Towards Real-world Application of Timeline Summarization in a Local Media

Hayata Shinjo[1,*], Israel Mendonça[1] and Masayoshi Aritsugi[1]

[1]Kumamoto University, 2-39-1 Kurokami, Chuo-ku, Kumamoto 860-8555 Japan

Abstract

Timeline Summarization (TLS) is a task that organizes past articles chronologically to help readers grasp the long-term progression of events. While Large Language Models (LLMs) can significantly improve the performance of TLS, real-world deployment of what faces the challenge of on-premise operation with limited computational resources to avoid data security concerns. This constraint necessitates a mechanism to appropriately limit the input length for LLMs. In this study, we propose a practical framework that integrates prompt compression and topic modeling. This method overcomes context window limitations by compressing inputs while explicitly supplementing them with extracted topic words, thereby preserving global contextual information. Furthermore, we implemented a post-processing filtering function to ensure summary quality. Experiments using a dataset from a local Japanese newspaper and evaluations by professional reporters suggest that this approach effectively maintains timeline quality and eliminates inappropriate summaries, even under limited resources.

Keywords

Timeline Summarization, Newspaper Archive, Real-world Application, Large Language Models

1. Introduction

With the rapid advancement of internet technology in recent years, the volume of text information available to the public has increased explosively. In particular, news information, such as newspaper articles that are updated continuously, serves as a crucial resource for understanding social trends and the progression of long-term events. However, given this sheer volume of information, it is difficult for readers to grasp the overall picture and the sequence of events. Against this background, research focused on extracting and summarizing essential information from massive textual data to present it effectively to users has garnered significant attention.

Among these approaches, "Timeline Summarization" (TLS) [1, 2], which summarizes the transition of events over an extended period, is particularly useful. TLS is an approach that facilitates the organization and structuring of significant events by extracting relevant past documents and arranging them chronologically. By generating summaries that illustrate the progression of events, it becomes possible to deepen readers' understanding of news topics.

Recently, the advent of Large Language Models (LLMs) has dramatically improved performance across various tasks in the field of Natural Language Processing [3, 4, 5, 6]. In the field of TLS as well, utilizing the advanced contextual understanding capabilities of LLMs is enabling the generation of high-quality timeline summaries [1, 7]. Leveraging these technologies, this study aims to implement timeline summary generation in a real-world setting within a newspaper organization. In such real-world settings, security concerns often make it difficult to utilize massive LLMs via external APIs [8]. Therefore, the use of models capable of operating within the on-premise environments of general companies, where computational resources are limited, is essential. Handling long contexts, however, requires substantial GPU memory [9]. Since TLS addressed in this research targets a massive volume of newspaper articles accumulated over long periods, execution is difficult in a general on-premise

IRCDL 2026: 22nd Conference on Information and Research Science Connecting to Digital and Library Science, February 19-20, 2026, Modena, Italy

*Corresponding author.

shinjo@st.cs.kumamoto-u.ac.jp (H. Shinjo); israel@cs.kumamoto-u.ac.jp (I. Mendonça); aritsugi@cs.kumamoto-u.ac.jp (M. Aritsugi)

0000-0001-6819-4305 (I. Mendonça); 0000-0003-0861-849X (M. Aritsugi)

environment with limited computational resources. Consequently, when considering on-premise operation, a mechanism is required to appropriately limit the input length given to the LLM.

To address this issue, "prompt compression" techniques, which remove redundant words based on token self-information, are considered effective. However, compressing documents based solely on generation probabilities risks omitting information critical to the overall topic of the timeline, potentially causing the resulting timeline to lose thematic consistency. Therefore, this research proposes an input prompt design that explicitly complements the global context of the document corpus—through the integration of a Topic Model—while simultaneously overcoming the physical constraint of input token length via prompt compression. This methodology aims to ensure comprehensive topical coverage in timeline construction, even when operating under severe computational resource limitations.

Furthermore, ensuring the quality of generated summaries is a critical practical challenge. Presenting low-quality summaries or outputs containing erroneous information undermines the reliability of the service and, conversely, reduces user convenience. Therefore, rather than outputting a summary for every input, it is necessary to implement a decision-making process where the system opts "not to present" a summary if the quality does not meet specific standards. To this end, this study introduces filtering based on quality evaluation as a post-processing step for timeline summarization. We verify a mechanism that automatically determines the quality of generated summaries and selects only those that meet the standard for presentation to readers.

To verify the effectiveness of the proposed method, we conducted an evaluation involving newspaper reporters, who are information professionals. Timelines constructed without applying the proposed method were evaluated to serve as a baseline. Based on this expert evaluation, we discuss the effectiveness of timeline construction using the proposed method and the validity of filtering for quality control.

The structure of this paper is as follows: Section 2 introduces the technologies used in this study. Section 3 describes the proposed method for integrating prompt compression with topic modeling and the post-processing filtering mechanism. Section 4 explains the experimental setup and evaluates the proposed method based on the experimental results. Section 5 provides the conclusion and future work of this paper.

2. Related Work

2.1. Timeline Summarization

Timeline Summarization (TLS) is a task that organizes and summarizes the transitions of events over a long period in chronological order [10, 11]. Conventional research has primarily adopted Extractive approaches. The main methods were those that treat date selection and summary sentence selection as separate subtasks [10, 12], or methods that use clustering to detect events and extract representative sentences [11]. Subsequently, Abstractive approaches began to be proposed to generate more fluent summaries [13, 14]. In recent years, the emergence of Large Language Models (LLMs) has advanced TLS [1, 7]. In particular, Hu et al. [1] proposed a method that leverages the powerful annotation capabilities of LLMs to perform incremental clustering and summarization on stream data, achieving high-quality timeline generation.

However, these state-of-the-art methods face challenges when deployed in real-world scenarios due to constraints such as limited computational resources in on-premise environments or security restrictions that prevent the use of external APIs. Furthermore, the practical aspect of deciding whether or not to actually present the generated summary has not been addressed. Therefore, this research aims to develop a method for achieving practical timeline summarization even under these operational constraints.

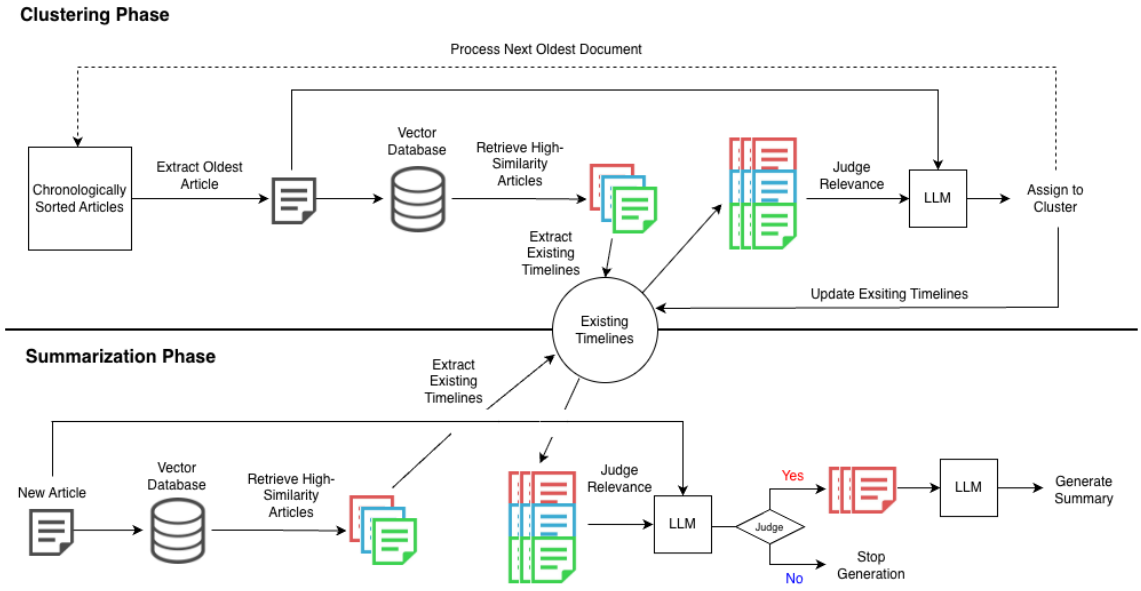

Figure 1: Overview of the timeline summarization framework used in this study.

2.2. Prompt Compression

Prompt compression technology is being actively researched to address the input length limitations of LLMs. Representative approaches include the removal of unnecessary words based on a token's self-information or Perplexity.

Li et al. [15] proposed "Selective Context," which deletes redundant tokens based on self-information, and demonstrated a significant reduction in input length while maintaining task performance. Furthermore, Jiang et al. [16] proposed "LLMLingua," which uses a smaller model to identify important parts of a prompt and performs compression incrementally from coarse to fine granularity. These methods are effective for reducing LLM inference costs and processing long contexts.

However, when these general-purpose compression methods are applied to timeline summarization alone, the fragmentation caused by sentence-level or word-level deletion risks undermining the global context of the entire article and the consistency of the story along the timeline. Therefore, this research investigates a method to reduce the input length of the LLM while capturing the core topic by combining the prompt-compressed text with topic words extracted by LDA (Latent Dirichlet Allocation) [17], a representative topic model.

3. Method

This section describes our proposed procedure for generating timeline summaries. Our framework consists of a baseline Timeline Summarization Framework, and two other components that enhance the produced timeline: The first enhancer tries to extract core information related to articles by compressing it and extracting its topics; The second one does quality control by filtering timelines that are subpar.

3.1. Baseline Timeline Summarization Framework

This section explains the baseline method we created for generating timeline summaries. Our method is based on the method by Hu et al. [1] and consists of two phases: Clustering phase and Summarization phase. The clustering phase tries to group related articles into a single group, whilst the Summarization phase tries summarize the contents of related articles into a timeline. Figure 1 shows the framework used in this study.

3.1.1. Clustering Phase

In the clustering phase, documents are classified into appropriate timelines by determining whether a new document belongs to an existing timeline or should form a new one. This is to guarantee that unrelated documents do not form a wrong timeline.

First, upon receiving a new document, the framework converts it into a vector representation using a pre-trained encoder. Next, this vector is used as a query to calculate the cosine similarity with vectors of past articles, retrieving the top N past articles with the highest similarity. Subsequently, a Large Language Model (LLM) is utilized to determine whether the new document is related to an existing timeline. Specifically, all documents belonging to the timeline of the new document and the retrieved past documents are concatenated and input into the LLM to determine the relevance of the new document to the existing timeline. The prompt used is shown below.

> You are given a list of articles, in chronological order, about some event in a timeline below:
>
> {timeline_articles}
>
> Consider the following new article:
>
> {new_article}
>
> If this new article follows the given list of articles in the same timeline (i.e., the new article is about the same event), then reply with "Yes".
>
> If the new article is not relevant to the event in the given timeline, then reply with "No".
>
> # Answer

In this process, the LLM may determine that a single input document is related to multiple timelines. If the new document is determined to be related to multiple timelines, the framework selects the one containing the document that is temporally closest to the input document. If it is determined not to belong to any timeline, a new timeline is created and the document is assigned to it. This process is executed by inputting all documents in chronological order, from oldest to newest.

3.1.2. Summarization Phase

In this phase, the framework generates a timeline summary for an article based on the articles related to its timeline. An LLM is prompted with the input article (*new_article*) and its related articles (*timeline_articles*) to generate a summary. The prompt utilized by this framework is described below. Needless to say, the framework cannot generate a timeline summarization if there be related clusters were found.

> Chronologically listed under ##Timeline are past articles related to the ##New Article. Summarize the ##Timeline to illustrate the evolution of the event.
>
> ## New Article
> { New_article }
>
> ## Timeline
> { Timeline_articles }
>
> ## Summary

3.2. Integration of Compression and Topic Extraction

As it will be shown in the Experimental section, the Baseline Summarization Framework can produce strong results, but this comes at the cost of long prompts containing full articles, which consume more resources. In this section, we describe how we reduce prompt size. We apply two methods to improve the baseline timeline-summary framework: prompt compression [15] and topic extraction using Latent Dirichlet Allocation (LDA) [17].

3.2.1. Text Compression

This method compresses the prompt by utilizing the self-information of tokens at the sentence level. First, the input document D is segmented into a sequence of m sentences, denoted as $D = \{S_1, S_2, \ldots, S_m\}$. For each sentence $S_j = \{w_1, w_2, \ldots, w_n\}$, a Small Language Model (SLM) is used to calculate the self-information value $I(w_i)$ for each token w_i, conditional on the preceding tokens within the sentence:

$$I(w_i) = -\log P(w_i \mid w_1, \ldots, w_{i-1}) \tag{1}$$

Subsequently, a filtering threshold τ_p is determined corresponding to the reduction ratio $p\%$. Tokens within the sentence whose self-information value falls below this threshold (i.e., $I(w_i) < \tau_p$) are removed, while the remaining tokens are retained. The compressed sentence S'_j is formally defined as the sequence of retained tokens:

$$S'_j = \{w_i \mid I(w_i) \geq \tau_p, 1 \leq i \leq n\} \tag{2}$$

Finally, the compressed document D' is obtained by concatenating all the compressed sentences in their original chronological order:

$$D' = \text{Concat}(S'_1, S'_2, \ldots, S'_m) \tag{3}$$

This compressed document is then input to the LLM to be used in the relevance judgment between documents.

3.2.2. Topic Extraction

This method utilizes article topics extracted via a topic model to determine the relevance between documents. First, LDA is used to detect the latent topics within the dataset. While the number of topics requires manual configuration, we found empirically that 500 yield the best results, so it is the one that was set to this study. Next, when inputting a document into the LLM, the topics constituting that document are inferred using LDA. Then, from the most dominant topic constituting the document, the top k words are extracted to serve as the topic representation of the article. The prompt input to the LLM consists of this extracted list of topic words and the article body.

3.2.3. Prompt Integration

To predict the timeline, the compressed text and extracted topics are integrated into a single prompt for the LLM. The extracted topic words are inserted immediately before the new article and each candidate article in the timeline. Furthermore, the body of each article is replaced with the compressed text. This hybrid structure allows the model to leverage both global topic context and condensed semantic information.

3.3. Quality Filtering Post-processing

After generating the timeline summary for a news article, the framework determine whether it is suitable for presentation; if not, the summary is discarded. The rationale is that a timeline with low relevance to the article is not appropriate. Both the article and its timeline summary are converted

into vector representations using a pre-trained encoder, and their cosine similarity is computed. If the similarity falls below a predefined threshold, the summary is excluded from the final output.

4. Experiments

In this study, we conduct three different experiments: (1) An experiment to evaluate the difference in quality between timelines constructed using the baseline framework introduced in Section 3.1 and timelines generated using compressed summaries introduced in Section 3.2.1. (2) An experiment comparing the baseline framework and timelines generated using compression and topic extraction, as described in Section 3.2. (3) An experiment to evaluate the effectiveness of the filtering mechanism, as described in Section 3.3.

4.1. Experimental Setup

4.1.1. Dataset

The dataset used for the experiments consists of online news articles collected from the digital edition of Kumamotonichinichi Shimbun[1], a local Japanese newspaper company. Each news article includes information such as "URL", "Headline", "Body Text", and "Date". This experiment targeted news articles spanning 13 months, from July 2023 to July 2024. Of the targeted period, news articles from July 2023 to June 2024 were used as data for clustering, and news articles from July 2024 were used as the test data for performing timeline prediction and summary generation for new articles. The clustering data consisted of 21,971 articles, and the test data consisted of 2,250 articles.

4.1.2. Implementation Details

In this experiment, we employed three pre-trained language models fine-tuned for the Japanese language:

- Article Embedding Generation (for past article retrieval and post-processing): Ruri [18], which shows excellent performance in Japanese text embedding benchmarks.
- Relevance Judgment and Timeline Summary Generation: Llama-3-ELYZA-JP-8B [19].
- Self-information calculation: rinna/japanese-gpt2-medium [20].

The hyperparameters used in the experiment are as follows:

- Number of related articles retrieved: $N = 10$.
- Number of topic words extracted: $k = 5$.
- Prompt reduction rate : $p = \{35, 50, 65, 80\}$.

All the experiments were run in a machine with an AMD Ryzen 9 5950X 3.40 GHz processor and 128GB RAM, also equipped with a GeForce RTX 3090 GPU with 24GB of RAM.

4.1.3. Expert Evaluation

To evaluate the practical utility of our approach, we commissioned a newspaper reporter to conduct an evaluation. The evaluation target was the timeline summary generated when the full text of the article was used in the prompt without using prompt compression or the topic model, serving as our baseline. We evaluated 120 randomly sampled data points from the test set. Based on this baseline evaluation, we qualitatively verify the effectiveness of the proposed method. The evaluation scores and their meaning are as follows:

- 4: The flow is easily understandable (Completely suitable for business use).

[1] https://kumanichi.com/

- 3: Generally good, but room for minor adjustments (Likely usable in business as is).
- 2: Selected slightly off-topic articles, or the summary is not what was expected, though not necessarily wrong (Some revisions would be needed before business use).
- 1: Contains clear mistakes, selected irrelevant articles (Problematic for business use as is).

4.2. Impact of Compression on Timeline Quality

In this section, we describe the impact of the prompt reduction rate on the quality of timeline generation. Figure 2 shows an example of the results when prompt compression was applied at various reduction rates to generate a timeline summary. A timeline is predicted for a single new article from the test data.

The new article in Figure 2 concerns a "velodrome" (bicycle racetrack). The predicted timeline at the baseline (0% reduction rate) is composed entirely of articles related to the "velodrome," and the evaluation score is high, at 3. Examining the predicted timelines at each reduction rate, it was confirmed that articles unrelated to the "velodrome" began to be included, showing the intrusion of indirectly related articles.

However, many of these intruded articles were related to "reconstruction," such as "disaster-affected areas" or "restoration." Since the new article also discusses a "velodrome rebuilt after a disaster," it is believed that the inclusion of articles related to disasters like "earthquakes" and "reconstruction" is the underlying cause. This suggests that simply reducing words based on self-information volume, while preserving local consistency of the context, may impair the global consistency of the topic running through the entire article, thereby increasing the risk of mixing in irrelevant noise articles.

Thus, because text compression involves reducing the amount of information, it may potentially lead to a degradation in the quality of the constructed timeline. To prevent the quality of the timeline from decreasing due to such text compression, this study attempts to incorporate the topic words extracted from the document as a factor for determining the relevance between the new article and the existing timeline.

In many of the timelines constructed in this study, results did not change significantly across reduction rates ranging from 35% to 65%. Thus, for the integrated method described in the section 4.3, we adopted the 65% reduction rate, which provides the maximum compression within this stable range.

4.3. Impact of Text Compression and Topic Words Addition

In this section we evaluate how using input compression and adding topic words impacts the timeline summarization process. We evaluate two cases: One of an article concerning "fare increases for public transportation", and another article concerning "food poisoning".

4.3.1. Public Transportation fare increase article

First, we will examine Figure 3. For the new article concerning "fare increases for public transportation," the baseline predicted timeline consists of articles related to this topic such as railways, Shinkansen, and city trams, achieving a high *expert evaluation score of 4*. When only text compression was applied, articles concerning railway facilities other than those related to fare increases were mixed into the predicted timeline, confirming a degradation in the quality of the timeline. In contrast, when the article's topic words were added, the noisy articles were eliminated, and the timeline was once again composed solely of articles related to fare increases. This demonstrates that the quality of the timeline, which had degraded due to compression, was improved by the addition of topic words.

To specifically confirm the cause of this phenomenon, we compared the extracted topic words. Topics such as [‘average’, ‘survey’, ‘rate’, ‘factor’, ‘egg’] were extracted from the new article, and similarly, the same topics—[‘average’, ‘survey’, ‘rate’, ‘factor’, ‘egg’]—were extracted from the related articles retrieved using the integrated method. On the other hand, for most of the noisy articles unrelated to the fare increase, different topics were extracted compared to the new article. A clear difference was confirmed in the coherence/incoherence of the topic words, leading to the conclusion that topic words

New article
2024-07-20: 2016年の熊本地震で被災し、熊本市が再建した熊本競輪場（中央区水前寺）で20日、8年ぶりにレースが開催された。選手は新設の1周400メートルのバンク（競走路）を駆け抜け、観客の声援が場内に響いた。...
2024-07-20: Racing resumed for the first time in eight years on the 20th at the Kumamoto Velodrome (Suizenji, Chuo-Ku), which was damaged in the 2016 Kumamoto Earthquake and rebuilt by Kumamoto City. Riders sped around the newly constructed 400-meter track (bank), and spectators' cheers echoed throughout the venue. ...

Predicted timeline
Baseline: (Evaluation score: 3)
2024-03-30: 2016年の熊本地震で被災し、再建が進む熊本市の熊本競輪場（中央区）のレースが今夏、再開される。一部が近隣住民の避難所としても活用された施設の復旧に、地元住民は「災害時に心強い」「地域のにぎわいにつながる」と歓迎する一方、 ...
2024-04-02: 熊本市は2日、7月に再開する熊本競輪場（中央区水前寺）で、GIレースの「全日本選抜競輪」を2026年2月20日から4日間開催すると発表した。25年度のグランドオープンに合わせて誘致した。...
2024-06-08: 2016年の熊本地震で被災した熊本市中央区の熊本競輪場の競走路（バンク）やスタンドが完成し、地域住民らを招いた内覧会が8日、現地であった。7月20日から8年ぶりに競技を再開する見通しで、 ...
2024-03-30: Racing is set to resume this summer at the Kumamoto Velodrome (Chuo-Ku, Kumamoto City), which was damaged in the 2016 Kumamoto Earthquake and is currently undergoing reconstruction. Local residents welcome the facility's restoration, parts of which were utilized as an evacuation center for nearby residents, stating it is "reassuring during a disaster" and will "lead to regional revitalization," while ...
2024-04-02: Kumamoto City announced on the 2nd that the GI race, the "All-Japan Selected Keirin," will be held for four days starting February 20, 2026, at the Kumamoto Velodrome (Suizenji, Chuo-Ku), which will reopen in July. The event was attracted to coincide with the grand opening in the 2025 fiscal year. ...
2024-06-08: The track (bank) and stands of the Kumamoto Velodrome in Kumamoto City's Chuo-Ku, which was damaged in the 2016 Kumamoto Earthquake, have been completed. An open house inviting local residents was held on-site on the 8th. Competition is expected to resume on July 20th after an eight-year hiatus, and ...

Compressed text (p=35)
2024-03-20: 球磨川の氾濫が甚大な被害をもたらした2020年7月の熊本豪雨。被災地では、住民の住まい再建や治水対策と並んで、街のにぎわい復活や観光の立て直しが大きな課題だ。...
2024-03-30: 2016年の熊本地震で被災し、再建が進む熊本市の熊本競輪場（中央区）のレースが今夏、再開される。一部が近隣住民の避難所としても活用された施設の復旧に、地元住民は「災害時に心強い」「地域のにぎわいにつながる」と歓迎する一方、...
2024-04-02: 熊本市は2日、7月に再開する熊本競輪場（中央区水前寺）で、GIレースの「全日本選抜競輪」を2026年2月20日から4日間開催すると発表した。25年度のグランドオープンに合わせて誘致した。...
2024-05-22: 人吉市の別府弘江さん（85）は今月、約50年間住んだ下青井町の自宅を立ち退いた。2020年の熊本豪雨で自宅は2階まで浸水したが、「ついのすみか」と決めていたため、親族の支援を受けて修復した。...
2024-06-08: 2016年の熊本地震で被災した熊本市中央区の熊本競輪場の競走路（バンク）やスタンドが完成し、地域住民らを招いた内覧会が8日、現地であった。...
2024-03-20: The Kumamoto heavy rain in July 2020 caused severe damage due to the overflowing of the Kuma River. In the affected areas, alongside rebuilding residents' homes and implementing flood control measures, the revival of community vitality and the reconstruction of tourism are major challenges. ...
2024-03-30: Racing is set to resume this summer at the Kumamoto Velodrome (Chuo-Ku, Kumamoto City), which was damaged in the 2016 Kumamoto Earthquake and is currently undergoing reconstruction. Local residents welcome the facility's restoration, parts of which were utilized as an evacuation center for nearby residents, stating it is "reassuring during a disaster" and will "lead to regional revitalization," while...
2024-04-02: Kumamoto City announced on the 2nd that the GI race, the "All-Japan Selected Keirin," will be held for four days starting February 20, 2026, at the Kumamoto Velodrome (Suizenji, Chuo-Ku), which will reopen in July. The event was attracted to coincide with the grand opening in the 2025 fiscal year. ...
2024-05-22: Ms. Hiroe Beppu (85) of Hitoyoshi City moved out of her home in Shimoaoi Town this month, where she had lived for about 50 years. Her home was submerged up to the second floor in the 2020 Kumamoto heavy rain, but because she considered it her "final resting place," she had it repaired with the support of relatives. ...
2024-06-08: The track (bank) and stands of the Kumamoto Velodrome in Kumamoto City's Chuo-Ku, which was damaged in the 2016 Kumamoto Earthquake, have been completed. An open house inviting local residents was held on-site on the 8th. ...

Compressed text (p=65)
2023-09-01: 阿蘇市は1日、屋根の老朽化で利用を休止している複合スポーツ施設「アゼリア21」について、黒字化には料金の大幅な値上げが必要になるとの見通しを示した。同日の市議会全員協議会で報告した。...
2023-09-20: 熊本市は、来年6月に再開予定の熊本競輪場の入場料を無料にする方針を固めた。全国的に競輪場の無料化が進んでいる現状を踏まえ、来場者増につなげるのが狙い。12月定例市議会に市自転車競走条例の改正案を提出する。 ...
2023-11-06: 熊本市は6日、国特別史跡・熊本城の天守閣前広場や二の丸芝生広場のイベント利用について、文化財保護の観点から熊本市や県の主催事業のみとしている現行制度を見直し、...
2024-03-16: 熊本市役所本庁舎の建て替えに関する聞き取り調査で、建て替え後に跡地を利用する場合、建物の高さ制限を緩和するよう求める民間事業者の意見が多く寄せられた。調査は市が昨年9月から今年1月末にかけて実施。 ...
2024-04-05: 熊本城がイベント会場になろうとしている。 先に熊本市が市文化財保護委員会などで明らかにした「熊本城の管理に関する取扱要領改正の目的」を見て、少なからずショックを受けた。...
2024-06-08: 2016年の熊本地震で被災した熊本市中央区の熊本競輪場の競走路（バンク）やスタンドが完成し、地域住民らを招いた内覧会が8日、現地であった。 ...
2024-06-25: 熊本県は25日の県議会一般質問で、所有する旧宿泊施設「火の国ハイツ」（熊本市東区）に関し、民間への売却を前提として利活用を進める考えを示した。 ...
2023-09-01: Aso City announced on the 1st that the multi-purpose sports facility "Azalea 21," which has suspended use due to roof aging, would require significant price increases to become profitable. This was reported at an all-member city council meeting that day. ...
2023-09-20: Kumamoto City has decided on a policy to make admission free for the Kumamoto Velodrome, which is scheduled to reopen next June. The aim is to increase attendance, reflecting the nationwide trend toward free admission at velodromes. A bill to amend the City Bicycle Racing Ordinance will be submitted to the December regular city council session. ...
2023-11-06: Kumamoto City announced an outline plan on the 6th to revise the current system, which limits the use of the plaza in front of Kumamoto Castle's main keep and the Ninomaru lawn area for events to only those organized by Kumamoto City or Kumamoto Prefecture, due to cultural property protection concerns, and ...
2024-03-16: An opinion survey regarding the rebuilding of the Kumamoto City Hall main building received numerous requests from private businesses to ease building height restrictions if the site were to be used after reconstruction. The survey was conducted by the city from September of last year to the end of January this year. ...
2024-04-05: Kumamoto Castle is set to become an event venue. We were considerably shocked after viewing the "Purpose of Revising the Guidelines for the Management of Kumamoto Castle," which Kumamoto City recently revealed to the City Cultural Property Protection Committee and others. ...
2024-06-08: The track (bank) and stands of the Kumamoto Velodrome in Kumamoto City's Chuo-Ku, which was damaged in the 2016 Kumamoto Earthquake, have been completed. An open house inviting local residents was held on-site on the 8th. ...
2024-06-25: Kumamoto Prefecture indicated at the general interpellation session of the Prefectural Assembly on the 25th that it intends to promote the utilization of its former accommodation facility, "Hinokuni Heights" (Higashi-Ku, Kumamoto City), based on the premise of selling it to the private sector. ...

Figure 2: Comparison of timeline construction results based on different prompt reduction rates (Case study: Velodrome). Words in red indicate key topics (manually selected) of the new article essential for timeline construction.

effectively compensate for the semantic information lost due to text compression, thus contributing to the filtering of noisy articles.

4.3.2. Food Poisoning article

Next, we describe the example in Figure 4. For the new article concerning food poisoning, the baseline timeline was consistently composed of articles related to the same food poisoning topic, and the *expert evaluation score* was high at 3.

Figure 3: Example of improved timeline construction using the proposed integrated method (Case study: Public transportation fare increase). Words in red indicate key topics (manually selected) of the new article essential for timeline construction, while words in blue represent other topics.

However, unlike the example in Figure 3, the timeline constructed after only applying text compression also consisted solely of food poisoning articles. Conversely, when topic words were added to the compressed text, articles concerning health damage from supplements, in addition to food poisoning, were mixed into the timeline. This result suggests that the LLM judged seemingly unrelated articles as "relevant" due to the presence of topic words.

The likely cause of this issue is the extraction of "noisy topic words" that are unrelated to the article's core essence. By checking the topic words, we found that articles related to food [‘cooking’, ‘birth’], while articles related to "supplement health damage" contained [‘period’, ‘weekday’, ‘anonymity’, ‘Shiga’, ‘crisis’]. Although the composition of the topic words in both timelines differs significantly, which would seemingly suggest low relevance, focusing on the topic words in the food poisoning article reveals terms such as "week", "nurse", and "birth", which are loosely related to food poisoning itself. Specifically, the term "nurse" strongly evokes contexts like "health", "patient". It is hypothesized that this led to the intrusion of articles related to supplement health damage.

The results above indicate that while the proposed method is effective in compensating for contextual information lost due to text compression using topic words, thereby suppressing the degradation of timeline quality, its performance is significantly influenced by the quality of the extracted topic words. Therefore, simply applying a topic model is insufficient, and improving the accuracy of topic extraction—such as revising the weighting of words extracted as topics or appropriately filtering out

New article
2024-07-25: 熊本市は25日、介護施設などで弁当を食べた35人が食中毒になったとして、弁当を製造した「彩りごはんJian」（同市中央区）を同日から2日間の営業停止処分にした。...
2024-07-25: Kumamoto City announced on the 25th that it had ordered the bento (lunch box) manufacturer "Irodori Gohan Jian" (Chuo-Ku, Kumamoto City) to suspend operations for two days, effective from that day, after 35 people who consumed its bentos at nursing care facilities and other locations suffered from food poisoning. ...

Predicted timeline
Baseline: (Evaluation score: 3)
2023-09-22: 熊本市は22日、サケの刺し身を食べた同市東区の50代男性が寄生虫アニサキスによる食中毒になったと発表した。...
2024-03-15: 熊本県は15日、自宅で調理したイワシの刺し身を食べた県内の70代女性が寄生虫アニサキスによる食中毒になったと発表した。...
2024-06-29: 熊本市は29日、市内の障害者支援施設で食中毒が発生したとして、調理業務を受託する「東臣」（同市東区）を30日まで2日間の営業停止処分にした。...
2023-09-22: Kumamoto City announced on the 22nd that a man in his 50s from Higashi-Ku of the city suffered from food poisoning caused by the parasite Anisakis after eating salmon sashimi. ...
2024-03-15: Kumamoto Prefecture announced on the 15th that a woman in her 70s in the prefecture suffered from food poisoning caused by the parasite Anisakis after eating sardine sashimi that she prepared at home. ...
2024-06-29: Kumamoto City announced on the 29th that it had ordered "Toshin" (Higashi-Ku, Kumamoto City), a company contracted to provide cooking services at a facility for people with disabilities in the city, to suspend operations for two days until the 30th, following a food poisoning incident at the facility. ...

Compressed text (p=65)
2024-05-02: 熊本市は2日、ヒラメの刺し身などを食べた市内の40〜70代の男女8人が寄生虫クドア・セプテンプンクタータ（クドア）による食中毒になったと発表した。...
2024-06-07: 熊本市は7日、提供した鶏料理で食中毒が発生したとして、中央区下通の飲食店「焼鳥集家［しゅうや］」を8日まで2日間の営業停止処分にした。...
2024-05-02: Kumamoto City announced on the 2nd that eight men and women in their 40s to 70s in the city suffered from food poisoning caused by the parasite Kudoa septempunctata (Kudoa) after eating flounder sashimi and other items. ...
2024-06-07: Kumamoto City announced on the 7th that it had ordered the restaurant "Yakitori Shuya" in Shimotori, Chuo-Ku, to suspend operations for two days until the 8th, following a food poisoning incident caused by the chicken dishes it served. ...

Compressed text (p=65) + Topic words:
2024-02-14: 熊本市は14日、提供した焼き鳥が食中毒の原因になったとして、中央区下通の飲食店「鳩恋［とりこ］」を16日まで3日間の営業停止処分にした。...
2024-03-15: 熊本県は15日、自宅で調理したイワシの刺し身を食べた県内の70代女性が寄生虫アニサキスによる食中毒になったと発表した。...
2024-04-03: 熊本市は3日、アジのにぎりずしなどを食べた市内の50代女性が寄生虫アニサキスによる食中毒になったと発表した。...
2024-04-03: 熊本市は3日、小林製薬の「紅こうじ」を使った機能性表示食品のサプリメントを摂取し、健康被害が出た疑いのある患者1人を確認したと発表した。...
2024-04-04: 熊本県は4日、小林製薬の「紅こうじ」を使った機能性表示食品のサプリメントを摂取し、健康被害が出た疑いのある患者4人を新たに確認したと発表した。...
2024-04-05: 「国民目線の消費者行政を確立する」。消費者庁が2009年に発足した際、政府はこんなキャッチフレーズを掲げてアピールした。 ...
2024-04-08: 熊本市は8日、小林製薬の紅こうじを使った機能性表示食品を摂取し、健康被害が出た疑いのある市内在住の患者1人を新たに確認したと発表した。...
2024-04-10: 熊本県は10日、マダニが媒介する日本紅斑熱の患者1人を確認したと発表した。...
2024-04-10: 熊本県と熊本市は10日、小林製薬の紅こうじを使った機能性表示食品のサプリメントを摂取し、健康被害が出た疑いのある患者をそれぞれ1人ずつ確認したと発表した。...
2024-04-12: 熊本県は12日、小林製薬の紅こうじを使った機能性表示食品のサプリメントを摂取し、健康被害が出た疑いのある患者4人を新たに確認したと発表した。...
2024-06-29: 熊本市は29日、市内の障害者支援施設で食中毒が発生したとして、調理業務を受託する「東臣」（同市東区）を30日まで2日間の営業停止処分にした。...
2024-02-14: Kumamoto City announced on the 14th that it had ordered the restaurant "Toriko" in Shimotori, Chuo-Ku, to suspend operations for three days until the 16th, stating that the yakitori (grilled chicken skewers) it served was the cause of food poisoning. ...
2024-03-15: Kumamoto Prefecture announced on the 15th that a woman in her 70s in the prefecture suffered from food poisoning caused by the parasite Anisakis after eating sardine sashimi that she prepared at home. ...
2024-04-03: Kumamoto City announced on the 3rd that a woman in her 50s in the city suffered from food poisoning caused by the parasite Anisakis after eating horse mackerel nigiri-zushi (sushi) and other items. ...
2024-04-03: Kumamoto City announced that it had confirmed one patient suspected of experiencing health damage after taking supplements of Food with Function Claims that used "Beni-koji" (red yeast rice) from Kobayashi Pharmaceutical. ...
2024-04-04: Kumamoto Prefecture announced on the 4th that it had newly confirmed four patients suspected of experiencing health damage after taking supplements of Food with Function Claims that used "Beni-koji" from Kobayashi Pharmaceutical. ...
2024-04-05: "Establish consumer administration from the perspective of the citizenry." The government promoted this slogan when the Consumer Affairs Agency was launched in 2009. ...
2024-04-08: Kumamoto City announced on the 8th that it had newly confirmed one patient residing in the city suspected of experiencing health damage after taking Food with Function Claims products that used Beni-koji from Kobayashi Pharmaceutical. ...
2024-04-10: Kumamoto Prefecture announced on the 10th that it had confirmed one patient with Japanese spotted fever, which is transmitted by ticks. ...
2024-04-10: Kumamoto Prefecture and Kumamoto City announced on the 10th that they had each confirmed one patient suspected of experiencing health damage after taking supplements of Food with Function Claims that used Beni-koji from Kobayashi Pharmaceutical. ...
2024-04-12: Kumamoto Prefecture announced on the 12th that it had newly confirmed four patients suspected of experiencing health damage after taking supplements of Food with Function Claims that used Beni-koji from Kobayashi Pharmaceutical. ...
2024-06-29: Kumamoto City announced on the 29th that it had ordered "Toshin" (Higashi-Ku, Kumamoto City), a company contracted to provide cooking services at a facility for people with disabilities in the city, to suspend operations for two days until the 30th, following a food poisoning incident at the facility. ...

Figure 4: Example of timeline construction where noise was introduced due to the influence of topic words in the proposed method (Case study: Food poisoning). Words in red indicate key topics (manually selected) of the new article essential for timeline construction, while words in blue represent other topics.

noisy vocabulary—is an important future challenge.

4.4. Evaluation of Quality Filtering

In this section, we describe the results of the verification conducted to see whether the low-quality timeline summaries could be effectively removed from the generated set using the filtering method proposed in Section 3.3.

To set the threshold, this study incrementally varied the overall discard ratio of the data to 25%, 50%, and 75%, and analyzed the distribution of the remaining timeline summaries' evaluation scores.

Table 1
Number of remaining data points and high-quality rate by filtering threshold.

Discard Ratio	Total	Score 4	Score 3	Score 2	Score 1	High Quality Rate (Score 3–4)
0% (Baseline)	120	9	70	29	12	65.8%
25%	90	9	60	17	4	76.7%
50%	60	6	44	10	0	83.3%
75%	30	5	22	3	0	90.0%

New article
2024-07-23: 熊本市は23日の市議会特別委員会で、市電の新たな延伸を改めて検討する方針を示した。現在、東区で計画している健軍町電停から市民病院までのルート（仮称・東町線）に加え、東部、南熊本駅、田崎の3方面の計5ルートを軸に延伸の是非を検討する。. . .
2024-07-23: Kumamoto City announced its policy to re-examine the possibility of a new extension of the streetcar (tram) system during a special committee meeting of the City Council on the 23rd. In addition to the currently planned route in Higashi-Ku, which runs from Kengunmachi streetcar stop to the City Hospital (tentative name: Higashimachi Line), the city will also consider the feasibility of extensions centered on a total of five routes across three other directions: the eastern area, Minami-Kumamoto Station, and Tasaki. ...

Generated summarization
Baseline: (Evaluation score: 4)
2024年4月以降、熊本市電の延伸や公共交通の課題についての議論や提案が行われています。具体的には、市電の快速化やバスとの連携強化、延伸ルートの検討、外国語対応の必要性などが話し合われています。
Since April 2024, discussions and proposals have been made regarding the extension of the Kumamoto streetcar and issues related to public transportation. Specifically, topics such as increasing the speed of the streetcar, strengthening cooperation with buses, examining extension routes, and the need for foreign language support have been discussed.

Figure 5: Example of a timeline summary retained by quality filtering.

New article
2024-07-17: 国土交通省は、港を核とした地域づくりに生かす「みなとオアシス」に、熊本市西区の熊本港一帯を認定した。. . .
2024-07-17: The Ministry of Land, Infrastructure, Transport and Tourism certified the Kumamoto Port area in Nishi-Ku, Kumamoto City, as a "Minato Oasis" to be utilized for community development centered on the port. ...

Generated summarization
Baseline: (Evaluation score: 2)
2024年5月31日、熊本港の2023年のコンテナ取扱量が前年比17.5%増の1万698個と4年ぶりに増加したことが報告された。同年6月25日、九州運輸局が発表した九州の5月の訪日客数は前年同月比60.5%増の35万9822人で、5月としては過去最多を記録した。
On May 31, 2024, it was reported that Kumamoto Port's container throughput for 2023 increased for the first time in four years, reaching 10,698 units, marking a 17.5% increase compared to the previous year. On June 25 of the same year, the Kyushu District Transport Bureau announced that the number of foreign visitors to Kyushu in May was 359,822, which was a 60.5% increase year-on-year and the highest figure ever recorded for the month of May.

Figure 6: Example of a timeline summary discarded by quality filtering.

Table 1 shows the number of test data points for each evaluation category and the number of test data points remaining at each threshold set for the cosine similarity between the new article and its timeline summary. It can be observed that when the threshold was set to discard up to 50% of the overall test data, all summaries with the lowest score were eliminated. Furthermore, the number of low-quality summaries (scores of 1 and 2) was reduced to approximately one-quarter of the non-filtered baseline. Conversely, while the number of high-quality summaries (scores of 3 and 4) decreased, more than half were retained.

Figure 5 shows an example of a timeline summary that was retained, and Figure 6 shows an example of a timeline summary that was discarded at a 50% discard ratio. The retained timeline summary has an evaluation score of 4, and the summary clearly explains what has been discussed previously regarding the new article about the extension of the city tram line. In contrast, the discarded timeline summary has an evaluation score of 2. For the new article concerning regional development around the port, the summary discusses container handling volume and the number of visitors to Japan, which is largely irrelevant to the new article. These results suggest that even the simple cosine similarity filtering approach employed here can effectively remove low-scoring timeline summaries while retaining many high-scoring summaries.

Finally, by shifting the threshold to discard up to 75% of the test data, it is confirmed that the number

of low-scoring summaries is reduced to approximately one-tenth of the non-filtered baseline. However, the number of summaries with high scores (3 and 4) is also significantly reduced. In real-world operation, it is considered important to set an appropriate threshold tailored to the specific objectives of the service.

5. Conclusion

In this study, we proposed a framework to automatically generate timelines for news article that was able to reach high *expert evaluation scores*. We also introduced means to address the specific constraints of computational resources and security inherent in on-premise environments, with the aim of realizing real-world operation of timeline summarization at a local newspaper company. Specifically, to overcome the input length limitation of Large Language Models (LLMs), we introduced an approach that combines prompt compression based on self-information with Topic Extraction using Latent Dirichlet Allocation (LDA).

Experiments conducted using a local newspaper company's dataset confirmed that by explicitly appending topic words to the compressed text generated by the proposed method, it is possible to compensate for lost contextual information and prevent a decline in timeline quality. Furthermore, the filtering method introduced for quality control with a view toward practical operation suggested that even a simple cosine similarity-based approach can effectively exclude low-quality summaries and secure a certain level of reliability.

On the other hand, it became clear that the quality of the extracted topic words significantly affects the accuracy of timeline construction. Improving topic extraction accuracy is a crucial challenge to enhance the effectiveness of this method for the future. Currently, we fixed the top 5 words constituting the topic, but words with little relevance to the article's context were also being extracted. Therefore, we are considering introducing weighting based on the relevance of the words to the article. This will prevent the extraction of words that do not contribute to the article's core topic. In addition, the creation of ground truth data, including annotations by experts, is essential for conducting a more rigorous evaluation.

Declaration on Generative AI

We clearly wrote all the places where and the ways in which we used generative AI software tools in this study in the body of this paper; Otherwise we did not use generative AI software tools except for text translation (Gemini).

Acknowledgments

We would like to express our sincere gratitude to Kumamotonichinichi Shimbun for providing the data in this study and for their expert evaluation of the experiments.

References

[1] Q. Hu, G. Moon, H. T. Ng, From moments to milestones: Incremental timeline summarization leveraging large language models, in: L.-W. Ku, A. Martins, V. Srikumar (Eds.), Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Association for Computational Linguistics, Bangkok, Thailand, 2024, pp. 7232–7246. URL: https://aclanthology.org/2024.acl-long.390/. doi:10.18653/v1/2024.acl-long.390.

[2] H. Rajaby Faghihi, B. Alhafni, K. Zhang, S. Ran, J. Tetreault, A. Jaimes, CrisisLTLSum: A benchmark for local crisis event timeline extraction and summarization, in: Y. Goldberg, Z. Kozareva, Y. Zhang (Eds.), Findings of the Association for Computational Linguistics: EMNLP 2022, Association for

Computational Linguistics, Abu Dhabi, United Arab Emirates, 2022, pp. 5455–5477. URL: https://aclanthology.org/2022.findings-emnlp.400/. doi:10.18653/v1/2022.findings-emnlp.400.

[3] S. Wang, X. Sun, X. Li, R. Ouyang, F. Wu, T. Zhang, J. Li, G. Wang, C. Guo, GPT-NER: Named entity recognition via large language models, in: L. Chiruzzo, A. Ritter, L. Wang (Eds.), Findings of the Association for Computational Linguistics: NAACL 2025, Association for Computational Linguistics, Albuquerque, New Mexico, 2025, pp. 4257–4275. URL: https://aclanthology.org/2025.findings-naacl.239/. doi:10.18653/v1/2025.findings-naacl.239.

[4] Y. Zhu, L. Wu, Q. Guo, L. Hong, J. Li, Collaborative large language model for recommender systems, in: Proceedings of the ACM Web Conference 2024, WWW '24, Association for Computing Machinery, New York, NY, USA, 2024, pp. 3162–3172. URL: https://doi.org/10.1145/3589334.3645347. doi:10.1145/3589334.3645347.

[5] M. Lewis, Y. Liu, N. Goyal, M. Ghazvininejad, A. Mohamed, O. Levy, V. Stoyanov, L. Zettlemoyer, BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension, in: D. Jurafsky, J. Chai, N. Schluter, J. Tetreault (Eds.), Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, Online, 2020, pp. 7871–7880. URL: https://aclanthology.org/2020.acl-main.703/. doi:10.18653/v1/2020.acl-main.703.

[6] X. Han, W. Zhao, N. Ding, Z. Liu, M. Sun, PTR: Prompt tuning with rules for text classification, AI Open 3 (2022) 182–192. URL: https://www.sciencedirect.com/science/article/pii/S2666651022000183. doi:https://doi.org/10.1016/j.aiopen.2022.11.003.

[7] D. Sojitra, R. Jain, S. Saha, A. Jatowt, M. Gupta, Timeline summarization in the era of LLMs, in: Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '24, Association for Computing Machinery, New York, NY, USA, 2024, pp. 2657–2661. URL: https://doi.org/10.1145/3626772.3657899. doi:10.1145/3626772.3657899.

[8] F. Dennstädt, J. Hastings, P. M. Putora, M. Schmerder, N. Cihoric, Implementing large language models in healthcare while balancing control, collaboration, costs and security, npj Digital Medicine 8 (2025) 143. URL: https://doi.org/10.1038/s41746-025-01476-7. doi:10.1038/s41746-025-01476-7.

[9] G. Xiao, J. Tang, J. Zuo, J. Guo, S. Yang, H. Tang, Y. Fu, S. Han, Duoattention: Efficient long-context LLM inference with retrieval and streaming heads, in: Y. Yue, A. Garg, N. Peng, F. Sha, R. Yu (Eds.), International Conference on Representation Learning, volume 2025, 2025, pp. 37228–37253. URL: https://proceedings.iclr.cc/paper_files/paper/2025/file/5c1ddd2e59df46fd2aa85c833b1b36ed-Paper-Conference.pdf.

[10] S. Martschat, K. Markert, A temporally sensitive submodularity framework for timeline summarization, in: A. Korhonen, I. Titov (Eds.), Proceedings of the 22nd Conference on Computational Natural Language Learning, Association for Computational Linguistics, Brussels, Belgium, 2018, pp. 230–240. URL: https://aclanthology.org/K18-1023/. doi:10.18653/v1/K18-1023.

[11] D. Gholipour Ghalandari, G. Ifrim, Examining the state-of-the-art in news timeline summarization, in: D. Jurafsky, J. Chai, N. Schluter, J. Tetreault (Eds.), Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, Online, 2020, pp. 1322–1334. URL: https://aclanthology.org/2020.acl-main.122/. doi:10.18653/v1/2020.acl-main.122.

[12] G. Binh Tran, M. Alrifai, D. Quoc Nguyen, Predicting relevant news events for timeline summaries, in: Proceedings of the 22nd International Conference on World Wide Web, WWW '13 Companion, Association for Computing Machinery, New York, NY, USA, 2013, pp. 91–92. URL: https://doi.org/10.1145/2487788.2487829. doi:10.1145/2487788.2487829.

[13] J. Steen, K. Markert, Abstractive timeline summarization, in: L. Wang, J. C. K. Cheung, G. Carenini, F. Liu (Eds.), Proceedings of the 2nd Workshop on New Frontiers in Summarization, Association for Computational Linguistics, Hong Kong, China, 2019, pp. 21–31. URL: https://aclanthology.org/D19-5403/. doi:10.18653/v1/D19-5403.

[14] X. Chen, Z. Chan, S. Gao, M.-H. Yu, D. Zhao, R. Yan, Learning towards abstractive timeline summarization, in: Proceedings of the Twenty-Eighth International Joint Conference on Artificial

Intelligence, IJCAI-19, International Joint Conferences on Artificial Intelligence Organization, 2019, pp. 4939–4945. URL: https://doi.org/10.24963/ijcai.2019/686. doi:10.24963/ijcai.2019/686.

[15] Y. Li, B. Dong, F. Guerin, C. Lin, Compressing context to enhance inference efficiency of large language models, in: H. Bouamor, J. Pino, K. Bali (Eds.), Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, Singapore, 2023, pp. 6342–6353. URL: https://aclanthology.org/2023.emnlp-main.391/. doi:10.18653/v1/2023.emnlp-main.391.

[16] H. Jiang, Q. Wu, C.-Y. Lin, Y. Yang, L. Qiu, LLMLingua: Compressing prompts for accelerated inference of large language models, in: H. Bouamor, J. Pino, K. Bali (Eds.), Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, Singapore, 2023, pp. 13358–13376. URL: https://aclanthology.org/2023.emnlp-main.825/. doi:10.18653/v1/2023.emnlp-main.825.

[17] D. M. Blei, A. Y. Ng, M. I. Jordan, Latent dirichlet allocation, J. Mach. Learn. Res. 3 (2003) 993–1022.

[18] H. Tsukagoshi, R. Sasano, Ruri: Japanese general text embeddings, arXiv preprint arXiv:2409.07737 (2024).

[19] M. Hirakawa, S. Horie, T. Nakamura, D. Oba, S. Passaglia, A. Sasaki, elyza/llama-3-elyza-jp-8b, 2024. https://huggingface.co/elyza/Llama-3-ELYZA-JP-8B.

[20] K. Sawada, T. Zhao, M. Shing, K. Mitsui, A. Kaga, Y. Hono, T. Wakatsuki, K. Mitsuda, Release of pre-trained models for the Japanese language, in: N. Calzolari, M.-Y. Kan, V. Hoste, A. Lenci, S. Sakti, N. Xue (Eds.), Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024), ELRA and ICCL, Torino, Italia, 2024, pp. 13898–13905. URL: https://aclanthology.org/2024.lrec-main.1213/.