DEEP: A Discourse Evolution Engine for Predictions about Social Movements

Valerio La Gatta[1*], Marco Postiglione[1*], Jeremy Gilbert[2], Daniel W. Linna Jr.[3], Morgan Manella Greenfield[4], Aaron Shaw[5], V.S. Subrahmanian[1]

[1]Department of Computer Science and Buffett Institute for Global Affairs, Northwestern University
[2]Medill School of Journalism, Media, Integrated Marketing Communications, Northwestern University
[3]Northwestern Pritzker School of Law and Department of Computer Science, Northwestern University
[4]The Wall Street Journal
[5]Department of Communication Studies, Northwestern University
vss@northwestern.edu

Abstract

Numerous social movements (SMs) around the world help support the UN's Sustainable Development Goals (SDGs). Understanding how key events shape SMs is key to the achievement of the SDGs. We have developed SMART (Social Media Analysis & Reasoning Tool) to track social movements related to the SDGs. SMART was designed by a multidisciplinary team of AI researchers, journalists, communications scholars and legal experts. This paper describes SMART's transformer-based multivariate time series Discourse Evolution Engine for Predictions about Social Movements (DEEP) to predict the volume of future articles/posts and the emotions expressed. DEEP outputs probabilistic forecasts with uncertainty estimates, providing critical support for editorial planning and strategic decision-making. We evaluate DEEP with a case study of the *#MeToo* movement by creating a novel longitudinal dataset (433K Reddit posts and 121K news articles) from September 2024 to June 2025, which is publicly available for research purposes.

Datasets —
https://sites.northwestern.edu/nsail/projects/smart/

Introduction

Social movements (SMs) play an important role in supporting the UN's Sustainable Development Goals (SDGs) (Salles et al. 2024). The #MeToo movement alone has influenced the reporting of gender-related crimes (Levy and Mattsson 2023) and health effects linked to sexual harassment (O'Neil et al. 2018). SMs involved with the fossil fuel energy have had some success in the UK, Netherlands, and Poland (Hielscher, Wittmayer, and Dańkowska 2022), and the rise of community choice movements in California (Smith 2019). In the Philippines, social movements supporting the environment have helped reduce pollution by companies (Magno 2017) and protect indigenous peoples' rights (Theriault 2011). Even small progress supporting the UN's SDGs can have a significant impact on millions of people, and SMs supportive of the SDGs attempt to do just that.

Our Social Media Analysis and Reasoning Tool (SMART)[1] tracks SMs that are supportive of the UN's SDGs and makes information about them available to journalists. The 8 journalists when SMART/DEEP were started had four requirements. (i) They wanted to know about SMs that were rising in popularity in terms of volume (number of posts mentioning the SM). (ii) For such SMs, they wanted to understand the sentiments and emotions expressed in those posts and how they evolve over time. (iiii) They were interested in the relationship between key events (e.g., political events such as elections) and these SMs. (iv) As journalists want to publish articles on timely topics before other journalists do or the public broadly knows, they wanted to identify such SMs early. *For these reasons, we designed our Discourse Evolution Engine for Predictions about Social Movements (DEEP) to forecast volume and emotional intensity so that journalists could look at forecasts about SMs of interest (e.g., a journalist in the Philippines might care more about SMs relating to fishing issues in her country than a journalist in Paris).* This way, a journalist can use the forecasts to hone in on SMs that are much talked about, and write about them *early*, before others do. *This paper is not about SMART, but it is discussed briefly for context.*

The main contributions of this paper are as follows. (I) We define the Discourse Evolution Prediction problem to predict the volume of posts about a specific SM *S*, the sentiment about *S*, and the emotional intensity of several emotions in posts about *S*, Δ time units into the future. We then propose the DEEP framework to solve this problem. (II) We describe experiments that show how DEEP performs on data about the #MeToo movement. DEEP achieves consistently higher precision, recall, and F1 on news than on Reddit, reflecting the clearer emotional signals in formal reporting. However, while short-horizon forecasts are most reliable for news, Reddit predictions improve with longer horizons, revealing platform-specific dynamics that enable accurate near-term detection in news and strong medium-term forecasting in social media discussions. (III) We provide a #MeToo dataset

[*]These authors contributed equally.

[1]https://buffett.northwestern.edu/research/global-working-groups/ai-and-social-movements.htmlhttps://buffett.northwestern.edu/research/global-working-groups/SMART

consisting of 433,016 Reddit posts and 121,849 news articles from September 1, 2024 and June 28, 2025, along with an analysis involving key events from the Sean "Diddy" Combs case. The SMART/DEEP framework has been developed with consistent feedback from journalists, demonstrated to over 20 journalists to gain feedback, including at a small workshop in March 2025 to fine-tune the system[2]. A list of the organizations involved is provided in the online Appendix. We expect to make DEEP available to journalists in early 2026. As such, this paper falls within IAAI's Emerging Applications Track.

DEEP Problem Formulation

We formalize the DEEP problem as follows. The *discourse state* $\mathcal{S}_t$ at time t is a vector $\mathcal{S}_t = (\mathbf{V}_t, \mathbf{E}_t, \mathbf{T}_t)$ where: (i) $\mathbf{V}_t \in \mathbb{R}^h$ represents the volume of posts/articles gathered at time t, (ii) $\mathbf{E}_t \in \mathbb{R}^d$ denotes emotional intensity across $d = 28$ emotional dimensions averaged across all posts at time t, and (iii) $\mathbf{T}_t \in \mathbb{R}^m$ represents the distribution over m topics of interest (e.g., politics, sports, entertainment).

The *Historical Trajectory* $\mathcal{H}_t$ is a sequence of discourse states up to time t: $\mathcal{H}_t = (\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_t)$.

Key events $\mathcal{K}_t$ are discrete occurrences that the journalist believes will significantly influence the dynamics of discourse about the phenomenon the journalist is interested in. Formally, $\mathcal{K}_t = k_1, k_2, \ldots, k_n$ where $k_i = (\epsilon_i, t_i, \alpha_i)$. Each key event k_i has a type $\epsilon_i \in \mathcal{E}$ (from a predefined event taxonomy), occurrence time t_i, and impact magnitude $\alpha_i \in \mathbb{R}^+$ (provided by a journalist).

The Discourse Evolution Engine Prediction (DEEP) Problem is to estimate the future discourse state $\mathcal{S}_{t+\Delta}$ given the historical trajectory $\mathcal{H}_t$ and key events $\mathcal{K}_t$, i.e. to find a function f such that:

$$\mathcal{S}_{t+\Delta} = f(\mathcal{H}_t, \mathcal{K}_{t:t+\Delta}, \boldsymbol{\theta}), \tag{1}$$

where $\mathcal{K}_{t:t+\Delta}$ represents key events the journalist expects to happen within the prediction window Δ, and $\boldsymbol{\theta}$ denotes model parameters learned from historical data.

Materials & Method

Our system uses a multi-stage pipeline that integrates data collection, feature engineering, and transformer-based forecasting. Figure 1 presents an overview of our pipeline, illustrating the flow from raw data collection through feature extraction to predictive modeling. We now discuss these modules in depth[3].

Data Collection

We identify data about a social movement through a corpus of content selected via keyword-based search and information retrieval processes that capture both explicit references to the movement and semantically related discussions. While our system is designed to support journalists predicting discourse evolution around specific social movements, this abstraction makes it sufficiently general to support other applications reducible to similar search and retrieval structures, including topic-based forecasting and thematic content analysis. For the #MeToo movement, we extract data daily using the hashtag "#MeToo" alongside related search keywords identified by journalists (e.g., gender equality, women's voices, equal opportunities; the full list is reported in the online Appendix).

Layer	Description	Reddit	News
L_0	Direct #MeToo mentions	$3,103$	$3,286$
L_1	Documents containing ≥30% of top 1% co-occurring keywords	$81,885$	$4,748$
L_2	Documents containing ≥20% of top 1% co-occurring keywords	$73,828$	$1,785$
L_3	Documents containing ≥10% of top 1% co-occurring keywords	$274,200$	$112,030$
Total Documents		$433,016$	$121,849$

Table 1: Summary of Multi-layer Data Extraction Results

The resulting comprehensive dataset consists of $433,016$ Reddit posts and $121,849$ news articles collected between September 1, 2024 and June 28, 2025. We now describe a multi-layer data extraction methodology we developed that captures both explicit mentions to the movement and related discourse.

Layer L_0 The first layer L_0 identifies documents that explicitly reference the target social movement sm (i.e., sm =#MeToo). Documents in this layer either were extracted via the "#MeToo" keyword or contain it in their title or body text. Layer L_0 contains $3,103$ Reddit posts and $3,286$ news articles.

Layers L_1, L_2, $\ldots, L_N$ To expand the dataset beyond documents containing direct mentions of sm =#MeToo, we implemented a systematic approach to identify semantically related content. First, we extracted keywords and keyphrases for each document using KeyBERT (Sharma and Li 2019) and Amazon Comprehend[4], respectively. We then selected all keywords that co-occur with sm across our corpus and computed their frequency distribution. From this distribution, we identified keywords falling within the 99th percentile of co-occurrence, creating a keyword set that captures the core vocabulary associated with sm. We progressively relaxed thresholds to capture documents at varying degrees of semantic relatedness to the movement: layers L_1, L_2 and L_3 include documents containing at least 30%, 20% and 10% of the high-salience keywords, respectively.

This tiered approach balances precision and recall, enabling the inclusion of documents that may not explicitly mention the target SM but contribute meaningfully to the broader discourse surrounding the movement. Table 1 shows the number of Reddit posts and news articles retrieved within each layer.

[2]https://buffett.northwestern.edu/news/2025/empowering-journalists-to-decode-social-movements-the-buffett-global-working-group-behind-new-ai-system.html

[3]While the description focuses on the #MeToo movement as the target SM, the methodology can be easily adapted to any social movement.

[4]https://aws.amazon.com/comprehend/

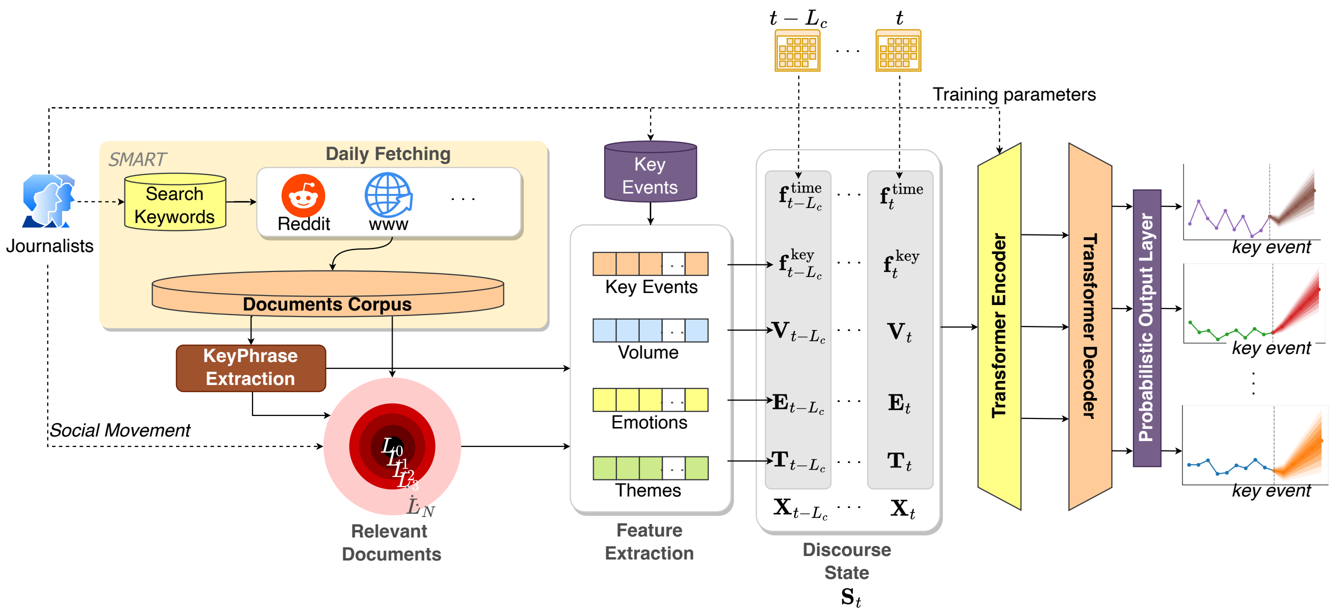


Figure 1: **Overview of DEEP.** Data is collected from Reddit and News using hashtag and semantic keyword extraction. Keyphrases determine document relevance to the input social movement. Feature extraction transforms text into structured representations across volume, emotions, themes, and key events. A TimeSeriesTransformer processes historical discourse through an encoder-decoder framework to generate probabilistic forecasts of future discourse states $S_{t+\Delta}$ with uncertainty quantification via Student-t distributions.

Feature Extraction

DEEP transforms raw social movement discourse into structured temporal features across three complementary dimensions. Our feature extraction pipeline processes content from news sources and Reddit to generate time-indexed feature vectors that capture the evolving nature of social movement discourse. The complete details and mathematical formulation are provided in the online Appendix.

Volume Features ($\mathbf{V}_t$). We quantify discourse intensity through platform-specific volume measurements that account for both content quantity and engagement patterns. Each content item is weighted by relevance scores derived from our multi-layer extraction methodology and engagement metrics. Volume features include raw counts, temporal derivatives (velocity and acceleration), and cross-platform distribution indices to capture how discourse spreads across news sources and Reddit channels.

Emotion Features ($\mathbf{E}_t$). We characterize the affective landscape using a 28-dimensional emotion space based on the GoEmotions dataset (Demszky et al. 2020). Content-level emotion intensities are computed using a fine-tuned RoBERTa model[5], then aggregated into weighted bin distributions across five intensity levels (Absent to Very High). This captures both the dominant emotional themes and their distributional patterns over time.

Thematic Features ($\mathbf{T}_t$). We capture the topical landscape through distance-based binning that leverages extracted keyphrases to quantify content alignment with predefined themes. Using a taxonomy of journalist-defined thematic categories (e.g., Gender Equality, Human Rights, Violence), we compute semantic distances between content keyphrases and topic centroids using DistilBERT embeddings (Sanh et al. 2019), then discretize these distances into six bins representing alignment levels.

Key Event Features ($\mathbf{f}_t^{\text{key}}$). To model external catalysts, we incorporate journalist-identified events that may influence discourse patterns. Events are categorized across thematic dimensions and encoded with impact assessments relative to social movement objectives, generating binary indicator features for category-impact combinations.

Transformer-based Forecasting

DEEP leverages a time series transformer architecture (Wen et al. 2023) to model the complex temporal dependencies inherent in social movement discourse evolution. Transformers are particularly well-suited for this task as they can capture long-range dependencies and handle multivariate time series with heterogeneous feature types. DEEP's forecasting component follows an encoder-decoder framework where the encoder processes historical discourse trajectories over L_c context steps, and the decoder generates probabilistic predictions for Δ future time steps. The input representation at time t integrates all discourse state components:

$$\mathbf{X}_t = [\mathbf{V}_t; \mathbf{E}_t; \mathbf{T}_t; \mathbf{f}_t^{\text{time}}; \mathbf{f}_t^{\text{key}}] \quad (2)$$

where $\mathbf{V}_t \in \mathbb{R}^h$ encodes volume features, $\mathbf{E}_t \in \mathbb{R}^d$ captures emotional intensity distributions, $\mathbf{T}_t \in \mathbb{R}^m$ represents thematic distributions, $\mathbf{f}_t^{\text{time}} \in \mathbb{R}^2$ includes temporal encodings

[5]https://huggingface.co/SamLowe/roberta-base-go_emotions

(day-of-week, month), and $\mathbf{f}_t^{\text{key}} \in \mathbb{R}^{Q\times(|\mathcal{I}|+1)}$ encodes key event information from $\mathcal{K}_t$.

To maintain computational efficiency while preserving predictive information, we apply mutual information-based feature selection to reduce dimensionality before feeding inputs to the transformer layers.

Social movement discourse exhibits both immediate reactions and delayed cascading effects. To capture this multi-scale behavior, we incorporate explicit lag sequences $\mathcal{L} = \{1, 2, 3\}$ that provide the model direct access to recent historical states.

Rather than point predictions, DEEP generates full probability distributions over future discourse states. We use a Student-t output distribution parameterized by location $\boldsymbol{\mu}_{t+\Delta}$, scale $\boldsymbol{\sigma}_{t+\Delta}$, and degrees of freedom $\boldsymbol{\nu}_{t+\Delta}$. This supports journalists' needs by providing both point forecasts and confidence intervals, enabling more informed reporting decisions when covering evolving social movements.

Experiments

Evaluation Metrics

We evaluate DEEP's ability to produce accurate forecasts for journalists by predicting directional changes in target variables (e.g., volume of articles $\mathbf{V}_{t+\delta}$, or intensity of specific emotions within $\mathbf{E}_{t+\delta}$) over the δ-day prediction horizon. For each target variable Y_t, we classify predictions into three categories using a statistical significance framework:

$$\text{Increase: } Y_{t+\delta} > \mu_Y + 2\sigma_Y \tag{3}$$

$$\text{Decrease: } Y_{t+\delta} < \mu_Y - 2\sigma_Y \tag{4}$$

$$\text{Stable: } \mu_Y - 2\sigma_Y \leq Y_{t+\delta} \leq \mu_Y + 2\sigma_Y, \tag{5}$$

where μ_Y and σ_Y represent the rolling mean and standard deviation computed over a 28-day historical window. The 2σ threshold corresponds to approximately 95% confidence intervals, ensuring that predicted "significant" changes represent genuinely newsworthy developments rather than normal discourse fluctuation. Performance is measured using standard classification metrics across the three-class prediction task: accuracy, precision, recall and F1 scores, computed both per-class and macro-averaged.

Implementation Details

The TimeSeriesTransformer was configured with parameters chosen to balance journalists' needs and computational efficiency: (1) A 4-week historical window captures monthly discourse patterns while providing sufficient context for detecting emerging trends. (2) A 7-day prediction horizon ($\Delta = 7$ days) aligns with the weekly news cycle, enabling journalists to plan coverage, allocate resources, and prepare investigative pieces within standard editorial workflows. (3) A dropout rate of 0.3 mitigates overfitting, while a batch size of 2 accommodates the sequential nature of time-series data. (4) A learning rate of 3×10^{-4} with a weight decay of 1×10^{-5} supports stable convergence during training according to our empirical results.

Key events $\mathcal{K}_t$ were extracted from a comprehensive public source[6] covering the study period from September 2024 to February 2025. Each event $k_i = (\epsilon_i, t_i, \alpha_i)$ was manually annotated by our team. *The resulting annotated events table will be provided as part of the dataset when this paper is published.* All experiments are conducted using three NVIDIA RTX A6000 GPUs with 48GB VRAM each.

[6]https://www.onthisday.com

DEEP System Performance

Next-day Prediction Performance. Table 2 reports macro-average metrics at $\Delta = 1$ day for the three-way *increase / stable / decrease* DEEP forecasting task on Reddit and news sources.

DEEP performs markedly better on news than on Reddit. Across emotions, precision varies from 0.829 on Reddit to 0.95 on news, recall from 0.786 to 0.916, F1 from 0.797 to 0.93, accuracy from 0.818 to 0.944, and AUC from 0.904 to 0.981. This gap likely reflects the more formal and topically focused language of news articles compared to the noisier and more informal expressions found on social media.

DEEP's Performance varies by emotion but is generally strong. Highly predictable emotions include *Grief* (F1: 0.939 Reddit, 0.935 news) and *Optimism* (0.894, 0.964). Others improve dramatically from Reddit to news: *Joy* (0.779 to 0.961), *Relief* (0.804 to 0.957), and *Amusement* (0.755 to 0.956). Some emotions *Disapproval* are almost perfectly predicted on news (precision = 0.974, F1 = 0.948).

For some emotions, especially on Reddit, DEEP achieves high precision but lower recall. For *Nervousness*, DEEP achieves 0.920 precision but 0.738 recall on Reddit, indicating reliable but conservative predictions. This high-precision benefits journalists by minimizing false positives that could lead to investigating trends that never develop.

One-Week Prediction Performance. We now examine DEEP's performance changes when Δ increases from 1 to 7 days. Figures 2a and 2b report precision trends for the *Increase* and *Decrease* classes on Reddit, while Figures 2c and 2d show the corresponding results for news.

Reddit benefits more from longer horizons than news. Reddit precision improves with increasing Δ for both *Increase* and *Decrease* classes, suggesting emotional shifts are more predictable over medium-term intervals.

News shows variable, emotion-specific horizon effects. For *Increase* predictions, some emotions have high precision (*Admiration*) while others improve (*Approval*, *Sadness*). For *Decrease* predictions, results are mixed: *Approval* benefits from longer horizons while *Admiration* declines, reflecting rapid news cycle shifts in emotional framing.

Overall, our findings indicate that DEEP achieves high precision on News even at short horizons, enabling journalists to act promptly on emerging narratives with confidence. Short-term forecasts are less reliable on Reddit, but predictability improves over longer horizons, making multi-day predictions valuable early indicators of slower-moving, community-driven trends. By leveraging both, journalists can combine immediate, trustworthy signals from news with horizon-informed insights from social media, allowing them to track and predict emotion shifts about SMs such as #MeToo.

	Target	Precision@1		Recall@1		F1@1		Accuracy@1		AUC@1	
		Reddit	News	Reddit	News	Reddit	News	Reddit	News	Reddit	News
Emotion	Admiration	0.881	0.982	0.819	0.952	0.835	0.966	0.854	0.975	0.916	0.983
	Amusement	0.786	0.968	0.749	0.946	0.755	0.956	0.765	0.964	0.885	0.989
	Anger	0.801	0.954	0.759	0.929	0.768	0.940	0.776	0.954	0.887	0.989
	Annoyance	0.843	0.954	0.818	0.947	0.823	0.951	0.826	0.964	0.917	0.996
	Approval	0.701	0.778	0.664	0.776	0.670	0.774	0.680	0.783	0.834	0.917
	Caring	0.884	0.958	0.846	0.923	0.860	0.939	0.872	0.957	0.927	0.988
	Confusion	0.858	0.968	0.815	0.934	0.830	0.950	0.854	0.975	0.951	0.988
	Curiosity	0.779	0.901	0.742	0.871	0.755	0.883	0.754	0.886	0.902	0.970
	Desire	0.830	0.981	0.765	0.929	0.786	0.951	0.819	0.975	0.900	0.992
	Disappointment	0.813	0.961	0.794	0.877	0.794	0.911	0.811	0.936	0.892	0.970
	Disapproval	0.876	0.974	0.817	0.929	0.833	0.948	0.872	0.964	0.944	0.992
	Disgust	0.802	0.976	0.725	0.907	0.736	0.936	0.783	0.961	0.869	0.984
	Embarrassment	0.849	0.978	0.834	0.964	0.839	0.970	0.851	0.972	0.936	1.000
	Excitement	0.810	0.974	0.787	0.958	0.794	0.965	0.794	0.972	0.901	0.998
	Fear	0.829	0.951	0.801	0.932	0.811	0.941	0.819	0.950	0.936	0.994
	Gratitude	0.799	0.979	0.772	0.952	0.781	0.964	0.794	0.968	0.881	0.992
	Grief	0.936	0.965	0.942	0.912	0.939	0.935	0.940	0.947	0.986	0.989
	Joy	0.807	0.984	0.768	0.944	0.779	0.961	0.797	0.975	0.894	0.988
	Love	0.836	0.995	0.774	0.972	0.793	0.983	0.811	0.993	0.881	0.982
	Nervousness	0.853	0.962	0.840	0.845	0.843	0.884	0.861	0.947	0.916	0.987
	Optimism	0.901	0.975	0.888	0.956	0.894	0.964	0.893	0.964	0.957	0.999
	Pride	0.857	0.980	0.738	0.890	0.763	0.925	0.858	0.972	0.864	0.986
	Realization	0.781	0.856	0.729	0.830	0.744	0.840	0.765	0.868	0.864	0.931
	Relief	0.883	0.981	0.781	0.937	0.804	0.957	0.865	0.972	0.890	1.000
	Remorse	0.854	0.972	0.764	0.954	0.786	0.962	0.829	0.964	0.904	0.988
	Sadness	0.793	0.984	0.758	0.979	0.769	0.981	0.794	0.975	0.892	0.996
	Surprise	0.772	0.962	0.732	0.964	0.741	0.963	0.765	0.968	0.854	0.992
	Average	$0.829_{0.047}$	$0.950_{0.057}$	$0.786_{0.054}$	$0.916_{0.058}$	$0.797_{0.052}$	$0.930_{0.057}$	$0.818_{0.051}$	$0.944_{0.057}$	$0.904_{0.033}$	$0.981_{0.027}$
Volume	Raw	0.884	0.864	0.833	0.783	0.852	0.811	0.856	0.817	0.869	0.803
	Velocity	0.656	0.799	0.597	0.779	0.612	0.788	0.612	0.791	0.756	0.811
	Average	$0.770_{0.114}$	$0.832_{0.032}$	$0.715_{0.118}$	$0.781_{0.002}$	$0.732_{0.120}$	$0.799_{0.012}$	$0.734_{0.122}$	$0.804_{0.013}$	$0.813_{0.021}$	$0.807_{0.015}$

Table 2: Macro-averaged performance metrics (Precision, Recall, F1) and overall Accuracy and AUC across Reddit and News sources. The final row shows the mean and standard deviation across all emotions.

Real-World Case Study

The allegations and legal proceedings involving Sean "Diddy" Combs represent a high-profile intersection between celebrity culture and #MeToo[7]. The timeline spans from Cassie Ventura's initial lawsuit in November 2023, alleging years of abuse and coerced participation in drug-fueled "freak-offs", through a cascade of lawsuits, the emergence of video evidence, and Combs' arrest on September 17, 2024. In the subsequent months, the trial brought forward multiple testimonies, including Ventura's detailed account on May 13, 2025, which described coercion, violence, and psychological abuse. While the case ended with Combs being cleared of the most serious charges of racketeering and sex trafficking, he was convicted on two counts of transportation to engage in prostitution.

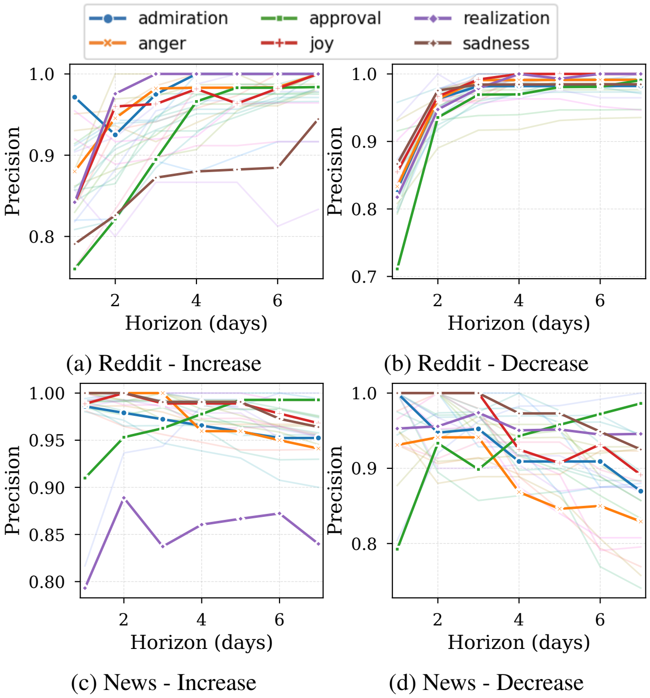
For this brief illustrative case study, we selected two pivotal key events: (KE_1) September 17, 2024 (Combs' arrest) and (KE_2) May 13, 2025 (Ventura's testimony). These events were chosen due to their potential to significantly alter both the volume and emotional tone of public conversation, particularly on social media platforms.

Figure 3 shows the real and DEEP forecast trajectories of two emotions (curiosity and confusion) in Reddit discussions about #MeToo SM (other emotions are in the online Appendix). DEEP used historical data up to each key event (i.e., KE_1 and KE_2) to predict the subsequent week's emotional trends. Ground truth values for the post-event period are plotted alongside predictions for comparison. Significant directional changes in emotion levels during the forecast windows are highlighted to emphasize shifts in discourse following each event.

The results demonstrate DEEP's ability to predict short-term emotional shifts in response to high-impact developments. Both key events were followed by a notable increase in *Curiosity* within the discourse. DEEP's forecasts captured these upward trends, even though the absolute predicted values differed from the ground truth. From a journalist's perspective, correctly predicting the *direction* of change is more important than matching exact magnitudes, as it signals whether public interest is likely to grow or wane in days following a key event. Hence, DEEP's provides timely, accurate trend prediction for newsroom decision-making.

For *confusion*, a more nuanced pattern emerged. Following the arrest on September 17, 2024, confusion levels decreased, potentially reflecting a perceived clarification of the case's trajectory with the formal filing of charges. In contrast, Ventura's testimony on May 13, 2025 was followed by a marked increase in confusion, likely tied to the emotionally complex and contradictory narratives presented during trial. In both cases, DEEP accurately predicted the direction of change, underscoring its ability to model how the interplay of new evidence, legal framing, and public perception shapes online emotional dynamics.

[7]https://www.bbc.com/news/articles/c869qd5j09xo



(a) Reddit - Increase

(b) Reddit - Decrease

(c) News - Increase

(d) News - Decrease

Figure 2: **Precision trends for forecasting emotional changes at varying horizons ($\Delta \in [1, 7]$ days).** (a) Precision for Reddit, *increase* class; (b) Precision for Reddit, *decrease* class; (c) Precision for News, *increase* class; (d) Precision for News, *decrease* class.

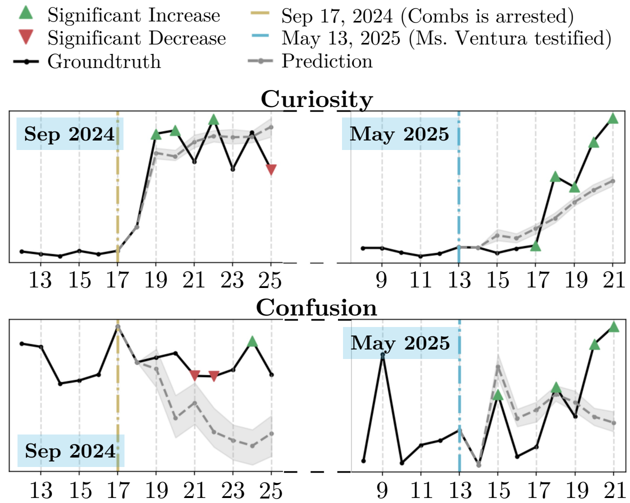


Figure 3: **Case study on Sean "Diddy" Combs**. The top panel shows the temporal trends for Curiosity and the bottom panel for Confusion. In both panels, the solid black line represents the ground truth, while the dashed gray line corresponds to DEEP's forecast. Green and red markers indicate significant increases and decreases in the forecast period, respectively. Vertical lines denote the two key events: Combs' arrest on September 17, 2024 (yellow) and Ventura's courtroom testimony on May 13, 2025 (cyan).

Related Work

Prior research (Ng et al. 2023; Guo et al. 2024b; He et al. 2021) on forecasting social-media discourse has largely focused on predicting activity volume, often leveraging exogenous real-world event signals such as news, GDELT[8], or ACLED[9], but rarely extending to sentiment or discrete emotion forecasting, particularly in a multi-platform setting involving Reddit and news sources. Examples include *ChatterNet* (Dutta et al. 2020), which predicts subreddit "chatter" volume by combining contemporaneous news-derived features with Reddit activity, and the *TAP* family of models (Ng, Horawalavithana, and Iamnitchi 2021, 2022), which use LSTMs to forecast topic-specific daily volumes over short horizons. Similar approaches on Twitter (Yates, Joselow, and Goharian 2021) model news-driven spikes and demonstrate the benefits of exogenous signals for event-aligned bursts in volume. Work on affect dynamics (Hamaker et al. 2015; Sener, Akpinar, and Ataman 2023; Nguyen et al. 2012) constructs discrete emotion time series and uses transformer-based topic modeling to explain shifts (e.g., during Black Lives Matter (Guo et al. 2024a)), but stops short of forecasting. Sentiment is often used as an input feature (Saleiro and Soares 2016; Iamnitchi et al. 2023) but predicting future sentiment has been rare. Across this literature, forecasting is dominated by RNN/GRU/LSTM architectures, and transformer use is limited to topic modeling rather than temporal prediction. No prior work jointly forecasts volume and discrete emotions across Reddit and news in a coordinated, multi-output framework, a gap that our DEEP framework addresses.

Conclusion

This paper presents the Discourse Evolution Engine for Predictions about Social Movements (DEEP) system that enables journalists to forecast volume and emotional changes in social movement discourse across Reddit and news sources. Through experiments on #MeToo data, DEEP achieves strong performance (e.g., F1 scores averaging 0.912 on news, 0.745 on Reddit for emotions classification) with complementary predictive patterns: high precision on news at short horizons and improved Reddit predictions over longer horizons. While designed for social movements, our keyword-based framework generalizes to other topic-based forecasting applications. Developed through engagement with over 20 journalists worldwide, DEEP addresses the need to identify emerging trends before mainstream attention, enabling journalists to anticipate discourse shifts and make more informed editorial decisions.

[8]https://www.gdeltproject.org

[9]https://acleddata.com

Acknowledgements

This research was supported by the Roberta Buffett Institute for Global Affairs at Northwestern University. Additional support was provided by Amazon Web Services (AWS) through the AWS Cloud Credits for Research program.

References

Demszky, D.; Movshovitz-Attias, D.; Ko, J.; Cowen, A.; Nemade, G.; and Ravi, S. 2020. GoEmotions: A Dataset of Fine-Grained Emotions. In Jurafsky, D.; Chai, J.; Schluter, N.; and Tetreault, J., eds., *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 4040–4054. Online: Association for Computational Linguistics.

Dutta, S.; Masud, S.; Chakrabarti, S.; and Chakraborty, T. 2020. Deep Exogenous and Endogenous Influence Combination for Social Chatter Intensity Prediction. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, KDD '20, 1999–2008. New York, NY, USA: Association for Computing Machinery. ISBN 9781450379984.

Guo, S.; He, Z.; Rao, A.; Jang, E.; Nan, Y.; Morstatter, F.; Brantingham, J.; and Lerman, K. 2024a. Measuring Online Emotional Reactions to Events. In *Proceedings of the 2023 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, ASONAM '23, 315–319. New York, NY, USA: Association for Computing Machinery. ISBN 9798400704093.

Guo, S.; He, Z.; Rao, A.; Morstatter, F.; Brantingham, J.; and Lerman, K. 2024b. The Pulse of Mood Online: Unveiling Emotional Reactions in a Dynamic Social Media Landscape. *ACM Trans. Web*. Just Accepted.

Hamaker, E. L.; Ceulemans, E.; Grasman, R. P. P. P.; and Tuerlinckx, F. 2015. Modeling Affect Dynamics: State of the Art and Future Challenges. *Emotion Review*, 7(4): 316–322.

He, L.; Shen, C.; Mukherjee, A.; Vucetic, S.; and Dragut, E. 2021. Cannot Predict Comment Volume of a News Article before (a few) Users Read It. *Proceedings of the International AAAI Conference on Web and Social Media*, 15(1): 173–184.

Hielscher, S.; Wittmayer, J. M.; and Dańkowska, A. 2022. Social movements in energy transitions: The politics of fossil fuel energy pathways in the United Kingdom, the Netherlands and Poland. *The Extractive Industries and Society*, 10: 101073.

Iamnitchi, A.; Hall, L. O.; Horawalavithana, S.; Mubang, F.; Ng, K. W.; and Skvoretz, J. 2023. Modeling information diffusion in social media: data-driven observations. *Frontiers in Big Data*, 6: 1135191.

Levy, R.; and Mattsson, M. 2023. The effects of social movements: Evidence from# MeToo. *Available at SSRN 3496903*.

Magno, F. A. 2017. Environmental movements in the Philippines. In *Asia's Environmental Movements*, 143–175. Routledge.

Ng, K. W.; Horawalavithana, S.; and Iamnitchi, A. 2021. Social-Media Activity Forecasting with Exogenous Information Signals. *CoRR*, abs/2109.11024.

Ng, K. W.; Horawalavithana, S.; and Iamnitchi, A. 2022. Social media activity forecasting with exogenous and endogenous signals. *Social Network Analysis and Mining*, 12(1): 102.

Ng, K. W.; Mubang, F.; Hall, L. O.; Skvoretz, J.; and Iamnitchi, A. 2023. Experimental evaluation of baselines for forecasting social media timeseries. *EPJ Data Science*, 12(1): 8.

Nguyen, L. T.; Wu, P.; Chan, W.; Peng, W.; and Zhang, Y. 2012. Predicting collective sentiment dynamics from timeseries social media. In *Proceedings of the First International Workshop on Issues of Sentiment Discovery and Opinion Mining*, WISDOM '12. New York, NY, USA: Association for Computing Machinery. ISBN 9781450315432.

O'Neil, A.; Sojo, V.; Fileborn, B.; Scovelle, A. J.; and Milner, A. 2018. The# MeToo movement: an opportunity in public health? *The Lancet*, 391(10140): 2587–2589.

Saleiro, P.; and Soares, C. 2016. Learning from the news: Predicting entity popularity on twitter. In *International Symposium on Intelligent Data Analysis*, 171–182. Springer.

Salles, D. M.; Giordani, A. C.; Biagi, A.; de Paiva Affonso, I.; and Fernandes, V. 2024. Social movements and the 2030 Agenda: the correlation between the progressist agendas and the Sustainable Development Goals. *Revista Brasileira de Ciências Ambientais*, 59: e2054–e2054.

Sanh, V.; Debut, L.; Chaumond, J.; and Wolf, T. 2019. DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. *CoRR*, abs/1910.01108.

Sener, B.; Akpinar, E.; and Ataman, M. B. 2023. Unveiling the dynamics of emotions in society through an analysis of online social network conversations. *Scientific Reports*, 13(1): 14997.

Sharma, P.; and Li, Y. 2019. Self-supervised contextual keyword and keyphrase retrieval with self-labelling.

Smith, I. D. 2019. Energy transition and social movements: the rise of a community choice movement in California. In *Renewable Energy: International Perspectives on Sustainability*, 91–129. Springer.

Theriault, N. 2011. The micropolitics of indigenous environmental movements in the Philippines. *Development and Change*, 42(6): 1417–1440.

Wen, Q.; Zhou, T.; Zhang, C.; Chen, W.; Ma, Z.; Yan, J.; and Sun, L. 2023. Transformers in Time Series: A Survey. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence, IJCAI 2023, 19th-25th August 2023, Macao, SAR, China*, 6778–6786. ijcai.org.

Yates, A.; Joselow, J.; and Goharian, N. 2021. The News Cycle's Influence on Social Media Activity. *Proceedings of the International AAAI Conference on Web and Social Media*, 10(1): 735–738.