Symbolic Framing Meets Diffusion Modeling: How Tariff Narratives Spread on Social Media

Nitin Agarwal[1,2,*], Bishwa Prakash Subedi[1]

[1]COSMOS Research Center, UA - Little Rock, Arkansas, USA
[2]International Computer Science Institute, University of California, Berkeley, USA

Abstract

The 2025 U.S. tariff expansion triggered widespread public discussion on social media platforms. This study examines how social, cultural, economic, and political (SCEP) symbols shape the spread of tariff-related narratives on X/Twitter, Instagram, and TikTok. We combined large language model–based clustering to identify dominant narratives, vision–language models to detect symbolic cues, and an epidemiological diffusion model (SEIZ) to analyze how narratives propagated over time. Our findings show that within the dominant clusters, economic symbols were more prominent in Instagram narratives (that focused on markets and finance), while political symbols played a stronger role on X/Twitter and TikTok, where tariffs were framed through trade negotiations and geopolitical tensions. We also find that narratives containing multiple symbolic cues exhibited higher transmission rates (β) on Instagram and TikTok, while persistence ($\mathcal{R}_0$) diverged across platforms and categories. These results highlight how symbolic framing influences the diffusion of policy-related discussions online.

Keywords

Narrative Diffusion, Large Language Models, Epidemiological Modeling, Semiotics

1. Introduction

At the beginning of 2025, the U.S. government raised tariffs on a broad range of goods, products, and services, with increases ranging from 10% to 25% [1]. This policy shift quickly became a central topic of public debate, influencing consumer prices and business supply chains. As news of the tariff increases spread, social media platforms became active spaces where individuals, policymakers, and interest groups shared opinions, voiced concerns, and shaped competing narratives.

Social media platforms play a key role in turning policy actions into stories. On platforms such as X/Twitter, Instagram, and TikTok, users react to tariff news by sharing posts, images, and videos that express opinions and emotions. They often rely on symbols, such as images or visual cues, to quickly communicate meaning. These symbols help simplify complex issues and make them easier to understand and share. We define symbols as textual or visual elements that signal social, cultural, economic, or political meaning within a post. These may include keywords, imagery, emojis, gestures, objects, or references that evoke broader collective interpretations.

We examine how social, cultural, economic, and political (SCEP) symbols shape how tariff-related discussions spread on social media. We used the term narrative to refer how users interpret and discuss events across posts, rather than single individual posts or isolated topics. Symbols play an important role in this process because they allow users to express ideas and positions without lengthy explanations. When symbols are repeated and shared across posts, they may help certain narratives gain visibility and influence.

This paper makes two main contributions. First, it provides an analysis of how SCEP symbols shape the spread of tariff-related discussions across different social media platforms. Second, it examines whether posts that combine multiple symbolic elements tend to spread more widely than those with fewer symbols.

In: R. Campos, A. Jorge, A. Jatowt, S. Bhatia, M. Litvak (eds.): Proceedings of the Text2Story'26 Workshop, Delft (The Netherlands), 29-March-2026
nxagarwal@ualr.edu (N. Agarwal); bsubedi@ualr.edu (B. P. Subedi)
0000-0002-5612-4753 (N. Agarwal)

RQ1. How do SCEP symbolic cues influence the diffusion of tariff-related discussions across different social media platforms?

RQ2. Does the presence of multiple symbolic cues within a post contribute to faster diffusion of tariff-related discussions on social media?

The remainder of this paper is organized as follows. Section 2 reviews related work on symbolic communication, and information diffusion. Section 3 describes the data collection process and methodological framework. Section 4 presents the results of the narrative and diffusion analysis. Section 5 concludes with a discussion of limitations and directions for future research.

2. Related Work

Symbolic Interaction Theory emphasizes that human action is guided by the meanings people assign to symbols, and that these meanings emerge from social interaction and interpretive processes [2, 3]. These shared interpretations shape behavior and continue to evolve through ongoing social experiences. Symbols will continue to take on new interpretations with the change in the social environment because the meanings are never fixed [4]. The variety of symbols people use has grown significantly with the rise of digital platforms. The new forms of symbolic expression are acted out by hashtags, emojis, and memes, which help people to communicate their ideas, emotions, and identities in quick and culturally recognizable ways. People from different cultural backgrounds and communities are now tied by modern semiotic tools, i.e., social media platforms, highlighting how symbolic interaction continues to adapt within present communication systems [5].

Media Richness Theory [6] argues that communication channels differ in how effectively they convey rich information. Our work builds on this theory by showing that symbols in digital content act as cognitive shortcuts that increase informational richness, helping users make sense of complex messages more efficiently within algorithm-driven platforms.

Recent advances in Large Language Models (LLMs) make it possible to study symbolic content at scale. Prior work has shown that vision-language models can identify visual and textual features in social media content with high accuracy [7, 8]. Other studies have demonstrated that symbols embedded in images influence user engagement and emotional response, particularly in short-form video platforms such as TikTok [9]. These findings suggest that automated models can support large-scale analysis of symbolic storytelling in online discourse.

Epidemiological models, originally developed to study disease spread, have also been applied to information diffusion on social media. Models such as SIR and SEIZ describe how individuals encounter, adopt, question, or stop sharing information over time [10, 11, 12]. These models are well suited for studying how stories circulate, persist, or fade within online communities. More recent work combines diffusion modeling with machine learning to improve prediction and interpretation of online spread patterns [13].

Prior research has examined how narratives and information spread on social media using computational approaches such as diffusion modeling and large-scale analysis of online discussions. Studies on information diffusion and rumor propagation have used epidemiological models to understand how stories circulate within online communities [11, 12]. Building on this line of work, our study focuses on symbolic cues embedded in social media posts and examines how these cues influence the diffusion of tariff-related discussions across platforms.

3. Methodology

This section describes how we collected the data, identified dominant discussions, detected symbolic cues, and analyzed how content spread across social media platforms. Figure 1 provides an overview of the overall workflow.

We first collected and pre-processed data from X/Twitter, Instagram, and TikTok related to the 2025 U.S. tariff expansion. We then used a Large Language Model (LLM) to group posts into clusters and

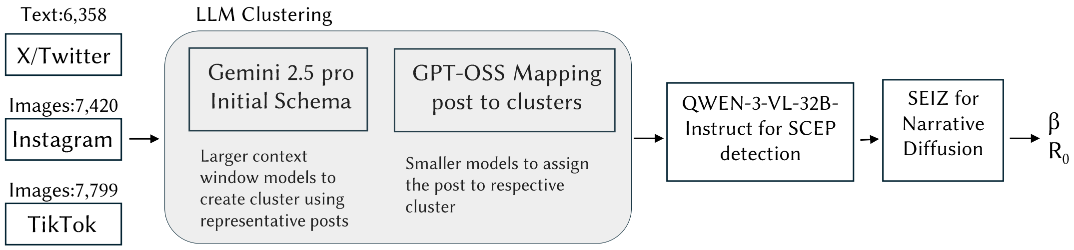

Figure 1: Workflow for clustering social-media content and modeling narrative spread using LLM-based clustering, SCEP symbol detection, and SEIZ epidemiological framework.

identify the dominant discussion on each platform. From the dominant cluster, we detected SCEP symbols. Finally, we applied an epidemiological diffusion model to study how posts spread across symbolic categories.

3.1. Dataset Collection and Preprocessing

We analyzed public discussions related to the 2025 U.S. tariff expansion across X/Twitter, Instagram, and TikTok. Data were collected using the Apify platform using the set of tariff-related keywords (e.g., TradeWar, TrumpTariffs, TariffImpact, TradeWars, TradeNegotiations, USChinaTrade) The full list of keywords is provided in appendix.. These keywords were selected by reviewing online news coverage of tariff-related discussions [1, 14]. The data were collected from January 2025 to May 2025. The collected dataset reflects globally posted public content containing the selected keyword rather than a discourse tied to a specific country, language, or user group, as we didn't target any specific language, geography, or user groups. During preprocessing, duplicate entities were removed using platform-specific post identifiers and URL matching. Posts that shared the same identifier were treated as duplicates, while the reposts, retweets, and reshares with distinct post identifiers were retained as separate observations. Table 1 shows the total number of collected posts for each platform, along with the number of unique user accounts contributing to these posts.

Our study analyzes public social media content, including text, images, and video frames from X/Twitter, Instagram, and TikTok, respectively. The data is collected in compliance with the platforms' terms. The study is conducted with strict adherence to human subject research under the supervision of the University of Arkansas' Institutional Review Board and U.S. Army's Human Research Protocol Office (HRPO) to maintain privacy and security of the individuals that may be referred or present in the social media data. No personally identifiable information (PII) is retained. All unique identifiers (post IDs, etc.) have been replaced with arbitrary sequence of characters. The data is stored securely, with physical access limited. Digital access is granted to project members authorized by the principal investigator of the project, i.e., the first author of this study. Only macro-level analysis is retained (i.e., presence/absence of a symbol). No images are retained.

Table 1
Dataset distribution across platforms, along with unique user accounts

Platform	Total unique posts	Number of unique accounts
X/Twitter	6,358	4,667
Instagram	7,420	3,403
TikTok	7,799	4,502

3.2. Identifying Dominant Discussions Using LLMs

Identifying the dominant narrative allows us to understand the main ways tariff-related issues were framed on each platform and to compare how these framings differed across X/Twitter, Instagram, and TikTok. This also provides insight into broader public perceptions surrounding the tariff expansion.

To identify dominant discussions, we used an LLM-based clustering approach. This approach was chosen because LLMs are well-suited to handling informal language, mixed modalities, and domain-specific expressions commonly found in social media content, which are often challenging for traditional topic modeling methods [15]. We first used the Gemini 2.5 Pro model to generate an initial clustering schema. Its large context window allowed us to process large batches of posts and derive high-level discussion categories without losing important context [16]. Once the schema was developed, clustering became a comparatively simple task. So, we switched to the GPT-OSS model, which, despite having a smaller context window, still possesses good reasoning ability and is well-suited for applying the predefined schema to the individual records [17]. Figure 2 shows the prompt used for schema generation and clustering.

We used a two-stage clustering pipeline across X/Twitter, Instagram, and TikTok. First, for each platform, we generated an initial cluster schema using Gemini 2.5 Pro by providing a stratified sample of posts up to 500 posts in a single multimodal prompt (Text-only posts for X/Twitter; Images + Captions for Instagram and TikTok). The prompt incorporated a domain-context string (referred to in Figure 2 as domain_hint), supplying background context on tariff policies and common debate themes to improve the interpretation of short and informal social-media content; the full domain_hint text is provided in the Appendix. The model was instructed to generate up to 20 mutually exclusive clusters, and return a structured JSON schema containing cluster ID, topic, name, description, and inclusion criteria. This schema was used as a fixed taxonomy for subsequent assignment.

Second, posts were assigned to clusters in batches using the GPT-OSS model. For each batch (Tweets - 50; Images - 20), the model received (a) the current taxonomy (cluster IDs, topic names, and inclusion criteria), (b) the domain-context string, (c) a small set of recent labeled examples to encourage consistency across batches, and (d) the content of each post (Text-only posts for X/Twitter; Images + Captions for Instagram and TikTok). The model was instructed to assign each post to only one existing cluster based on inclusion criteria. If it finds a new post that doesn't fit into the existing clusters, then it can propose a new cluster, which will be appended to the cluster taxonomy and passed to the subsequent batches. As the assignment was performed sequentially, each posts were mapped to mutually exclusive clusters in the evolving unified schema by the end of the operation. After clustering, we evaluated cluster quality using embedding-based intra-cluster and inter-cluster cosine distances. Across the three platforms, intra-cluster distances ranged from 0.266 to 0.315, while inter-cluster distances ranged from 0.085 to 0.161, indicating moderate cluster compactness and reasonable separation. Such geometric overlap is expected for discourse-level clusters, where distinctions often rely on stance, framing, or actor focus rather than completely distinct vocabulary. To further verify cluster coherence, we manually inspected 20 samples from each cluster and confirmed that the assigned posts were consistent with the cluster definitions and inclusion criteria.

Schema Prompt

You are a research assistant creating a taxonomy for Instagram posts about 'Trump Tariffs 2025'.", f"DOMAIN CONTEXT: {CONFIG['domain_hint']}", "Analyze the following sample images and captions.", "Create a list of 10-20 mutually exclusive clusters.", "Return JSON ONLY

System Prompt

"You are an AI analyzing Instagram posts about US Tariffs. " "Task: Assign the post to an existing cluster OR propose a new one if it fits nowhere."
f"Context:{CONFIG['domain_hint']}" f"Current Schema:\n{schema_json}\"
"Format: Cluster_id : "text ; Cluster_topic : "text"

Figure 2: LLM Prompt for Clustering

3.3. Symbol Detection

We used the QWEN-3-VL-32B-Instruct [18] model to identify four symbolic cues, S, C, E, and P, embedded in text, image, or both from the dominant posts.We selected this model because it supports multimodal (text + image) input and offers superior visual perception and reasoning alongside strong text comprehension, making it well-suited for extracting symbolic cues from social media posts that include both captions and images [18]. Furthermore, it provides structured outputs in response to guided JSON prompts, enabling consistent SCEP classification across platforms. Each post received binary labels (1/0) to specify whether the symbolic cues were present or not in a JSON output format: Social: 0/1, Cultural:0/1, Economic: 0/1, Political: 0/1.

Since text remains the most common form of content shared on X/Twitter [19], we applied our symbol-detection method directly to tweet text rather than images. For TikTok posts, we analyzed the initial video frame, and for Instagram posts, we used images. We employed a structured prompting strategy combining role specification, in-context learning, and constrained JSON output formatting. The system prompt defined the model's role as a symbolic content analyst specialized in tariff-related discourse. A small number of labeled examples (few-shot in-context learning) were included to guide classification boundaries for SCEP categories. The model was instructed to return structured binary outputs (0/1) in JSON format to ensure consistency across platforms. During development, we iteratively tested variations of zero-shot and few-shot prompts and observed that few-shot prompting improved category consistency, particularly for cultural and social symbols, which are more context-dependent. We did not employ chain-of-thought prompting or complex reasoning frameworks, as preliminary testing indicated that structured classification with constrained outputs produced more stable and interpretable results for this task. Figure 3 shows the designed prompt used to analyze the tweets and it was modified slightly to analyze the images for the other two platforms.

To evaluate how our model performed in symbol detection, we tested our model on labelled datasets of images and texts, and the model demonstrated strong symbol detection performance across most categories, as shown in Table 2. The images for social symbols were analysed from the USED dataset [20], whereas cultural symbols were evaluated using datasets containing religious imagery, pilgrimage sites, and Indian temples [21, 22, 23]. Political symbols were assessed using images of political parties [24], while for the economic symbols, we used dataset containing pictures of coins and currencies [25, 26].

Table 2
Dataset and Performance of the model for Symbol Detection on Images and Text

Symbols	Datasets (Images)	QWEN-3-VL-32B-Instruct (Images)	Datasets (Texts)	QWEN-3-VL-32B-Instruct (Text)
Social	USED	100%	Social-Chem-101	91%
Cultural	Religious, Pilgrims, Temple	99.09%	HateXplain	72%
Economic	Coins, Currencies	87.72%	twitter-financial-news-sentiment	95%
Political	Political Parties	96.39%	Political tweets	82%

For text-based evaluation, the social category was used from the test split of the Social-Chem-101 dataset [27], which includes everyday scenarios and examples of norm-breaking behaviors. This makes it well-suited for identifying social norms, social issues, and community-related cues. For the cultural category, we analyzed a sampled subset of the HateXplain dataset that contains tweets targeting cultural or ethnic identities, providing clear cultural and ethnic references [28]. For the political category, we constructed a sample of 10,000 tweets by combining datasets from the 2020 U.S. election, Indian politics, and the 2025 Philippine election. These datasets cover a broad range of political discussions and include explicit mentions of political actors, events, ideologies, and public issues, making them suitable for assessing political symbol detection [29, 30, 31]. Similarly, the economic symbols were

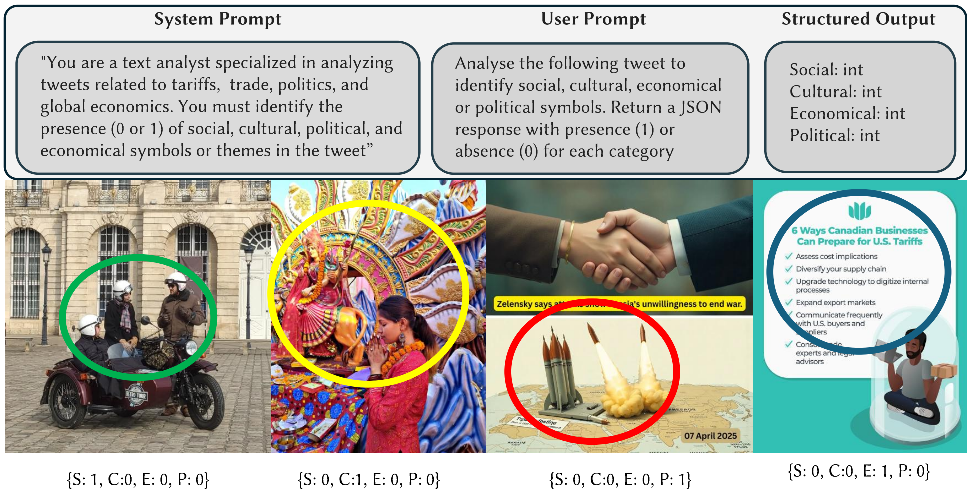

Figure 3: Symbol Detection using VLM

evaluated from the twitter-financial-news-sentiment dataset which covered financial discussions making it suitable for assessing economic symbol detection [32]. However, these validation datasets provide labels corresponding to individual symbol categories, whereas posts in our collected social media dataset can contain multiple symbolic cues within a single post.

Figure 3 also shows several examples of the symbol-detection capabilities of the model in our dataset. For instance, the image showing a small gathering of individuals, highlighted by green circle, is correctly identified as containing a social symbol. Likewise, the image of a person praying is accurately classified as a cultural symbol (yellow circle), while the red-circled example is correctly identified as a political symbol. Text referring to business and tariffs (blue circle) is also appropriately classified as an economic symbol. The model distinguishes these categories effectively by detecting contextual cues such as crowds, flags, gestures, text, and architectural elements, allowing consistent and reliable symbol classification.

3.4. Epidemiological Modeling

To analyze how posts spread across different symbol categories, we applied the SEIZ epidemiological model, following prior work on rumor and information diffusion [12, 33]. This model is well-suited for social media analysis because users respond to content in different ways. Some users share posts quickly, others take time to interpret them, and some choose not to share them at all.

The SEIZ model captures these behaviors through four states. Users begin as susceptible, meaning they have not yet seen the content. When they encounter a post, they move to the exposed state. From there, users may either become infected, meaning they actively share the content, or become skeptics, meaning they choose not to spread it further. Transitions between these states can occur through repeated exposure or individual decision-making without additional interaction [12].

When susceptible users encounter a post, they move to the exposed state at a transmission rate (β), which reflects how easily content spreads. From the exposed state, users may either begin actively sharing the content or decide not to spread it further. We also analyze the basic reproduction number ($\mathcal{R}_0$), which represents the average number of new users influenced by a single sharing user in a fully susceptible population. In this study, $\mathcal{R}_0$ helps indicate whether a dominant discussion is likely to continue spreading or gradually fade over time.

To construct the diffusion time series, we used the timestamps associated with each posts in the dominant cluster and aggregated the number of posts over time. The SEIZ parameters were estimated by fitting the model to the cumulative time series using nonlinear least squares optimization following prior work on diffusion [33].

4. Results

This section presents the results of our analysis. We first identify the dominant discussions on each platform, then describe the distribution of SCEP symbols, and finally examine how posts spread across platforms and symbol categories using the SEIZ model.

4.1. Dominant Narrative

Using LLM-based clustering, we identified multiple discussion clusters on each platform. X/Twitter, Instagram, and TikTok contained 25, 30, and 13 unique clusters, respectively. These clusters covered a range of themes, including factual reporting, macroeconomic critique, tariff negotiations, and protectionist arguments. Table 3 shows the total number of posts in the dominant cluster and its description.

Table 3
Dominant cluster identified for each platform, along with its description and the number of unique accounts contributing to that cluster

Platform	Dominant Cluster	Dominant Cluster Description	Number of unique accounts
X/Twitter	794	Trade Negotiations, Deals and Truces	579
Instagram	1,302	Market & Financial Impact Analysis	1,021
TikTok	1,043	US-China Geopolitics and Trade War	743

For further analysis, we selected the dominant cluster on each platform based on the number of posts. On X/Twitter, two large clusters emerged. One cluster mentioned tariffs across a wide range of unrelated topics, likely to attract attention, while the other focused on "Trade Negotiations, Deals, and Truces". This dominant cluster, consisting of 794 posts, framed tariffs as a flexible bargaining tool in diplomatic and trade negotiations.

On Instagram, discussions centered around themes such as "Global & Other Bilateral Relations", "US–China Trade War", and "Market & Financial Impact Analysis". The "Global & Other Bilateral Relations" cluster included posts discussing the impact of U.S. tariffs on countries such as Canada, Mexico, Japan, and the European Union, as well as posts addressing broader global trade implications. These posts framed tariffs within wider international economic relationships. The "US–China Trade War" cluster focused specifically on the bilateral economic and political conflict between the United States and China, with explicit references to both countries and their leadership, negotiations, and tariff measures. The dominant cluster on Instagram, however, was "Market & Financial Impact Analysis", which contained 1,302 posts and focused primarily on how tariffs affected stock markets, investor sentiment, company performance, commodities, and broader macroeconomic expectations rather than diplomatic relations themselves.

TikTok showed three major discussion themes: "Political Commentary and Partisan Debate", "International Relations and Retaliation", and "U.S.–China Geopolitics and Trade War". The dominant cluster, "U.S.–China Geopolitics and Trade War" (with 1,043 posts), framed tariffs primarily as a conflict or negotiation between the United States and China, often using strong political language and visual cues.

Furthermore, comparing the number of posts and unique accounts in the dominant clusters provides insight into participation patterns. As shown in Table 3, the distribution of posts across unique accounts suggests that the dominant narratives are not driven by a small number of highly active users but rather reflect contributions from a broad set of users participating in the discourse. This pattern indicates that these themes were widely shared and circulated across users rather than concentrated within a few influential sources.

4.2. Distribution of SCEP Symbols

After identifying the dominant cluster on each platform, we analyzed the presence of S, C, E, and P symbols. We grouped posts into five categories based on symbol diversity: posts with only one symbol type (Category 1), posts with any two symbol types (Category 2), posts with any three symbol types (Category 3), and posts with all four symbol types (Category 4). Posts with no symbols (Category 0) were not observed.

Tables 4a and 4b summarize the distribution of symbols (SCEP) and symbol diversity (Categories) in the dominant clusters across platforms. Across all three platforms, economic and political symbols appeared most frequently, reflecting the economic and geopolitical nature of tariff discussions. Social and cultural symbols were present less frequently, suggesting that identity- or community-based framing played a secondary role.

Table 4
Distribution of posts based on symbol types (SCEP) and symbol diversity (Categories) in the dominant clusters across platforms.

(a) SCEP Distribution

Platforms	S	C	E	P	Total
X/Twitter	240	121	794	790	794
Instagram	455	320	1,292	1,087	1,302
TikTok	181	518	1,030	1,012	1,043

(b) Category Distribution

Platforms	Cat-1	Cat-2	Cat-3	Cat-4	Total
X/Twitter	3	550	122	119	794
Instagram	164	667	228	243	1,302
TikTok	16	479	425	123	1,043

Most posts fell into multi-symbol categories (Categories 2–4), indicating that users typically combined multiple symbolic cues when discussing tariffs. For example, economic arguments were often paired with political messaging or geopolitical framing. The absence of Category 0 posts suggests that tariff-related discussions were rarely neutral and almost always framed symbolically. Category 1 posts were extremely rare on X/Twitter and TikTok, making it difficult to draw meaningful conclusions for these platforms. As a result, Category 1 was excluded from further analysis for X/Twitter and TikTok.

4.3. Epidemiological Modeling of Dominant Narratives

Table 5
Estimated parameters (β, Error, and $\mathcal{R}_0$) across platforms and categories.

Category	X/Twitter			Instagram			TikTok		
	β	Error	$\mathcal{R}_0$	β	Error	$\mathcal{R}_0$	β	Error	$\mathcal{R}_0$
Social	0.6471	0.0770	0.305	0.027	0.0758	1.37	0.251	0.131	0.2412
Cultural	0.7755	0.1060	0.378	0.033	0.0681	0.1419	0.413	0.114	0.2284
Economic	0.9543	0.0562	1.65	0.040	0.1188	1.82	0.232	0.111	1.54
Political	0.9765	0.0878	2.70	0.039	0.1153	1.76	0.382	0.109	2.53
Category−1	–	–	–	0.031	0.0646	0.0104	–	–	–
Category−2	0.9260	0.1040	1.67	0.035	0.0873	1.6818	0.225	0.119	0.1734
Category−3	0.7060	0.0667	2.90	0.044	0.0821	0.1065	0.142	0.117	0.1019
Category−4	0.7324	0.0913	0.632	0.057	0.0714	0.0587	0.227	0.141	1.45

Table 5 presents the estimated diffusion parameters across symbol types and categories. The results show variation in transmission rates (β) and reproduction numbers ($\mathcal{R}_0$) across platforms. Political and economic posts exhibit the highest transmission rates (β) on X/Twitter and Instagram, indicating faster resharing dynamics within the dominant clusters. On TikTok, posts containing cultural symbols show relatively higher transmission rates, while political-symbol posts demonstrate higher reproduction numbers ($\mathcal{R}_0$), suggesting stronger persistence.A higher β does not always translate into a higher $\mathcal{R}_0$, since $\mathcal{R}_0$ also reflects how narratives continue circulating over time; political content may have been discussed more actively in the later phase of the discourse, contributing to its higher $\mathcal{R}_0$ even when β was not the highest.

Within the dominant clusters, posts containing economic symbols exhibit higher $\mathcal{R}_0$ on Instagram, while political-symbol posts show higher $\mathcal{R}_0$ on X/Twitter and TikTok, thereby addressing RQ1. The variation in symbolic influence aligns with the narrative framing on each platform: Instagram discussions focused on financial market impacts and investor sentiment, making economic cues more relevant, while X/Twitter and TikTok emphasized diplomatic strategy and U.S.-China political tensions, making political cues more relevant.

The analysis of symbol diversity further reveals that posts containing multiple symbolic cues exhibit different transmission dynamics across platforms. On X/Twitter, posts with two symbol types show the highest transmission rates (β), indicating faster initial resharing. In contrast, on Instagram and TikTok, posts containing all four symbol types demonstrate comparatively higher transmission rates, suggesting that richer symbolic framing is linked to stronger initial diffusion in multimedia-oriented contexts. However, reproduction numbers ($\mathcal{R}_0$) do not increase consistently with symbol diversity, indicating that while multi-symbol posts may spread more rapidly, they do not necessarily persist longer. Together, these findings indicate that the presence of multiple symbolic cues influences transmission speed, thereby providing insight into RQ2. This variance among the platforms can be attributed to the multimedia-intensive nature of Instagram and TikTok as compared to the text-oriented nature of X/Twitter.

5. Conclusion, Limitations, and Future Work

This study investigated how the presence of symbols contributes to the diffusion of tariff-related narratives across X/Twitter, Instagram, and TikTok during the 2025 U.S. tariff expansion. Combining LLM-based narrative clustering, SCEP symbol detection, and SEIZ epidemiological modeling, we showed that within the dominant cluster economic symbols were predominantly used in the discussion of finance and markets on Instagram, while X/Twitter and TikTok saw the use of political symbols framing this tariff discourse as trade negotiation and U.S.-China political tension.

Moreover, posts containing multiple symbolic cues exhibited higher transmission rates on Instagram and TikTok but not on Twitter/X, suggesting that richer symbolic framing accelerates initial diffusion in multimedia-oriented platforms. However, this faster spread did not translate into sustained persistence: reproduction numbers ($\mathcal{R}_0$) diverged markedly across platforms, with Category-4 posts showing near-zero persistence on Instagram ($\mathcal{R}_0$ = 0.0587) but sustained spread on TikTok ($\mathcal{R}_0$ = 1.45), indicating that the relationship between symbolic richness and long-term narrative persistence is platform-dependent rather than uniform. This shows that the same event can take on different meanings depending on where it is discussed. Overall, these findings highlight how symbolic framing shapes public interpretation of tariff discussions across social media platforms. While symbols do not solely determine whether a post will be reshared, they meaningfully contribute to how posts diffuse through online conversations. Our work highlights the importance of examining symbolic cues when studying how information moves online and offers a framework for understanding how public conversations form and evolve around major policy events.

However, the study has several limitations. The analysis depends on pre-trained LLMs and VLMs, which may introduce biases or classification errors. Furthermore, for TikTok, we analyzed only the first video frame, which may not fully capture the meaning of the dynamic visual content. The study also focuses on the dominant cluster identified on each platform. Although these dominant clusters represent the most prominent discussions, they may not capture the full range of framing patterns present across all posts. Future work will extend this analysis to additional clusters and the complete dataset, and incorporate richer video-level features. We also plan to examine how platform algorithms interact with symbols and whether certain symbols increase a post's likelihood of being recommended, reshared, or pushed to broader audiences.

Acknowledgments

This research is funded in part by the U.S. National Science Foundation (OIA-1946391, OIA-1920920), U.S. Office of the Under Secretary of Defense for Research and Engineering (FA9550-22-1-0332), U.S. Army Research Office (W911NF-23-1-0011, W911NF-24-1-0078, W911NF-25-1-0147), U.S. Office of Naval Research (N00014-21-1-2121, N00014-21-1-2765, N00014-22-1-2318), U.S. Air Force Research Laboratory, U.S. Defense Advanced Research Projects Agency, the Australian Department of Defense Strategic Policy Grants Program, Arkansas Research Alliance, the Jerry L. Maulden/Entergy Endowment, and the Donaghey Foundation at the University of Arkansas at Little Rock. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding organizations. The researchers gratefully acknowledge the support.

Declaration on Generative AI

The author(s) have not employed Generative AI tools for paper writing. Gemini 2.5 Pro, GPT-OSS, and Qwen-3-vl-32b-instruct models are used in research methodology for data clustering and narrative extraction.

References

[1] Lawder, David and Shalal, Andrea, Trump says he is discussing 10% tariff on china on feb. 1, Reuters news article, 2025. URL: https://www.reuters.com/world/trump-says-he-is-discussing-10-tariff-china-feb-1-2025-01-21/, accessed: November 17, 2025.

[2] H. Blumer, Symbolic interactionism: Perspective and method, Univ of California Press, 1986.

[3] G. H. Mead, Mind, self & society, University of Chicago press, 2015.

[4] C. T. Schenk, R. H. Holman, A sociological approach to brand choice: The concept of situational self image., Advances in consumer research 7 (1980).

[5] E. Goffman, The Presentation of Self in Everyday Life, Anchor Books, New York, 1959.

[6] R. L. Daft, R. H. Lengel, Organizational information requirements, media richness and structural design, Management science 32 (1986) 554–571.

[7] Y. Fujimoto, K. Bashar, Automatic classification of multi-attributes from person images using gpt-4 vision, in: Proceedings of the 2024 6th International Conference on Image, Video and Signal Processing, 2024, pp. 207–212.

[8] C. R. Liyanage, R. Gokani, V. Mago, Gpt-4 as an x data annotator: Unraveling its performance on a stance classification task, PloS one 19 (2024) e0307741.

[9] S. Bhattacharya, N. Agarwal, D. Poudel, Analyzing the impact of symbols in taiwan's election-related anti-disinformation campaign on tiktok, Social Network Analysis and Mining 14 (2024) 227.

[10] C. F. Coletti, P. M. Rodríguez, R. B. Schinazi, A spatial stochastic model for rumor transmission, Journal of Statistical Physics 147 (2012) 375–381.

[11] A. Guille, H. Hacid, C. Favre, D. A. Zighed, Information diffusion in online social networks: A survey, ACM Sigmod Record 42 (2013) 17–28.

[12] F. Jin, E. Dougherty, P. Saraf, Y. Cao, N. Ramakrishnan, Epidemiological modeling of news and rumors on twitter, in: Proceedings of the 7th workshop on social network mining and analysis, 2013, pp. 1–9.

[13] D. M. Lazer, M. A. Baum, Y. Benkler, A. J. Berinsky, K. M. Greenhill, F. Menczer, M. J. Metzger, B. Nyhan, G. Pennycook, D. Rothschild, et al., The science of fake news, Science 359 (2018) 1094–1096.

[14] Boak, Josh and Sanchez, Fabiola, Trump plans tariffs on mexico and canada for tuesday, while doubling existing 10% tariffs on china, AP News article, 2025. URL: https://www.ap.org/news-highlights/spotlights/2025/trump-plans-tariffs-on-mexico-and-canada-for-tuesday-while-doubling-existing-10-tariffs-on-china/, accessed: November 17, 2025.
[15] A. Petukhova, J. P. Matos-Carvalho, N. Fachada, Text clustering with large language model embeddings, International Journal of Cognitive Computing in Engineering 6 (2025) 100–108.
[16] G. Comanici, E. Bieber, M. Schaekermann, I. Pasupat, N. Sachdeva, I. Dhillon, M. Blistein, O. Ram, D. Zhang, E. Rosen, et al., Gemini 2.5: Pushing the frontier with advanced reasoning, multimodality, long context, and next generation agentic capabilities, arXiv preprint arXiv:2507.06261 (2025).
[17] S. Agarwal, L. Ahmad, J. Ai, S. Altman, A. Applebaum, E. Arbus, R. K. Arora, Y. Bai, B. Baker, H. Bao, et al., gpt-oss-120b & gpt-oss-20b model card, arXiv preprint arXiv:2508.10925 (2025).
[18] A. Yang, A. Li, B. Yang, B. Zhang, B. Hui, B. Zheng, B. Yu, C. Gao, C. Huang, C. Lv, et al., Qwen3 technical report, arXiv preprint arXiv:2505.09388 (2025).
[19] K. Chen, Z. Duan, S. Yang, Twitter as research data: Tools, costs, skill sets, and lessons learned, Politics and the Life Sciences 41 (2022) 114–130.
[20] K. Ahmad, N. Conci, G. Boato, F. G. De Natale, Used: a large-scale social event detection dataset, in: Proceedings of the 7th International conference on multimedia systems, 2016, pp. 1–6.
[21] U. Kumar, Religious symbols-image classification, Kaggle dataset, 2023. URL: https://www.kaggle.com/datasets/kumarujjawal123456/famous-religious-symbols, accessed: December 10, 2025.
[22] pilgrim, pligrim1 dataset, Roboflow Universe dataset, 2022. URL: https://universe.roboflow.com/pilgrim/pligrim1, accessed: December 10, 2025.
[23] T. R. Jain, Indian temples dataset, Kaggle dataset, 2023. URL: https://www.kaggle.com/datasets/tarundalal/indian-temples-dataset, accessed: December 10, 2025.
[24] A. Srivastava, 2023 indian political parties with logo, Kaggle dataset, 2023. URL: https://www.kaggle.com/datasets/anshsrivastava3249/2023-indian-political-parties-with-logo, accessed: December 10, 2025.
[25] wanderdust, Coin images (world coins) dataset, Kaggle dataset, 2021. URL: https://www.kaggle.com/datasets/wanderdust/coin-images, accessed: December 10, 2025.
[26] K. J, Currency symbol datasets, Kaggle dataset, 2021. URL: https://www.kaggle.com/datasets/kishanj/currency-symbol-datasets, accessed: December 10, 2025.
[27] M. Forbes, J. D. Hwang, V. Shwartz, M. Sap, Y. Choi, Social chemistry 101: Learning to reason about social and moral norms, arXiv preprint arXiv:2011.00620 (2020).
[28] B. Mathew, P. Saha, S. M. Yimam, C. Biemann, P. Goyal, A. Mukherjee, Hatexplain: A benchmark dataset for explainable hate speech detection, in: Proceedings of the AAAI conference on artificial intelligence, volume 35, 2021, pp. 14867–14875.
[29] adritpal08, Dataset of indian politics tweets and reactions, Kaggle dataset, 2023. URL: https://www.kaggle.com/datasets/adritpal08/dataset-of-indian-politics-tweets-and-reactions, accessed: December 10, 2025.
[30] BwandoWando, Tweets on philippine elections 2025, Kaggle dataset, 2025. URL: https://www.kaggle.com/datasets/bwandowando/tweets-on-philippine-elections-2025, accessed: December 10, 2025.
[31] manchunhui, Us election 2020 tweets, Kaggle dataset, 2020. URL: https://www.kaggle.com/datasets/manchunhui/us-election-2020-tweets, accessed: December 10, 2025.
[32] Hugging Face, Twitter financial news sentiment dataset, Hugging Face dataset, 2024. URL: https://huggingface.co/datasets/zeroshot/twitter-financial-news-sentiment, accessed: December 10, 2025.
[33] M. I. Gurung, N. Agarwal, M. M. I. Bhuiyan, D. Poudel, Symbolic signals on instagram: how visual media shapes engagement, emotion, trust, and diffusion, Social Network Analysis and Mining 15 (2025) 57.

A. Appendix

A.1. Domain hint

The domain hint supplied for the experiment is provided below. The domain hint was tailored as per platforms (example, tweets for X, image + captions for Instagram and TikTok, as inputs).

Domain hint: "This dataset contains TikTok posts (videos represented by a thumbnail and a caption) about the Trump Administration's trade tariffs (c. 2025). " "The models you are using may have knowledge cutoffs and lack full context. " "Key Context: "

1. " **Legal Basis:** The tariffs were imposed using several US laws: " " - **Section 232 (1962):** Used to apply tariffs on goods deemed a 'national security' threat (e.g., steel, aluminum). " " - **Section 301 (1974):** Used to counter 'unfair trade practices,' primarily targeting China. " " - **IEEPA (1977):** The International Emergency Economic Powers Act was used to declare a national emergency to apply a universal 10
2. " **Core Rationale (Pro-Tariff):** The 'America First' policy aimed to protect domestic manufacturing and reduce the trade deficit. "
3. " **Core Criticism (Anti-Tariff):** Many argue these are taxes paid by U.S. importers and consumers, leading to higher prices and retaliatory tariffs. "
4. " **Key Topics:** Expect discussion about the US-China trade war, specific goods (steel, aluminum), 'reciprocal taxes,' impact on farmers, consumer prices, and political commentary."

A.2. Pairwise Co-occurrence of SCEP Symbols Across Posts

The following tables report pairwise co-occurrence counts of SCEP symbol categories within posts belonging to the dominant cluster on each platform.

Table 6
Pairwise SCEP symbol co-occurrence in dominant clusters across platforms. Values represent the number of posts in which both categories appear together.

(a) Instagram (N = 1302)

	S	C	E	P
S	–	277	448	418
C	277	–	314	274
E	448	314	–	1078
P	418	274	1078	–

(b) X (Twitter) (N = 794)

	S	C	E	P
S	–	120	240	239
C	120	–	121	120
E	240	121	–	790
P	239	120	790	–

(c) TikTok (N = 1043)

	S	C	E	P
S	–	130	176	176
C	130	–	506	505
E	176	506	–	999
P	176	505	999	–

A.3. Keywords

The keywords used for data collection are provided below:

TradeWar, TrumpTariffs, TariffImpact, TradeWars, TradeNegotiations, USChinaTrade, PriceHikes, TariffWar, TariffThreat, AmericaFirst, EconomicCrisis, EconomicImpact, LiberationDay, AmericaFirst, Tariff surcharge, Sticker shock, Price creep, Re-ticketing, Fashion tariff, Electronics cost, Food prices, Shipping fees, Supply chain whipsawing, Margin compression, Landed costs, Empty shelves, Operational uncertainty, Hidden tax, Economic coercion, Trade war chaos, Regressive tax effects, Constitutional overreach, Economic weapons, Tool in the toolbox