

Storyline Extraction of Document-Level Events Using Large Language Models

Ziyang Hu¹, Yaxiong Li^{2*}

¹School of Foreign Studies, Zhongnan University of Economics and Law, Wuhan, China

²Information Center, Hubei University of Science and Technology, Wuhan, China

Email: huziyang@stu.zuel.edu.cn, *liyaxiong@hbust.edu.cn

How to cite this paper: Hu, Z.Y. and Li, Y.X. (2024) Storyline Extraction of Document-Level Events Using Large Language Models. *Journal of Computer and Communications*, 12, 162-172.

<https://doi.org/10.4236/jcc.2024.1211012>

Received: October 22, 2024

Accepted: November 24, 2024

Published: November 27, 2024

Copyright © 2024 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

This article proposes a document-level prompt learning approach using LLMs to extract the timeline-based storyline. Through verification tests on datasets such as ESCv1.2 and Timeline17, the results show that the prompt + one-shot learning proposed in this article works well. Meanwhile, our research findings indicate that although timeline-based storyline extraction has shown promising prospects in the practical applications of LLMs, it is still a complex natural language processing task that requires further research.

Keywords

Document-Level Storyline Extraction, Timeline, Large Language Models, Topological Structure of Storyline, Prompt Learning

1. Introduction

The explosive growth of information in cyberspace has surrounded people with information from news websites, media, and other sources every day. It isn't easy to understand the development and evolution of a news event in a comprehensive, accurate, and timely manner. Most users obtain fragmented news information, and it is difficult to grasp and understand the overall development process of the event, which has become a thorny problem of information access in the era of big data. In this context, Makkonen [1] regarded storyline mining as a subtopic of topic detection and tracking, using graph structures to represent storylines and event evolution relationships. Later, Nallapati *et al.* [2] proposed the concept of Event Threading, laying the foundation for storyline mining. Storyline mining is the process of analyzing the relationship between events in a text and subsequent related events, constructing a storyline, and conducting event evolution analysis

based on this, to uncover the stages and patterns of event evolution.

Through storyline mining, it is possible to obtain a clearly structured storyline and the lifecycle of event evolution from massive, fragmented information, thereby enhancing the application value of web information.

Due to the lack of predefined event types and explicit fixed scenarios in open domain event extraction, the text used comes from many complex corpora such as social media, making the extraction laborious and the related research not yet in-depth enough. At present, academic research is more profound regarding event extraction in restricted domains. This article attempts, for the first time, to use large language models to propose a timeline-based method for extracting open domain storylines through prompt learning.

The subsequent content is arranged as follows: Chapter 2 reviews the current status, definition, and topological structure of storyline extraction, related technologies, and timeline summarization-based methods for constructing a storyline. Chapter 3 elaborates on the practical content of timeline-based storyline extraction based on large language models, including experimental results, evaluation system, and result analysis. Chapter 4 discusses the prospects for future research.

The experimental results show that using the method mentioned above, GPT-4 achieves accuracies of 100% and 99.4% on the ESCv1.2 and TimeLine17 datasets, respectively, and verifies that prompt + one-shot Learning performs more stable than base prompt, with higher accuracy and recall. This article can provide a reference for subsequent research on the in-depth application of large models in NLP.

2. Literature Review

The systematic research on storyline mining can be traced back to the Text Retrieval Conference (TREC) held in 2013. The TREC 2013 conference held the first evaluation of timeline summarization, requiring participants to extract the timeline of events from news corpora, which was an early approach to storyline mining. Subsequently, multiple conferences also conducted similar evaluations. In 2014, TAC (Text Analysis Conference) conducted an event tracking evaluation; In 2015, SemEval (International Workshop on Semantic Evaluation) held a competition to rank cross-document event timelines; In 2017, ACL (Annual Meeting of the Association for Computational Linguistics) held the first seminar on “Events and Stories in the News”; In 2024, EAACL (Annual Conference of the European Chapter of Association for Computational Linguistics) held The 7th Workshop on Challenges and Applications of Automated Extraction of Socio-political Events from Text. So far, traditional machine learning and deep learning methods have not been able to achieve satisfactory experimental results. Based on this background, this article proposes a prompt learning method based on large language models, which achieved 100% accuracy and opened up a good beginning for the subsequent application of LLMs in this direction.

2.1. Related Definitions

There is currently no clear definition in academic research related to storyline mining. Domestic researchers generally use “event storyline”, “timeline”, and “storyline” to describe the evolution of events over time, while English literature mostly uses “timeline summarization” and “storyline”. Based on these definitions, this article proposes the concept of “branch” to represent subevents in a complete storyline.

Definition 1 Event: An event that occurs at a specific time and place, involving one or more objects and consisting of one or more actions [2].

Definition 2 Branch: Several subevents on an event timeline connected by event relationships, composed of multiple events with the same theme [3] [4].

Definition 3 Storyline: Composed of one or more related branches, representing the topological structure of the evolutionary relationship of events over time.

Definition 4 Event Evolution: The process of events resembling the development of things in philosophy, with a similar lifecycle of germination, growth, peak, decline, and demise [5]-[7].

2.2. Topological Structure of Storyline

In event development, subevents are interdependent and follow specific evolutionary patterns. Wang *et al.* [8] believed that there is a chain reaction of subevents in emergencies. They constructed a network topology of the chain reaction of emergencies using the causal relationship of subevents. In the storyline discussed in this article, the evolution of subevents over time mainly unfolds from the structure of the storyline. We categorize the evolutionary patterns into three types: chain structure, tree structure, and reticular structure.

Chain structure refers to the correlation between the cause of a subevent and the occurrence of the previous subevent, with the subevents arranged sequentially on the timeline, as shown in **Figure 1(a)**. Alonso *et al.* [9]-[11] proposed a method based on similarity analysis, which only considers the correlation between subevents when constructing the storyline and arranges subevents according to temporal characteristics; Mishra *et al.* [12]-[14] proposed a method based on timeline summarization, which considers the storyline construction as a problem of multi-document summarization.

Tree structure is composed of multiple chain structures, where branches evolve in their directions, as shown in **Figure 1(b)**. Researchers [15] [16] applied Bayesian Modeling and found different branches where documents are located. Ansah *et al.* [17] [18] used clustering-based methods to link events to distinct branches. Yuan *et al.* [4] [19] [20]. Proposed a propagation model-based method that uses optimization strategies to generate trees from the graph and thus storylines. These methods often generate multi-cue storylines, *i.e.*, tree structures.

Reticular structure refers to the relationship between a single branch and other branches in the storyline, or the mutual relationship between multiple branches, as shown in **Figure 1(c)**. Zhang *et al.* [21] analyzed the event relationships between

different branches and obtained the connections between events in different branches, which promote and interweave with each other to form a reticular structure of a storyline.

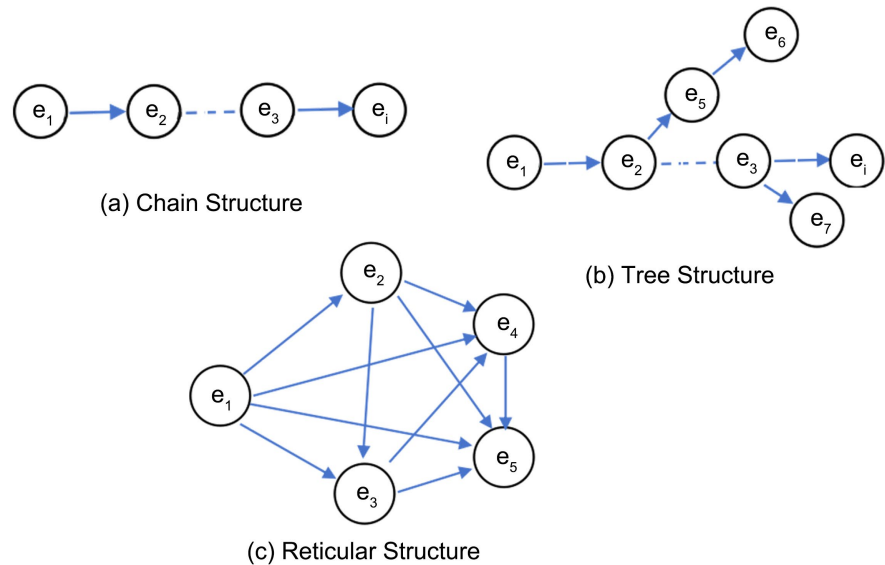


Figure 1. Topological structure of storyline.

2.3. Technologies Related to Storyline Mining

Due to the fragmentation and disorder of Internet news, it is difficult for people to directly obtain the whole process of news event development from the Internet. Compared with using the current fragmented way of information acquisition on the Internet, people prefer to directly get the complete storyline of the whole news event, rather than clear the cause and effect of the event from a pile of disordered data. Although the storyline generated by manual editing is relatively accurate, it requires a lot of human resources. Therefore, using machine algorithms to automatically construct the storyline [11] [20]-[24] is an essential task in storyline mining research.

2.4. Timeline-Based Construction of Storyline

The timeline-based method considers storyline generation as a problem of multi-stage document summarization on the timeline. By arranging subevents in chronological order on the timeline and using summarization techniques to select representative sentences (or generate sentences) to summarize the development of events in each stage, the logical coherence and completeness of storyline information are ensured, and the storyline is ultimately obtained. According to different technology roadmaps, Guo *et al.* [17] [25] modeled the event extraction as a sequence labeling problem, using the “event type-parameter role” as the parameter combination label. While completing the event parameter extraction, the role and event type corresponding to the parameter are identified through the parameter label. In addition, in response to the multi-event parameters coupling in the

dataset, Liu *et al.* [26] [27]. The parameters were grouped into several 2-tuples (core parameters, marginal parameter), and the model was used to determine whether the two belonged to the same event. If so, they were combined into the same event instance. With the conceptual hierarchies of “text → topic → event → storyline”, the research on text retrieval based on news topic (text → topic), clustering based on news event (topic → event), and online construction technology of storyline (event → storyline) is carried out step by step.

3. Methodology

3.1. Dataset

The mining of storyline requires rich social event data. In recent years, with the rapid development of the Internet, social news has spread over the Internet, forming an abundant repository of social event resources. The commonly used datasets for storyline mining include two parts: the author’s self-built dataset and the public dataset.

This article uses ESCv1.2 and TimeLine17 as the datasets for the extraction task. ESCv1.2 [28] is a dataset for causal and temporal relation extraction. ESCv1.2 covers 22 news topics and 7778 storylines, and the corpus data has been annotated by experts. TimeLine17 [29] is one of the earliest datasets in the field of timeline-based storylines, consisting of 9 different themes and 19 timeline-based storylines. Its content comes from CNN, BBC, and NBC News, with multiple dates and sentences for each storyline.

3.2. Language Model

In this experiment, we used three LLMs: Vicuna-7b-v1.5, GLM4-9B chat (128K), and GPT-4. Vicuna-7b is an open-source model fine-tuned from Llama 2, with a capacity of 7 billion parameters, capable of continuously outputting responses in the format required for automated tasks [30]. GLM4-9B-chat (128K) is also an open-source model distilled and fine-tuned from GLM4, which even surpasses GPT-4-turbo-2022-04-09 in terms of aware reasoning and comprehensive capabilities of NLP. Its long text inference supports a maximum context length of 128K [31]. GPT-4 is at the forefront of current LLM technology, providing the most advanced functions [32].

3.3. Prompt & One-Shot Learning

Unlike traditional machine learning methods, this method uses conceptual hierarchies of “text → topic → event → storyline” to calculate step by step. We envision using prompt to guide the large language model to read language materials and conduct content reasoning in order to obtain the expected results.

1) Using prompt Learning alone (basic prompt), the prompt template is as follows:

You are a natural language processing expert, and please follow the steps

below to extract the event process from the text.

Step 1: Read the target news and identify the main topic.

Step 2: Read each news paragraph to determine if they are relevant to the topic identified in Step 1. Only consider news that is directly related to the target news or has contextual meaning.

Step 3: Extract all date information from the text and summarize the events that occurred on each date in a concise sentence based on the context of the date information. Use the format [YYYY-MM-DD] or [YYYY-MM] as a time format reference.

Step 4: Please analyze the input text step by step according to the above steps, and finally output it as a JSON object in the following format.

{“[YYYY-MM-DD]”: “Statement describing the event”, or “[YYYY-MM]”: “Statement describing the event”}

2) Using prompt + one-shot Learning, the prompt template is as follows.

The template is divided into two parts, the first part is the Chain-of-thought (CoT) of prompt, and the second part is a single instance for One Shot Learning. The first part is prompt’s Chain-of-thought (CoT) as follows.

Step 1: Identify the theme event of the news.

Step 2: Extract the time sequence of the same events as the first step from the input text. The specific process is to first extract all the specific date information in the input text and summarize the events that occurred on each date in a concise sentence based on the context of the date information.

Step 3: Input format: {“[YYYY-MM-DD]”: “Statement describing the event”, or “YYYY-MM”: “Statement describing the event”}

The second part of one-shot Learning is as follows:

“On April 13, 2021, the Japanese government held a cabinet meeting and officially decided to filter and dilute millions of tons of nuclear contaminated water from the Fukushima Daiichi nuclear plant and discharge it into the sea, which will begin about two years later.

On November 27, 2021, Tokyo Electric Power Company (TEPCO) plans to begin underwater geological surveys in preparation for the construction of an underwater tunnel for discharging contaminated water from the Fukushima Daiichi nuclear plant. In December of the same year, researchers from the Japan Agency for Marine-Earth Science and Technology (JAMSTEC) found that the radioactive substance leaked from the Fukushima Daiichi Nuclear Disaster in 2011 had spread into the Arctic Ocean.

TEPCO plans to conduct four rounds of nuclear contaminated water discharge into the sea within 2023, with the first two rounds taking place from August 24th to September 11th and October 5th to 23rd. The third round will run from November 2nd to November 20th.

The final JSON format obtained is as follows:

{“[2021-04-13]”: “*The Japanese government held a cabinet meeting and officially decided to filter and dilute millions of tons of nuclear contaminated water from the Fukushima Daiichi nuclear plant and discharge it into the sea, which will begin about two years later.*”

“[2021-11-27]”: “*TEPCO is preparing to construct an underwater tunnel for the discharge of Fukushima nuclear contaminated water.*”

“[2021-12]”: “*JAMSTEC has found that the radioactive substance leaked from the Fukushima Daiichi Nuclear Disaster in 2011 had spread into the Arctic Ocean.*”

“[2023-08-24]”: “*TEPCO began the first round of nuclear contaminated water discharge into the sea.*”

“[2023-10-05]”: “*TEPCO began the second round of nuclear contaminated water discharge into the sea.*”

“[2023-11-02]”: “*TEPCO began the third round of nuclear contaminated water discharge into the sea.*”}

3.4. Evaluation Index

Researchers have different evaluation indexes for evaluating the branches in the storyline [5]. This article uses precision (P), recall (R), and F value (F1-score, F1) as evaluation indicators. The formulas are as follows:

$$P = \frac{|E_p \cap E_s|}{E_s} \quad R = \frac{|E_p \cap E_s|}{E_p} \quad F_1 = \frac{2 \times P \times R}{P + R}$$

Formula 1 Evaluation Index

Among them, E_p is the generated event cluster, and E_s is the standard event cluster.

4. Research Results

Table 1. Relevance check performance for basic prompt and prompt + one-shot Learning.

Model	Prompt	TimeLine17			ESC		
		P	R	F1	P	R	F1
Vicuna-7B	Basic prompt	0.785	0.480	0.596	0.805	0.475	0.597
	Prompt + one-shot learning	0.930	0.676	0.783	0.941	0.630	0.755
GLM4-9B-chat (128K)	Basic prompt	0.872	0.592	0.705	0.972	0.618	0.756
	Prompt + one-shot learning	1.00	0.631	0.774	0.985	0.639	0.775
GPT-4	basic prompt	1.00	0.749	0.856	0.990	0.721	0.834
	Prompt + one-shot learning	1.00	0.758	0.862	0.994	0.737	0.846

Table 1 shows the comparative performance of two prompt templates on two different datasets, TimeLine17 and ESCv1.2. In the investigated datasets, both

basic prompt and prompt + One Shot Learning show high F1 scores. Although the basic prompt has high accuracy, its recall is low. This indicates that although the predictions are accurate, they may miss many related articles. In contrast, prompt + One Shot Learning has a more stable performance, improves accuracy and recall, and ultimately achieves higher F1 scores in all models. This trend suggests that involving LLM in storyline extraction may promote a more detailed evaluation of article publishing, resulting in a fairer selection of news articles. Compared with basic prompts, the performance improvement of prompt + One Shot Learning is more significant.

5. Conclusions and Implications

In our research, we propose a prompt + one-shot Learning method based on LLM to extract news timelines and storylines, which dramatically improves the accuracy of extracting storylines. We hope that this research will not only arouse people's interest but also stimulate further exploration in the field of event timeline construction and refinement of LLM prompt strategies.

The limitation of the research is that it only clarifies the storylines based on the timeline dimension, while improving the interpretability and accuracy of derivative events is also an important issue. In future research, how to use external knowledge to enhance the comprehensibility and accuracy of derivative events, how to integrate human knowledge with machine learning, reinforcement learning, and other technologies to improve the performance of event analysis, and how to fully utilize the link relations of complex networks to enhance the accuracy of event evolution relationship recognition, are also issues worth exploring in depth.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Makkonen, J. (2003) Investigations on Event Evolution in TDT. *Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology Proceedings of the HLT-NAACL 2003 Student Research Workshop—NAACL'03, Canada, 27 May-1 June 2003*, 43-48. <https://doi.org/10.3115/1073416.1073424>
- [2] Nallapati, R., Feng, A., Peng, F. and Allan, J. (2004) Event Threading within News Topics. *Proceedings of the 13th ACM International Conference on Information and Knowledge Management, New York, 8-13 November 2004*, 446-453. <https://doi.org/10.1145/1031171.1031258>
- [3] Dehghani, N. and Asadpour, M. (2018) SGSG: Semantic Graph-Based Storyline Generation in Twitter. *Journal of Information Science*, **45**, 304-321. <https://doi.org/10.1177/0165551518775304>
- [4] Li, Y., Ma, S., Jiang, H., Liu, Z., Hu, C. M. and Li, X. (2018) An Approach for Storytelling by Correlating Events from Social Networks. *Journal of Computer Research and Development*, **55**, 1972-1986.

- <https://doi.org/10.7544/issn1000-1239.2018.20180155>
- [5] Alonso, O., Kandylas, V., Tremblay, S., Hofman, J.M. and Sen, S. (2017) What's Happening and What Happened. *Proceedings of the 2017 ACM on Web Science Conference*, New York, 25-28 June 2017, 191-200. <https://doi.org/10.1145/3091478.3091484>
- [6] Mu, L., Jin, P., Zheng, L. and Chen, E. (2018) Eventsys: Tracking Event Evolution on Microblogging Platforms. In: *Lecture Notes in Computer Science*, Springer, 797-801. https://doi.org/10.1007/978-3-319-91458-9_51
- [7] Lu, X.S., Zhou, M., Qi, L. and Liu, H. (2019) Clustering-Algorithm-Based Rare-Event Evolution Analysis via Social Media Data. *IEEE Transactions on Computational Social Systems*, **6**, 301-310. <https://doi.org/10.1109/tcss.2019.2898774>
- [8] Wang, J.W. and Rong, L.L. (2008) Research on the Chain-Reacting Network Model of Emergency Events. *Application Research of Computers*, **25**, 3288-3291. <https://ieeexplore.ieee.org/document/8667666>
- [9] Alonso, O., Tremblay, S. and Diaz, F. (2017) Automatic Generation of Event Timelines from Social Data. *Proceedings of the 2017 ACM on Web Science Conference*, New York, 25-28 June 2017, 207-211. <https://doi.org/10.1145/3091478.3091519>
- [10] Guo, B., Ouyang, Y., Zhang, C., Zhang, J., Yu, Z., Wu, D., et al. (2017) Crowdstory: Fine-Grained Event Storyline Generation by Fusion of Multi-Modal Crowdsourced Data. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, **1**, 1-19. <https://doi.org/10.1145/3130920>
- [11] Nomoto, T. (2010) Two-Tier Similarity Model for Story Link Detection. *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*, New York, 26-30 October 2010, 789-798. <https://doi.org/10.1145/1871437.1871539>
- [12] Mishra, A. and Berberich, K. (2016) Event Digest. *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*, New York, 17-21 July 2016, 493-502. <https://doi.org/10.1145/2911451.2911526>
- [13] AlNoamany, Y., Weigle, M.C. and Nelson, M.L. (2017) Generating Stories from Archived Collections. *Proceedings of the 2017 ACM on Web Science Conference*, New York, 25-28 June 2017, 309-318. <https://doi.org/10.1145/3091478.3091508>
- [14] Wang, H. and Koh, J.L. (2017) Timeline Summarization for Event-Related Discussions on a Chinese Social Media Platform. In: *Advances in Artificial Intelligence: From Theory to Practice. 30th International Conference on Industrial Engineering and Other Applications of Applied Intelligent Systems*, Springer, 579-594. https://link.springer.com/chapter/10.1007/978-3-319-60042-0_64
- [15] She, Y.X. and Xiong, Y. (2018) Storyline Mining Algorithm Based on Bayesian Network. *Computer Engineering*, **44**, 55-59. <https://doi.org/10.3969/j.issn.1000-3428.2018.03.009>
- [16] Guo, L., Zhou, D., He, Y. and Xu, H. (2020) Storyline Extraction from News Articles with Dynamic Dependency. *Intelligent Data Analysis*, **24**, 183-197. <https://doi.org/10.3233/jida-184448>
- [17] Ansah, J., Liu, L., Kang, W., Kwashie, S., Li, J. and Li, J. (2019). A Graph Is Worth a Thousand Words: Telling Event Stories Using Timeline Summarization Graphs. *The World Wide Web Conference*, New York, 13-17 May 2019, 2565-2571. <https://doi.org/10.1145/3308558.3313396>
- [18] Goyal, P., Kaushik, P., Gupta, P., Vashisth, D., Agarwal, S. and Goyal, N. (2020) Multilevel Event Detection, Storyline Generation, and Summarization for Tweet Streams.

- IEEE Transactions on Computational Social Systems*, **7**, 8-23.
<https://doi.org/10.1109/tcss.2019.2954116>
- [19] Yuan, R., Zhou, Q. and Zhou, W. (2018) dTexSL: A Dynamic Disaster Textual Storyline Generating Framework. *World Wide Web*, **22**, 1913-1933.
<https://doi.org/10.1007/s11280-018-0640-8>
- [20] Fan, X.B., Rao, Y., Wang, S., Li, R.X. and Liu, X.H. (2021) Named Entity Sensitive Generation of Hierarchical News Storyline. *Journal of Chinese Information Processing*, **35**, 113-124. <http://jcip.cipsc.org.cn/CN/Y2021/V35/I1/113>
- [21] Zhang, H., Li, G.H., Sun, B.L. and Jia, L. (2013) Modeling News Event Evolution. *Journal of National University of Defense Technology*, **35**, 166-170.
http://journal.nudt.edu.cn/gfkjdxxb/ch/reader/view_abstract.aspx?file_no=201304029&flag=1
- [22] Zhou, D., Guo, L. and He, Y. (2018) Neural Storyline Extraction Model for Storyline Generation from News Articles. *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, **1**, 1727-1736.
<https://doi.org/10.18653/v1/N18-1156>
- [23] Chen, L.M., Huang, R.Z., Qin, Y.B. and Chen, Y.P. (2020) Story Tree Construction Approach for News Events. *Computer Engineering and Design*, **41**, 1910-1919.
<https://doi.org/10.16208/j.issn1000-7024.2020.07.018>
- [24] Hua, T., Zhang, X., Wang, W., Lu, C. and Ramakrishnan, N. (2016) Automatic Storyline Generation with Help from Twitter. *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*, New York, 24-28 October 2016, 2383-2388. <https://doi.org/10.1145/2983323.2983698>
- [25] Zhao, T.Z., Duan, L., Yue, K., Qiao, S.J. and Ma, Z.J. (2021) Generating News Clues with Biterm Topic Model. *Data Analysis and Knowledge Discovery*, **5**, 1-13.
<https://doi.org/10.11925/INFOTECH.2096-3467.2020.1025>
- [26] Mu, L., Jin, P., Zheng, L., Chen, E. and Yue, L. (2018) Lifecycle-Based Event Detection from Microblogs. *Companion of the Web Conference 2018 on the Web Conference 2018—WWW'18*, Lyon, 23-27 April 2018, 283-290.
<https://doi.org/10.1145/3184558.3186338>
- [27] Liu, G.W. and Cheng, Q. (2018) Research on Topic Evolution of Microblog Hot Events Based on Life Cycle of Network Public Opinion. *Information Research*, **1**, 11-19. <http://www.qbts.org/CN/abstract/abstract7542.shtml>
- [28] Caselli, T. and Inel, O. (2018) Crowdsourcing Story Lines: Harnessing the Crowd for Causal Relation Annotation. In: *Events and Stories in the News*, Association for Computational Linguistics, 44-54. <https://aclanthology.org/W18-4306/>
- [29] Pratapa, A., Small, K. and Dreyer, M. (2023) Background Summarization of Event Timelines. *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, Singapore, 8111-8136.
<https://doi.org/10.18653/v1/2023.emnlp-main.505>
- [30] Zheng, L., Chiang, W.L., Sheng, Y., Zhuang, S., Wu, Z., Zhuang, Y. and Stoica, I. (2024) Judging LLM-as-a-Judge with MT-Bench and Chatbot Arena. *Advances in Neural Information Processing Systems*, **36**, 1-12.
- [31] Zeng, A., Xu, B., Wang, B., Zhang, C., Yin, D. and Wang, Z. (2024) ChatGLM: A Family of Large Language Models from GLM-130B to GLM-4 All Tools.
<https://doi.org/10.48550/arXiv.2406.12793>
- [32] Anand, Y., Nussbaum, Z., Duderstadt, B., Schmidt, B. and Mulyar, A. (2023) GPT4All:

Training an Assistant-Style Chatbot with Large Scale Data Distillation from GPT-3.5-Turbo.

https://s3.amazonaws.com/static.nomic.ai/gpt4all/2023_GPT4All_Technical_Report.pdf