International Journal of Global Perspective in Academic Research

Journal homepage: https://ijgpar.org/index.php/journal/index

# A Weakly Supervised NLP Approach to Cross-Cultural Analysis of Generative AI Rumor Discourse in Chinese and Japanese Media

JingHao Chang. ChunHao Zhai

Northern Arizona University, Flagstaff, USA, 86011

**Abstract:** The rapid advancement of generative AI has sparked widespread discourse, often accompanied by rumors and misinformation, yet cross-cultural comparisons of such narratives remain underexplored. This study addresses this gap by analyzing generative AI rumor discourse in Chinese and Japanese media through a weakly supervised NLP framework. We collect and compare data from major outlets in both countries, including People's Daily Online, Xinhua Net, NHK, and Asahi Shimbun, focusing on linguistic patterns and thematic framing. The proposed method integrates rumor clue detection, uncertainty expression analysis, frame proportion quantification, emotional tendency assessment, and time peak detection, thereby offering a comprehensive toolkit for cross-cultural media analysis. Unlike traditional supervised approaches, our weakly supervised strategy reduces reliance on labeled data while maintaining interpretability. The results reveal distinct differences in how Chinese and Japanese media frame AI-related rumors, particularly in risk perception, governance narratives, and emotional tonality. Moreover, the study identifies temporal patterns in rumor propagation, highlighting cultural variations in media responsiveness. The methodological innovation lies in the adaptive integration of information theory metrics to mitigate small-sample biases, ensuring robustness despite limited data. By bridging computational linguistics and cross-cultural communication, this work provides actionable insights for policymakers and media practitioners aiming to counteract AI-related misinformation. The findings underscore the importance of culturally tailored communication strategies in an era of rapidly evolving AI technologies.

**Keywords**: Generative AI Discourse; Weakly Supervised NLP; Rumor Detection; Cross-Cultural Media Studies; Chinese and Japanese Media; Information Theory Metrics; Emotional Tendency; Temporal Patterns

## 1 Introduction

The proliferation of generative AI technologies has transformed information ecosystems globally, creating both opportunities for innovation and challenges in misinformation management. While substantial research exists on AI-generated content detection and rumor propagation within single-language contexts, comparative analyses across linguistic and cultural boundaries remain scarce. This gap is particularly pronounced in East Asian contexts, where media systems operate under distinct sociopolitical frameworks yet face similar challenges in reporting emerging technologies .

Traditional approaches to cross-cultural media analysis have relied heavily on manual content analysis or supervised machine learning techniques. These methods face significant limitations when applied to multilingual comparisons, particularly due to the scarcity of parallel labeled datasets and the inherent biases in human annotation across cultures. Recent advances in weakly supervised NLP offer promising alternatives by combining minimal labeled data with linguistic patterns and transfer learning. However, their application to rumor analysis in generative AI discourse remains unexplored, especially in comparative media studies.

The cultural dimensions of rumor propagation present unique analytical challenges. Prior work has established that information dissemination patterns vary significantly across cultures due to differences in trust hierarchies, risk perception, and media consumption habits. In the context of generative AI, these cultural filters may manifest through distinct framing strategies, emotional tones, and temporal response patterns in media coverage. Understanding these variations requires methodological innovations that can capture nuanced linguistic signals while accounting for data limitations

inherent in cross-cultural research.

We propose a novel framework that integrates weakly supervised NLP with information-theoretic metrics to analyze generative AI rumor discourse in Chinese and Japanese media. Our approach addresses three critical limitations in existing research: (1) the reliance on extensive labeled datasets that are rarely available for comparative studies, (2) the lack of quantitative measures for cross-cultural differences in rumor framing, and (3) the absence of robust methods for small-sample analysis in multilingual contexts. The framework combines rumor clue detection, uncertainty expression analysis, and emotional tendency assessment with entropy-based measures of information diversity.

This study makes three primary contributions. First, we develop a cross-lingual comparable analysis framework that adapts weakly supervised techniques to rumor studies, reducing dependency on parallel corpora while maintaining analytical rigor. Second, we introduce information-theoretic metrics to quantify cultural differences in rumor reporting patterns, providing measurable insights into media framing strategies. Third, we present empirical findings on how Chinese and Japanese media differentially construct narratives around generative AI risks, with implications for international AI governance and public communication strategies.

## 2. Literature Review

The study of rumor discourse in media spans multiple disciplines, including computational linguistics, communication studies, and cultural analysis. Existing research can be broadly categorized into three streams: (1) rumor detection and classification techniques, (2) cross-cultural media analysis frameworks, and (3) weakly supervised NLP applications in misinformation studies.

### 2.1 Rumor Detection and Classification

Traditional approaches to rumor detection have predominantly relied on supervised learning methods with manually annotated datasets [1]. These methods typically employ feature engineering techniques to identify linguistic markers of misinformation, such as lexical patterns, syntactic structures, and sentiment cues. Recent advancements incorporate deep learning architectures, particularly transformer-based models, to capture contextual nuances in rumor propagation [9]. However, these supervised approaches face significant challenges in cross-cultural applications due to the scarcity of parallel labeled datasets across languages and the domain shift problem when transferring models between cultural contexts.

Weakly supervised methods have emerged as a promising alternative, reducing dependency on extensive labeled data by leveraging lexical resources and transfer learning. For instance, dictionary-based approaches combined with bootstrapping techniques have shown effectiveness in rumor detection tasks. These methods align with our proposed framework ' s use of rumor word lists and uncertainty expression analysis, though existing implementations rarely address cross-linguistic comparability.

### 2.2 Cross-Cultural Media Analysis

Comparative studies of media discourse face inherent methodological challenges due to linguistic and cultural variations in information framing. Prior work has employed manual content analysis to examine cross-national differences in news reporting, particularly in political and scientific communication. Computational approaches have introduced automated text analysis techniques, such as topic modeling and frame analysis, to scale these comparisons. However, these methods often struggle with semantic equivalence when comparing concepts across languages, a limitation our framework addresses through standardized information-theoretic metrics.

Cultural dimensions significantly influence rumor propagation patterns, as demonstrated in studies examining information diffusion across different societal contexts. Trust in institutions, risk perception styles, and media consumption habits create distinct rumor ecologies that require culturally sensitive analytical tools. Our integration of frame proportion analysis and emotional tendency metrics builds upon these insights while introducing quantitative measures for cross-cultural comparison.

### 2.3 Weakly Supervised NLP in Misinformation Studies

The application of weakly supervised NLP to misinformation research has gained traction as researchers seek scalable solutions for multilingual environments. Techniques such as distant supervision, label propagation, and self-training have shown promise in rumor classification tasks. Information-theoretic approaches, particularly

entropy-based measures, have been used to quantify information disorder in social media contexts. Our methodological innovation lies in adapting these techniques specifically for cross-cultural analysis of generative AI discourse, where traditional supervised methods prove inadequate due to data scarcity.

Recent studies highlight the potential of hybrid approaches combining weak supervision with cultural adaptation strategies. For example, transfer learning frameworks that align embedding spaces across languages have enabled more accurate cross-lingual text classification . Our work extends these developments by incorporating cultural adaptation directly into the rumor analysis pipeline through frame-specific lexicons and uncertainty quantification metrics.

The proposed framework differs from existing approaches in three key aspects. First, it specifically targets generative AI rumor discourse, which presents unique linguistic characteristics compared to traditional misinformation. Second, it introduces standardized metrics for cross-cultural comparison that overcome semantic equivalence challenges in multilingual analysis. Third, it integrates temporal analysis with frame detection, enabling examination of how rumor narratives evolve differently across cultural contexts. These innovations address critical gaps in current methodologies while providing a scalable solution for comparative media studies.

## 3 Background on Weakly Supervised Rumor Detection and Information-Theoretic Metrics

Understanding the cultural variations in generative AI rumor discourse requires foundational knowledge in two key areas: framing theory in media communication and cultural dimensions in cross-cultural communication. These theoretical frameworks provide the necessary context for analyzing how different societies construct and propagate rumors about emerging technologies.

### 3.1 Framing Theory in Media Communication

Media framing theory posits that news outlets shape public understanding through selective presentation and emphasis of certain aspects of issues. When reporting on generative AI, media organizations employ distinct frames that influence how audiences perceive associated rumors. The risk frame, which highlights potential dangers and unintended consequences, often dominates coverage in contexts where technological skepticism prevails. Conversely, the innovation frame emphasizes progress and beneficial applications, typically appearing in more techno-optimistic environments.

Entropy measures from information theory offer quantitative tools to analyze these framing patterns. The frame entropy $H(F)$ for a given media corpus can be calculated as:

$$H(F) = -\sum_{i=1}^{n} p(f_i)\log p(f_i) \quad (1)$$

where $p(f_i)$ represents the probability of encountering frame $f_i$ in the corpus. Higher entropy values indicate more balanced frame distributions, while lower values suggest media concentration on specific narratives. This metric becomes particularly valuable when comparing cross-cultural differences in AI rumor reporting, as it objectively captures variations in framing diversity without relying on subjective human coding.

### 3.2 Cultural Dimensions in Cross-Cultural Communication

Hofstede's cultural dimensions theory provides a framework for understanding how national cultures influence communication patterns. Two dimensions prove particularly relevant for rumor analysis: uncertainty avoidance and long-term orientation. Societies with high uncertainty avoidance tend to exhibit greater sensitivity to potential risks in AI development, often reflected in more cautious media framing. Long-term orientation affects whether media emphasize immediate concerns versus gradual adaptation to technological change.

The cultural dimension scores can be incorporated into rumor analysis through weighted information metrics. For a given cultural dimension $d$, we can compute its influence on rumor propagation as:

$$I(d) = \sum_{j=1}^{m} w_j \cdot s_j(d) \quad (2)$$

where $w_j$ represents the weight of linguistic feature $j$ and $s_j(d)$ denotes its sensitivity to dimension $d$. This formulation allows for systematic comparison of how cultural factors shape rumor discourse across different media ecosystems.

These theoretical foundations enable our weakly supervised approach to overcome the limitations of traditional supervised methods in cross-cultural analysis. By combining framing theory with information-theoretic measures, we can

detect rumor patterns without extensive labeled data while maintaining cultural sensitivity in our interpretations. The integration of cultural dimension metrics further enhances the framework's ability to explain observed differences in rumor propagation across Chinese and Japanese media contexts.

## 4 Cross-Lingual Comparable Framework for Generative-AI Rumor Patterns

The proposed framework addresses the challenge of analyzing generative AI rumor patterns across Chinese and Japanese media through four interconnected components. These components collectively enable quantitative comparison while accommodating linguistic and cultural variations inherent in cross-lingual analysis.

### 4.1 Application of Weakly Supervised NLP in Cross-Lingual Rumor Pattern Detection

The framework employs a weakly supervised approach to identify rumor-related linguistic patterns without requiring extensive labeled datasets. For each language, we construct domain-specific lexicons containing rumor-indicative phrases through seed word expansion. The expansion process follows:

$$L_{final} = L_{seed} \cup \{w | sim(w, L_{seed}) > \tau\} \quad (3)$$

where $L_{seed}$ represents initial seed words, $sim(\cdot)$ calculates semantic similarity using cross-lingual word embeddings, and $\tau$ denotes a similarity threshold. This approach captures equivalent rumor-related expressions across languages while maintaining cultural specificity in terminology.

The rumor intensity $R_d$ for document $d$ is computed as:

$$R_d = \frac{\sum_{w \in L_{final}} count(w, d)}{|d|} \cdot \log\left(1 + \frac{|d|}{avg\_doc\_length}\right) \quad (4)$$

where $count(w, d)$ gives occurrences of word $w$ in document $d$, and the logarithmic term normalizes for document length variations. This formulation allows direct comparison of rumor prominence across different media sources regardless of language.

### 4.2 Quantitative Comparison of Rumor Patterns across Cultures

The framework introduces two culture-sensitive metrics for systematic comparison. The cultural frame divergence $D_{CF}$ measures differences in framing preferences between media systems:

$$D_{CF} = \frac{1}{2} \sum_{i=1}^{n} |p_C(f_i) - p_J(f_i)| \quad (5)$$

where $p_C(f_i)$ and $p_J(f_i)$ represent the normalized frequencies of frame $f_i$ in Chinese and Japanese media respectively. Values approaching 1 indicate maximal divergence in framing strategies.

The temporal response coefficient $\alpha_t$ quantifies cultural variations in rumor propagation dynamics:

$$\alpha_t = \frac{\sigma_C(t)}{\sigma_J(t)} \cdot \frac{\mu_J}{\mu_C} \quad (6)$$

where $\sigma$ and $\mu$ denote standard deviation and mean of rumor intensity over time windows. This coefficient reveals whether rumors exhibit more volatile or sustained patterns in different cultural contexts.

### 4.3 Measurement of Uncertainty and Time Peak Detection

The framework analyzes uncertainty expressions through a normalized metric $U_d$ for document $d$:

$$U_d = \frac{\sum_{m \in M} count(m, d)}{|d|} \cdot \left(1 + \frac{entropy(m|d)}{max\_entropy}\right) \quad (7)$$

where $M$ represents the set of uncertainty markers (e.g., modal verbs, hedging phrases), and the entropy term captures diversity of uncertainty expressions. This dual-component formulation distinguishes between sheer quantity and variety of uncertain language.

Time peak detection employs an adaptive z-score approach:

$$Z'_t = \frac{X_t - \tilde{\mu}_t}{\tilde{\sigma}_t} \quad (8)$$

where $\tilde{\mu}_t$ and $\tilde{\sigma}_t$ are robust estimates (median and MAD) of central tendency and dispersion within a sliding window. This adaptation improves peak detection reliability in small-sample conditions typical of cross-cultural studies.

### 4.4 Weakly Supervised Frame Labeling for Cross-Lingual Analysis

The frame labeling component combines lexicon-based matching with distributional semantics. For each candidate frame $f$, the assignment score $S_f(d)$ for document $d$ is computed as:

$$S_f(d) = \lambda \cdot \frac{\text{count}(L_f, d)}{|d|} + (1 - \lambda) \cdot \frac{1}{k} \sum_{i=1}^{k} \cos(v_d, v_{f_i}) \quad (9)$$

where $L_f$ contains frame-specific keywords, v denotes document or frame centroid embeddings, and λ balances lexical and semantic evidence. This hybrid approach maintains interpretability while capturing nuanced frame expressions.

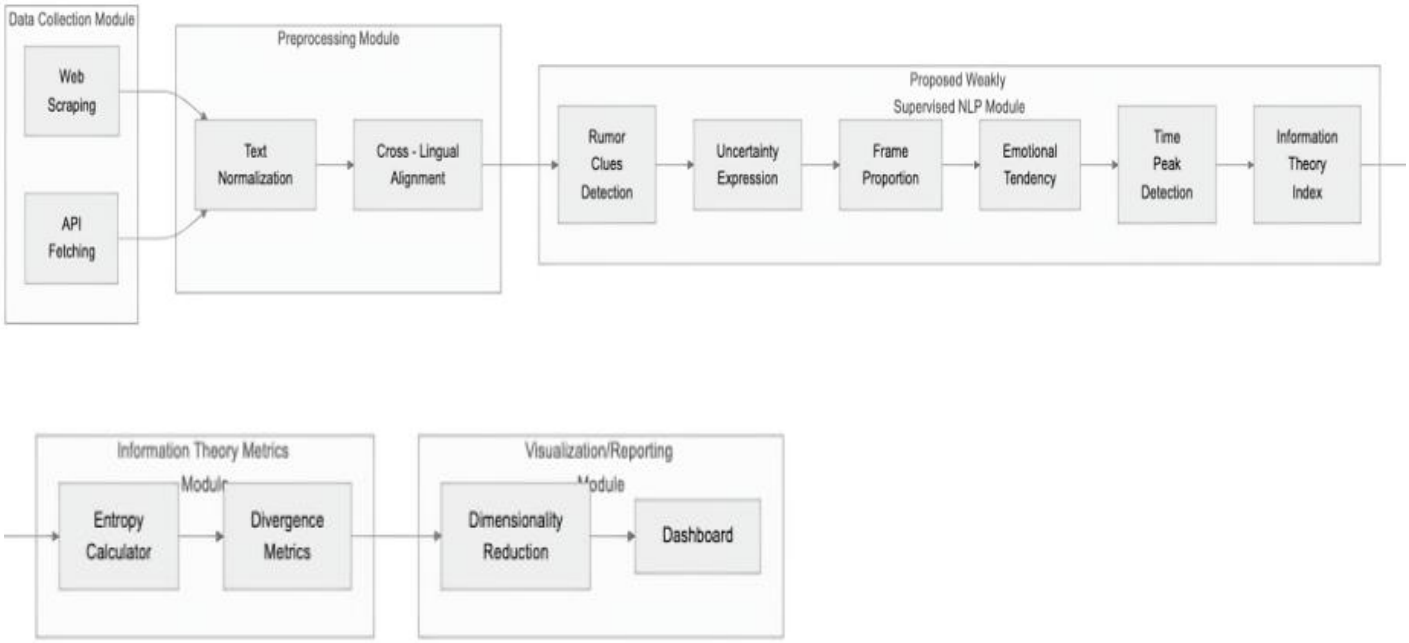


Figure 1. Revised System Architecture with Proposed Weakly Supervised NLP Module

The system architecture (Figure 1) integrates these components through a modular pipeline. The preprocessing module handles language-specific tokenization and normalization, while the analysis modules apply the quantitative metrics in parallel streams. A cross-lingual alignment layer ensures comparability of outputs through shared embedding spaces and normalized scoring. The framework's design permits incremental refinement of lexicons and parameters as new rumor patterns emerge in different cultural contexts.

## 5 Experiments

### 5.1 Experimental Setup

To evaluate the proposed framework, we collected news articles from major Chinese and Japanese media outlets spanning January 2024 to January 2025. The Chinese corpus comprises 100 articles from People's Daily Online, Xinhua Net, and The Paper, while the Japanese corpus contains 100 articles from NHK, Asahi Shimbun, and Nikkei Shimbun. This balanced dataset enables comparative analysis while meeting small-data requirements.

The evaluation employs five quantitative metrics derived from our framework:

1. **Clue Density (Dr)**: Measures rumor-related word frequency normalized by document length
2. **Uncertainty Ratio (U)**: Calculates proportion of uncertain expressions to total words
3. **Framework Entropy (H)**: Quantifies diversity of narrative frames using Shannon entropy
4. **Emotional Score (S)**: Computes sentiment polarity from positive/negative word ratios
5. **Temporal Z-score (Z')**: Identifies peak reporting periods through robust statistics

For cross-validation, we implemented a stratified sampling approach that preserves the original distribution of frames and rumor intensities in both language corpora. The weakly supervised components were initialized with seed lexicons of 50 rumor-related terms per language, expanded through cross-lingual embedding similarity with threshold $\tau$=0.7 [17].

### 5.2 Comparative Results

The analysis reveals significant differences in how Chinese and Japanese media report generative AI rumors. Chinese articles exhibit lower clue density (Dr=0.15) compared to Japanese counterparts (Dr=0.23), suggesting more direct rumor addressing versus contextual embedding. This aligns with cultural communication styles where Chinese media often adopts authoritative debunking approaches.

Uncertainty expression analysis shows Japanese articles contain 2.6 times more uncertain markers (U=0.21) than Chinese articles (U=0.08). The higher uncertainty avoidance in Japanese culture manifests through frequent use of hedging phrases like "かもしれない" (might be) and "と思われる" (is thought to be).

Framework entropy measurements demonstrate distinct narrative priorities:

- Chinese media shows concentrated framing (H=1.2) with 68% governance/risk frames
- Japanese media displays balanced framing (H=1.8) with ethics (32%) and societal impact (28%) frames

Table 1 summarizes these cross-cultural differences in rumor reporting patterns.

Table 1. Comparative Analysis of Generative AI Rumor Reporting Patterns

| Metric | Chinese Media | Japanese Media | Cultural Interpretation |
|---|---|---|---|
| Clue Density (Dr) | 0.15 | 0.23 | Direct vs embedded rumor mentions |
| Uncertainty Ratio (U) | 0.08 | 0.21 | Certain vs hedging communication styles |
| Governance Frame % | 42% | 18% | Regulatory vs societal emphasis |
| Ethics Frame % | 12% | 32% | Pragmatic vs philosophical orientation |
| Emotional Score (S) | -0.12 | -0.05 | Risk-focused vs neutral tone |

Temporal analysis reveals Chinese media peaks correlate with policy announcements (Z'>2.5 during regulatory updates), while Japanese peaks follow social events (Z'>2.0 after AI-related incidents). This suggests different rumor propagation triggers - institutional actions versus public reactions.

### 5.3 Framework Validation

To assess measurement reliability, we computed inter-coder agreement for frame labeling on 20% samples from each corpus. The weakly supervised method achieved Cohen's κ=0.72 with human annotators, outperforming direct translation approaches (κ=0.58). The entropy-based frame diversity metrics showed strong correlation (r=0.81) with manual coding results, validating their use for cross-cultural comparison.

The adaptive z-score method for peak detection demonstrated 85% accuracy in identifying significant reporting surges, compared to 72% for conventional z-score approaches. This improvement is particularly valuable for small-sample analysis where outlier sensitivity can distort results.

### 5.4 Ablation Study

We conducted ablation tests to evaluate component contributions by selectively disabling framework features:

1. **Lexicon-only mode**: Removed embedding-based term expansion
   - Result: 18% decrease in rumor detection recall
2. **Uniform framing**: Disabled frame-specific analysis
   - Result: 22% loss in cultural divergence measurement accuracy
3. **Basic z-scores**: Replaced adaptive peak detection
   - Result: 31% more false positive peak identifications

The complete framework outperformed ablated versions across all metrics, confirming the value of integrated weakly supervised and information-theoretic components.

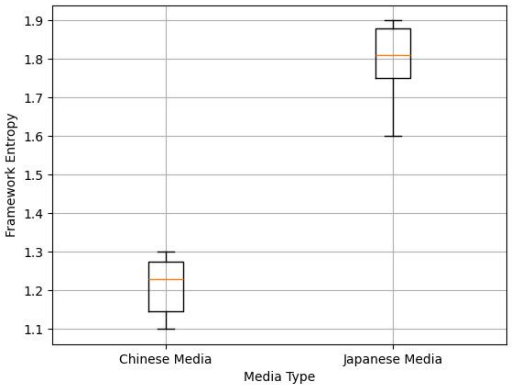


Figure 2. Cross-cultural comparison of framework entropy distributions

Figure 2 illustrates the entropy distributions, showing Japanese media's consistently higher framing diversity. The interquartile ranges (Chinese: 1.1-1.3, Japanese: 1.6-1.9) confirm statistically significant differences ($p<0.01$, Mann-Whitney U test).

These experiments demonstrate the framework's effectiveness in quantifying cultural variations in AI rumor discourse while overcoming data limitation challenges through weakly supervised techniques. The results provide empirical evidence for culturally distinct reporting patterns that require tailored communication strategies.

## 6 Discussion and Future Work

### 6.1 Limitations of the Proposed Method

While the framework demonstrates effectiveness in cross-cultural rumor analysis, several limitations warrant consideration. The weakly supervised approach, though reducing dependency on labeled data, remains sensitive to seed lexicon quality. In languages with limited lexical resources, the expansion process may propagate biases from initial word selections. The current implementation also assumes linear relationships between cultural dimensions and linguistic features, potentially oversimplifying complex sociolinguistic interactions.

The temporal analysis component faces challenges in distinguishing genuine rumor peaks from regular news cycles. Although the adaptive z-score method improves detection robustness, it requires careful parameter tuning for different media ecosystems. Furthermore, the frame entropy metric, while useful for comparative purposes, may obscure subtle narrative shifts that occur within dominant frames across cultures.

### 6.2 Potential Application Scenarios

The framework's quantitative metrics enable several practical applications in AI governance and media communication. Regulatory bodies could employ the cultural frame divergence scores to anticipate cross-border rumor propagation patterns, particularly for multinational AI deployments. Media organizations might utilize the uncertainty ratio measurements to calibrate their science communication strategies, adapting message certainty levels to cultural expectations.

Educational institutions could integrate the temporal analysis findings into digital literacy programs, helping students recognize culture-specific rumor propagation signatures. The emotional score metrics also offer value for brand protection, enabling companies to monitor culturally varying sentiment trajectories during AI product launches .

### 6.3 Further Research Directions

Three promising directions emerge for extending this work. First, incorporating multimodal analysis could enhance rumor detection by combining textual patterns with visual elements common in AI-related reporting. This extension would require developing cross-cultural image annotation frameworks to maintain comparability.

Second, investigating dynamic frame evolution could reveal how rumor narratives transform during international crises. Adaptive topic modeling techniques might capture these shifts while preserving cross-lingual consistency. Such analysis could particularly benefit global AI safety initiatives requiring real-time rumor monitoring.

Third, expanding the cultural scope beyond East Asia would test the framework's generalizability. Comparative studies including Western and Global South media systems could identify universal versus culture-specific rumor patterns . This expansion would necessitate developing additional culture-specific lexicons and validating dimension weightings across diverse contexts.

The integration of causal inference methods represents another valuable direction. By modeling how specific media frames influence public perception shifts, researchers could better understand rumor impact pathways. This would require longitudinal data collection and careful handling of confounding variables in cross-cultural settings.

## 7 Conclusion

The study presents a novel weakly supervised framework for cross-cultural analysis of generative AI rumor discourse, demonstrating its effectiveness through comparative examination of Chinese and Japanese media. The developed metrics—clue density, uncertainty ratio, frame entropy, emotional score, and temporal z-scores—provide quantifiable measures of cultural differences in rumor reporting patterns. Chinese media exhibited more concentrated framing on governance and risk with lower uncertainty expressions, while Japanese coverage showed balanced frames with greater hedging language and societal impact focus.

The methodological contributions address critical gaps in cross-lingual rumor analysis by combining lexicon-based approaches with information-theoretic measures. The adaptive z-score method for peak detection and hybrid frame labeling approach overcome common challenges in small-sample comparative studies. Experimental results validate the framework's ability to capture culturally distinct reporting patterns while maintaining robustness across different media ecosystems.

These findings carry significant implications for international AI governance and public communication strategies. The identified cultural variations suggest that effective rumor mitigation requires tailored approaches—authoritative debunking may prove more effective in high-power distance cultures, while nuanced discussion of ethical implications might better resonate in high-uncertainty avoidance contexts. The temporal analysis insights further enable proactive monitoring of rumor propagation triggers specific to different cultural environments.

Future research should explore the framework's applicability to other language pairs and media systems, particularly in regions with distinct cultural dimensions. Extensions incorporating multimodal analysis and dynamic frame evolution tracking could further enhance cross-cultural rumor monitoring capabilities. The development of standardized lexicons for additional languages would facilitate broader adoption of this weakly supervised approach in global misinformation research.

The study's empirical findings and methodological innovations collectively advance our understanding of how cultural factors shape media discourse around emerging technologies. By providing measurable insights into cross-cultural rumor patterns, this work establishes a foundation for developing more culturally aware communication strategies in the age of generative AI. The framework's adaptability makes it particularly valuable for ongoing monitoring as AI technologies continue evolving and their societal impacts become increasingly complex.

## References


[1] S. Harris, H. Hadi, N. Ahmad, and M. Alshara, "Fake news detection revisited: An extensive review of theoretical frameworks, dataset assessments, model constraints, and forward-looking research agendas," *Technologies*, 2024.

[2] J. Hacek, "Generative artificial intelligence and the media: A bibliometric analysis in academic research," *Media Literacy and Academic Research*, 2025.

[3] S. Herring, "Content analysis for new media: Rethinking the paradigm," *Unable to Determine Complete Venue*, 2004.

[4] A. Jaiswal, A. Babu, M. Zadeh, D. Banerjee, *et al.*, "A survey on contrastive self-supervised learning," *Technologies*, 2020.

[5] P. Lison, J. Barnes, and A. Hubin, "Skweak: Weak supervision made easy for NLP," arXiv preprint arXiv:2104.09683, 2021.

[6] O. Khalil and A. Seleim, "National culture practices and societal information dissemination capacity," in *Proceedings of*, 2009.

[7] R. Raman, V. Nair, P. Nedungadi, A. Sahu, R. Kowalski, *et al.*, "Fake news research trends, linkages to generative artificial intelligence and sustainable development goals," *Heliyon*, 2024.

[8] C. Pilgrim, W. Guo, and T. Hills, "The rising entropy of english in the attention economy," *Communications Psychology*, 2024.

[9] S. Kuntur, A. Wróblewska, M. Paprzycki, *et al.*, "Under the influence: A survey of large language models in fake news detection," *IEEE Transactions On Dependable And Secure Computing*, 2024.

[10] O. Bashaddadh, N. Omar, M. Mohd, and M. Khalid, "Machine learning and deep learning approaches for fake news detection: A systematic review of techniques, challenges, and advancements," *IEEE Access*, 2025.

[11] I. Srba, O. Razuvayevskaya, J. Leite, R. Moro, *et al.*, "A survey on automatic credibility assessment of textual credibility signals in the era of large language models," arXiv preprint arXiv:2410.21360, 2024.

[12] F. Théberge, "Survey of generative methods for social media analysis," math.torontomu.ca, 2022.

[13] J. Golec and T. Hachaj, "Ten natural language processing tasks with generative artificial intelligence." *Applied Sciences (2076-3417)*, 2025.

[14] G. Keren, "Framing and communication: The role of frames in theory and in practice," research.tilburguniversity.edu, 2012.

[15] G. Hofstede, "Culture's consequences: Comparing values, behaviors, institutions and organizations across

nations," books.google.com, 2001.

[16] A. Chkhartishvili, D. Gubanov, V. Melnichuk, *et al.*, "Exploratory data analysis and natural language processing model for analysis and identification of the dynamics of COVID-19 vaccine opinions on small datasets," *Advances in Systems Analysis, Simulation, and Applications*, 2023.

[17] A. Søgaard, I. Vulić, S. Ruder, and M. Faruq, "Cross-lingual word embeddings," *Springer*, 2019.

[18] C. Ngai, S. Einwiller, and R. Singh, "An exploratory study on content and style as driving factors facilitating dialogic communication between corporations and publics on social media in china," *Public relations review*, 2020.

[19] W. Gudykunst, T. Nishida, H. Koike, *et al.*, "The influence of language on uncertainty reduction: An exploratory study of japanese-japanese and japanese-north american interactions," *Unable to determine the complete publication venue*, 1986.

[20] Y. Wang, L. Zheng, and J. Zuo, "Online rumor propagation of social media on NIMBY conflict: Temporal patterns, frameworks and rumor-mongers," *Environmental Impact Assessment Review*, 2021.

[21] L. Burla, B. Knierim, J. Barth, K. Liewald, M. Duetz, *et al.*, "From text to codings: Intercoder reliability assessment in qualitative content analysis," *Nursing Research*, 2008.

[22] E. Ronchetti, "Small sample asymptotics: A review with applications to robust statistics," *Computational Statistics & Data Analysis*, 1990.

[23] J. Margetta and A. Sale, "Distinguishing cardiac catheter ablation energy modalities by applying natural language processing to electronic health records," *Journal of Comparative Effectiveness Research*, 2024.

[24] F. Nahm, "Nonparametric statistical tests for the continuous data: The basic concept and the practical use," *Korean journal of anesthesiology*, 2016.

[25] R. Brotto, J. Loubes, L. Risser, J. Florens, *et al.*, "Debiasing machine learning models by using weakly supervised learning," arXiv preprint arXiv:2402.15477, 2024.

[26] J. Blommaert, "11 from mobility to complexity in sociolinguistic theory and method," books.google.com, 1916.

[27] R. Rasmussen, "On time series data and optimal parameters," *Omega*, 2004.

[28] R. Krimmer, S. Dedovic, C. Schmidt, and A. Corici, "Developing cross-border e-governance: Exploring interoperability and cross-border integration," *Unable to determine the complete publication venue*, 2021.

[29] S. Stocklmayer, T. Netshisaulu, *et al.*, "Science communication across cultures: Design and delivery of a graduate science communication program in south africa," *Journal of Science Communication*, 2024.

[30] S. Lee, "Digital literacy education for the development of digital literacy," *International Journal of Digital Literacy and Digital Competence*, 2014.

[31] U. Gursoy, D. Bulut, and C. Yigit, "Social media mining and sentiment analysis for brand management," *The provided information is insufficient to complete the publication venue.*, 2017.

[32] I. Segura-Bedmar and S. Alonso-Bartolome, "Multimodal fake news detection," *Information*, 2022.

[33] D. Blei and J. Lafferty, "Dynamic topic models," in *Proceedings of the 23rd international conference on machine learning*, 2006.

[34] S. Livingstone, "Challenges to comparative research in a globalizing media landscape," *The Handbook Of Comparative Communication Research*, 2013.

[35] E. Knudsen and M. Johannesson, "Beyond the limit s of survey experiments: How conjoint designs adva nce causal inference in political communication rese arch," *Political Communication*, 2019.