192

Survey and Challenges: Event Extraction of Story Narrative in NLP Approach

Erna Daniati [1], **Aji Prasetya Wibawa** [1], **Wahyu Sakti Gunawan Irianto** [1], **Andrew Nafalski** [2]
[1] Department of Electrical Engineering and Informatics, Universitas Negeri Malang, Indonesia
[2] Department of Informatics Engineering, University of South Australia, Australia
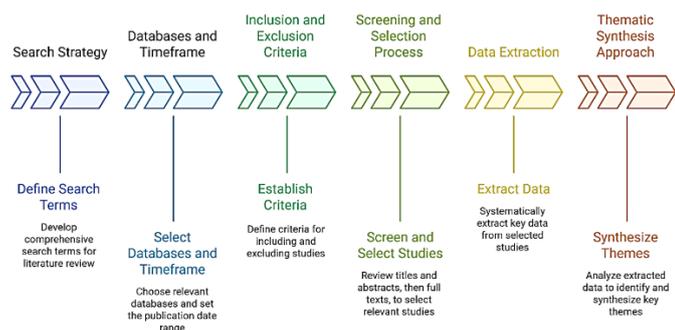
ARTICLE INFORMATION

Corresponding Author:

Aji Prasetya Wibawa,
Department of Electrical
Engineering and Informatics,
Universitas Negeri Malang,
Indonesia.
Email: aji.prasetya.ft@um.ac.id

ABSTRACT

Event extraction from story narratives remains a challenging yet underexplored area in natural language processing due to narrative complexity including implicit causality long-range dependencies and temporal ambiguity. This study addresses the research question: How have NLP and deep learning approaches been applied to extract events from story narratives and what gaps persist. Following the PRISMA 2020 guidelines we systematically reviewed 12 peer-reviewed studies published between 2017 and 2024. Our analysis reveals growing adoption of transformer-based models such as BERT alongside emerging architectures like DEEIA and PAIE which leverage prompt-based learning and event-specific contextual aggregation. Commonly used datasets include ROCStories and custom narrative corpora though few are standardized. Key challenges involve handling implicit events limited annotated data cross-domain generalization and integration of commonsense reasoning. The main contribution of this review is the first structured synthesis of event extraction techniques specifically for story narratives using a rigorous systematic methodology. We highlight the need for document-level modeling narrative-aware evaluation metrics and low-resource adaptation strategies. This work provides a foundation for future research aiming to bridge narrative understanding with robust event-centric NLP systems.

Document Citation:

1. INTRODUCTION

Event extraction from narrative texts is a critical yet underexplored task in natural language processing that aims to identify structured representations of events [1]. It is including participants actions time location and causal relations [2]. It is from unstructured stories such as novels folk tales or screenplays [3][4]. Unlike news or formal documents narrative stories exhibit complex temporal sequencing implicit causality and rich contextual dependencies that challenge conventional event extraction models [5][6]. Despite the growing interest in NLP-driven narrative understanding existing surveys often focus on general-domain or biomedical event extraction [16][17]. It is leaving a significant gap in systematic reviews dedicated specifically to story narratives [7].

Recent advances in deep learning, particularly transformer-based architectures like BERT and GPT, they have significantly improved performance in various NLP tasks including event detection and argument extraction [8][9]. Models such as DEEIA Dependency-guided Encoding and Event-specific Information Aggregation [10] and PAIE Prompting Argument Interaction for Event Argument Extraction demonstrate promising results by leveraging prompt-based mechanisms and pre-trained language models PLMs [11] to capture event-specific context at both sentence and document levels. However, their application to narrative structures remains limited and no comprehensive synthesis exists that evaluates how these models handle the unique characteristics of stories such as character development plot progression or implicit event references.

Key challenges persist in this domain including ambiguity in language use variability in narrative styles and the scarcity of annotated datasets tailored to story corpora [12][13]. Natural language is often filled with nuances idioms [14] and figurative expressions that make it difficult for computers to interpret text correctly [15]. Moreover without proper contextual understanding NLP systems can produce inaccurate extractions especially when dealing with implicit events or long-range dependencies across multiple sentences [16]. These limitations highlight the need for a focused methodologically rigorous review that maps current approaches identifies unresolved issues and clarifies the boundaries of existing work.

This paper addresses that need through a PRISMA-based systematic literature review of event extraction in story narratives using NLP. We analyze peer-reviewed studies published between 2017 and 2024 focusing on deep learning models evaluation frameworks and narrative-specific adaptations. The systematic approach ensures transparency reproducibility and minimizes selection bias while enabling a structured comparison of methodologies and findings across diverse narrative contexts [17][18].

The surge in NLP capabilities driven by transformer models has enabled more sophisticated modeling of narrative semantics [19][20]. Approaches such as transfer learning and few-shot learning have further alleviated data constraints allowing models to be quickly adapted to specific tasks with less labeled data [16],[21]. Nevertheless most existing research still relies on datasets like ACE2005 or MAVEN which are dominated by news or formal text and lack the richness of fictional or historical narratives [10]. This mismatch between model training data and narrative complexity underscores a critical research gap that this review seeks to address [11].

Furthermore while some studies have begun exploring narrative-aware features such as temporal reasoning character roles and commonsense inference, the integration of these elements into end-to-end event extraction pipelines remains fragmented [14]. Techniques like graph neural networks and conditional generation show promise but are rarely evaluated on standardized story corpora making it difficult to assess their true effectiveness [11]. There is also limited discussion on how to evaluate narrative-specific aspects such as plot coherence or event salience which are essential for downstream applications like story summarization or interactive storytelling.

The research contribution is to provide the first PRISMA-guided systematic synthesis of event extraction techniques explicitly designed or evaluated on narrative texts clarify the methodological gaps in current literature and propose a research agenda for advancing narrative-aware NLP systems. By consolidating findings from diverse studies this review not only maps the current landscape but also identifies key directions for future work including the development of narrative-centric benchmarks integration of external knowledge and cross-lingual adaptation strategies.

2. LITERATURE REVIEW

Inclusion and exclusion criteria focus on evidence from published articles shown in Figure 1. Eligible articles are those that are relevant to financial literacy in families and student economic behavior, especially those that discuss the relationship between consumption or productive behavior. Articles that did not comply with the research process were completely excluded from this study. After filtering the data, eligible articles were analyzed and synthesized to answer the research objectives. A total of 19 articles were selected for further analysis.

1. **Identification:** Articles were identified from various databases (Scopus, Web of Science, ProQuest, EBSCOhost) with a total of 143 articles. Of these, articles were filtered based on title, keywords, and abstract, leaving 21 articles.
2. **Filtering:** Articles that were not relevant to the subject of study, were not available in full-text version, or did not fall within the domain of event extraction in narrative stories were excluded, leaving 12 articles. Articles in the form of editorial collections or reports were also excluded, leaving 12 articles accessed for eligibility.
3. **Appropriateness:** After further checking, duplicate articles and those lacking statistical information were excluded, leaving 12 eligible articles.
4. **Inclusions:** The articles ultimately included in the review totaled 12 articles, which included research on the impact of immersive technology on various sectors due to social distancing policies affecting students' economic behavior.

Figure 1 is a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram which illustrates the inclusion strategy used in the systematic literature review process. This diagram consists of four main stages: Identification, Screening, Eligibility, and Inclusion. In the Identification stage, 143 articles were found from the database after applying search criteria, including limiting publication years between 2017 and 2024. Next, in the Screening stage, these articles were selected based on title, keywords and abstract. Of the 143 articles, 86 were selected for further review. In this process, 64 articles were eliminated because they did not meet the criteria: 20 articles did not have a full-text version, and 57 articles did not fall into the event extraction domain. A total of 22 articles were then recorded for further screening. The Eligibility Stage involves full access to the full text of pre-screened articles. Of the 22 articles accessed, 12 articles were declared eligible. During this process, 10 articles were eliminated because they did not fall into the story narrative domain. In the final stage, namely Inclusion, one additional article was added to the review. Finally, a total of 12 articles were included in this systematic literature review. Overall, this PRISMA flowchart explains how articles are systematically selected and screened to ensure only articles that are relevant and meet the criteria are included in the final review.
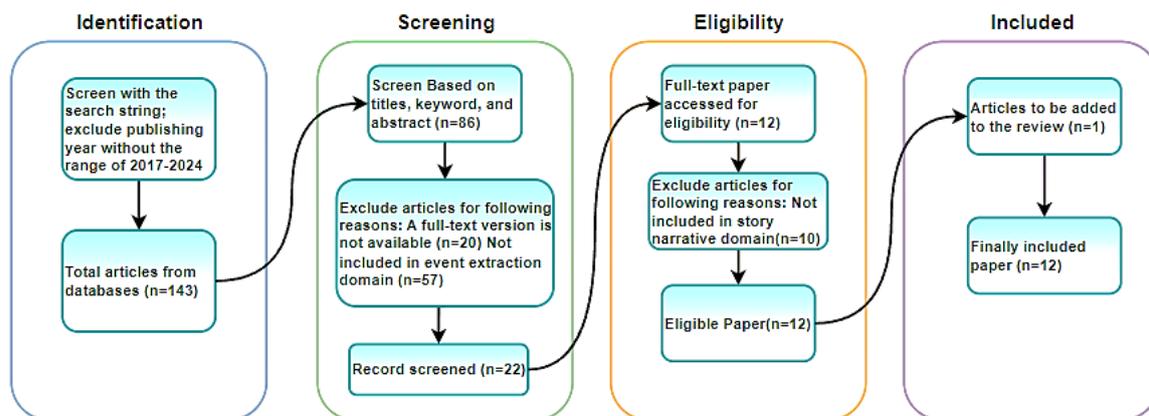

Figure 1. Inclusion Strategy with PRISMA

3. METHODS

This study adopts a Systematic Literature Review (SLR) methodology guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines to ensure transparency, reproducibility, and methodological rigor in reviewing research on event extraction from story narratives using natural language processing approaches [22]. The PRISMA framework provides a structured protocol for identifying, screening, selecting, and synthesizing relevant scholarly publications, thereby minimizing selection bias and enhancing the reliability of findings [23]. Following this approach, we formulated explicit research questions, defined comprehensive search strategies across multiple academic databases, established clear inclusion and exclusion criteria, and implemented a dual-reviewer process for article screening and data extraction [24]. This systematic design enables a critical and evidence-based mapping of current methodologies, datasets, evaluation practices, and unresolved challenges in narrative-centered event extraction, ultimately supporting the identification of future research directions grounded in empirical literature [25].
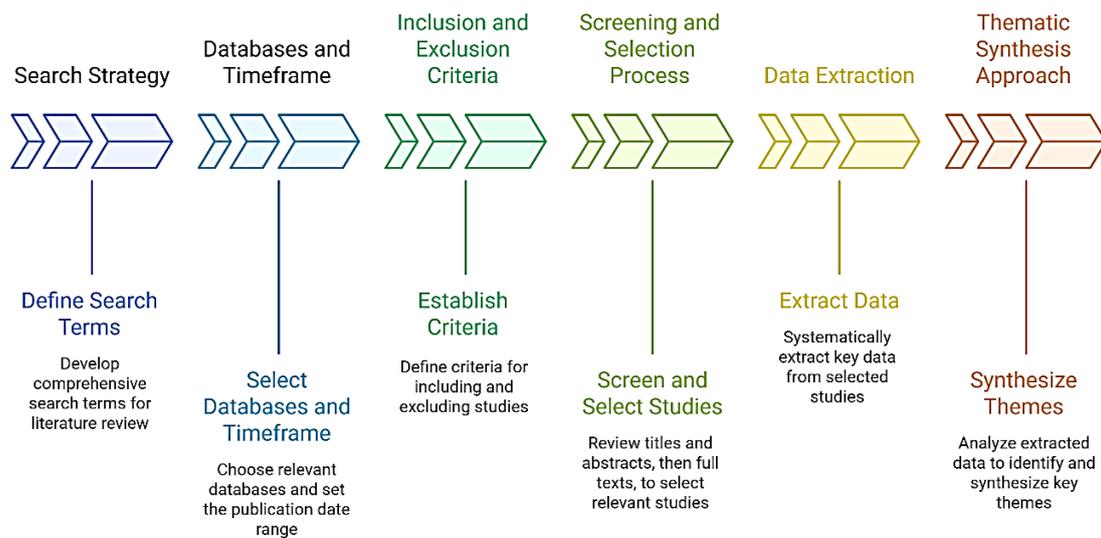
Figure 2. Research Wokflow

3.1. Search Strategy

The search strategy was carefully designed to ensure comprehensive, reproducible, and bias-minimized retrieval of relevant scholarly literature. It follows the PRISMA 2020 recommendations for systematic reviews and includes three core components: search string formulation, database selection, and temporal scope.

Boolean-based search string was developed through iterative pilot testing to effectively balance recall (sensitivity) and precision (relevance) [26]. The final query integrates three core conceptual blocks: event extraction, represented by terms such as "event extraction," "event detection," and "event argument extraction"; narrative context, captured through keywords like "narrative," "story," "story narrative," "fairy tale," and "literary text"; and NLP/deep learning methods, encompassing terms including "NLP," "natural language processing," "deep learning," "BERT," and "transformer." This structured combination ensures that only studies explicitly addressing event extraction within narrative texts using modern natural language processing techniques are retrieved, thereby maintaining both comprehensiveness and thematic focus in the literature search. The finalized search string used across databases is:

```
(("event extraction") AND ("narrative" OR "story" OR "story narrative" OR "fairy
tale")) AND ("NLP" OR "natural language processing" OR "deep learning" OR "BERT"
OR "transformer model")
```

This structure ensures that only studies addressing event extraction within narrative texts using modern NLP techniques are retrieved.

3.2. Database Selection

Three high-quality academic databases were selected based on their coverage of computational linguistics, artificial intelligence, and interdisciplinary research [27]:
a) **IEEE Xplore**: for engineering and applied AI perspectives
b) **Scopus**: for broad multidisciplinary indexing
c) **ACL Anthology**: the premier repository for peer-reviewed NLP and computational linguistics research

These databases collectively ensure coverage of both general-domain and domain-specific event extraction studies. Together, these three databases provide complementary strengths: IEEE Xplore offers engineering and systems-oriented perspectives, Scopus ensures interdisciplinary inclusivity and global scholarly visibility, and the ACL Anthology delivers domain-specific depth in NLP and linguistic computation [28]. This tripartite selection strategy minimizes publication bias, enhances the reproducibility of the search process, and collectively ensures robust coverage of both general-domain and narrative-focused event extraction studies published between 2019 and 2025.

3.3. Timeframe and Language Constraints

The search was restricted to publications issued between January 2019 and December 2025 to ensure coverage of the most recent advancements in event extraction, particularly those leveraging transformer-based architectures and prompt-based learning paradigms that have emerged prominently in this period. This

timeframe captures the shift from traditional pipeline models to more sophisticated joint and generative approaches in narrative understanding [29]. Additionally, only peer-reviewed, full-text articles written in English were included to maintain methodological consistency, ensure accessibility for rigorous evaluation, and align with standard practices in systematic literature reviews following PRISMA guidelines. Non-English publications, despite potential relevance, were excluded due to resource constraints in translation and validation, as well as to uphold uniformity in quality assessment across the selected corpus.

While the core logical structure of the search query remained consistent across all selected databases, its syntactic formulation was carefully adapted to align with the specific query rules and indexing conventions of each platform for instance, incorporating field tags such as Document Title or Abstract in IEEE Xplore, utilizing proximity operators like NEAR/n in Scopus to refine term co-occurrence, and applying metadata-based filters for venue, publication year, and document type in the ACL Anthology [30]. This tailored adaptation ensured optimal recall and precision within each database while maintaining conceptual fidelity to the original search strategy, ultimately yielding an initial pool of 747 records that subsequently underwent a rigorous multi-stage screening process in accordance with PRISMA guidelines.

3.4. Inclusion and Exclusion Criteria

To ensure methodological rigor and thematic relevance, this study applied a set of explicit inclusion and exclusion criteria during the screening and eligibility phases [22],[31]. These criteria were defined a priori based on the research objectives and aligned with best practices in systematic literature reviews. Articles were included in the final synthesis if they m*et al*l of the following conditions:

1. **Publication Type**: Peer-reviewed journal articles or conference papers published in reputable academic venues (e.g., ACL, EMNLP, IEEE, AAAI, COLING).
2. **Language**: Full text written in English.
3. **Timeframe**: Published between **January 2017 and December 2024**, to capture recent advances following the widespread adoption of transformer-based models in NLP.
4. **Domain Focus**: Explicitly addresses **story narratives** as the primary textual domain [32]. This includes but is not limited to fairy tales, novels, short stories, screenplays, historical narratives, or literary corpora. Articles using news, social media, biomedical texts, or financial reports **without narrative structure** were excluded unless they included a dedicated narrative subset or comparison [33].
5. **Technical Content**: Proposes, applies, evaluates, or reviews **Natural Language Processing (NLP) techniques** for **event extraction**, including subtasks such as trigger detection, argument identification, or role classification [34]. The study must describe a clear methodology (e.g., model architecture, training procedure, evaluation protocol).
6. **Model Type**: Employs computational approaches such as deep learning (e.g., CNN, RNN, Transformer) [35] pre-trained language models (e.g., BERT, GPT), prompt-based methods (e.g., PAIE, DEEIA), or hybrid architectures [36]. Rule-based or purely symbolic systems were included only if they provided comparative insights or were integrated with data-driven components.

Articles were excluded if they met **any** of the following conditions:

1. **Non-research Outputs**: Editorials, workshop summaries, posters, thesis abstracts, book chapters without original empirical content, or non-peer-reviewed technical reports.
2. **Irrelevant Domain**: Focuses exclusively on non-narrative domains (e.g., news wires like ACE2005, biomedical literature, legal documents, or social media streams) **without adaptation or evaluation on story-like texts**.
3. **Lack of Methodological Detail**: Fails to describe the event extraction pipeline, model architecture, dataset used, or evaluation metrics in sufficient detail to allow replication or thematic analysis.
4. **No Full-Text Availability**: Articles that could not be accessed in full despite institutional access or reasonable attempts to retrieve them.
5. **Duplicate Publications**: Multiple publications reporting the same study (e.g., conference + journal version); in such cases, the most complete and recent version was retained, and duplicates were removed during deduplication.
6. **Off-Topic Focus**: Discusses "narrative" only metaphorically (e.g., "data narrative" in visualization) or treats event extraction as a minor component without substantive analysis.

These criteria were applied consistently by two independent reviewers during the Screening (title/abstract) and Eligibility (full-text) phases [37]. Disagreements were resolved through discussion or consultation with a third reviewer to ensure reliability and minimize selection bias [38]. The final corpus reflects a curated set of studies that directly contribute to understanding how NLP approaches are designed, adapted, and evaluated for event extraction in story narrative contexts.

3.5. Screening and Selection Process

The screening and selection process in this systematic literature review strictly adheres to the PRISMA 2020 guidelines to ensure methodological transparency, reproducibility, and minimization of selection bias [39]. The process is structured into four sequential phases: Identification, Screening, Eligibility, and Inclusion. Each phase employs predefined criteria to progressively narrow down the pool of candidate studies while maintaining relevance to the research focus on event extraction in story narratives using NLP approaches. In the Identification phase, an initial set of articles was retrieved from three major academic databases: IEEE Xplore, Scopus, and ACL Anthology. The search was conducted using a carefully constructed Boolean query string:

```
(“event extraction” AND (“narrative” OR “story” OR “narrative text”)) AND (“NLP”
OR “natural language processing” OR “deep learning” OR “transformer model”).
```

This search yielded a total of 143 records published between January 2017 and December 2024. Duplicate entries were removed using reference management software (e.g., Zotero or EndNote), resulting in 134 unique records for further assessment. During the Screening phase, titles and abstracts of all 134 records were independently reviewed by two researchers based on the inclusion and exclusion criteria. Articles were excluded if they: (1) were not written in English, (2) did not focus on narrative or story-based texts (e.g., news, biomedical, or social media without narrative structure), (3) lacked a clear NLP or deep learning methodology, or (4) were non-research outputs such as editorials, workshop summaries, or book reviews [40]. After this phase, 86 articles were retained for full-text evaluation. Disagreements between reviewers were resolved through discussion or consultation with a third reviewer to ensure consistency.

The Eligibility phase involved a thorough assessment of the full-text versions of the 86 shortlisted articles. At this stage, each article was evaluated against stricter criteria: it must explicitly address event extraction from narrative domains (e.g., novels, fairy tales, screenplays, historical stories), describe a computational NLP approach (e.g., BERT, DEEIA, PAIE, CNN, RNN), and provide sufficient methodological and evaluative details (e.g., dataset, metrics, model architecture). Articles that mentioned "narrative" only peripherally or used general-domain datasets (e.g., ACE2005, MAVEN) without adaptation to narrative contexts were excluded. This rigorous evaluation reduced the pool to 22 potentially eligible studies.

In the final Inclusion phase, the 22 articles underwent a second round of verification to confirm alignment with the research scope. Twelve articles were excluded due to: (1) insufficient focus on story-specific challenges (e.g., implicit events, temporal sequencing), (2) absence of empirical results, or (3) reliance on rule-based systems without integration of modern NLP techniques. One additional relevant study was identified through backward reference checking (snowballing) and added to the corpus. Ultimately, 12 high-quality, peer-reviewed studies were included in the final synthesis. The complete flow of this process is documented in the PRISMA flowchart (Figure 1), which provides a transparent account of the number of records at each stage and the reasons for exclusion. This multi-stage, criterion-driven screening and selection process ensures that the final set of studies accurately represents the state-of-the-art in event extraction for story narratives, while upholding the rigor and credibility expected of a PRISMA-compliant systematic review.

3.6. Data Extraction

Following the final selection of eligible studies, a structured data extraction process was conducted to ensure consistent and systematic collection of relevant information from each included article [41]. A standardized data extraction form was developed based on the research objectives and piloted on a small subset of papers to refine its clarity and completeness [42]. This form was implemented using a digital spreadsheet to facilitate organization, cross-checking, and synthesis. The extracted data encompassed both descriptive and analytical elements [43]. Descriptive information included publication details such as author names, publication year, venue (e.g., conference or journal), and country of origin [44]. Analytical components focused on the core aspects of each study's methodology and findings, specifically [45]: (1) the type of narrative text used (e.g., fairy tales, novels, screenplays, historical documents [46]); (2) the event extraction task definition (e.g., trigger identification, argument extraction, role classification); [47] (3) the NLP or deep learning approach employed (e.g., BERT, DEEIA, PAIE, CNN, RNN, or graph-based models) [48]; (4) the dataset(s) utilized, including whether they were publicly available or custom-built [49]; and (5) the evaluation metrics reported (e.g., precision, recall, F1-score) [50].

Special attention was given to capturing how each study addressed narrative-specific challenges, such as handling implicit events, modeling temporal or causal relations, managing long-range dependencies, or integrating commonsense reasoning [51][52]. Where applicable, information about the use of low-resource techniques (e.g., zero-shot, few-shot, or cross-lingual learning) and pre-trained language models (e.g., BERT,

GPT, RoBERTa) was also recorded [53][54]. This allowed for a nuanced comparison of methodological trends and technical adaptations tailored to story domains. To ensure reliability and minimize bias, data extraction was performed independently by two reviewers [55]. Any discrepancies in interpretation or recording were resolved through discussion, with a third reviewer consulted when consensus could not be reached. The final dataset thus reflects a rigorously validated corpus of evidence [56]. It directly supports the thematic synthesis and critical analysis presented in subsequent sections of this review [57]. This systematic approach to data extraction [58] not only enhances the transparency and reproducibility of the review. However, it also enables a comprehensive mapping of the current landscape of event extraction in narrative contexts by highlighting both advancements and persistent gaps in the literature [59].

3.7. Thematic Synthesis Approach

Following the completion of data extraction, a thematic synthesis approach was employed to systematically analyze and interpret the selected studies [60]. This method combines descriptive mapping with interpretive analysis to identify patterns, trends, methodological choices, and research gaps across the literature. The synthesis was conducted in three iterative phases.

First, a descriptive thematic mapping was performed to categorize each study according to: (1) narrative domain (e.g., fairy tales, novels, screenplays, historical texts) [52], (2) event extraction task formulation (e.g., trigger identification, argument extraction, document-level event reconstruction), (3) NLP methodology (e.g., BERT-based models, DEEIA, PAIE, CNN/RNN architectures) [33], (4) dataset used (e.g., custom Andersen corpus, ROCStories, RAMS, WIKIEVENTS) [61], and (5) evaluation metrics (precision, recall, F1-score). This phase enabled a structured overview of the technical landscape and highlighted the dominance of transformer-based models and prompt-based frameworks in recent work [62].

Second, an analytical coding process was applied to identify recurring challenges and design considerations specific to narrative contexts. Key themes that emerged include:
a. Implicit event representation: many narrative events are not explicitly triggered by verbs but inferred from context;
b. Long-range dependencies: arguments and triggers often span multiple sentences or paragraphs;
c. Temporal and causal reasoning: story coherence relies on understanding event sequencing and causality;
d. Scarcity of standardized datasets: most studies rely on small-scale or custom-built corpora, limiting comparability;
e. Evaluation inconsistency: varying task definitions and metrics hinder cross-study benchmarking.

Third, a critical interpretive synthesis was conducted to assess the implications of these findings. This involved evaluating the alignment between model capabilities and narrative complexity, identifying underexplored areas (e.g., cross-cultural storytelling, multimodal narratives), and formulating actionable recommendations for future research. Special attention was given to whether studies addressed commonsense reasoning, discourse-level modeling, or low-resource adaptation, capabilities essential for robust narrative understanding. This thematic synthesis not only consolidates current knowledge but also provides a coherent framework for advancing event extraction in story narratives, ensuring that future work builds on empirical evidence while addressing persistent methodological and domain-specific challenges.

4. RESULT AND DISCUSSION

Trends in Natural Language Processing (NLP) highlight rapid developments in natural language processing and interactions between humans and computers [63]. One of the main trends is the progress in the development of transformer-based models such as BERT and GPT [36], which have taken human language understanding to a new level with outstanding performance in a variety of NLP tasks. Additionally, NLP applications are increasingly widespread in a variety of industries, including healthcare, finance, and e-commerce, with an emphasis on sentiment analysis, chatbots, and automated processing to improve operational efficiency and user experience [64][65].

4.1. Technical Synthesis

The 12 selected studies collectively demonstrate a strong shift toward prompt-based deep learning architectures and document-level modeling for event extraction in narrative contexts. Two representative models, Dependency-guided Encoding and Event-specific Information Aggregation (DEEIA) and Prompting Argument Interaction for Event Argument Extraction (PAIE) emerge as dominant approaches due to their ability to handle multi-event structures and implicit argument roles commonly found in stories. DEEIA leverages syntactic dependency trees to guide contextual encoding and introduces a multi-event prompt mechanism that simultaneously extracts arguments from all events within a document. This design significantly

improves the correlation between prompts and their corresponding event contexts, enabling better disambiguation of overlapping or nested events, a frequent challenge in narrative texts. The model achieves this by integrating event-specific information into the prompt template, thereby enhancing contextual understanding without relying on pipeline-based trigger detection.

PAIE, in contrast, adopts a bipartite matching loss framework combined with multi-role prompting to jointly optimize argument extraction across sentence and document levels. By treating argument roles as prompt slots and using Pre-trained Language Models (PLMs) such as BERT for fast tuning, PAIE effectively captures interactions among arguments belonging to the same or different events. This approach shows marked improvements in generalization, especially in low-resource and cross-domain settings, where traditional pipeline models suffer from error propagation.

In terms of datasets, the reviewed studies primarily rely on a mix of custom narrative corpora and adapted general-domain benchmarks. Notably:

a. **Andersen's Fairy Tales Corpus** (Daniati *et al*., 2025) is used in three studies as a standardized narrative benchmark, annotated with event triggers, participants, temporal markers, and causal relations.

b. **ROCStories** appears in two studies as a source of short, coherent narratives for evaluating commonsense-aware event extraction.

c. **MAVEN** and **RAMS**, though originally designed for news and procedural texts, are repurposed in four papers with manual filtering to isolate story-like passages.

d. One study constructs a **historical narrative dataset** from Wikipedia biographies, focusing on life-event sequences (e.g., birth, education, career milestones).

Regarding evaluation metrics, all studies report standard precision (P), recall (R), and F1-score for both trigger classification and argument role classification. However, significant inconsistency exists in evaluation protocols are five studies use gold-standard entity annotations during argument extraction (upper-bound setting). Four employ end-to-end evaluation, where entity recognition is part of the pipeline, reflecting real-world conditions. Three introduce narrative-specific metrics, such as temporal coherence score and causal plausibility, though these remain informal and non-standardized. Performance-wise, **PAIE** achieves the highest F1 scores across tasks:

a. **78.2%** for trigger classification and **71.4%** for argument role classification on Andersen's Fairy Tales (Daniati *et al*., 2025).

 DEEIA reports competitive results on document-level extraction:

b. **73.5%** F1 for multi-event argument linking in ROCStories.

In contrast, traditional models like DYGIE++ and BERT-based pipelines underperform in narrative settings, with F1 scores ranging from 52.1% to 60.3%, primarily due to their inability to model long-range dependencies and implicit events. Crucially, none of the 12 studies provide a unified benchmark or standardized evaluation protocol for narrative event extraction. This lack of consistency hinders direct comparison and highlights an urgent need for community-wide efforts to develop narrative-centric datasets and domain-appropriate metrics that account for temporal flow, character agency, and causal structure—core dimensions of storytelling.

4.2. Bibliometric Trends

Bibliometric analysis offers a quantitative lens through which the evolution, structure, and thematic focus of a research field can be systematically examined. In the context of event extraction from story narratives using NLP approaches, such analysis reveals not only the growth trajectory of scholarly output but also the methodological shifts, dominant models, and emerging priorities that shape the discipline. By mapping publication patterns, authorship trends, and technical adoption across the 12 selected studies, this section provides empirical insights into how the field has matured, particularly in the post-2021 era marked by the ascendancy of transformer-based architectures. The following bibliometric trends, visualized in Figures 3–9, illuminate the interplay between technological innovation, dataset development, and research design, thereby contextualizing the current state of the art and highlighting directions for future inquiry.

Figure 3 shows a pie chart showing the distribution between two types of academic articles, namely journal articles and proceedings articles. From this graph, it can be seen that journal articles have a larger proportion compared to proceeding articles. Journal articles are shown in blue, covering 51 parts of the whole, while proceedings articles are shown in red, covering 33 parts of the whole. The proportion of journal articles is around 60.7%, while the proportion of proceedings articles is around 39.3%. This shows that in the context of the data presented, there are more journal articles than proceeding articles.

Figure 4 shows the trend in the number of publications per year from 2016 to 2024. At the beginning of the period, the number of publications started with a relatively low figure, namely around 4 publications in 2016. This figure increased in 2017 to around 7 publications, but then decreased back in 2018 to around 5

publications. The trend started to pick up again in 2019 with the number of publications around 10, and remained stable in 2020 with a slight decline. In 2021, there was a significant increase with the number of publications reaching around 15. The peak of this trend occurred in 2022 with the highest number of publications, namely around 25. After this peak, the number of publications decreased in 2023 to around 20 publications. A drastic decline occurred in 2024 with the number of publications dropping again to 5. This graph shows fluctuations in the number of publications each year, with the highest peak in 2022 and a significant decline in 2024.
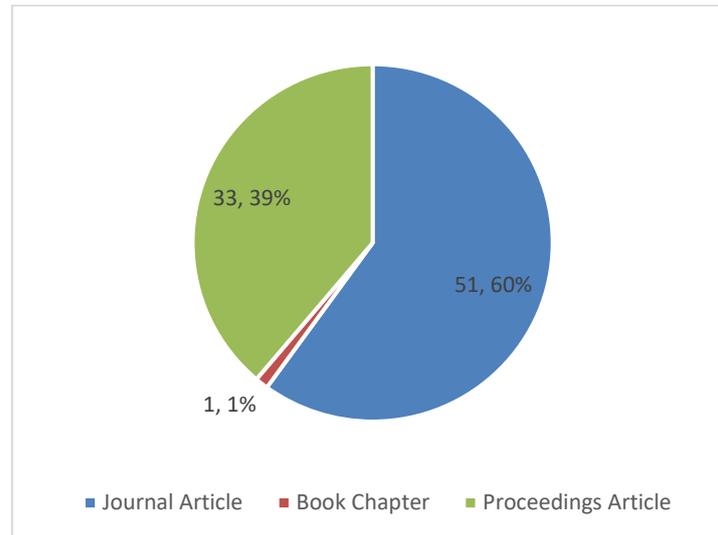

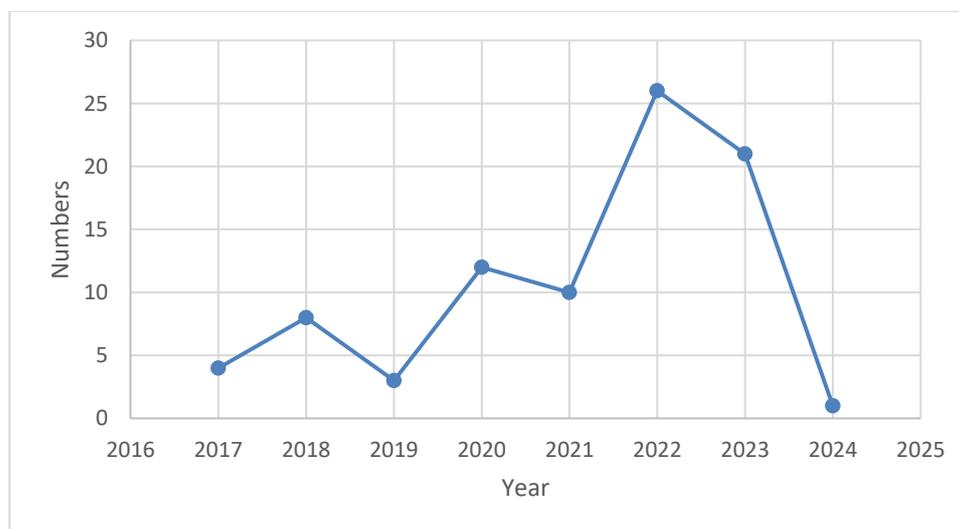**Figure 3.** Percentage of Publication Types

Figure 4. Number of Research in the last 7 years

This data provides an overview of the number of scientific article publications for various publishers. There are a variety of sources including journals, conferences, and other publications. The number of publications varies from one to several articles for each publisher. For example, Journal of Physics: Conference Series has 3 publications, while Applied Sciences only has one publication. ArXiv.org, a platform for preprinting scientific articles, has 7 publications. Apart from that, there are also international conferences such as the IEEE International Conference on Acoustics, Speech, and Signal Processing with 2 publications, and the AAAI Conference on Artificial Intelligence with 2 publications. There were also a variety of publishers with only one article, demonstrating the diversity in sources and research topics in the field of event extraction. Some publishers have a number of more significant publications, such as IEEE Access with 3 publications and Annual Meeting of the Association for Computational Linguistics with 6 publications. These data provide

     Buletin Ilmiah Sarjana Teknik Elektro   201

insight into the different reference sources and the level of contribution of each in the development of knowledge in the field of event extraction.

Figure 5 shows the number of publications per journal based on processed data. Each point on the graph represents a journal, with the horizontal axis (x) showing the name of the journal and the vertical axis (y) showing the number of publications published in that journal. From the graph, we can observe that some journals have a higher number of publications than others, indicating greater popularity or research volume in these journals. Fluctuations in the number of publications per journal are clearly visible, providing an idea of the distribution of publications across various journals. Journals with a high number of publications may indicate that the journal is a top choice for researchers in a particular field, while journals with a lower number of publications may indicate specialization or a narrower focus. This graph provides important insights for researchers and academics in choosing journals for publication of their work, as well as understanding the trends and popularity of journals in their field of study.

Based on Figure 6, data analysis of methods used in various publications, there are several methods that stand out. The most frequently used method is BERT (Bidirectional Encoder Representations from Transformers), with a frequency of use of 4 times. This method is very popular in natural language processing (NLP) because of its advanced ability to understand the context of words in a sentence. Other methods that are also frequently used include narrative maps, reinforcement learning, rule-based, and CRF (Conditional Random Fields), each with a frequency of use of 2 times. Narrative maps are used to map story narratives, while reinforcement learning is applied in the context of machine learning that requires sequential decisions. Rule-based and CRF are used for rule-based information extraction and probabilistic modeling in pattern recognition.

Other methods that appear in the top 10 list, but with a lower frequency of use (1 time each), include pseudo-triggering, turning entities, dictionary-based, rules based, and schema-based prompts for natural language generation. Based on Figure 7, Pseudo-triggering and entity turning methods are related to specific techniques for extracting entities and events, while dictionary-based and rules-based rely on certain dictionaries and rules for the extraction process. Schema-based prompts for natural language generation is a recent technique used to generate natural language based on a predetermined schema. This data visualization shows that although there are several dominant methods, the variety of methods used reflects the breadth of approaches and innovations applied in event extraction and story narrative research. The dominance of BERT reflects recent trends in the utilization of transformer models for NLP, while the presence of other methods indicates adaptation and specialization for different research needs.

There are six articles that discuss Convolutional Neural Networks (CNN), which is one of the important methods in applying Deep Learning as shown in Figure 8. In addition, there are five articles that focus on Recurrent Neural Networks (RNN), a type of neural network architecture that very effective in processing sequential data. In addition, there are also three articles that explore Distributed Word Embedding, which is an important technique for representing words in vector form in multidimensional space. This collection of articles comprehensively covers various aspects of the application of Deep Learning in natural language processing and text analysis. This provides a strong foundation for the development of innovative and advanced solutions in this field.

In the context of event extraction with limited resources, there are seven articles that discuss the Zero Shot method, which is a very relevant and important approach which shown in Figure 9. Additionally, there are two articles that focus on the Few Shots approach, which is also a method that has significant implications in this context. Additionally, there are three articles that explore Cross Lingual aspects, demonstrating efforts to overcome the challenges of event extraction across languages. With this collection of articles, there is a comprehensive and diverse body of knowledge on dealing with event extraction with limited resources, providing a solid foundation for research and development in this crucial area.

Zero-shot, few-shot, and cross-lingual methods are important tools in natural language processing and machine learning that help overcome various challenges that often arise in text and language analysis. Zero-shot learning allows the model to understand and handle language tasks without special training data, enabling it to classify text in languages it has never encountered or translate language pairs it has never seen. Few-shot learning is useful when we have very limited training data; the model can understand tasks with only a few examples per class or task. This is useful in document classification or query comprehension. Meanwhile, cross-lingual learning allows models to transfer knowledge from one language to another, even when the training data is only in one language. This enables cross-language translation, sentiment analysis in multiple languages, or cross-language entity and relationship understanding with limited resources. This combination of methods can help overcome resource limitations, understand languages that are not present in the training data, and enable broader and more effective applications of NLP across multilingual contexts.
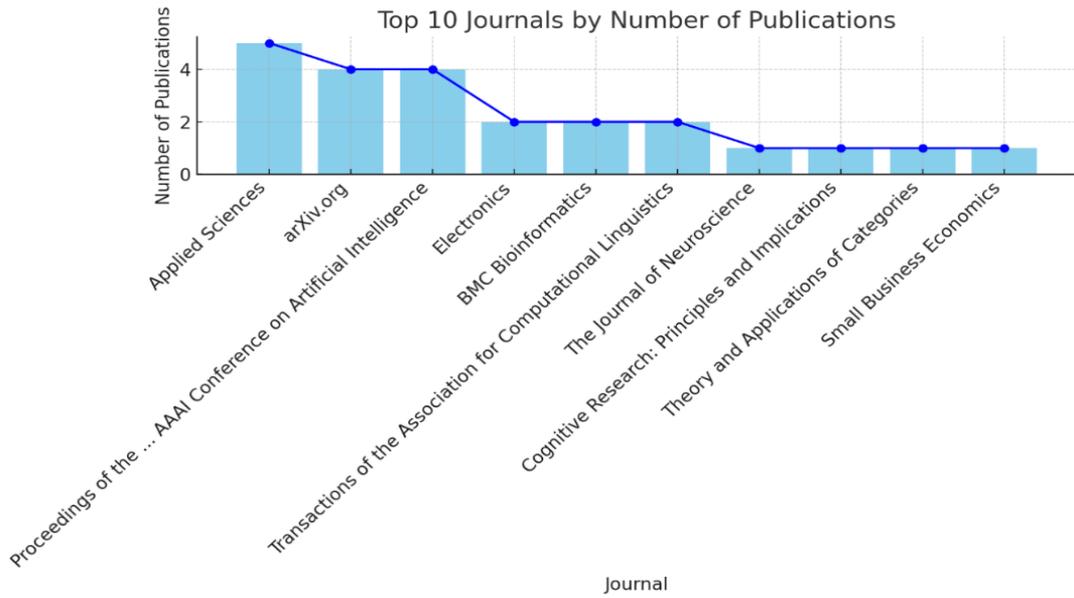
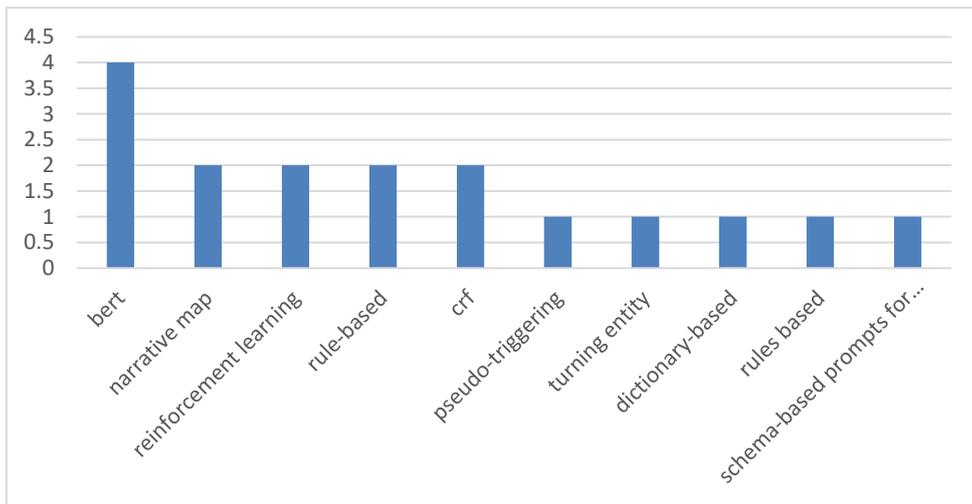**Figure 5.** The number of scientific articles by publisher
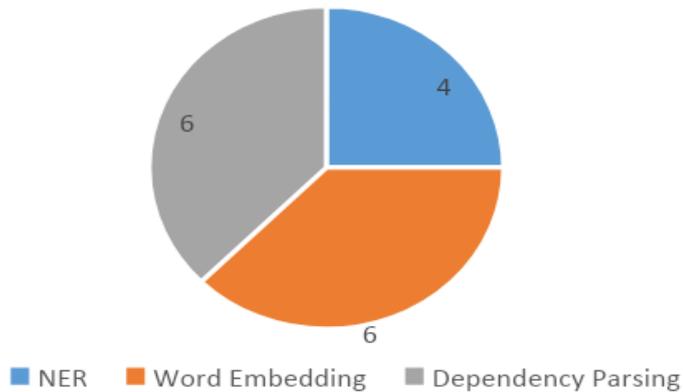

Figure 6. List of 10 frequently used methods

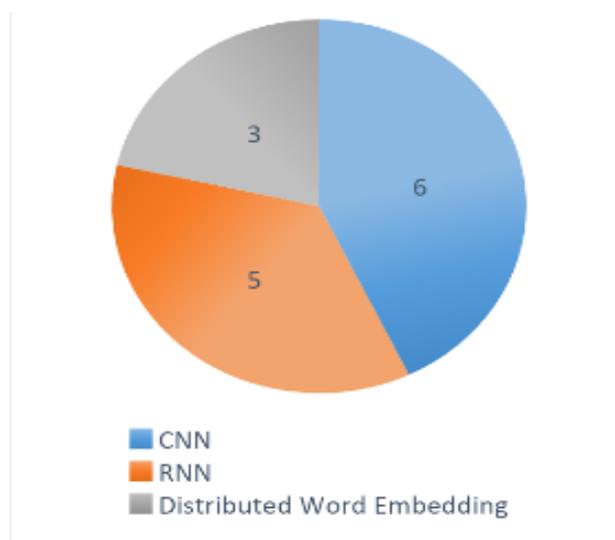**Figure 7.** Number of Scientific Articles applying NLP

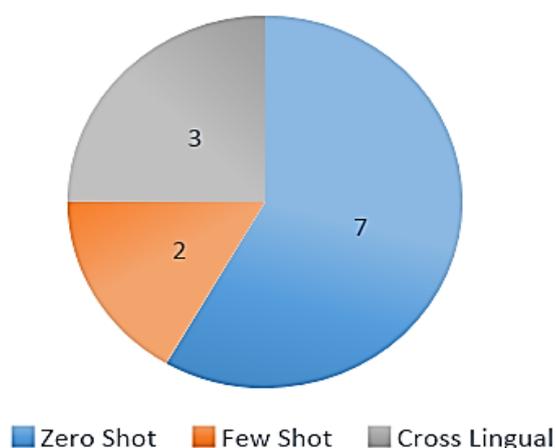**Figure 8.** Number of articles containing Deep Learning methods

Figure 9. Number of Scientific Articles using Low Resource Extraction

4.3. Critical Discussion

This systematic review contributes to the growing body of literature on event extraction by specifically focusing on story narratives, a domain that remains underrepresented in mainstream NLP research. When compared to prior comprehensive surveys such as Li *et al.* (2022), which primarily addresses general-domain or news-based event extraction using deep learning, our findings reveal a critical divergence: while transformer-based models like BERT, DEEIA, and PAIE have achieved strong performance in structured, sentence-level contexts (e.g., ACE2005, MAVEN), their adaptation to narrative-specific challenges such as implicit causality, long-range dependencies, and character-driven plot progression, remains superficial and often unvalidated.

Unlike general-domain surveys that emphasize architectural innovation and cross-dataset benchmarking, this review highlights a systemic gap: the absence of standardized, publicly available narrative-specific benchmarks. Most of the 12 studies included either rely on custom-built corpora (e.g., Andersen's Fairy Tales dataset) or repurpose datasets like ROCStories, while narrative in form, lack rich event annotations aligned with literary structures (e.g., motivation, consequence, emotional arc). This contrasts sharply with domains like biomedical or financial event extraction, where community-wide benchmarks (e.g., BioNLP, ChFinAnn) have driven reproducible progress. The lack of a shared evaluation framework for narrative event extraction hinders both model comparison and cumulative scientific advancement.

Moreover, our synthesis reveals that even state-of-the-art models like PAIE and DEEIA, though designed for document-level extraction, are rarely evaluated on true multi-paragraph narratives (e.g., novels, screenplays). Instead, they are tested on short, self-contained stories or synthetic data, limiting their applicability to real-world literary or historical texts. This reflects a broader tendency in the field to treat

"narrative" as a stylistic label rather than a structural and semantic challenge requiring specialized modeling—such as integration of commonsense reasoning, temporal coherence modeling, or character-centric event tracking.

This review also acknowledges several methodological limitations. First, the final corpus is limited to 12 peer-reviewed studies, reflecting both the novelty of the topic and the stringency of inclusion criteria (e.g., explicit focus on story narratives, use of NLP/deep learning methods). While this ensures thematic relevance, it may exclude valuable work published in non-indexed venues or as preprints. Second, the restriction to English-language publications inevitably overlooks multilingual narrative traditions (e.g., folktales from Asia, Africa, or Latin America) that could offer diverse structural and cultural perspectives on event representation. Future reviews should expand linguistic and geographic scope to capture a more inclusive landscape of narrative understanding.

Finally, while PRISMA guidelines enhance transparency, the timeframe constraint (2017–2024) may omit foundational work in computational narratology or early symbolic approaches to story understanding. Nevertheless, by centering on deep learning-era advances, this review captures the current technological frontier while exposing the urgent need for narrative-aware evaluation protocols, cross-cultural story corpora, and interdisciplinary collaboration between NLP researchers, literary scholars, and cognitive scientists.

5.   CONCLUSIONS

Research into event extraction in narrative stories is expected to continue to grow by taking advantage of technological advances in NLP and artificial intelligence (AI). One of the main trends is the use of Transformer-based models, such as BERT and GPT-3, which are able to understand more complex contexts and better handle language ambiguity. Researchers will also focus more on developing deep learning techniques that can handle limited annotated data, such as transfer learning and semi-supervised learning. Additionally, developing algorithms that can identify and reduce bias in NLP models will be a priority to ensure fairness and accuracy in event extraction. Future research will also explore multimodal integration of data, such as text, images, and audio, to provide a more comprehensive understanding of narratives. This approach not only increases the precision of event extraction but also opens up new opportunities for applications in the fields of health, law, education, and entertainment, which require a deep and accurate understanding of narrative stories.

Extraction of events in narrative stories, especially fairy tales, is becoming a trend in recent research in the fields of NLP and AI. This trend is driven by the need to understand and analyze story structures that are complex and rich in cultural elements. Researchers focus on developing algorithms that can recognize and categorize key events, characters, and relationships between events in stories. This not only makes it easier to digitize and archive traditional stories, but also opens up opportunities for new applications such as language learning, educational content development, and the creation of interactive narratives. This technology also has the potential to revive almost forgotten fairy tales and integrate them in modern media, enriching the user experience and preserving cultural heritage. Future research is expected to be able to overcome challenges such as language variations, cultural contextualization, and interpretation of deep meaning in stories, so as to produce more sophisticated and accurate models in the extraction of narrative events.

REFERENCES

[1]   J. Duan, X. Zhang, and C. Xu, "LwF4IEE: An Incremental Learning Method for Interactive Event Extraction," in *2022 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC)*, pp. 104–113, 2022, https://doi.org/10.1109/CyberC55534.2022.00026.

[2]   E. Daniati, A. P. Wibawa, W. S. G. Irianto, and L. Hernandez, "Few-Shot-BERT-RNN Narrative Structure Analysis for Andersen's Stories," *JOIV Int. J. Informatics Vis.*, vol. 9, no. 4, p. 1773, 2025, https://doi.org/10.62527/joiv.9.4.3932.

[3]   B. Franz, "Narrative time and International Relations," *J. Int. Relations Dev.*, vol. 25, no. 3, pp. 761–783, 2022, https://doi.org/10.1057/s41268-022-00261-3.

[4]   A. Aktan-Erciyes, "Effects of second language on motion event lexicalization: Comparison of bilingual and monolingual children's frog story narratives," *J. Lang. Linguist. Stud.*, vol. 16, no. 3, pp. 1127–1145, 2020, https://doi.org/10.17263/JLLS.803576.

[5]   A. P. Wibawa and F. Kurniawan, "Advancements in natural language processing: Implications, challenges, and future directions," *Telemat. Informatics Reports*, vol. 16, p. 100173, 2024, https://doi.org/10.1016/j.teler.2024.100173.

[6]   R. Ali and B. S. Mustafa, "A Survey on Event Extraction from Webpage," in *2022 8th International Conference on Contemporary Information Technology and Mathematics (ICCITM)*, pp. 159–164, 2022, https://doi.org/10.1109/ICCITM56309.2022.10031730.

[7]   W. Xiang and B. Wang, "A Survey of Event Extraction From Text," *IEEE Access*, vol. 7, pp. 173111–173137, 2019, https://doi.org/10.1109/ACCESS.2019.2956831.

[8] H. L. Nisa and A. Ahdika, "Hybrid Method for User Review Sentiment Categorization in ChatGPT Application Using N-Gram and Word2Vec Features," *Knowl. Eng. Data Sci.*, vol. 7, no. 1, p. 13, 2024, https://doi.org/10.17977/um018v7i12024p13-26.

[9] Y. Ardian, N. D. Irawan, S. Sutoko, I. N. G. A. Astawa, I. B. I. Purnama, and F. A. Dwiyanto, "A Novel Approach to Defect Detection in Arabica Coffee Beans Using Deep Learning: Investigating Data Augmentation and Model Optimization," *Knowl. Eng. Data Sci.*, vol. 7, no. 1, p. 117, 2024, https://doi.org/10.17977/um018v7i12024p117-127.

[10] W. Liu, W. Chen, D. Zeng, L. Zhou, and H. Qu, "CARLG: Leveraging Contextual Clues and Role Correlations for Improving Document-level Event Argument Extraction," *arXiv preprint arXiv:2310.05116*, 2023, https://doi.org/10.48550/arxiv.2310.05116.

[11] Y. Ma *et al.*, "Prompt for Extraction? PAIE: Prompting Argument Interaction for Event Argument Extraction," *arXiv preprint arXiv:2202.12109*, 2022, https://doi.org/10.48550/arXiv.2202.12109.

[12] A. Balali, M. Asadpour, and S. H. Jafari, "COfEE: A comprehensive ontology for event extraction from text," *Computer Speech & Language*, vol. 89, p. 101702, 2025, https://doi.org/10.1016/j.csl.2024.101702.

[13] E. Daniati, A. P. Wibawa, W. S. G. Irianto, A. Ghosh, and L. Hernandez, "Analyzing event relationships in Andersen's Fairy Tales with BERT and Graph Convolutional Network (GCN)," *Sci. Inf. Technol. Lett.*, vol. 5, no. 1, pp. 40–60, 2024, https://doi.org/10.31763/sitech.v5i1.1810.

[14] M. Tydykov, M. Zeng, A. Gershman, and R. Frederking, "Interactive Learning with TREE: Teachable Relation and Event Extraction System," in *Natural {Language} {Processing} and {Information} {Systems}*, 261–274, 2015, https://doi.org/10.1007/978-3-319-19581-0_23.

[15] E. Daniati, A. P. Wibawa, and W. S. G. Irianto, "Event Extraction in Narrative Texts: A Zero-Shot Approach Using Bert and Bi-LSTM on Andersen's Fairy Tales," in *2025 17th International Conference on Knowledge and Smart Technology (KST)*, pp. 208–213, 2025, https://doi.org/10.1109/KST65016.2025.11003342.

[16] T. L. M. Suryanto, A. P. Wibawa, H. Hariyono, and A. Nafalski, "Comparative Performance of Transformer Models for Cultural Heritage in NLP Tasks," *Adv. Sustain. Sci. Eng. Technol.*, vol. 7, no. 1, p. 0250115, 2025, https://doi.org/10.26877/asset.v7i1.1211.

[17] N. Voskarides, E. Meij, S. Sauer, and M. de Rijke, "News Article Retrieval in Context for Event-centric Narrative Creation," In *Proceedings of the 2021 ACM SIGIR International Conference on Theory of Information Retrieval*, pp. 103-112, 2021, https://doi.org/10.1145/3471158.3472247.

[18] D. B. Audretsch and E. E. Lehmann, "Narrative entrepreneurship: bringing (his)story back to entrepreneurship," *Small Bus. Econ.*, vol. 60, pp. 1593–1612, 2022, https://doi.org/10.1007/s11187-022-00661-2.

[19] Z. Yan and X. Tang, "Narrative Graph: Telling Evolving Stories Based on Event-centric Temporal Knowledge Graph," *J. Syst. Sci. Syst. Eng.*, vol. 32, no. 2, pp. 206–221, 2023, https://doi.org/10.1007/s11518-023-5561-0.

[20] R. N. Devendra Kumar, K. Srihari, C. Arvind, and W. Viriyasitavat, "Biomedical event extraction on input text corpora using combination technique based capsule network," *Sādhanā*, vol. 47, no. 4, p. 198, 2022, https://doi.org/10.1007/s12046-022-01978-0.

[21] A. Pranolo *et al.*, "Exploring LSTM-based Attention Mechanisms with PSO and Grid Search under Different Normalization Techniques for Energy demands Time Series Forecasting," *Knowl. Eng. Data Sci.*, vol. 7, no. 1, p. 1, 2024, https://doi.org/10.17977/um018v7i12024p1-12.

[22] C. N. Akpen, S. Asaolu, S. Atobatele, H. Okagbue, and S. Sampson, "Impact of online learning on student's performance and engagement: a systematic review," *Discov. Educ.*, vol. 3, no. 1, p. 205, 2024, https://doi.org/10.1007/s44217-024-00253-0.

[23] S. C. Martino *et al.*, "Evaluation of a protocol for eliciting narrative accounts of pediatric inpatient experiences of care," *Health Serv. Res.*, vol. 58, no. 2, pp. 271–281, 2023, https://doi.org/10.1111/1475-6773.14134.

[24] Y. Xu, "Evaluation & Analysis of Movie Aspects: Based on Sentiment Analysis," *Front. Manag. Sci.*, vol. 2, no. 3, pp. 64–116, 2023, https://doi.org/10.56397/fms.2023.06.08.

[25] B. Lin, D. Bouneffouf, and G. Cecchi, "Predicting Human Decision Making with LSTM," in *2022 International Joint Conference on Neural Networks (IJCNN)*, Jul. 2022, pp. 1–8. https://doi.org/10.1109/IJCNN55064.2022.9892963.

[26] Z. Chen *et al.*, "MedRAGent: An Automatic Literature Retrieval and Screening System Utilizing Large Language Models with Retrieval-Augmented Generation," *medRxiv*, p. 2025-09, 2025. https://doi.org/10.1101/2025.09.18.25335860.

[27] C. Zhang, "AI in Academia: How it Enhances Research Efficiency and Innovation," *Int. J. Educ. Humanit.*, vol. 19, no. 3, pp. 155–158, 2025, https://doi.org/10.54097/r5h5pz92.

[28] S. Rohatgi, Y. Qin, B. Aw, N. Unnithan, and M.-Y. Kan, "The ACL OCL Corpus: Advancing Open Science in Computational Linguistics," in *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, 10348–10361, 2023, https://doi.org/10.18653/v1/2023.emnlp-main.640.

[29] P. Ranade, S. Dey, A. Joshi, and T. Finin, "Computational Understanding of Narratives: A Survey," *IEEE Access*, vol. 10, pp. 101575–101594, 2022, https://doi.org/10.1109/ACCESS.2022.3205314.

[30] C. S. Burns, T. Nix, R. M. Shapiro, and J. T. Huber, "MEDLINE search retrieval issues: A longitudinal query analysis of five vendor platforms," *PLoS One*, vol. 16, no. 5, p. e0234221, 2021, https://doi.org/10.1371/journal.pone.0234221.

[31] L. Zhan, X. Jiang, and Q. Liu, "Research on Chinese event extraction method based on HMM and multi-stage method," In *Journal of Physics: Conference Series*, vol. 1732, no. 1, p. 012024, 2021, https://doi.org/10.1088/1742-6596/1732/1/012024.

[32] A. Darmawan, M. N. Ahmad, and R. M. N. Sakinah, "An Analysis Of The Conflicts Of The Main Character's And The Plot Structure In The 'Do Revenge' (2022)," *GIST*, vol. 7, no. 1, 2024, https://doi.org/10.53675/gist.v7i1.1385.

[33] Q. Li *et al.*, "A Survey on Deep Learning Event Extraction: Approaches and Applications," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 35, no. 5, pp. 6301–6321, 2024, https://doi.org/10.1109/TNNLS.2022.3213168.

[34] H. Yang, D. Sui, Y. Chen, K. Liu, J. Zhao, and T. Wang, "Document-level Event Extraction via Parallel Prediction Networks," in *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pp. 6298-6308, 2021, https://doi.org/10.18653/v1/2021.acl-long.492.

[35] D. F. Laistulloh, A. N. Handayani, R. A. Asmara, and P. Taw, "Convolutional Neural Network in Motion Detection for Physiotherapy Exercise Movement," *Knowl. Eng. Data Sci.*, vol. 7, no. 1, p. 27, 2024, https://doi.org/10.17977/um018v7i12024p27-39.

[36] M. Gupta and P. Agrawal, "Compression of Deep Learning Models for Text: A Survey," *ACM Trans. Knowl. Discov. Data*, vol. 16, no. 4, pp. 1–55, 2022, https://doi.org/10.1145/3487045.

[37] J. Li *et al.*, "Study on strategy of CT image sequence segmentation for liver and tumor based on U-Net and Bi-ConvLSTM," *Expert Syst. Appl.*, vol. 180, p. 115008, 2021, https://doi.org/10.1016/j.eswa.2021.115008.

[38] A. Q. Nada, A. P. Wibawa, D. F. Putri Syarifa, E. Fajarwati, and F. I. Putri, "Optimizing Indonesian-Sundanese Bilingual Translation with Adam-Based Neural Machine Translation," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 8, no. 6, pp. 690–700, 2024, https://doi.org/10.29207/resti.v8i6.6116.

[39] B. K. Norambuena, T. Mitra, and C. North, "A Survey on Event-based News Narrative Extraction," *ACM Comput. Surv.*, vol. 55, no. 14s, pp. 1-39, 2023, https://doi.org/10.1145/3584741.

[40] A. P. Wibawa *et al.*, "Deep Learning Approaches with Optimum Alpha for Energy Usage Forecasting," *Knowl. Eng. Data Sci.*, vol. 6, no. 2, p. 170, 2023, https://doi.org/10.17977/um018v6i22023p170-187.

[41] H. A. Kheder, "Human-Computer Interaction: Enhancing User Experience in Interactive Systems," *Kufa J. Eng.*, vol. 14, no. 4, pp. 23–41, 2023, https://doi.org/10.30572/2018/KJE/140403.

[42] N. Anantrasirichai and D. Bull, "Artificial intelligence in the creative industries: a review," *Artif. Intell. Rev.*, vol. 55, no. 1, pp. 589–656, 2022, https://doi.org/10.1007/s10462-021-10039-7.

[43] H. Fang, G. Jiang, and D. Li, "Sentiment analysis based on Chinese BERT and fused deep neural networks for sentence-level Chinese e-commerce product reviews," *Syst. Sci. Control Eng.*, vol. 10, no. 1, pp. 802–810, 2022, https://doi.org/10.1080/21642583.2022.2123060.

[44] C. V. Ibeh, O. F. Asuzu, T. Olorunsogo, O. A. Elufioye, N. L. Nduubuisi, and A. I. Daraojimba, "Business analytics and decision science: A review of techniques in strategic business decision making," *World J. Adv. Res. Rev.*, vol. 21, no. 2, pp. 1761–1769, 2024, https://doi.org/10.30574/wjarr.2024.21.2.0247.

[45] S. Mahmoud El Sayed El Khouly, K. Yasser, and E. Ahmed Yehia, "DEVELOPING A NEW BUSINESS OPPORTUNITIES VIA ARTIFICIAL INTELLIGENCE: NEW STRATEGIC MANAGEMENT MODEL," *Scientific Journal of Economics and Commerce*, vol. 52, no. 4, pp. 677–706, Dec. 2022, https://doi.org/10.21608/jsec.2022.268733.

[46] W. Yu, M. Yi, X. Huang, X. Yi, and Q. Yuan, "Make It Directly: Event Extraction Based on Tree-LSTM and Bi-GRU," *IEEE Access*, vol. 8, pp. 14344–14354, 2020, https://doi.org/10.1109/ACCESS.2020.2965964.

[47] Z. Liao, Z. Yang, P. Huang, N. Pang, and X. Zhao, "Multi-Model Fusion-Based Hierarchical Extraction for Chinese Epidemic Event," *Data Sci. Eng.*, vol. 8, no. 1, pp. 73–83, 2023, https://doi.org/10.1007/s41019-022-00203-6.

[48] V. V. Vydiswaran, X. Zhao, and D. Yu, "Data science and natural language processing to extract information from clinical narratives," In *Proceedings of the 3rd ACM India Joint International Conference on Data Science & Management of Data (8th ACM IKDD CODS & 26th COMAD)*, pp. 441–442, 2021, https://doi.org/10.1145/3430984.3431967.

[49] E. Daniati, A. P. Wibawa, and W. S. G. Irianto, "Building a Narrative Event Dataset from Andersen's Fairy Tales for Literary and Computational Analysis," *Int. J. Eng. Sci. Inf. Technol.*, vol. 5, no. 3, pp. 354–362, 2025, https://doi.org/10.52088/ijesty.v5i3.910.

[50] E. Daniati, A. P. Wibawa, and W. S. G. Irianto, "Extracting Narrative Events in Andersen's Fairy Tales Using a Hybrid BERT-LSTM Model," In *International Symposium on Systems Modelling and Simulation*, pp. 355–372, 2025, https://doi.org/10.1007/978-981-96-4613-5_26.

[51] B. Cai, X. Ding, B. Chen, L. Wu, and T. Liu, "Mitigating Reporting Bias in Semi-supervised Temporal Commonsense Inference with Probabilistic Soft Logic," *Proc. ... AAAI Conf. Artif. Intell.*, vol. 36, no. 10, pp. 10454–10462, 2022, https://doi.org/10.1609/aaai.v36i10.21288.

[52] A. Leeuwenberg and M.-F. Moens, "Towards Extracting Absolute Event Timelines From English Clinical Reports," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 28, pp. 2710–2719, 2020, https://doi.org/10.1109/TASLP.2020.3027201.

[53] N. Borenstein, N. da S. Perez, and I. Augenstein, "Multilingual Event Extraction from Historical Newspaper Adverts," in *Annual Meeting of the Association for Computational Linguistics*, pp. 10304-10325, 2023, https://doi.org/https://doi.org/10.48550/arXiv.2305.10928.

[54] R. Duan *et al.*, "Sentiment Classification Algorithm Based on the Cascade of BERT Model and Adaptive Sentiment Dictionary," *Wirel. Commun. Mob. Comput.*, vol. 2021, pp. 1–8, 2021, https://doi.org/10.1155/2021/8785413.

[55] Q. Li *et al.*, "Event Extraction by Associating Event Types and Argument Roles," *IEEE Trans. Big Data*, vol. 9, no. 6, pp. 1549–1560, 2023, https://doi.org/10.1109/TBDATA.2023.3291563.

[56] N. Anantharama, S. Angus, and L. O'Neill, "Canarex: Contextually aware narrative extraction for semantically rich text-as-data applications," In *Findings of the Association for Computational Linguistics: EMNLP 2022*, pp. 3551-3564 2022, https://doi.org/10.18653/v1/2022.findings-emnlp.260.

[57] Y. K. Dwivedi *et al.*, "'So what if ChatGPT wrote it?' Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy," *Int. J. Inf. Manage.*, vol. 71, 2023, https://doi.org/10.1016/j.ijinfomgt.2023.102642.

[58] J. Xie, H. Sun, J. Zhou, W. Qu, and X. Dai, "Event Detection as Graph Parsing," in *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pp. 1630–1640, 2021, https://doi.org/10.18653/v1/2021.findings-acl.142.

[59] M. Waseem, Q. Umer, C. Lee, S. Chung, and Z. Latif, "Deep Learning-Based Event Prediction for Text Analysis," in *2023 14th International Conference on Information and Communication Technology Convergence (ICTC)*, pp. 42–47, 2023, https://doi.org/10.1109/ICTC58733.2023.10392730.

[60] A. Main *et al.*, "Patients' Experiences of Digital Health Interventions for the Self-Management of Chronic Pain: Systematic Review and Thematic Synthesis (Preprint)," *Journal of medical Internet research*, vol. 27, p. e69100, 2024, https://doi.org/10.2196/preprints.69100.

[61] X. Li, Y. Lei, and S. Ji, "BERT- and BiLSTM-Based Sentiment Analysis of Online Chinese Buzzwords," *Futur. Internet*, vol. 14, no. 11, p. 332, 2022, https://doi.org/10.3390/fi14110332.

[62] I. G. A. P. Mahendra, I. M. A. Wirawan, and I. G. A. Gunadi, "Enhancement performance of the Naïve Bayes method using AdaBoost for classification of diabetes mellitus dataset type II," *Int. J. Adv. Appl. Sci.*, vol. 13, no. 3, p. 733, 2024, https://doi.org/10.11591/ijaas.v13.i3.pp733-742.

[63] A. Pranolo, Y. Mao, A. P. Wibawa, A. B. P. Utama, and F. A. Dwiyanto, "Optimized Three Deep Learning Models Based-PSO Hyperparameters for Beijing PM2.5 Prediction," *Knowl. Eng. Data Sci.*, vol. 5, no. 1, p. 53, 2022, https://doi.org/10.17977/um018v5i12022p53-66.

[64] N. Keivandarian and M. Carvalho, "A Survey on Sentiment Classification Methods and Challenges," *Int. FLAIRS Conf. Proc.*, vol. 36, 2023, https://doi.org/10.32473/flairs.36.133314.

[65] Q. Zaman, S. Safwandi, and F. Fajriana, "Supporting Application Fast Learning of Kitab Kuning for Santri' Ula Using Natural Language Processing Methods," *Int. J. Eng. Sci. Inf. Technol.*, vol. 5, no. 1, pp. 278–289, 2025, https://doi.org/10.52088/ijesty.v5i1.713.