complex natural language phenomena, 2018.

[91] 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.

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

[93] Sophia Henn, Abigail Sticha, Timothy Burley, Ernesto Verdeja, and Paul Brenner. Visualization techniques to enhance automated event extraction, 2021.

[94] Felix Hamborg, Corinna Breitinger, and Bela Gipp. Giveme5w1h: A universal system for extracting main events from news articles, 2019.

[95] Mingyu Derek Ma, Xiaoxuan Wang, Po-Nien Kung, P. Jeffrey Brantingham, Nanyun Peng, and Wei Wang. Star: Boosting low-resource information extraction by structure-to-text data generation with large language models, 2024.

[96] 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.

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

[98] Tin Kuculo. Comprehensive event representations using event knowledge graphs and natural language processing, 2023.

[99] Yi-Fan Lu, Xian-Ling Mao, Bo Wang, Xiao Liu, and Heyan Huang. Exceeds: Extracting complex events as connecting the dots to graphs in scientific domain, 2024.

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

[101] Hao Peng, Xiaozhi Wang, Feng Yao, Kaisheng Zeng, Lei Hou, Juanzi Li, Zhiyuan Liu, and Weixing Shen. The devil is in the details: On the pitfalls of event extraction evaluation, 2023.

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

[103] Chris Miller and Soroush Vosoughi. Big green at wnut 2020 shared task-1: Relation extraction as contextualized sequence classification, 2020.

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

[105] Yi Luan, Dave Wadden, Luheng He, Amy Shah, Mari Ostendorf, and Hannaneh Hajishirzi. A general framework for information extraction using dynamic span graphs. *arXiv preprint arXiv:1904.03296*, 2019.

[106] Bishan Yang and Tom Mitchell. Joint extraction of events and entities within a document context, 2016.

[107] Sander Bijl de Vroe, Liane Guillou, Miloš Stanojević, Nick McKenna, and Mark Steedman. Modality and negation in event extraction, 2021.

[108] Lars Klöser, Philipp Kohl, Bodo Kraft, and Albert Zündorf. Multi-attribute relation extraction (mare) – simplifying the application of relation extraction, 2021.

[109] Ying Lin, Heng Ji, Fei Huang, and Lingfei Wu. A joint neural model for information extraction with global features. In *Proceedings of the 58th annual meeting of the association for computational linguistics*, pages 7999–8009, 2020.

[110] Wanlong Liu, Shaohuan Cheng, Dingyi Zeng, and Hong Qu. Enhancing document-level event argument extraction with contextual clues and role relevance, 2023.

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

[112] Xuming Hu, Shiao Meng, Chenwei Zhang, Xiangli Yang, Lijie Wen, Irwin King, and Philip S. Yu. Gradient imitation reinforcement learning for general low-resource information extraction, 2022.

[113] Jiaju Lin and Qin Chen. Poke: A prompt-based knowledge eliciting approach for event argument extraction, 2022.

[114] Shun Zheng, Wei Cao, Wei Xu, and Jiang Bian. Doc2edag: An end-to-end document-level framework for chinese financial event extraction, 2019.

[115] Ying Zeng, Yansong Feng, Rong Ma, Zheng Wang, Rui Yan, Chongde Shi, and Dongyan Zhao. Scale up event extraction learning via automatic training data generation, 2017.

[116] 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.

[117] Chen Ling, Xujiang Zhao, Xuchao Zhang, Yanchi Liu, Wei Cheng, Haoyu Wang, Zhengzhang Chen, Takao Osaki, Katsushi Matsuda, Haifeng Chen, and Liang Zhao. Improving open information extraction with large language models: A study on demonstration uncertainty, 2023.

[118] Italo Luis da Silva, Hanqi Yan, Lin Gui, and Yulan He. Weak reward model transforms generative models into robust causal event extraction systems, 2024.

[119] Hongcheng Guo, Jian Yang, Jiaheng Liu, Liqun Yang, Linzheng Chai, Jiaqi Bai, Junran Peng, Xiaorong Hu, Chao Chen, Dongfeng Zhang, Xu Shi, Tieqiao Zheng, Liangfan Zheng, Bo Zhang, Ke Xu, and Zhoujun Li. Owl: A large language model for it operations, 2024.

[120] 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.

[121] 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.

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

[123] Lifu Huang, Heng Ji, Kyunghyun Cho, and Clare R. Voss. Zero-shot transfer learning for event extraction, 2017.

[124] Zhiqiang Wang, Yiran Pang, and Yanbin Lin. Large language models are zero-shot text classifiers, 2023.

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.