How Misunderstanding Can Influence a Journal Negatively: Public Opinion Analysis of the *TIV* Event Based on Collaboration of Large Language Models and Small Models

Shuai Guo, Weishan Zhang, Baoyu Zhang, Tao Chen, Hui Zhang, and Mohammad Kamrul Hasan

***Abstract*—The "on hold" event of *IEEE Transactions on Intelligent Vehicles* (*TIV*) is a typical case of misunderstanding of how a journal runs, and misleading public opinions from online media seriously harm a journal. Public opinions on this kind of case had never been investigated. The paper proposes a multidimensional quantitative regression framework (MQRF) that demonstrates how to quantify and predict public opinion influences for such cases through the collaboration of large language model and small model. The framework leverages large language models' comprehensive analytical capabilities for contextual understanding while employing specialized small models for precise time-series prediction, creating a synergistic approach that significantly outperforms single-model solutions. We analyze the evolving relationships among public opinions and examine how these changes prompt the journal to disclose and clarify its operational practices. Public misunderstandings cause lasting damage to the journal. Understanding how different types of papers run in the journal can eliminate misunderstanding though with a time lag. Through in-depth analysis of the *TIV* event and comparative experiments with six small models (Linear, DLinear, NLinear, TimesNet, Transformer, and PatchTST) and six large language models (ChatGPT-4.5, Claude3.7sonnet, Grok3, Deepseek-R1, Doubao, and Qwen2.5-72B-Instruct), the effectiveness of the MQRF is verified.**

***Index Terms*—Large language model and small model collaboration, public opinion analysis, influence quantification, public opinion prediction**

I. Introduction

DURING June to August 2024, some controversies emerged regardig the *IEEE Transactions on Intelligent Vehicles* (*TIV*) journal being "on hold". This event triggered a multidimensional online public opinion. During this process, cyberspace has evolved into a primary domain of discursive contention and emotional convergence among public opinions [1]. The dissemination of public opinion demonstrates distinctive characteristics including long-term duration, extensive demographic data of the participants, and marked polarization of points of view [2]. Powered by social media platforms, public attention to *TIV* has undergone a rapid escalation, accompanied by substantial emotional fluctuations and ideological divergences [3]. This phenomenon has precipitated a series of operational clarifications and self-verification measures from the journal, herein referred to as reality processes. The acce-lerated intensification of online sentiment not only exerts propulsive forces on these reality processes, but also unveils the multifaceted complexity inherent in public misunderstanding and affective communication.

The proliferation of social media platforms has created an unprecedented urgency for understanding public opinion dynamics in academic contexts [4]. Academic journals, as pillars of scholarly communication, are increasingly vulnerable to rapid-spreading misinformation that can cause irreversible reputational damage within hours [5]. The *TIV* case exemplifies this critical vulnerability, where a journal's decade-long reputation was severely compromised within weeks due to public misunderstanding of standard editorial practices. This represents a significant research gap, as existing literature lacks systematic approaches to quantify, predict, and mitigate such opinion-driven crises in academic publishing.

In such circumstances, a profound understanding of the mechanisms of dissemination and the dynamic evolution of public opinion is important [6, 7]. Quantitative assessment of sentiment influence can provide data-driven support to better mitigate the societal and institutional impacts of opinion storms [8]. To do this, deep learning technologies have become predominant methodologies in sentiment analysis [9, 10]. However, current research mainly focuses on the classification of textual emotions while lacking systematic investigations into influence prediction and procedural impacts on reality [11]. Moreover, though journal-related

Manuscript received: 18 June 2025; revised: 24 August 2025; accepted: 28 August 2025. *(Corresponding author: Weishan Zhang.)*

Citation: S. Guo, W. Zhang, B. Zhang, T. Chen, H. Zhang, and M. K. Hasan, How misunderstanding can influence a journal negatively: Public opinion analysis of the *TIV* event based on collaboration of large language models and small models, *Int. J. Intell. Control Syst.*, 2025, 30(3), 232–243.

Shuai Guo, Weishan Zhang, Baoyu Zhang, and Tao Chen are with Qingdao Institute of Software (College of Computer Science and Technology), China University of Petroleum (East China), Qingdao 266580, China (e-mail: bz24070010@s.upc.edu.cn; zhangws@upc.edu.cn; b22070002@s.upc.edu.cn; dongshou2015@s.upc.edu.cn).

Hui Zhang is with School of Computer Science and Technology, Beijing Jiaotong University, Beijing 100044, China (e-mail: huizhang1@bjtu.edu.cn).

Mohammad Kamrul Hasan is with Center for Cyber Security, The National University of Malaysia, Bangi 43600, Malaysia (e-mail: hasankamrul@ieee.org).

Digital Object Identifier 10.62678/IJICS202509.10223

opinion storms triggered by public misunderstandings frequently occur, the corresponding exploration for such issues remains scarce.

In recent years, large language models (LLMs) have been widely adopted in public opinion research [12]. LLMs are rich in prior knowledge and suitable for making comprehensive analyses [13]. Small models can achieve better results in specific tasks after training [14]. It is of high research significance to combine the advantages of both models to realize accurate analysis of public opinions. The integration of LLMs with task-specific insights from smaller models has shown promising results. However, there remains a notable literature gap on the collaborative application of large language models and small models to analyze public opinion [15]. To address this, we propose a novel method for quantifying and predicting public opinion influence (POI) through the synergistic use of large language models and small models. By analyzing the dynamic change of the "misunderstanding-public opinion-reality process" in the *TIV* event, this study provides new perspectives for understanding how online public opinion influences real-world processes.

(1) What inherent patterns govern the individual evolution of misunderstandings, public opinion, and reality processes?

(2) Through what mechanisms do public misunderstandings influence reality processes?

(3) Can data-driven methodologies effectively forecast future reality processes?

To answer these questions, this paper collects public opinion data related to the *TIV* event from various social media platforms. Integrating these data with observations of reality event developments, we construct a public opinion quantification and prediction framework based on the collaborative capabilities of large language models and small models. By performing a multidimensional analysis of indicators such as temporal progression, scale of dissemination, and semantic factors, this paper dynamically quantified the public opinion influence and examined its correlation with reality processes. Furthermore, a comparative evaluation of the performance of different large language models and small models was conducted, offering empirical evidence to support model selection.

The main contributions of this paper are as follows:

(1) Using the *TIV* event as a case, we analyze why public misunderstandings arise and how they can cause lasting damage to journal operations.

(2) A multidimensional method for quantifying the public opinion influence is proposed. This method evaluates the severity of public discourse from multiple perspectives, including development time, development scale, and sentiment factors, thereby laying the foundation for effective public opinion prediction.

(3) A collaborative framework for public opinion influence quantification and prediction is designed, leveraging the complementary strengths of large language models and small models. Specifically, large language models provide contextual interpretation and semantic analysis capabilities, while small models deliver precise temporal pattern recognition and numerical prediction accuracy. In this framework, small models perform specific public opinion prediction tasks, while LLMs analyze the results based on the prediction. This approach offers immediate practical value for academic journals by enabling proactive crisis detection before peak negative sentiment, allowing time for strategic intervention.

(4) Comprehensive experimental evaluations are designed to validate the effectiveness of multidimensional quantitative regression framework (MQRF) in the analysis of public opinion and prediction of the reality process, which can serve as scientific criteria for LLM model selection.

The subsequent sections are organized as follows. Section II reviews existing work on sentiment quantification and analysis. Section III delineates the chronology of the *TIV* event and the data collection methodology. Section IV details the MQRF architecture, including workflow design and quantification techniques. Section V presents analytical findings from the application of MQRF to the *TIV* event together with model comparison experiments. Section VI concludes with research insights and future directions.

II. Related Work

The internet has become the primary platform for individuals to express their opinions [16, 17], significantly shaping the dissemination of public opinion events. Public opinion, in turn, can influence reality processes, making quantification and exploration of its impact crucial for practical applications [18, 19].

Currently, the literature on the quantification of public opinion remains limited. Thukral et al. utilized bidirectional long short-term memory (Bi-LSTM) networks to classify tweets on public opinions about climate change, constructing a framework to quantify the intensity of these opinions as a measure of public concern [20]. Ma et al. combined pretrained models with topic clustering and correlation analysis methods to quantitatively investigate temporal dynamics, content characteristics, and audience feedback during the COVID-19 pandemic [21]. Pinto et al. employed topic detection to compare media agendas with public agendas, using Shannon entropy and Jensen-Shannon divergence to measure agenda diversity and discrepancies [22].

LLMs have become prevalent in public opinion analysis. Qu and Wang used LLMs to reveal public thematic biases in political and environmental discussions [23]. Ren et al. applied the MiniBatchKMeans algorithm to group events based on the online heat index of public opinion, categorizing them into four distinct heat levels [24]. Lee et al. utilized LLMs to track public opinion trends and biases toward global warming, exploring the factors influencing these shifts [25]. Zhang et al. proposed an LLM-based method for the analysis of opinion about protests, using the case of the Nahel Merzouk protest in France to explore the impact of emerging social media on protest activities [26]. Chen et al. introduced a simplified large graph clustering algorithm (ALGC) to generate document and topic representations, analyzing topic evolution on social media during Pelosi's visit [27].

In recent years, the collaboration of large language model

and small model has gained traction as a research hotspot. By combining the prior knowledge of LLMs with the specialized expertise of small models, this collaboration has demonstrated enhanced performance in various tasks. Liu et al. proposed a brain-like cognitive-driven model factory (BC-MF), where LLMs act as supervisors, adaptively generating small-scale models tailored to heterogeneous data [28]. Fan et al. introduced a parameter-efficient federated mutual knowledge transfer framework that allows large language models to transfer knowledge to smaller models, enriching them with domain-specific insights while enhancing overall task performance [29]. Li et al. used LLMs to generate synthetic data for small models through demonstrations in small numbers and differential privacy, facilitating domain-specific knowledge transfer [30]. Furthermore, Fan et al. explored methods to improve small models by uploading perturbed commands to a server's LLM, improving multitask learning performance [31]. Yin et al. investigated the relationship between LLM performance and data selection, linking training loss and data compression ratio with LLM lightweighting and collaboration efficiency [32].

Despite these advancements, there remains a notable gap in the application of large language model and small model collaboration specifically for public opinion quantification, prediction, and analysis of public opinion's influence on reality processes.

III. Data Collection

A. Event Description

Fig. 1 illustrates the evolution of public opinion and the corresponding course of events in the *TIV* event. In terms of public opinion development, by mid-to-late June 2024, online discussions questioning the *TIV* journal's high self-citation rate began to surface. This sentiment intensified in early-to-mid July as negative opinions about the journal continued to gain attentions. Between 23 July and 6 August, public discourse experienced a significant surge, with discussions largely focused on alleged illegal paper acceptances, the journal's status of "on hold", and the investigation involving three ITSS commissioners. This period was an escalation of verbal attacks directed to the journal operations. Meanwhile, malicous rumors related to *TIV* journal operations and editors began to emerge containing personal attacks against researchers involved in the editoral process. These rumors attracted considerable attentions and caused even more discussions at various platform of social media and apps. By mid-to-late August, public opinion began to stabilize and subside.

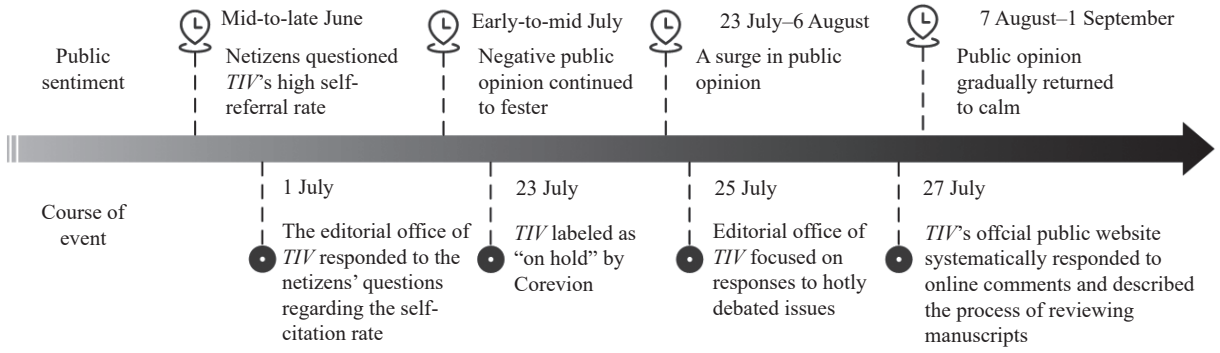


Fig. 1 Public opinion and course of event in the *TIV* event.

Regarding the developments in reality processes, on 1 July 2024, the editorial office of *TIV* addressed internet users' concerns regarding the journal's self-citation rate. On 23 July, Clarivate Analytics officially marked the *TIV* journal as "on hold". Two days later, on 25 July, the editorial office of *TIV* issued a comprehensive response to several key issues, including alleged illegal acceptance of articles, the investigation by ITSS commissioners, and the high rate of self-citation of the journal. In this response, the editorial office of *TIV* revealed the potential sources of the anonymous letters and provided an analysis of the reasons behind the investigation. On 27 July, the official *TIV* account systematically responded to various online criticisms, including those related to "on hold" status and self-citation concerns, and clarified the journal review process. Different types of papers in a journal have different focuses and review processes.

B. Data Source

To obtain more comprehensive public opinion data, we collected more than 10,000 pieces of data from four public opinion platforms, namely the public accounts of Letpub, Rednote, Zhihu, and WeChat, and the total amount of data available after screening was more than 4182 pieces. The data volume for each platform is summarized in Table 1. The data collection process spanned from 20 June 2024 to 1 September 2024 and included information on the content of the comments, the date, and the trend of evaluation. To ensure the quality of the dataset, we conducted several pre-processing steps. Specifically, we removed URL links and emoticons, retaining only textual content. Additionally, duplicate comments posted on the same day were eliminated to prevent the distortion of public opinion data caused by "comment flooding".

IV. Method

The general process of the MQRF is illustrated in Fig. 2. First, public opinion data are collected from various platforms, while reality process data are gathered from the WeChat

Table 1 Data volume from different Chinese social media platforms.

Data source	Data volume
Letpub	889
Rednote	1596
Zhihu	1633
WeChat	64

public account used by the editorial team. Next, we propose a method to quantify the public opinion influence, calculate the public opinion influence, and determine the sentiment factors. Simultaneously, we discretize the data of the reality process, dividing it into several distinct process nodes.

Subsequently, we perform a regression analysis of the sentiment factors, public opinion influence, and reality processes over time, generating emotional, public opinion, and reality lines. We then integrate the quantified public opinion influence at each process node and analyze how the accumulated public opinion drives the progression of reality processes, ultimately producing a line that represents the public opinion influence on reality.

Following this, time-series prediction models are used to forecast sentiment factors, public opinion influence, and the timing of future process nodes. Finally, the prediction results are fed into an LLM, which utilizes its prior knowledge to generate an analysis based on the various predicted outcomes.

A. Multi-Indicator Public Opinion Influence Quantification Method

Currently, the processing of public opinion is mainly based on analytical methods, with a lack of effective approaches to quantify the public opinion influence. This paper addresses this gap by extracting four key indicators from online public opinion: development time, development scale, sentiment factors, and cumulative scale. By integrating these indicators, we propose a multi-indicators public opinion influence quantification method to calculate the daily value of public opinion influence.

The development time of the public opinion is determined by calculating the time interval between the start date of the public opinion event and the calculation date, as shown in Eq. (1).

$$T = T_{\text{present}} - T_{\text{start}} \tag{1}$$

where T_{present} is the date for which public opinion influence is to be calculated, and T_{start} is the start date of public opinion.

The sentiment factor of public opinion is obtained by calculating the average of the emotional scores of all data. Its calculation formula is shown in Eq. (2).

$$\overline{S} = \frac{(S_1 + S_2 + S_3 + \cdots + S_N)}{N} \tag{2}$$

where S_i is the emotional score of the i-th data, and N is the total amount of data on the calculation date. If a data point has a positive evaluation, its emotional score is 1. In contrast, the emotional score is –1. If a data point has a neutral evaluation, its emotional score is 0.

The public opinion development scale is obtained by aggregating the total amount of public opinion data collected from all platforms on the calculation date. The cumulative scale of public opinion is the sum of the development scales of all days from the start date of public opinion to the calculation date, as shown in Eq. (3).

$$M = N_1 + N_2 + N_3 + \cdots + N_T \tag{3}$$

where N_i is the development scale on the i-th day.

Finally, the calculation formula for public opinion influence is shown in Eq. (4).

$$\text{POI} = \frac{\left(\left(1 - \mathrm{e}^{-0.05T}\right) + \frac{M + \varepsilon}{M + N} + \left|\overline{S}\right|\right)}{3} \times N \tag{4}$$

where ε represents a very small constant, used to avoid the numerator being 0. The time factor $1 - \mathrm{e}^{-0.05T}$ limits the effect of T between 0 and 1, and increases with increasing T. This ensures that the public opinion influence is positively correlated with development time and does not cause numerical inflation. The coefficient –0.05 ensures that the influence of development time is no longer significant after 60 days, which meets the requirements for the heat period of events of general public opinion. By performing a regression

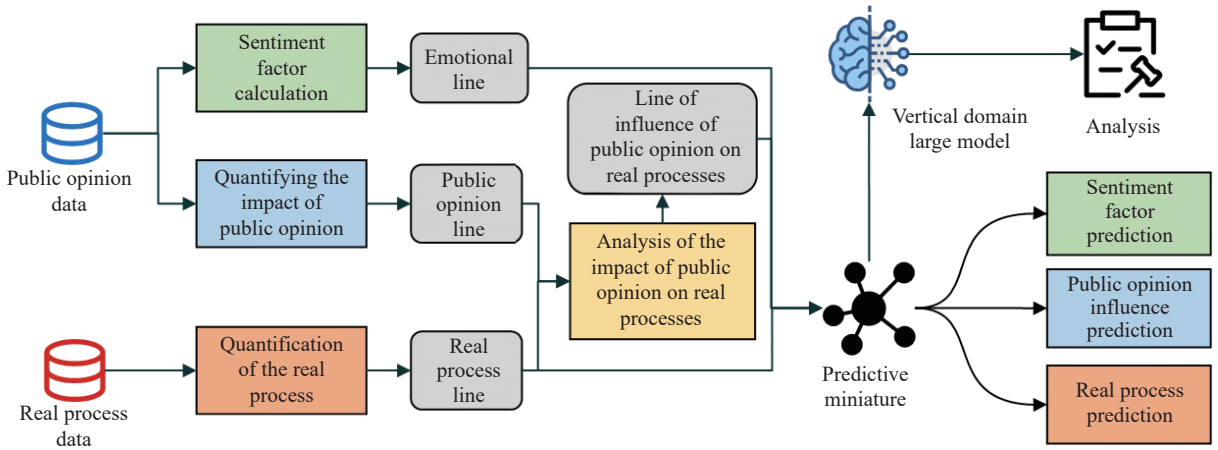


Fig. 2 General process of the MQRF.

analysis on each date and the daily public opinion influence value, the public opinion influence line (POIL) is obtained.

B. Opinion-Reality Correlation Analysis Method Based on Linear Regression

In order to quantify the process of the reality process and the role of public opinion on the reality process, a linear regression based correlation analysis method of perception and reality is proposed. First, for the event process $R = \{r_1, r_2, ..., r_m\}$, each process is quantified as $RQ = \{1, 2, ..., m\}$ in order of occurrence, and the time interval $TR = \{tr_1, tr_2, ..., tr_{m-1}\}$ between each process is recorded. Here, r_i represents the i-th reality process, m is the total number of processes, and tr_i represents the time interval between the i-th and $(i+1)$-th processes. By performing a regression analysis on each date and the quantified process, the reality line is obtained. Then, the integral of the public opinion line during each time interval is calculated, yielding the cumulative public opinion influence required to realize each reality process. Finally, a regression analysis is conducted on each cumulative public opinion influence and each quantified process, resulting in the line of public opinion influence on reality processes (POIRLs). Algorithm 1 provides a detailed explanation of our proposed method.

Algorithm 1 Algorithm for public opinion-reality correlation analysis method based on linear regression

Input: POIL and TR

Output: POIRL

$A \leftarrow \emptyset$ // Initialize the cumulative set of public opinion influence

for each $i \leftarrow 1$ to m **do**

  $RQ_i = i$ $tr_i = T_{r_{i+1}} - T_{r_i}$

  $\text{POIR}_i = \int_{T_{r_i}}^{T_{r_{i+1}}} \text{POIL}$ A.add(POIR_i)

end

for each $i \leftarrow 1$ to m **do**

  POIRL = fit(TR_i, A_i)

end

return POIRL

V. EXPERIMENT

A. Experimental Setup

To verify the effectiveness of the proposed methods, experiments were conducted using the *TIV* event as a case to quantify the public opinion influence and assess its impact on reality processes. Comparative experiments were also carried out to identify the most suitable models for both the small (time series prediction) and large (text generation) model components, using currently mainstream architectures in each category.

The experiments were conducted on a server equipped with a 128-core Intel(R) Xeon(R) Gold 6338 @ 2.00 GHz CPU, 512 GB of RAM, and two Nvidia RTX A6000 GPUs. The software environment included PyCharm as the development interface and PyTorch as the primary deep learning framework.

B. Evaluation Metric

For the time series prediction task involved in small model selection, three commonly used evaluation metrics were adopted: mean squared error (MSE), mean absolute error (MAE), and relative squared error (RSE).

MSE assesses prediction accuracy by computing the average of squared differences between predicted and actual values. MAE evaluates the average magnitude of the prediction errors by calculating the mean of the absolute differences between the predicted and true values. RSE, on the other hand, measures the relative error of the model by comparing the sum of squared prediction errors with the total variance of the ground truth values.

The computational formulas of MSE, MAE, and RSE are provided in Eqs. (5)–(7).

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{5}$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{6}$$

$$\text{RSE} = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{7}$$

where y_i represents the true value, $\hat{y}_i$ represents the predicted value, $\bar{y}$ is the mean of the true values, and n is the number of samples.

For the text generation task (LLM selection), the evaluation metrics were set as Rouge-1, Rouge-L, and BLEU. Rouge-1 measures the overlap between generated analyses and reference analyses at the word level. Rouge-L focuses on the longest common subsequence between generated analyses and reference analyses, reflecting the model's ability to generate consistent long sequences. BLEU evaluates the similarity between generated analyses and reference analyses through n-gram precision. The calculation principles of precision (P), recall (R), and the F1 score for Rouge-L are shown in Eqs. (8)–(10).

$$\text{RL}(P) = \frac{\text{LCS}(l, g)}{\text{len}(g)} \tag{8}$$

$$\text{RL}(R) = \frac{\text{LCS}(l, g)}{\text{len}(l)} \tag{9}$$

$$\text{RL(F1)} = \frac{2 \times \text{RL}(P) \times \text{RL}(R)}{\text{RL}(P) + \text{RL}(R)} \tag{10}$$

where l represents the reference analysis, g represents the generated analysis, and $\text{LCS}(l, g)$ is the length of the longest common subsequence between l and g.

C. Analysis of Public Opinion Influence

The quantified results of the public opinion influence throughout the event are illustrated in Fig. 3. Before 17 July, the influence level remained low and stable, consistently below 1. During this period, only limited discussions emerged around the issue of high self-citation rates and public attention was minimal.

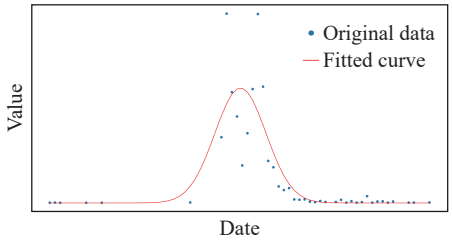


Fig. 3 Quantitative results of public opinion influence in the *TIV* event.

Between 17 July and 23 July, the public opinion influence exhibited a sharp upward trend, increasing from 0.8 to 130.2. This surge was driven by public misconceptions about the journal's operational mechanisms and was amplified by herd behavior, causing the discourse to rapidly intensify. From 23 July to 1 August, the influence values fluctuated between 225.7 and 374.4. The announcement of the journal being "on hold" marked the peak of the public reaction, as it was widely interpreted as confirmation of the perceived issues. In response, *TIV* issued multiple statements during this period, further fueling public attention and accelerating the reality progression of the event.

From 1 August to 7 August, the public opinion influence began to decline sharply, dropping from 83.4 to 6.4, as the intensity of the discussions waned and the overall interest in the event started to decrease. Between 7 August and 1 September, the influence value stabilized once again below 1, with only few discussions and limited attention.

In general, the temporal pattern of public opinion influence approximately followed a normal distribution. A regression analysis was conducted to model this trend, and the resulting functional relationship between public opinion influence and time is presented in Eq. (11).

$$\mathrm{POI} \approx 227.91 \times \mathrm{e}^{-\frac{(T-36.66)^2}{2\times 4.91^2}} \tag{11}$$

D. Analysis of Sentiment Changes and Reality Process

The evolution of sentiment factors during the event is shown in Fig. 4. Before 17 July, public opinion fluctuated between negative and neutral with large amplitudes. This volatility was primarily due to the limited scope of public discourse, resulting in a smaller sample size to calculate sentiment factors. During this stage, the editorial office of *TIV* issued an initial response regarding the high self-citation issue; however, public opinions remained largely skeptical.

From 17 July to 23 July, as the breadth of public discussion expanded and the initial response began to gain partial acceptance, the negative sentiment gradually decreased. The sentiment factor increased, reflecting a delayed public response to the journal's clarifications, taking approximately 15 days after initial clarification.

A brief decline in the sentiment factor occurred between 23 July and 25 July, driven by the announcement that *TIV* had been placed "on hold". From 25 July to 4 August, the sentiment factor initially dipped slightly before rising significantly. This recovery was primarily driven by the journal's series of concentrated responses issued on 25 July and 27 July, which shortened the public acceptance delay to approximately two days.

From 5 August to 1 September, the sentiment factor remained consistently positive, as the public was gradually overcoming misunderstandings about the journal, and support for the journal was increasingly becoming the prevailing sentiment.

To analyze the relationships between public opinion and reality developments, we discretize the reality process of the

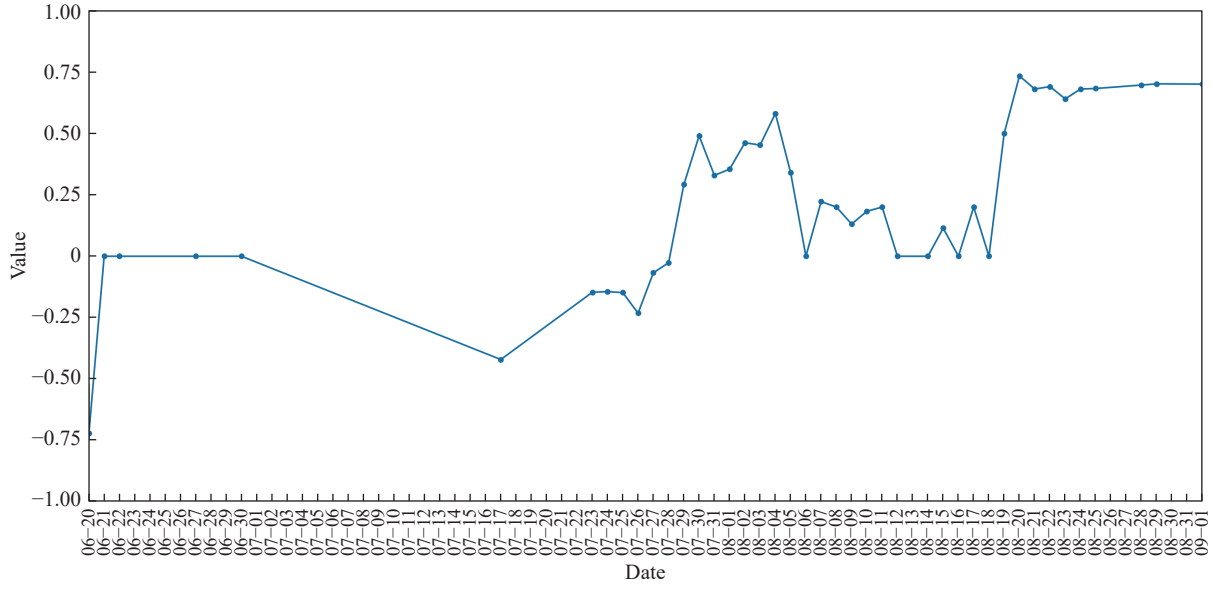


Fig. 4 Quantitative results of sentiment factor in the *TIV* event.

TIV event into five key stages. The specific time frames and the corresponding developments are summarized in Table 2.

The relationships between public opinion influence and the progression of reality events are illustrated in Fig. 5. The horizontal axis denotes the cumulative value of public opinion influence, while the vertical axis represents the stages of the event process, with stages 0 through 4 corresponding to the five reality nodes defined in Table 2.

The curve shows a non-linear pattern, where the cumulative public opinion influence required to trigger a transition from one stage to the next initially decreases and then increases. This suggests that, at the peak of public attention, even a relatively small additional amount of opinion influence was sufficient to drive the event forward. In other words, higher levels of public opinion influence exerted a stronger driving force on the progression of reality developments. In contrast, in the early and late stages, a greater accumulation of influence was needed to affect the change, indicating a relatively weaker impact during those periods. To further investigate the impact of public opinion on the progression of reality events, two key dates in the reality process, 25 July and 27 July, were selected for detailed analysis. The analysis covered a two-day time window before and after each node, with the results presented in Figs. 6 and 7, respectively.

Table 2 The reality process of the *TIV* event.

Time	Reality process
7.1 (initial)	Editorial office responds to the issue of *TIV*'s high self-citation rate. Editorial office of *TIV* further responds to the issue of impact factor.
7.23 (rising)	*TIV* is marked as "on-hold".
7.25 (climax)	Editorial office of *TIV* responds to the issue of questioned papers, and reveals possible sources of anonymous letters. Editorial office of *TIV* responds to questions about self-citation rates and analyzes the reasons for the journal being investigated. Editorial office of *TIV* responds to the alleged investigation by three ITSS commissioners, as well as the issue of high self-citation rates.
7.27 (declining)	The journal's official public account systematically responds to issues such as "on hold", self-citation rates, false online statements, and the transparency of the review process.
8.1 (ending)	The journal's official public account publishes the response to online questions about TITS publishing PPT. The journal's official public account systematically responds to issues such as Letters/Perspectives being indexed as Articles, replacing or adding reviewers, self-citation rates, and distributed hybrid seminars.

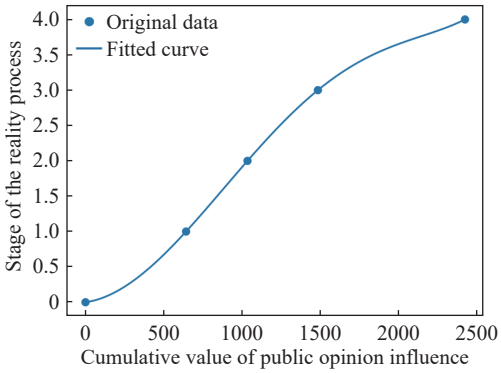


Fig. 5 Quantitative results of the public opinion influence on reality processes in the *TIV* event.

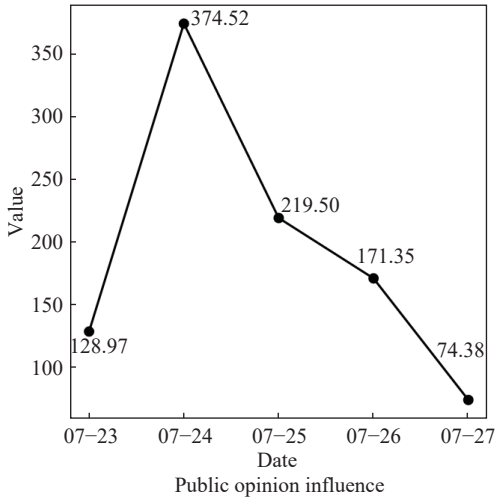
During the period 23 July to 25 July, both public opinion influence and comment volume exhibited a pattern of rising and then falling. The sentiment factor remained relatively stable. This suggests that the announcement of the status of "on hold" caused a surge in public discourse and intensified cognitive biases.

From 25 July to 27 July, public opinion influence and comment volume showed a declining trend, while the sentiment factor experienced a rapid rebound after a brief dip. This indicates that the *TIV* response during this period helped mitigate public misconceptions about the operational mechanisms of the journal.

From 25 July to 29 July, the influence of public opinion and the volume of comments both initially declined before rising again. The sentiment factor changed from negative to positive on 29 July, signaling a gradual change in public opinion. Following the additional response of *TIV* on 27 July, the discussions around the event intensified once more and public attitudes began to improve. Finally, positive sentiment supporting the journal emerged as dominant, public misunderstandings were resolved, and discussions about the *TIV* event gradually subsided.

The discussions of the *TIV* event diminished, but its damage to the journal continued. As of this writing, the journal is still "on hold". High self-citation rates led the public to question journal compliance and develop misunderstandings about editorial practices. Anonymous allegations triggered an investigation of *TIV*, negatively impacting the journal. During the past three years, *TIV* has shown a big increase in both the number of submissions and the impact factor, with the growth over the years shown in Figs. 8 and 9, respectively. The public misunderstanding is causing a huge blow to a rising journal.

The correlation analysis reveals critical temporal patterns between official responses and public opinion changes. Early single-issue responses (1 July) showed delayed public acceptance of 15 days, while concentrated multi-issue responses (25 July to 27 July) achieved public opinion reversal within 2–3 days. Multi-channel disclosure approaches demonstrated superior effectiveness, generating 56% positive sentiment shift within 48 h compared to 12% improvement from isolated responses over two weeks. The analysis establishes that comprehensive responses addressing multiple



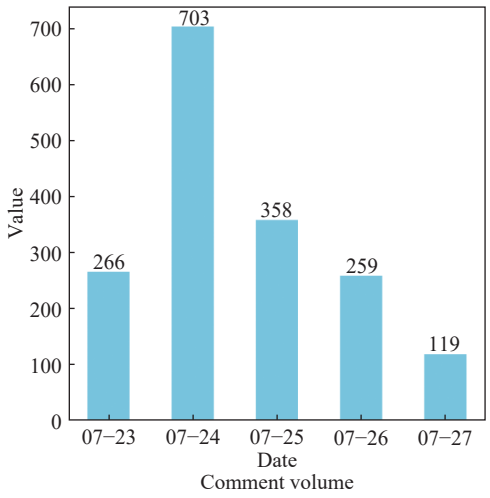
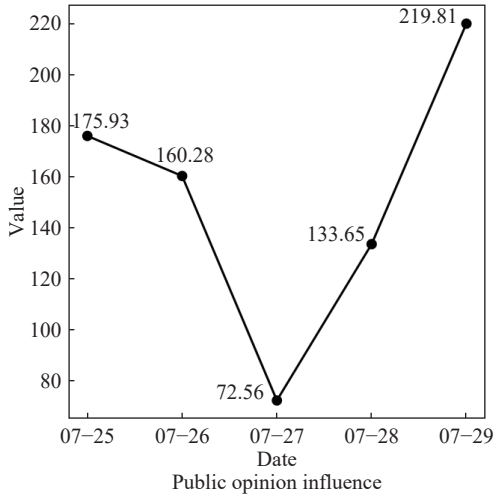


Fig. 6 25 July comprehensive analysis results in the *TIV* event.



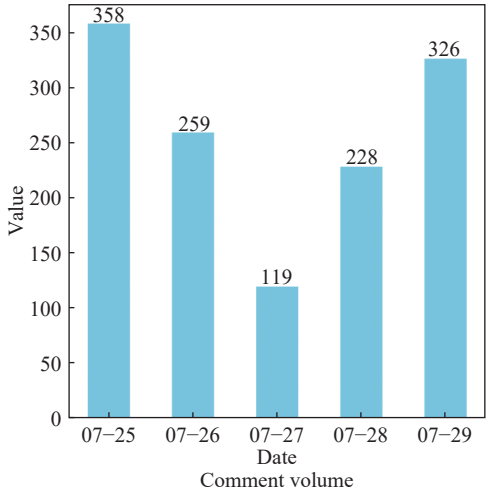


Fig. 7 27 July comprehensive analysis results in the *TIV* events.

issues simultaneously reduce public acceptance lag from 10–15 days to 3–5 days, with cumulative official engagement achieving 67% reduction in peak negative influence levels, providing quantitative evidence for optimal crisis communication timing within 48–72 h of negative sentiment emergence.

In summary, through the analysis of the *TIV* events, we have reached the following conclusions.

(1) Public misunderstandings can generate negative opinion that damages journal operations. This damage will continue for a long time after the public opinion stops.

(2) Public misunderstanding and herd mentality play a crucial role in triggering public opinion events. A rapid surge in public opinion influence can significantly accelerate the progression of reality processes.

(3) Reality process responses have the potential to correct public misunderstanding, and this corrective effect exhibits a temporal delay. As the event evolves, the lag in public acceptance tends to be shorten.

(4) During the lifecycle of a public opinion event, the trajectory of public opinion influence generally follows a normal distribution, with noticeable fluctuations occurring near the peak phase.

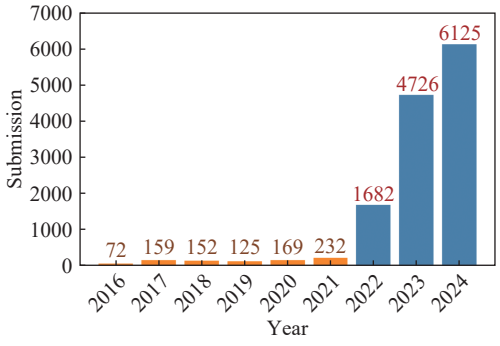


Fig. 8 Submission state of *TIV*.

E. Selecting Small Analysis Models

To identify the lightweight model most suitable for predicting public opinion, a series of experiments were conducted to evaluate the performance of six compact models in three prediction tasks: public opinion influence, sentiment, and reality process, under varying prediction lengths. The experimental results are summarized in Table 3.

The findings demonstrate that PatchTST consistently achieves strong performance in most tasks and evaluation

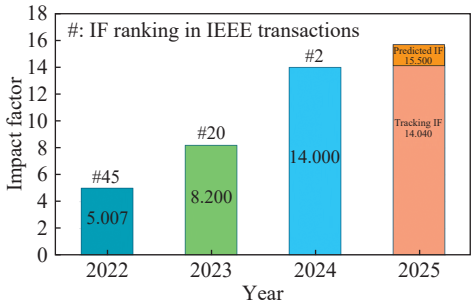


Fig. 9 Impact factor trend of *TIV*.

metrics, particularly excelling in short-term prediction scenarios. Across all models, a slight decrease in performance is observed as the prediction length increases, indicating a negative correlation between predictive accuracy and the prediction horizon. This aligns with the general challenge that long-term forecasting is inherently more difficult than short-term forecasting.

In terms of mean squared error, PatchTST shows particularly outstanding results on the POI prediction task. Its MSE remains exceptionally low, between 0.001 and 0.003 across different prediction lengths, significantly outperforming other models. For example, under prediction length 1, Linear reaches an MSE of 0.016, Transformer 0.044, and TimesNet up to 10.752.

Similarly, in the sentiment prediction task, PatchTST also delivers superior accuracy, achieving MSEs of 0.021 and 0.047 at prediction lengths 1 and 2, respectively, much lower than the corresponding values for DLinear (1.013 and 0.239).

In terms of the mean absolute error, PatchTST continues to demonstrate excellent performance. In the POI prediction task, it achieves the lowest MAE in all three prediction lengths: 0.011, 0.014, and 0.019, respectively. In the sentiment prediction task, its MAE at prediction length 1 is only 0.119, substantially outperforming Transformer (0.246) and DLinear (1.002).

Regarding the relative squared error, PatchTST also significantly surpasses other models across multiple tasks. In POI prediction, it maintains the lowest RSE values in all prediction lengths: 3.705, 4.253, and 5.185 compared to DLinear (17.632) and Transformer (145.014). These results suggest that PatchTST exhibits a smaller overall error relative to the variance of true values, highlighting its greater stability and predictive reliability.

In the sentiment prediction task, PatchTST achieves RSEs of 5.174 and 4.239 at prediction lengths 1 and 2, respectively, significantly outperforming models such as NLinear and TimesNet. This further demonstrates its strong modeling capability for sentiment-related indicators.

Although PatchTST's performance in the reality process prediction task is slightly less competitive on certain metrics compared to models like DLinear and NLinear, its overall MSE, MAE, and RSE values remain within a reasonable and acceptable range, reflecting good generalization ability.

Table 3 Result of small model comparison experiment.

Model	Type	Prediction length	MSE	MAE	RSE
Linear	POI	1	0.016	0.128	33.722
		2	0.363	0.587	155.831
		3	0.344	0.509	153.897
	Sentiment	1	0.196	0.438	61.454
		2	0.307	0.519	82.983
		3	0.370	0.597	101.469
	Real	1	0.634	0.531	1.295
		2	1.061	0.941	2.153
DLinear	POI	1	0.005	0.069	7.122
		2	0.639	0.608	209.162
		3	0.426	0.527	171.112
	Sentiment	1	0.239	0.478	63.265
		2	0.073	0.216	9.944
		3	0.032	0.194	3.457
	Real	1	0.100	0.352	0.558
		2	1.028	0.858	1.897
NLinear	POI	1	0.078	0.279	73.317
		2	0.124	0.343	92.208
		3	0.014	0.105	31.422
	Sentiment	1	0.058	0.237	8.503
		2	0.198	0.404	8.748
		3	0.125	0.303	12.404
	Real	1	0.078	0.243	0.584
		2	1.167	0.997	2.258
TimesNet	POI	1	10.752	1.962	6.854
		2	11.843	2.074	6.883
		3	13.224	2.275	7.283
	Sentiment	1	0.082	0.241	5.662
		2	0.091	0.263	6.314
		3	0.106	0.284	8.823
	Real	1	2.722	1.304	0.550
		2	66.572	6.011	1.202
Transformer	POI	1	0.044	0.158	145.010
		2	0.085	0.207	202.358
		3	0.082	0.246	228.067
	Sentiment	1	0.081	0.228	14.749
		2	0.065	0.198	23.673
		3	0.121	0.276	19.967
	Real	1	1.598	1.091	9.459
		2	0.841	0.809	6.862
PatchTST	POI	1	0.001	0.011	3.705
		2	0.002	0.014	4.253
		3	0.003	0.019	5.174
	Sentiment	1	0.021	0.119	5.174
		2	0.047	0.192	4.239
		3	0.086	0.292	3.421
	Real	1	1.809	1.328	2.812
		2	2.753	1.608	3.469

In conclusion, PatchTST achieves the best overall performance in the three evaluation metrics, MSE, MAE, and RSE, demonstrating high precision, robustness, and general applicability. These qualities make it a strong candidate for use as a lightweight model to predict the public opinion influence, sentiment factors, and the trajectory of the reality process.

F. Selecting Large Language Models

To identify the most suitable large model for public opinion analysis, we performed a comparative analysis of six large language models: ChatGPT-4.5, Claude 3.7 Sonnet, Grok 3, Deepseek-R1, Doubao, and Qwen2.5-72B-Instruct, with results presented in Table 4.

As shown in Table 4, Qwen2.5-72B-Instruct achieves the highest scores in both Rouge-L (0.2126) and BLEU (0.0223), and ranks second in Rouge-1 (0.3575), indicating that its generated analysis exhibits the highest similarity to the reference analysis. In particular, it achieves the highest precision in Rouge-L (0.2895) and a relatively balanced recall (0.1679), suggesting that its outputs are not only accurate but also concise, with minimal redundant information.

Doubao, on the other hand, achieves the highest F1 score in Rouge-L (0.4011) and the highest recall (0.2824), indicating a strong coverage of content relative to the reference. However, its precision is the lowest (0.1555), implying a tendency toward verbosity and inclusion of irrelevant information. Although its overall Rouge-L and BLEU scores (0.2005 and 0.0195, respectively) are moderate, the lower precision limits its effectiveness in generating refined analyses.

The remaining models demonstrate comparatively modest performance across all metrics, with less favorable balances of content similarity, coverage, and relevance. In conclusion, Qwen2.5-72B-Instruct stands out as the most effective model for generating public opinion analyses, offering the best overall performance in key metrics and striking an optimal balance between precision and recall. Although Doubao excels in content coverage, its low precision undermines the overall quality and conciseness of its outputs.

Table 4 Experimental result on the analysis generation performance of LLM.

Model	Rouge-1	Rouge-L	BLEU	Rouge-L		
				F1	*P*	*R*
ChatGPT-4.5	0.3113	0.1698	0.0102	0.1698	0.2222	0.1374
Claude3.7sonnet	0.2891	0.1562	0.0149	0.1562	0.1600	0.1527
Grok3	0.2844	0.1991	0.0117	0.1991	0.2625	0.1603
Deepseek-R1	0.2596	0.1635	0.0075	0.1635	0.2208	0.1298
Doubao	0.4011	0.2005	0.0195	0.2005	0.1555	0.2824
Qwen2.5-72B-Instruct	0.3575	0.2126	0.0223	0.2126	0.2895	0.1679

G. Practical Significance for Academic Publishing

The MQRF framework addresses critical needs in academic publishing crisis management. For journal editors and publishers, the framework enables: (1) early warning that detects emerging public opinion threats before they reach wide audience, (2) evidence-based response timing optimization, providing recommended response time for journals based on public opinion shifts, and (3) long-term damage assessment tools that quantify reputational impact duration, typically extending 4–6 months beyond initial discourse periods. These capabilities transform reactive crisis management into proactive reputation protection, offering measurable benefits for academic institutions worldwide. early warning that detect emerging public opinion threats before they reach wide audience.

VI. CONCLUSION

This paper investigated the *TIV* event emerged in mid 2024, leading to the journal being placed “on hold” by Clarivate. Through a comprehensive analysis of public opinion data collected from multiple Chinese social media platforms, public misunderstanding can generate damaging public opinions that continue to harm journal operations long after the event finished. It is critical to clarify the running of a typical journal to the public in order to eliminate potential misunderstandings of different types of papers. Such kind of event brings bad effect to potential researchers who want to submit papers to the journal. The proposed MQRF effectively quantified public opinion influence through four key indicators (development time, development scale, sentiment factors, and cumulative scale) and successfully established the correlation between public opinions and reality processes using linear regression analysis. The effectiveness of MQRF is confirmed by several sets of experiments, and the large language model and small model for the framework are experimentally selected. These findings provide valuable insights for understanding public opinion events in academic publishing, while the MQRF framework offers a practical tool

for predicting and responding to similar events in the future.

ACKNOWLEDGMENT

This work was supported by the Shandong Provincial Natural Science Foundation Major Basic Research Program (No. ZR2024ZD20), the Shandong Data Open Innovative Application Laboratory, and the National Natural Science Foundation of China (No. 62072469).

REFERENCES

[1] R. Widyatama and M. H. Mahbob, The potential hazards of fake accounts and buzzer behaviour on deliberative democracy, *J. Komun.: Malay. J. Commun.*, 2024, 40(1), 324–341.

[2] Y. Shen, X. Ma, H. Zhang, and J. Zhan, Fusion social network and regret theory for a consensus model with minority opinions in large-scale group decision making, *Inform. Fusion*, 2024, 112, 102548.

[3] T. McMurtry, *Dismissed: How Media Agendas and Judicial Bias Conspire to Undermine Justice*, Kimemasu, 2024.

[4] I. Abdulazeez, Z. Omale, and C. O. Florence, Implications of social media disinformation and false narratives for public opinion among Nigerian electorate, *Int. J. Sub-Saharan African Res.*, 2024, 2(4), 286–304.

[5] E. Alaslani, The transformation of Arab narrative from oral to virtual reality: A media ecology perspective, PhD dissertation, Duquesne University, USA, 2024.

[6] L. Zhang, P. Dong, L. Zhang, B. Mu, and A. Yang, A systematic literature review of crisis management in online public opinion: Evolutionary path and implications for China, *Kybernetes*, to be published.

[7] C. Yu, J. He, Q. Ma, and X. Liu, Dynamic evolution model of internet financial public opinion, *Information*, 2024, 15(8), 433.

[8] C. Bucur, B. Tudorica, J. V. Andrei, D. Dusmanescu, D. Paraschiv, and C. Teodor, Sentiment analysis of global news on environmental issues: Insights into public perception and its impact on low-carbon economy transition, *Front. Environ. Sci.*, 2024, 12, 1360304.

[9] M. Chen and W. Du, The predicting public sentiment evolution on public emergencies under deep learning and internet of things, *J. Supercomput.*, 2023, 79(6), 6452–6470.

[10] M. Gong, Analysis of internet public opinion popularity trend based on a deep neural network, *Comput. Intell. Neurosci.*, 2022, 2022(1), 9034773.

[11] M. Rodríguez-Ibánez, A. Casánez-Ventura, F. Castejón-Mateos, and P.-M. Cuenca-Jiménez, A review on sentiment analysis from social media platforms, *Expert Syst. Appl.*, 2023, 223, 119862.

[12] Z. Wang, D. Wang, Y. Xu, L. Zhou, and Y. Zhou, Intelligent computing social modeling and methodological innovations in political science in the era of large language models, *J. Chin. Polit. Sci.*, to be published.

[13] K. Sakurai, R. Togo, T. Ogawa, and M. Haseyama, LLM is knowledge graph reasoner: LLM's intuition-aware knowledge graph reasoning for cold-start sequential recommendation, in *Proc. 47th European Conference on Information Retrieval*, Lucca, Italy, 2025, 263–278.

[14] K. Kumar, T. Ashraf, O. Thawakar, R. M. Anwer, H. Cholakkal, M. Shah, M.-H. Yang, P. H. S. Torr, F. S. Khan, and S. Khan, LLM post-training: A deep dive into reasoning large language models, arXiv preprint arXiv: 2502.21321, 2025

[15] K. Miyazaki, T. Murayama, T. Uchiba, J. An, and H. Kwak, Public perception of generative AI on Twitter: An empirical study based on occupation and usage, *EPJ Data Sci.*, 2024, 13(1), 2.

[16] Q. Zhang, D. Zeng, F.-Y. Wang, R. Breiger, and J. A. Hendler, Brokers or bridges? Exploring structural holes in a crowdsourcing system, *Computer*, 2016, 49(6), 56–64.

[17] F.-Y. Wang, D. Zeng, J. A. Hendler, Q. Zhang, Z. Feng, Y. Gao, H. Wang, and G. Lai, A study of the human flesh search engine: Crowd-powered expansion of online knowledge, *Computer*, 2010, 43(8), 45–53.

[18] B. Wan, P. Wu, C. K. Yeo, and G. Li, Emotion-cognitive reasoning integrated BERT for sentiment analysis of online public opinions on emergencies, *Inform. Process. Manag.*, 2024, 61(2), 103609.

[19] R. Mao, Q. Lin, Q. Liu, G. Mengaldo, and E. Cambria, Understanding public perception towards weather disasters through the lens of metaphor, in *Proc. 33rd International Joint Conference on Artificial Intelligence*, Jeju Island, South Korea, 2024, 7394–7402.

[20] T. Thukral, A. Varshney, and V. Gaur, Intensity quantification of public opinion and emotion analysis on climate change, *Int. J. Adv. Technol. Eng. Explor.*, 2021, 8(83), 1351–1366.

[21] N. Ma, G. Yu, X. Jin, and X. Zhu, Quantified multidimensional public sentiment characteristics on social media for public opinion management: Evidence from the COVID-19 pandemic, *Front. Public Health*, 2023, 11, 1097796.

[22] S. Pinto, F. Albanese, C. O. Dorso, and P. Balenzuela, Quantifying time-dependent media agenda and public opinion by topic modeling, *Phys. A: Stat. Mech. Appl.*, 2019, 524, 614–624.

[23] Y. Qu and J. Wang, Performance and biases of large language models in public opinion simulation, *Humanit. Soc. Sci. Commun.*, 2024, 11(1), 1095.

[24] Y. Ren, T. Zhang, W. Li, D. Zhou, C. Qin, and F. Dong, Research on predicting public opinion event heat levels based on large language models, in *Proc. 2024 3rd International Conference on Computer Applications Technology (CCAT)*, Wenzhou, China, 2024, 49–54.

[25] S. Lee, T. Peng, M. H. Goldberg, S. A. Rosenthal, J. E. Kotcher, E. W. Maibach, and A. Leiserowitz, Can large language models estimate public opinion about global warming? An empirical assessment of algorithmic fidelity and bias, *PLoS Climate*, 2024, 3(8), e0000429.

[26] B. Zhang, T. Chen, X. Wang, Q. Li, W. Zhang, and F.-Y. Wang, Decoding activist public opinion in decentralized self-organized protests using LLM, *IEEE Trans. Comput. Soc. Syst.*, 2024, 11(4): 5525–5535.

[27] T. Chen, B. Zhang, X. Wang, W. Zhang, C. Hon, W. Di, L. Chen, Q. Li, and F.-Y. Wang, Public opinion evolution in cyberspace: A case analysis of Pelosi's visit to Taiwan, *IEEE Trans. Comput. Soc. Syst.*, 2024, 11(1), 319–329.

[28] Y. Liu, W. Zhang, Z. Bao, X. Chai, M. Gu, W. Jiang, Z. Zhang, Y. Tian, and F.-Y. Wang, Brain-like cognition-driven model factory for IIoT fault diagnosis by combining LLMs with small models, *IEEE Internet Things J.*, 2025, 12(16), 32296–32309.

[29] T. Fan, G. Ma, Y. Kang, H. Gu, Y. Song, L. Fan, K. Chen, and Q. Yang, FedMKT: Federated mutual knowledge transfer for large and small language models, in *Proc. 31st International Conference on Computational Linguistics*, Abu Dhabi, UAE, 2025, 243–255.

[30] H. Li, X. Zhao, D. Guo, H. Gu, Z. Zeng, Y. Han, Y. Song, L. Fan, and Q. Yang, Federated domain-specific knowledge transfer on large language models using synthetic data, arXiv preprint arXiv: 2405.14212, 2024

[31] T. Fan, Y. Kang, W. Chen, H. Gu, Y. Song, L. Fan, K. Chen, and Q. Yang, PDSS: A privacy-preserving framework for step-by-step distillation of large language models, arXiv preprint arXiv: 2406.12403, 2024

[32] M. Yin, C. Wu, Y. Wang, H. Wang, W. Guo, Y. Wang, Y. Liu, R. Tang, D. Lian, and E. Chen, Entropy law: The story behind data compression and LLM performance, arXiv preprint arXiv: 2407.06645, 2024

Shuai Guo received the MS degree from College of Computer Science and Engineering, Shandong University of Science and Technology, Qingdao, China, in 2024. He is currently pursuing the PhD degree at Qingdao Institute of Software (College of Computer Science and Technology), China University of Petroleum (East China), Qingdao, China. His research interests include social computing and process mining.

Weishan Zhang received the PhD degree in mechanical manufacturing and automation from Northwestern Polytechnical University, Xi'an, China, in 2001. Currently, he is a full professor at Qingdao Institute of Software (College of Computer Science and Technology), China University of Petroleum (East China), Qingdao, China. His research interests include big data platforms, pervasive cloud computing, service-oriented computing, and federated learning.

Baoyu Zhang received the MS degree in management science and engineering from School of Computer and Artificial Intelligence, Beijing Technology and Business University, Beijing, China, in 2022. He is currently pursuing the PhD degree in computer science at Big Data Intelligent Processing Team and Qingdao Institute of Software (College of Computer Science and Technology), China University of Petroleum (East China), China. His research interests include data mining, social computing, and nature language process.

Tao Chen received the MS degree in computer science and technology from College of Computer Science and Technology, China University of Petroleum (East China), Qingdao, China, in 2023. He is currently pursuing the PhD degree in computer science at Big Data Intelligent Processing Team and Qingdao Institute of Software (College of Computer Science and Technology), China University of Petroleum (East China), China. His research interests include data mining and social computing.

Hui Zhang received the PhD degree from State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, China, in 2020. She is currently an associate professor at School of Computer Science and Technology, Beijing Jiaotong University, China. Her research interests include computer vision, parallel vison, embodied intelligence, and intelligent transportation systems.

Mohammad Kamrul Hasan received the PhD degree in electrical and communication engineering from Faculty of Engineering, International Islamic University, Malaysia, in 2016. He is currently an associate professor and the head of Network and Communication Technology Research Laboratory, Center for Cyber Security, The National University of Malaysia, Malaysia. He is a certified professional technologist in Malaysia. He has published more than 300 indexed papers in ranked journals and conference proceedings. His research interests include cutting-edge information-centric networks, computer networks, data communication and security, mobile network and privacy protection, cyber-physical systems, the industrial IoT, transparent AI, smart grid networks and applications, and electric vehicle networks.