A Multi-Baseline Framework for Ranking Global Event Significance Using Google Trends and Large Language Models

Zenan Chen
School of Information, Journalism and Communication, University of Sheffield, UK
zchen249@sheffield.ac.uk

Abstract

Determining global event significance lacks standardized metrics for quantifying worldwide impact. While Google Trends has demonstrated utility in domain-specific studies, its application to global event ranking remains limited. This paper presents a framework combining Google Trends data with large language models for automated global event ranking. This study leverages Command R+ and Llama 3.3-70B-Instruct to generate contextually relevant event keywords and establishes significance through comparative search volume analysis against baseline keywords, incorporating temporal weighting mechanisms to address chronological biases. The proposed methodology identified globally significant events across technology, health, sports, and natural disasters from a dataset of 1,094 events (2020-2024) extracted from Wikipedia.

1 Introduction

Global events, from pandemics and natural disasters to geopolitical crises and economic disruptions, shape international relations and affect millions of lives across national boundaries. Understanding their relative significance is essential for effective resource allocation, risk assessment, media coverage and informed policymaking. However, quantification of global event significance presents fundamental challenges due to the absence of standardized international metrics and inconsistently formatted information across different sources. This capability is essential for news organizations prioritizing coverage, policy makers identifying emerging issues, and researchers analyzing global trends. Traditional approaches to event ranking often rely on expert judgment, media coverage analysis, or domain-specific indicators (Wang et al., 2008; Kong et al., 2012), each introducing scalability limitations and subjective biases.

This study addresses these constraints by introducing an automated analysis framework that leverages search behavior data to measure event significance through demonstrated public attention patterns. Building upon the established relationship between search volume and public interest, this work utilizes Google Trends (GT) data (Google Trends) to create standardized significance metrics through baseline comparison analysis. While previous applications of search data have focused on specific domains such as financial or health contexts (Husnayain et al., 2020; Knipe et al., 2020), the framework developed here extends across diverse event categories and geographic contexts. Critical temporal biases are addressed through weighted analysis, with cross-event comparability ensured via consistent baseline reference points.

The paper's primary contributions include: (1) a multi-baseline comparison framework ensuring consistency; (2) an automated keyword generation system eliminating manual selection biases; and (3) an cross-domain application to diverse event categories. To our knowledge, this represents the first systematic integration of GT and LLMs for global event significance ranking.

After reviewing existing approaches to event significance assessment and search data applications in event analysis research, the paper presents the ranking methodology in detail. The approach encompasses automated keyword generation through large language models (LLMs) and multi-baseline aggregation strategies.

2 Related Work

In this section, current methods used to measure and rank significance of global events will be described. Then, the usage of GT in measuring the significance of events will be discussed.

Proceedings of the Student Research Workshop associated with RANLP-2025, pages 1–9, Varna, Bulgaria, Sep 8-10, 2025.
https://doi.org/10.26615/issn.2603-2821.2025_001

2.1 Event assessment and ranking methods

Several approaches have been developed for global event assessment and ranking. AI-GlobalEvent (Sufi, 2022) analyzes and identifies global breaking news with sentiment extraction, yet it lacks mechanism for measuring event significance. News event ranking has been categorized into three distinct methods: ranking news streams, incorporating external sources, and employing query-based approaches (Setty et al., 2017).

Topic popularity forms the basis of one ranking system developed for incremental corpora, though this approach fails to incorporate historical features and may undermine assessments of long-lasting event significance (Corso et al., 2005). The Top Story Identification Task (Soboroff et al., 2010) evaluated news based on perceived popularity within the blogosphere. Query-based methods, by contrast, rank news relative to specific user queries rather than providing assessments of independent event significance (Setty et al., 2017).

User engagement metrics have gained prominence in recent news ranking frameworks. One approach integrates news diversity (measured through user-generated Twitter content), completeness, and speed metrics, effectively synthesizing external source-based ranking with independent event assessment (Karimi et al., 2021). Additionally, a more recent frameworks for extracting key news events from media streams through temporal trend analysis and unsupervised clustering techniques that identify events capturing significant public attention (Nakshatri et al., 2023).

Our approach employs GT as an external data source to measure public attention. Unlike query-dependent systems that rank events relative to specific user searches, our framework provides query-independent significance assessment by measuring inherent event importance through aggregated search behavior patterns.

2.2 Using Google Trends to evaluate the event significance

GT provides reports on search term popularity using the Google search engine (Cebrián and Domenech, 2022). Existing studies primarily use GT to evaluate health and financial event significance. GT search intensity correlates with various societal impact measures, including economic effects, policy changes, cultural discourse, and disease outbreaks (Liu et al., 2020; Simionescu and Cifuentes-Faura, 2022; Mavragani and Ochoa, 2019), making it a valuable indicator for assessing public interest in events or topics.

Unlike traditional media coverage metrics or expert assessments, search behavior reflects genuine public interest, providing objective measures of how events resonate with audiences. Recent work has explored combining LLMs with GT data for automated keyword generation in search engine optimization applications, demonstrating the potential of integrating language models with search trend analysis (Vadlapati, 2024).

3 Methods

The methodology consists of four primary stages (Figure 1). First, global events were extracted from Wikipedia's chronological pages. Second, LLMs generated contextually relevant search keywords for each event, as GT requires specific search terms rather than complete descriptions. Third, these keywords were used to collect GT search data with baseline comparisons over five years. Finally, composite significance scores were calculated to enable systematic event ranking based on global attention.

3.1 Wikipedia event extraction

Wikipedia was selected as the event source because contributors worldwide collaboratively summarize significant annual events in standardized chronological articles[1], providing global coverage with diverse perspectives and enabling systematic extraction. These chronological articles contain events categorized by month with occurrence times and brief descriptions (Hienert and Luciano, 2015). While 198 language versions exist, the English version was selected for practical methodological reasons: (1) English serves as a lingua franca for international news and global events coverage, making English search terms more likely to reflect international rather than purely regional significance, and (2) using a single language ensures consistency in keyword extraction and search query formulation throughout the research process.

This research extracted all 1,094 events and their descriptions appearing in Wikipedia from 2020 to 2024, including political elections, natural disasters, economic crises, sports competitions, technology breakthroughs and disruptions, health crises, and military conflicts worldwide.

[1] The page for 2020 events: https://en.wikipedia.org/wiki/2020

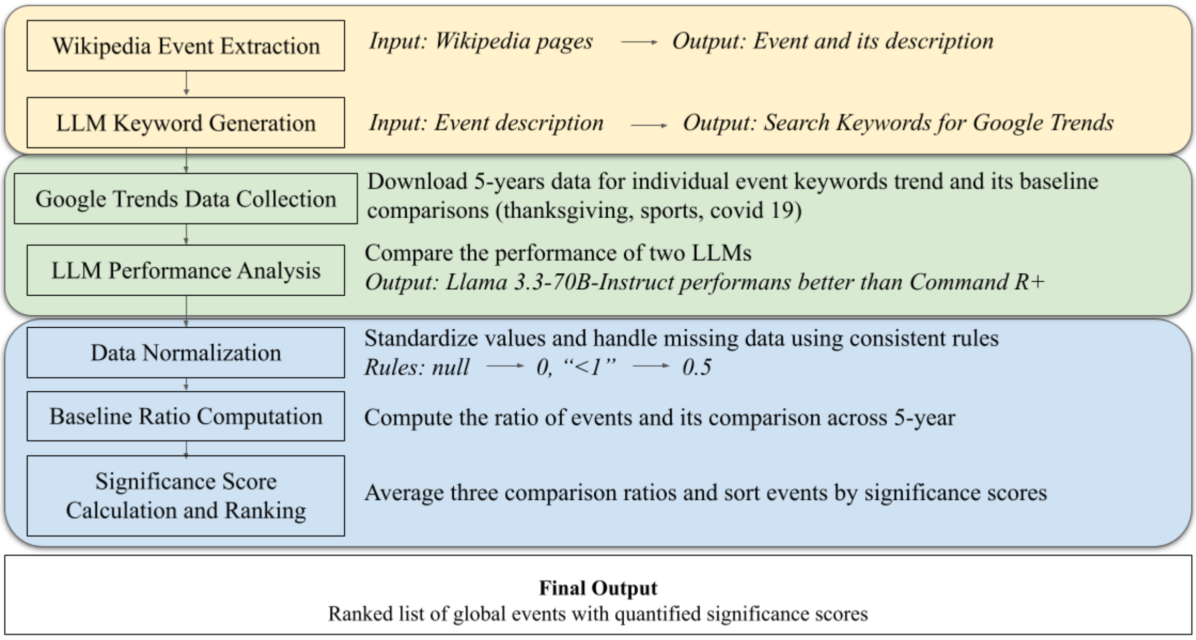

Figure 1: Event significance analysis workflow from data collection to ranking generation

3.2 LLM keywords generation

To identify the most significant global events, this study employed GT to rank events based on search intensity. The underlying assumption was that events more frequently searched on Google at the worldwide level correspond to greater influence at the global scale, as public search behavior serves as a reliable indicator of collective attention (Costola et al., 2021). Major sporting events, significant political development or breakthrough scientific discoveries can all generate high search volumes.

LLMs were employed to automatically extract keywords from Wikipedia's event descriptions. This process is essential because Wikipedia descriptions often contain detailed narratives unsuitable for direct GT searches, which require concise, targeted terms. For example, a Wikipedia entry about a natural disaster might contain extensive geological and casualty information, but effective GT keywords would be simplified terms like the disaster name and location

> **Source Events (Wikipedia)**: Flash floods struck Jakarta, Indonesia, killing 66 people in the worst flooding in over a decade.
> **Keywords extracted manually:** flash flood Jakarta

Specifically, Llama 3.3-70B-Instruct (Meta Llama, 2024) and Command R+ (Cohere, 2024) were selected based on their established capabilities in text generation tasks and availability. Identical prompts[2] for comparison purposes, examples and reasoning frameworks were applied to both LLMs.

The generated keywords were standardized to ensure GT compatibility by removing special characters and formatting according to GT query requirements. The quality of extracted keywords was initially evaluated based on:

- Peak timing match: whether search peaks of keywords generated align with event occurrence based on Wikipedia entry
- Delayed recognition: whether the keywords produced measurable search data on GT rather than data insufficiency warnings, even if the peak search occurred with some delay

The initial exploration on the 20 randomly sampled events revealed the key difference in Llama 3.3-70B-Instruct's consistent preservation of specific year information – a critical requirement for GT data collection methodology. This temporal specificity proves essential when analyzing events such as natural disasters, political elections, and other time-sensitive occurrences where precise chronological context directly impacts search behavior patterns.

> **Source Events (Wikipedia)**: The 2020 Serbian parliamentary election is held to elect all 250 members of the National Assembly of Serbia and

[2] https://github.com/Zenanc/Prompt_for_keywords_generation

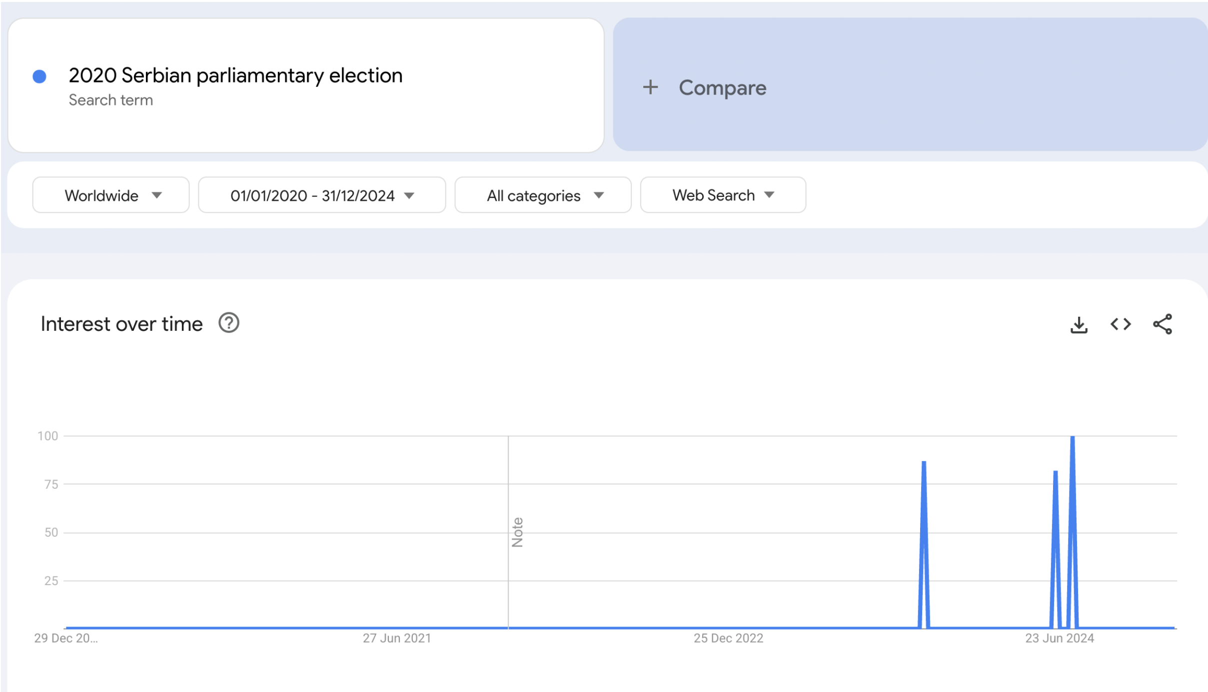

Figure 2: Google Trends search data for event keyword "2020 Serbian parliamentary election" (2020-2024) worldwide

the ruling For Our Children coalition won 188 out of 250 seats.
Keywords extracted by Command R+: Serbian parliamentar[sic] election results
Keywords extracted by Llama 3.3-70B-Instruct: 2020 Serbian parliamentary election

To evaluate the performance of these two LLMs, a secondary comparison was conducted. For events, where Llama 3.3-70B-Instruct generated keywords containing year information, the corresponding Command R+ keywords were augmented with the same temporal markers. This approach enabled a controlled comparison between the two LLMs' keyword generation capabilities while isolating the effect of temporal specificity. By standardizing the temporal component across both models' outputs, the analysis can determine whether the performance differences stem from temporal markers alone or from other qualitative aspects of keyword selection.

3.3 Google Trends data collection

GT provides relative search interest data where higher values indicate greater popularity of the search term within the chosen time frame and region (see Figure 2). However, this presents limitations for direct event comparison: GT normalizes data to each term's maximum within the selected timeframe rather than providing absolute search volumes—when comparing event keywords, the highest-searched keyword's peak becomes 100, with all others scaled proportionally, and restricts simultaneous comparisons to five keyword groups maximum (each describe one event). Therefore, events cannot be compared directly using raw GT data.

The GT methodology involves inputting keywords and selecting both geographic regions (worldwide, US, UK, etc.) and specific time periods. This study utilized worldwide geographic scope from January 1, 2020 to December 31, 2024.

To ensure comparability across all events, a two-step data collection method was employed. In the first step, individual GT data for each event keyword from 2020 to 2024 was downloaded to identify the peak search week for each event as it is impossible to obtain more fine-grained data than week intervals within a historical search period. Peak week analysis serves as a validation mechanism for LLM-generated keywords, where smaller temporal gaps between event occurrence and search peaks indicate superior keyword formulation that better reflects real-time public attention patterns.

In the second step, pairwise comparison data between each event keyword and a set of baseline keywords were downloaded across the full 2020-2024 period. This approach ensured consistent comparative scaling and enabled events to be ranked according to their significance scores against identical baseline keywords. The full five-year window was necessary as relative scaling between event keywords and baselines keywords varies dramatically depending on the temporal window selected.

Three baseline keywords were initially selected: “covid 19”, “weather”, and “black friday”. These represented consistent everyday search interest (weather), recurring seasonal events (black friday), and major global phenomena (covid 19), providing a comparative framework capturing various dimensions of human search behavior.

However, preliminary findings revealed these keywords exhibited very high search volumes, resulting in very few events producing sufficient aggregate significance scores. To address this limitation, a refined set was adopted: “sports” (everyday search interest), “thanksgiving” (seasonal events), and “covid 19” (retained for global significance). This adjustment maintained the comparative framework’s foundation while preserving detectability across a broader range of global events.

3.4 Peak week identification and LLM performance analysis

Historical GT data for each event keyword from 2020-2024 were downloaded via DataforSEO API[3], a third-party service that enables researchers to access the same publicly available trend data accessible through GT’s web interface.

Llama 3.3-70B-Instruct generated 1,024 valid keywords from 1,094 events, with 70 keywords excluded due to generation failures (producing no output or generating fewer than the required minimum of two keywords per event). Command R+ generated 1,069 keywords from 1,094 events, with 25 excluded for duplication across different events. After Google Trends data collection, 885 out of 1,024 Llama 3.3-70B-Instruct keywords and 842 out of 1,069 Command R+ keywords successfully returned valid peak data with temporal timeframes. Failed keywords either contained formatting errors incompatible with GT queries or represented events with insufficient search volume.

The duplication issue in Command R+ demonstrates overgeneralization when processing semantically similar events. For instance, two distinct Covid-19 events generated identical keywords, compromising specificity required for accurate trend analysis as distinct events become indistinguishable in search data.

For keywords exhibiting multiple search peaks throughout the five-year period, the first highest peak was selected, as this initial peak typically represents the moment of maximum public attention to an event, providing the most representative measure of initial global impact. Additionally, selecting the first peak minimizes potential confounding effects from anniversary coverage, follow-up events, or media retrospectives in subsequent years.

Following peak identification, temporal gaps between event occurrence and search peaks were computed. Since exact peak dates within each week are unavailable, gap calculations employed a standardized approach: events occurring within the peak search week were assigned a gap value of zero, indicating perfect temporal alignment; events preceding the peak week were measured from the event date to the peak week’s start date; events following the peak week were measured from the peak week’s end date to the event date.

The results showed that Llama 3.3-70B-Instruct maintains superior temporal accuracy even when Command R+ keywords were augmented with identical year information (Table 1). Specifically, Llama 3.3-70B-Instruct achieved relatively better temporal alignment (Gap = 0) for 256 events (28.92%), compared to 95 events (11.28%) for Command R+. This performance differential persisted across short-term temporal windows, with Llama 3.3-70B-Instruct capturing 327 events (36.94%) within a seven-day window versus 118 events (14.01%) for Command R+.

Notably, Command R+ exhibited a pronounced tendency towards extended temporal gaps, with 53.56% of events showing gaps exceeding 365 days, compared to 35.37% for Llama 3.3-70B-Instruct. Examination of specific cases revealed the nature of this temporal displacement. For instance, “Greece wildfires” reached its search peak on June 19-24, 2023, despite the actual event occurring on August 3, 2021, where the keywords registered an initial but smaller search peak (August 8-14, 2021). Similarly, “Abdallah Hamdock resignation protest” reached its only search peak on July 16-22, 2023, while the actual events occurred on January 2, 2022. These examples suggested that Command R+’s keyword formulation, despite temporal augmentation, generated search terms that failed to capture immediate public attention or pro-

[3] https://dataforseo.com/

Model Name	Command R+ (N=842)	Llama 3.3-70B-Instruct (N=885)
Gap = 0	95 (11.28%)	**256 (28.92%)**
$0 <$ Gap ≤ 7 days	23 (2.73%)	71 (8.02%)
$7 <$ Gap ≤ 30 days	33 (3.92%)	40 (4.52%)
$30 <$ Gap ≤ 365 days	241 (28.62%)	205 (23.16%)
Gap > 365 days	451 (53.56%)	313 (35.37%)

Table 1: The temporal gap comparison of Command R+ and Llama 3.3-70B-Instruct

duced keywords that aligned with later waves of interest rather than initial event-driven search.

The persistent performance gap indicated that temporal accuracy depends not only on the inclusion of year information, but on the underlying semantic and lexical choices that determine how effectively keywords match actual search behavior patterns during critical attention periods.

3.5 Data normalization

A consistent transformation was applied to account for the three value types returned by GT: (1) "null" indicates insufficient data for the event within the given time period, assigned a value of 0; (2) "$<$1" indicates searches below the minimum reporting threshold, assigned a value of 0.5 for numerical calculations to distinguish it from 0 and 1; and (3) numerical values 1-100 represent quantified relative search interest.

3.6 Baseline ratio computation

To ensure all the events were comparable across the five-year period, each event significance was measured by comparing search intensity against fixed baseline events. For each event keyword E and the set of baseline keywords $B = \{\text{covid 19}, \text{sports}, \text{thanksgiving}\}$, the significance ratio relative to each baseline was computed:

$$r_{E,B} = \frac{\sum_{t=2020}^{2024} V_{E,t} \cdot w_t}{\sum_{t=2020}^{2024} V_{B,t} \cdot w_t} \tag{1}$$

where:

- $V_{E,t}$ represents search volume for event E in year t
- $V_{B,t}$ represents search volume for baseline B in year t
- w_t represents time weight for year t

This cumulative approach captured the total social impact of events rather than just peak attention, recognizing that sustained or recurring interest might often indicate lasting influence on public consciousness.

3.7 Time weighting strategy

This weighting method attempted to address the temporal bias: events occurring in 2020 naturally had higher search volumes during 2020-2024 compared to events occurring in 2024. Higher weights for recent years corrected this systematic advantage of earlier events.

Given that the events span 2020-2024, but comparison requires fixed temporal windows for consistency, distance-based time weights were applied:

$$w_t = 0.05 * (t - 2020) + 0.1 \tag{2}$$

3.8 Significance score calculation

For event E with baseline ratio vector ($r_{E,covid}$, $r_{E,sports}$, $r_{E,thanksgiving}$), the significance scores are calculated using three different methods. Each method has distinct mathematical properties affecting how baseline performances are combined.

The arithmetic mean provides the most intuitive aggregation, testing all baseline comparison equally:

$$Score_{AM}(E) = \frac{1}{|B|} \sum_{b \in B} r_{E,b} \tag{3}$$

However, this method is disproportionately influenced by extreme outliers.

The geometric mean emphasizes proportional relationships and requires consistent performance across baselines for high scores:

$$Score_{GM}(E) = \left(\prod_{b \in B} r_{E,b} \right)^{1/|B|} \tag{4}$$

rank	keywords	weighted geometric score	peak week start	peak week end
1	US Open	29.666	2023-09-03	2023-09-09
2	Stock market	27.341	2020-03-08	2020-03-14
3	2022 FIFA World Cup	17.073	2022-11-27	2022-12-03
4	2023 Cricket World Cup	7.812	2023-10-29	2023-11-04
5	lunar eclipse	7.417	2022-05-15	2022-05-21
6	WHO COVID	6.557	2020-03-22	2020-03-28
7	2023 Rugby World Cup	5.303	2023-09-10	2023-09-16
8	James Webb Space Telescope	4.928	2022-07-10	2022-07-16
9	Cape Verde	4.736	2024-01-28	2024-02-03
10	2024 ICC T20 World	4.701	2024-05-26	2024-06-01

Table 2: Top 10 Events Ranked by Weighted Geometric Mean Score

However, this method may undervalue events with mixed significance patterns.

The harmonic mean takes a more conservative approach, penalizing low-performing baseline comparisons:

$$Score_{HE}(E) = \frac{|B|}{\sum_{b \in B} \frac{1}{r_{E,b}}} \tag{5}$$

This method is excessively conservative for events with mixed significance patterns.

To assess the ranking stability of the three aggregation methods, real events with diverse baseline ratio patterns were analyzed. These genuine cases revealed the practical advantages and limitations of each method in handling mixed-baseline performance and varying ratio distributions.

The real event data analysis demonstrated significant differences among aggregation methods in handling mixed baseline performance patterns. For instance, "Openai Chat" illustrated these ranking differences effectively. This event ranked 44th with arithmetic mean, 14th with geometric mean, and 8th with harmonic mean, despite having identical underlying data. The dramatic ranking variation (from 44th to 8th) demonstrated how different aggregation methods can fundamentally alter event prioritization. Arithmetic mean's low ranking (44th) suggested that OpenAI Chat's performance is diminished by extreme baseline ratios that skewed the average. Harmonic mean's high ranking (8th) indicated strong consistent performance across baselines, while geometric mean provided a moderate assessment (15th) that balanced both exceptional and weaker baseline performances.

This ranking instability highlighted that geometric mean offered the most reliable approach for event significance assessment, providing consistent evaluation that neither inflated nor unfairly penalized mixed performance patterns.

4 Result and discussion

This study successfully collected comparison data for 804 global events (with Llama 3.3-70B-Instruct), enabling significance scoring through comparison with three established baseline keywords. The proposed scoring methodology effectively ranked all 804 events, with the weighted geometric mean proving a solid measure of relative significance across diverse event categories. Table 2 presents the top-ranked events using keywords generated by Llama 3.3-70B-Instruct.

The ranking results demonstrated the methodology's effectiveness in capturing diverse event categories, including sports events ("US Open", "FIFA World Cup"), scientific achievements ("James Webb Telescope"), natural phenomena ("lunar eclipse"), and tragic events ("WHO COVID"). This diversity validated the methodology's capacity to identify various forms of global attention rather than only crisis-driven events.

The prominence of sporting events in the top rankings reflected their substantial capacity to generate global attention, with the "US Open" achieving the highest (score: 29.666) and the "FIFA World Cup" ranking third (score: 17.073). It might be because mega-events can capture widespread public engagement across diverse demographic and geographic segments.

The second ranked "Stock market" (score: 27.341), corresponded to the global stock market crash that began on February 20, 2020, following growing instability due to the Covid-19 pandemic. This crash ended on April 7, 2020, representing one of the most significant financial disruptions in recent history, and it validated the methodology's capacity to capture major adverse economic events.

The bottom events, such as "2024 Namibian election Netumbo" (score: 0.001487), "Falkland Islands land mine" (score: 0.001013), and "Sichuan earthquake Luding" (score: 0.000984), demonstrated clear differentiation from top global events. This ranking aligned with expected global significance patterns, as these represented region-specific political developments, geographically isolated incidents, and localized disasters respectively.

4.1 Validation

To validate the weighting method, this study assessed whether the time weighting scheme introduces systematic bias towards recent events. The validation examined event distribution across score bins by year, calculated expected versus actual numbers of top 10% events, and computed correlation between time weights and bias ratios.

It is revealed that most events (85%) fall within the 0-1 score range using geometric mean values (minimum score: 0.000053, maximum score: 0.979104), and this distribution pattern remains stable across years with 129-170 events annually. Very few events achieve scores above 5 (only 7 out of 804), with highest score categories containing just 3 events across all years.

The correlation coefficient of -0.9175 between time weights and bias ratios indicated that recent years actually produce fewer high-scoring events than expected, providing evidence that the time weighting methodology does not artificially inflate recent event significance. The slight over-representation in 2020-2021 likely reflects genuinely significant historical events (such as the Covid-19 pandemic and related global disruptions).

5 Conclusion

This study presents a novel framework for ranking global event significance through multi-baseline comparison using GT data. The methodology employs three distinct baseline keywords-"covid 19" (burst pattern), "sports" (stable pattern), and "thanksgiving" (seasonal pattern)-to provide robust comparative assessment across diverse temporal contexts. This diversified approach decreases the risk of biased significance scores from single-baseline fluctuations.

Key methodological findings demonstrate that Llama 3.3-70B-Instruct outperforms Command R+ in generating keywords from event descriptions that capture immediate public response to global events, achieving superior temporal alignment between event occurrence and search peaks (28.92% vs. 11.28% alignment). Among aggregation methods, geometric mean proves most effective, providing balanced significance assessment while avoiding the outlier sensitivity of arithmetic mean and excessive conservatism of harmonic mean.

This approach represents the first systematic integration of GT and LLMs for global event ranking, introducing a scalable, automated methodology that eliminates manual keyword selection biases while maintaining cross-temporal comparability. The study's novelty lies in combining automated keyword generation with multi-baseline aggregation, offering practical applications for news organizations, policy makers, and researchers requiring objective event significance assessment.

Limitation

There are few limitations of this proposed methodology. First is that major events often generate diverse search terminologies, potentially diluting their apparent significance. The outbreak of Covid-19 pandemic, for instance, might be searched as "Covid-19", "coronavirus", "covid-19 pandemic" or other variants, fragmenting the search signal.

Second, the API's data collection process occasionally returns null values despite the existence of actual search data, potentially leading to the systematic exclusion of some events.

Third, the methodology's reliance on English-language search terms may introduce geographic and linguistic bias, potentially undermining events of significance in non-English speaking regions. This limitation could affect the global representativeness of the event rankings.

Acknowledgments

This research was supported by funding from the Economic and Social Research Council (ESRC). I would like to express my sincere gratitude to Dr Judita Preiss and Professor Peter A Bath for their invaluable supervision and continuous support throughout this research. I also thank Nicholas Musembi for his technical assistance in downloading the Wikipedia events.

References

Eduardo Cebrián and Josep Domenech. 2022. Is Google Trends a quality data source? *Applied Economics Letters*, 30(6):811–815.

Cohere. 2024. Cohere's Command R+ Model. Accessed: 2025-08-10.

Gianna M. Del Corso, Antonio Gullí, and Francesco Romani. 2005. Ranking a stream of news. In *WWW '05: Proceedings of the 14th international conference on World Wide Web*, pages 97–106.

Michele Costola, Matteo Iacopini, and Carlo R.M.A. Santagiustina. 2021. Google search volumes and the financial markets during the COVID-19 outbreak. *Finance Research Letters*, 42.

Google Trends. Google Trends. Accessed: 2025-08-10.

Daniel Hienert and Francesco Luciano. 2015. Extraction of Historical Events from Wikipedia. In *The Semantic Web: ESWC 2012 Satellite Events*.

Atina Husnayain, Anis Fuad, and Emily Chia-Yu Su. 2020. Applications of Google Search Trends for risk communication in infectious disease management: A case study of the COVID-19 outbreak in Taiwan. *International Journal of Infectious Diseases*, 95:221–223.

Samaneh Karimi, Azadeh Shakery, and Rakesh Verma. 2021. Online News Media Website Ranking Using User Generated Content. *Journal of Information Science*, 47(3):340–358.

Duleeka Knipe, Hannah Evans, Mark Sinyor, Thomas Niederkrotenthaler, David Gunnell, and Ann John. 2020. Tracking online searches for emotional wellbeing concerns and coping strategies in the UK during the COVID-19 pandemic: a Google Trends analysis. *Wellcome Open Research*, 5:220.

Liang Kong, Shan Jiang, Rui Yan, Shize Xu, and Yan Zhang. 2012. Ranking news events by influence decay and information fusion for media and users. In *CIKM '12: Proceedings of the 21st ACM international conference on Information and knowledge management*, pages 1849–1853.

Ying Liu, Geng Peng, Lanyi Hu, Jichang Dong, and Qingqing Zhang. 2020. Using Google Trends and Baidu Index to analyze the impacts of disaster events on company stock prices. *Industrial Management & Data Systems*, 120(2):350–365.

Amaryllis Mavragani and Gabriela Ochoa. 2019. Google Trends in Infodemiology and Infoveillance: Methodology Framework. *JMIR Public Health and Surveillance*, 5(2).

Meta Llama. 2024. meta-llama/Llama-3.3-70B-Instruct. Accessed: 2025-08-10.

Nishanth Nakshatri, Siyi Liu, Sihao Chen, Dan Roth, Dan Goldwasser, and Daniel Hopkins. 2023. Using LLM for Improving Key Event Discovery: Temporal-Guided News Stream Clustering with Event Summaries. In *Findings of the Association for Computational Linguistics: EMNLP 2023*.

Vinay Setty, Abhijit Anand, Arunav Mishra, and Avishek Anand. 2017. Modeling Event Importance for Ranking Daily News Events. In *WSDM '17: Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, pages 231–240.

Mihaela Simionescu and Javier Cifuentes-Faura. 2022. Can unemployment forecasts based on Google Trends help government design better policies? An investigation based on Spain and Portugal. *Journal of Policy Modeling*, 44(1):1–21.

Ian M. Soboroff, Iadh Ounis, and Craig Macdonald. 2010. Overview of the TREC 2008 Blog Track.

Fahim K Sufi. 2022. AI-GlobalEvents: A Software for analyzing, identifying and explaining global events with Artificial Intelligence. *Software Impacts*, 11.

Praneeth Vadlapati. 2024. AutoTrendyKeywords: Real-Time AI-Driven Trend-Based SEO Using LLMs. *International Journal of Computer Engineering and Technology*, 16(3):241–255.

Canhui Wang, Min Zhang, Liyun Ru, and Shaoping Ma. 2008. Automatic online news topic ranking using media focus and user attention based on aging theory. In *CIKM '08: Proceedings of the 17th ACM conference on Information and knowledge management*, pages 1033–1042.