The Pulse of Mood Online: Unveiling Emotional Reactions in a Dynamic Social Media Landscape

SIYI GUO, Information Sciences Institute, Marina Del Rey, United States
ZIHAO HE, Information Sciences Institute, Marina Del Rey, United States
ASHWIN RAO, Information Sciences Institute, Marina Del Rey, United States
FRED MORSTATTER, Information Sciences Institute, Marina Del Rey, United States
JEFFREY BRANTINGHAM, University of California Los Angeles, Los Angeles, United States
KRISTINA LERMAN, Information Sciences Institute, Marina Del Rey, United States

The rich and dynamic information environment of social media provides researchers, policymakers, and entrepreneurs with opportunities to learn about social phenomena in a timely manner. However, using these data to understand social behavior is difficult due to heterogeneity of topics and events discussed in the highly dynamic online information environment. To address these challenges, we present a method for systematically detecting and measuring emotional reactions to offline events using change point detection on the time series of collective affect, and further explaining these reactions using a transformer-based topic model. We demonstrate the utility of the method by successfully detecting major and smaller events on three different datasets, including (1) a Los Angeles Tweet dataset between Jan. and Aug. 2020, in which we revealed the complex psychological impact of the BlackLivesMatter movement and the COVID-19 pandemic, (2) a dataset related to abortion rights discussions in USA, in which we uncovered the strong emotional reactions to the overturn of Roe v. Wade and state abortion bans, and (3) a dataset about the 2022 French presidential election, in which we discovered the emotional and moral shift from positive before voting to fear and criticism after voting. We further demonstrate the importance of disaggregating data by topics and populations to mitigate potential biases when studying collective emotions. The capability of our method allows for better sensing and monitoring of population's reactions during crises using online data.

CCS Concepts: • **Human-centered computing** → **Social media**.

Additional Key Words and Phrases: Emotional Reactions, Change Point Detection, Topic Modeling

1 Introduction

Social media platforms connect billions of people worldwide, enabling them to exchange information and opinions, express emotions, and to respond to others. Researchers, policy makers, and entrepreneurs have grown interested in learning what the unfettered exchange of information reveals about current social conditions, including using social media data to track public opinion on important issues [26, 33] and monitor the well-being of populations at an unprecedented spatial scale and temporal resolution [46].

Using social media data to learn about human behavior, however, poses significant challenges. Social media represents a heterogeneous, highly dynamic information environment where some topics are widely discussed

Authors' Contact Information: Siyi Guo, Information Sciences Institute, Marina Del Rey, California, United States; e-mail: siyiguo@usc.edu; Zihao He, Information Sciences Institute, Marina Del Rey, California, United States; e-mail: zihaoh@usc.edu; Ashwin Rao, Information Sciences Institute, Marina Del Rey, California, United States; e-mail: mohanrao@usc.edu; Fred Morstatter, Information Sciences Institute, Marina Del Rey, California, United States; e-mail: fredmors@isi.edu; Jeffrey Brantingham, University of California Los Angeles, Los Angeles, California, United States; e-mail: branting@ucla.edu; Kristina Lerman, Information Sciences Institute, Marina Del Rey, California, United States; e-mail: lerman@isi.edu.

ACM 1559-114X/2024/12-ART
https://doi.org/10.1145/3708513

while others are barely mentioned [17]. It includes people's self-reports of their own lives, as well as reactions to external events. Researchers have developed methods to detect events from online discussions [37, 44, 51, 59]. However, social media data provide evidence for learning about human behavior beyond shifts in topics. For example, it can also shed light on emotions and moral sentiments, which are important drivers of individual attitudes, beliefs, and psychological well-being [24, 58].

To study collective affect, researchers have investigated how social media content influences emotional user engagement [2, 6]. These works, however, leave a gap in our understanding of collective emotional and moral reactions to socio-political events, which impacts our understanding of the opinion dynamics, emergence of polarization, and even timely identification of online influence campaigns.

To bridge these gaps, we present a methodology for detecting, measuring and explaining the collective emotional reactions to offline events. Using transformer-based models, we construct the time series of aggregate affect from social media posts. We detect emotional reactions as discontinuities in these time series, and then explain the reactions using topic modeling. We demonstrate the utility of the methodology on three different datasets:

(1) *2020 Los Angeles Tweets*: a dataset collected between January and August 2020, a time span that represents a complex period in American history with important social, political and cultural changes. We successfully detect the simultaneous crises of the COVID-19 pandemic, a racial justice reckoning, and other important events like political primaries. We show how these developments had profound impact on the emotions of the population.
(2) *2022 Abortion Tweets*: a collection of tweets related to abortion rights and the overturning of Roe v. Wade, spanning the entire year of 2022. We uncovered key abortion-related events and discussions, such as the leak of SCOTUS ruling, the overturning of Roe v. Wade, state abortion bans, and the complex emotional reactions to the 2022 US Midterm elections.
(3) *2022 French Election Tweets*: a dataset about the 2022 French presidential election, the timeline of which coincided with other major events, such as the Russia-Ukraine war and the G7 summit. We detect the first and second rounds of voting, as well as the shift in population's emotional responses from positive before each voting round to negative afterwards.

The success of our method on three different datasets indicates its effectiveness and generalizability. We show several benefits of our emotional reaction detection method. First, the method is able to not only detect major events but also capture their complex and multifaceted emotional and moral impacts. Second, as an unsupervised method, it can detect smaller events often overlooked by mainstream media. Third, the method can differentiate events that happened closely in time or even on the same day based on their different emotional reactions. In addition, we demonstrate the value of disaggregating data by topics and subgroups. We show the different emotions expressed in the sub-topics during COVID-19 pandemic, the divergent perspectives to abortion-related events in liberal and conservative populations, and the distinct emotion reactions to different candidates in the French election. These findings emphasize that disaggregating heterogeneous data can help mitigate potential biases. While this study applies our method to Twitter (now called X) datasets, it is also applicable to other social media platforms and news sources. Our results suggest that studying the collective emotional reactions on social media can provide valuable insights into understanding people's opinions and responses to timely socio-political events, and aid policy makers in crafting messages that align with the values and concerns of the population.

2 Related Works

Previous research has utilized social media data in various applications, including predicting elections [28], detecting political manipulation [20], and analyzing social movements [15]. A key direction in this field is event detection, which allows for real-time tracking of unfolding situations and capturing unfiltered public opinions — crucial for responding to emergencies, natural disasters, or rapidly evolving political and social events. Additionally, understanding public reactions to these detected events is essential. Prior studies have examined public responses to the COVID-19 pandemic [32], political events such as elections [27], and wars and conflicts [13, 55], providing valuable insights for policymaking and interventions.

Event Detection: Researchers have developed various methods for event detection on online platforms, including topic detection techniques like Latent Dirichlet Allocation (LDA) [8] and Topic2Vec [45], clustering documents based on textual similarity [37], studying term co-occurrence and performing term frequency analysis [59], and detecting bursty terms [37, 44]. More recent approaches incorporate deep learning techniques. For instance, [4] trains a recurrent neural network to detect breaking news rumors on social media, while [51] employs a deep learning classifier combined with clustering to detect events in a semi-supervised manner. [11] utilizes graph neural networks with contrastive loss terms to adapt to a changing number of event classes, extending knowledge from unseen data. Similarly, [60] applies graph neural networks to model rumor propagation and understand event spread.

While these methods primarily focus on detecting events from social media by identifying abnormal content patterns, it is equally important to understand the dynamics of emotions and sentiments in the aggregate online population. For example, [48] combines event detection with sentiment analysis to improve the detection of event polarity on social media. Our work takes a step in this direction by detecting emotionally and morally charged events, which can significantly impact offline behaviors.

Sentiments and Emotions: Early research on quantifying online emotions largely relied on dictionary-based approaches, measuring sentiment by counting the occurrences of positive or negative words [21, 52]. These studies revealed that the sentiment of aggregated tweets exhibited distinct hourly, diurnal, and weekly mood patterns [18, 21]. Other work demonstrated the feasibility of monitoring population-level subjective well-being at an unprecedented temporal scale and resolution [31, 41]. Social media sentiment analysis has also been applied to study reactions to political campaigns [52], as an alternative to expensive public opinion polls [14], and even to predict stock prices [9] and election outcomes [57].

In studying emotional reactions to offline events, [25] used emojis to analyze online reactions to Brexit. More recent work has conducted emotion analysis of user responses to online news [6], exploring the relationship between news content, emotions, and user engagement. Another study investigated emotional reactions to news articles by predicting user responses before and after publication [2]. These studies focus on how social media content and news articles influence emotional engagement, using metrics like comments, likes, and shares.

Our research differs from these studies in its focus on understanding the online dynamics of emotional and moral sentiments in response to the continuous stream of real-life events. This is critical because affect is closely tied to opinions and often drives offline actions [40]. By combining event detection and emotion analysis, we aim to more effectively detect, measure, and explain emotionally and morally charged reactions on social media.

3 Methods and Materials

To understand the dynamics of affect, we propose a pipeline (Fig. 1) that *detects*, *measures* and *explains* online emotional and moral reactions to offline events. With a set of timestamped texts, e.g. tweets, we first perform emotion and morality detection from text. We then construct the time series of the aggregate affect on a daily basis. Next, to detect reactions, we perform change point detection on each emotion and morality time series. We measure the magnitude of the change at each detected change point and perform topic modeling to explain the offline event that triggered the specific online reaction.

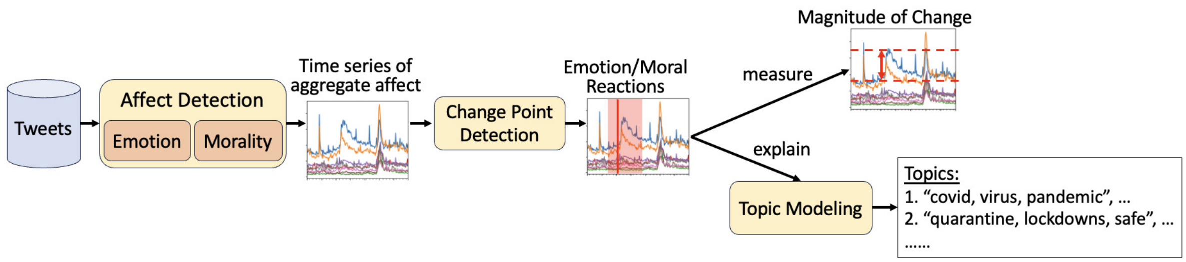

Fig. 1. Pipeline to detect and measure online emotional reactions.

3.1 Data

3.1.1 2020 Los Angeles Tweets. The data was collected using Twitter's Filter API by specifying a geographic bounding box over Los Angeles County. This method collects every tweet that is either geotagged within the bounding box (using the device's coordinates with the user's permission), or by using the Twitter "place" feature, where the user tags their location. We collected 17M tweets from 350K unique users. The dataset includes a wider range of topics and authors than data collected using keywords typically do.

3.1.2 2022 Abortion Tweets. The data comprises tweets about abortion rights in the U.S. and the overturning of Roe vs Wade [12]. We filter English tweets posted within U.S. over the period of an entire year of 2022. Each tweet contains at least one term from a list of keywords that reflect both sides of the abortion debate, such as "roevswade", "prochoice", "MyBodyMyChoice", "prolife", and "ValueLife". The data include 12M tweets generated by 1M users.

3.1.3 2022 French Election Tweets. This is a corpus of 6M tweets about the 2022 French presidential election. The tweets were collected by querying Twitter with a set of keywords related to the election: e.g., "election", "élection", "l'élection", "Elysee 2022", "Elysee2022", etc. 90% of the tweets were in French, and the rest were in English and other languages.

3.2 Emotion and Morality Detection

We first measure emotions and moral sentiments expressed in an individual tweet. For emotions, we use a transformer-based language model SpanEmo [3], fine-tuned on the SemEval 2018 1e-c data [42]. This model outperforms prior methods by learning the correlations among the emotions. It measures *anticipation*, *joy*, *love*, *trust*, *optimism*, *anger*, *disgust*, *fear*, *sadness*, *pessimism* and *surprise*.

We quantify the moral sentiments of tweets along five dimensions [24]: dislike of suffering (*care*/*harm*), dislike of cheating (*fairness*/*cheating*), group loyalty (*loyalty*/*betrayal*), respect of authority and tradition (*authority*/*subversion*), and concerns with purity and contamination (*purity*/*degradation*). We fine-tune a transformer-based model on diverse training data (see [23] for details). The large amount and the variety of topics in our training data helps mitigate the data distribution shift during inference. For both emotion and morality detection, we use the multilingual XLM-T [7] as our base transformer model on the 2022 French Election data. For other English datasets, we use "bert-base-uncased" as the base model. After labeling tweets, we calculate the daily fractions of tweets with different emotion and moral categories to construct the time series.

Evaluation. We evaluate the effectiveness of emotion and morality detection on a random subset of 850 tweets from the Los Angeles dataset. We asked five educated annotators to go through two training sessions, wherein they annotated 50 random tweets within one week and discussed how to improve the agreement on the definitions of emotions and morality. Then each of the five annotators individually labeled 510 randomly picked tweets, giving us three annotations per tweet. The duration of the final round annotation is 20 days, allowing enough time and flexibility for the annotators to pay their full attention. Given the subjectivity of moral and emotional judgments, the Fleiss's κ for emotion categories ranges in 0.42 ± 0.02. For morality categories, it ranges in 0.30 ± 0.03. Similar to prior works [30, 42], we have found the κ scores of some categories to be low. However, our agreement is still comparable to, and on some categories even better than, these prior works.

Table 1. Evaluation of Emotion and Morality Detection Methods.

Emotion	Fleiss's κ	F1-Score		Support	Morality	Fleiss's κ	F1-Score		Support
		Emolex	Ours				DDR	Ours	
anger	0.49	0.28	**0.45**	93	care	0.29	**0.54**	0.47	63
anticipation	0.32	0.12	**0.40**	43	harm	0.28	0.18	**0.36**	60
disgust	0.46	0.33	**0.53**	116	fairness	0.17	0.17	**0.25**	10
fear	0.22	0.00	**0.37**	113	cheating	0.28	0.20	**0.43**	31
joy	0.48	0.25	**0.37**	113	loyalty	0.18	0.00	**0.05**	8
love	0.66	N/A	**0.71**	122	betrayal	0.01	0.00	0.00	1
optimism	0.37	0.15	**0.34**	58	authority	0.45	0.00	**0.32**	24
pessimism	0.26	**0.20**	0.08	82	subversion	0.60	0.15	**0.37**	78
sadness	0.53	0.15	**0.45**	66	purity	0.49	0.27	**0.56**	35
surprise	0.51	0.06	**0.44**	33	degradation	0.22	0.21	**0.33**	23
trust	0.30	0.00	**0.43**	22					

We compare our emotion and morality detection methods with widely used dictionary-based methods, namely keyword matching using Emolex [43] (does not include "love" category) for emotions, and Distributed Dictionary Representations (DDR) [19] for morality. On the Los Angeles dataset, our methods outperforms baselines on ten out of 11 emotion categories and we outperform on nine out of ten moral categories (Table 1). On 2022 French Election Tweets, similar evaluation was performed based on human annotation. The emotions had an average F1-score of 0.66, and the moral categories had an average F1-score of 0.58. On the 2022 Abortion data, validation was given by Rao et al. [49]. We notice that the model performance inevitably varies with support for different categories, as also observed in previous studies [30, 56]. Despite some variation in model performance, prior research [46] has validated that when aggregating on the collective level, the time series of sentiments constructed with supervised deep learning detection have strong correlations with those from self-reports.

3.3 Detecting and Measuring Changes

3.3.1 Change point detection. The time series of emotions and morality reveal the complex dynamics of aggregate affect on social media. We define an emotional reaction as a change in the corresponding time series. To detect such change points, we combine two popular methods. The first, cumulative sum (CUSUM) method [29], detects a shift of means, and is good at detecting changes like the COVID-19 outbreak, which shifted the baseline emotion and moral sentiment. To detect multiple change points, we use a sliding window to scan the whole time series. We set the window size to be four weeks and slide it every three days for the best precision. Another type of event, such as Valentine's Day, creates a short surge of emotions, can be better detected with Bayesian Online Change Point Detection (BOCPD) [1]. It uses Bayesian inference to determine if the next data point is improbable, which is good at detecting sudden changes. We identify a change point to be significant when either CUSUM or BOCPD gives a significant confidence score, using 0.5 as the threshold. We perform change point detection separately for each time series of emotion and morality, because different types of events may elicit different reactions.

3.3.2 Interrupted time series analysis. For each detected change point, we quantify the magnitude of the collective reaction in two aspects, the short-term and the long-term changes. To calculate the short-term change, we use interrupted time series analysis [47, 53]. On the time series of daily fraction of tweets expressing each emotional and moral category, we select the time window from seven days before the event to three days after, as discussions on Twitter usually die down within a short time [35]. We perform linear regression for each change point on each time series:

$$y = \beta_0 + \beta_1 t + \beta_2 \mathbb{1}_{after\ event} + \beta_3 t * \mathbb{1}_{after\ event},$$

where y is the daily fraction of tweets in an emotion/morality category, t is the time variable, and $\mathbb{1}_{after\ event}$ is a binary indicator which equals 0 before a change point and 1 after the change point. To represent the change associated with an event, we use $\frac{\beta_3}{mean(y)}$, the coefficient for the interaction term normalized by the mean of this segment of time series, and convert it to percentage. We also report the p-value from the regression along with the change.

To measure long-term changes, we compute the baseline level as the mean of the time series over seven days before the change point. Then we compare the baseline to the time series value two weeks after the event (we take a five-day average around the two-week mark). The size of the window is empirically chosen to be two weeks so that enough observations are made, but it would not be affected by another event earlier or later. We report the accompanying p-value from Student's T-test.

3.4 Explaining Changes with Topic Modeling

We try to explain changes in emotions detected by our method using topic modeling. We choose BERTopic [22], a transformer-based language model that extracts highly coherent topics compared to traditional LDA. We evaluate both methods on a set of 10% randomly selected tweets from our data, using a different numbers of topics ranging from 10 to 50 in steps of 10. Over different runs, BERTopic gives higher NPMI [10] coherence scores (0.14 ± 0.01) compared to LDA (0.03 ± 0.01), and similar diversity [16] scores (0.75 ± 0.04) compared to LDA (0.76 ± 0.04).

For each emotional reaction, we extract the topics of tweets that are tagged with that emotion or morality category. We apply BERTopic to tweets within the three-day time window before and after the change, as discussions quickly die on social media [35]. For example, for the Black Lives Matter (BLM) protests starting on 2020-05-26, we extract the topics from tweets posted between 05-23 to 05-25 to develop a baseline and then separately extract the topics between 05-26 to 05-28. By comparing the top 10 baseline topics before the change point with those after the change point, we determine the new topics that emerged after the change points that are possibly relevant to the event. See Table 6 in Appendices for examples.

For preprocessing, we remove URLs and name mentions, transform emojis to their textual descriptions, and split hashtags into individual words. We use the Sentence-BERT "all-MiniLM-L6-v2" model [50] to directly embed the processed English tweets, and "sentence-camembert-large" [38] for French tweets. After topic modeling, we remove stopwords in the learned topic keywords. With each emerging topic, we manually verify if there is an associated offline event by examining the tweets belonging to this topic and by searching related news articles. Such manual verification is a necessary and common practice event detection literature [44].

3.5 Evaluation of the Proposed Method

Similar to previous work [36, 44], we evaluate our method using precision, Duplicate Event Rate (DERate), and the number of detected events. Precision refers to the fraction of detected events that correspond to real-world events [5]. We manually verify each detected event by searching for related news using topic keywords associated with each change point, a common practice in event detection research. A false positive is a change point that cannot be linked to any identifiable topic or real event. The second metric, Duplicate Event Rate (DERate), measures the percentage of duplicate events among all detected real-world events [36]. We define DERate as the fraction of emotion and morality categories (out of 21) that detect the same event. A higher DERate indicates greater confidence in the detection. Lastly, we do not use recall because it is impractical to annotate every tweet to create an exhaustive list of events. Instead, following [44], we use the number of detected events as a proxy for retrievability.

We first evaluate our method on the benchmark dataset Events2012 [39], a Twitter corpus consisting primarily of everyday events with varying degrees of impact. The dataset is not filtered by any specific theme (e.g., elections) and does not include any global event (e.g., a pandemic) which might dominate or bias the event detection results, making it an ideal evaluation dataset. Following [44], we select a subset from October 11 to October 17 and compare our method to SEDTWik. Our precision and DERate are comparable to SEDTWik (Table 2). Moreover, *we detect more events than SEDTWik*, ranging from major events such as Nobel Prize announcements to smaller events like football games, demonstrating strong retrievability even for minor events.

Table 2. Comparison of Proposed Method with SEDTWik on benchmark data Events2012 during Oct 11 - Oct 17, 2012.

	Precision	DERate	# Change Points	Change Point Confidence
Ours	0.87	0.12	86	0.97 ± 0.10
SEDTWik	0.88	0.14	79	N/A

Our method, however, differs fundamentally from traditional event detection approaches like SEDTWik, which rely on detecting spikes in word or phrase frequencies. While these methods aim to detect as many events as possible, they often capture numerous minor, mundane events that can introduce noise in downstream research. In contrast, our method excels at identifying emotionally and morally charged events, which are more relevant for studies focused on the relationship between offline events and online emotional expression (see Table 3 for examples detected by our method but not by SEDTWik).

We further evaluate our method on the three real-world datasets we analyze. Table 4 highlights its strong performance. When applied to these larger datasets which contain millions of tweets, our method tends to prioritize more impactful events, while smaller, mundane topics are excluded by BERTopic due to their lower frequency. This serves as an effective noise-filtering mechanism.

Table 3. Examples of Events Detected by Our Method But Not by SEDTWik.

Events	Changing Emotions/Moral Sentiments
10-11 The first private clinic to offer abortions to women in northern ireland was opening.	care, anticipation
10-14 Mauritania President Was Shot Accidentally by Soldier.	care, harm, betryal, subversion, anger, fear
10-15 The Philippines was to sign a peace plan with its largest Muslim rebel group AIME.	care, anticipatioin
10-17 Cambodia's former king Norodom Sihanouk die	care, authority, pessimism, sadness

Table 4. Evaluation of Our Proposed Method on 3 Real Datasets

Dataset	Precision	DERate	# Change Points	Change Point Confidence
2020 LA Data	0.84	0.18	54	0.94 ± 0.12
2022 French Election Data	0.83	0.16	40	0.85 ± 0.19
2022 Abortion Data	0.82	0.16	99	0.96 ± 0.12

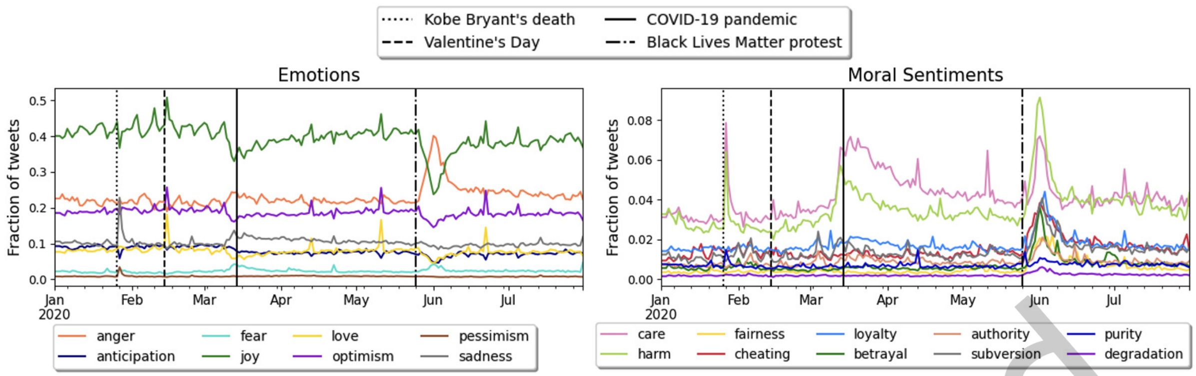

Fig. 2. Time series of emotions and moral sentiments in Los Angeles Tweets from January 1 to August 1, 2020. We show the daily fraction of tweets with different affect labels. The notable peaks and dips in the time series can be associated with the external events marked as vertical lines.

4 Disentangling Socio-political Events in 2020 Los Angeles

The year of 2020 was a particularly challenging period for the city of Los Angeles. In addition to the world-wide pandemic, which led to a national lockdown mid-March, political primaries were also taking place during this time period, which also saw one of the largest social justice protests triggered by the murder of George Floyd in police custody, as well as the death of a beloved sports icon. These developments had a profound impact, as demonstrated by the many rises and dips in emotions and moral sentiments. Time series of the aggregate affect from January to August 2020 (Fig. 2) shows complex dynamics with seasonal variation (weekly cycles in *joy*), short-term bursts (spike in *love* on Valentine's Day), and long-term changes in emotions and moral sentiments.

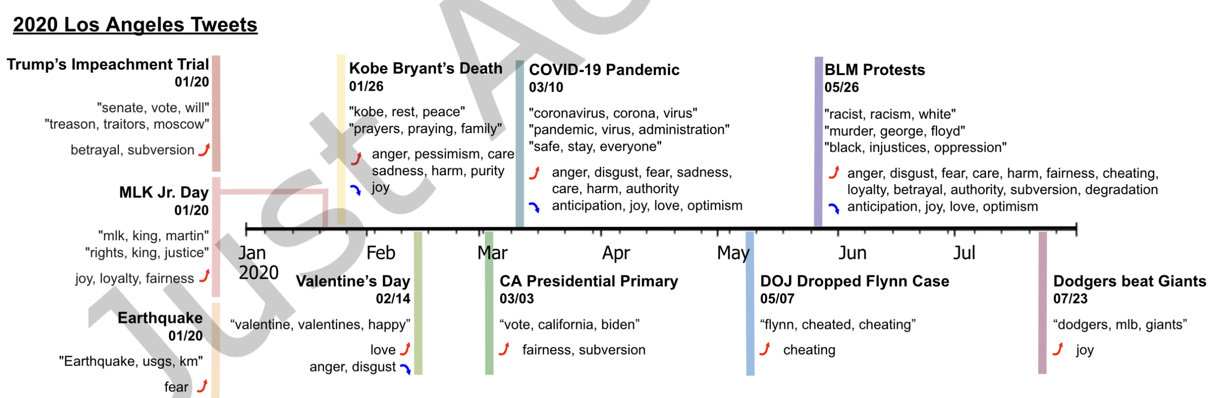

Fig. 3. Top events, their associated topics and corresponding emotional and moral reactions detected in 2020 Los Angeles data. See the full list of events in Table 7 in Appendices.

We ran the proposed pipeline to detect and explain the online emotional reactions to events (see Table 7 in Appendices for a full list of detected events). Figure 3 shows that our method is able to identify key events such as the COVID-19 pandemic and the BLM protests. We see the complex reactions to the pandemic along the multiple dimensions of emotions and morality. The unsupervised method also enables us to discover reactions to smaller events that might be easily missed, such as earthquakes and baseball playoffs. We also show that BERTopic reveals highly relevant discussion topics. In addition, because we detect changes separately in each emotion, we can disentangle events based on different emotional reactions, even when they take place on the same day: Trump's impeachment trial was associated with an increase in *betrayal* and *subversion*, MLK Day with *joy*, *loyalty* and *fairness*, and an earthquake with *fear*.

Our proposed method enables us to study collective reactions to events along multiple dimensions of affect. For example, the BLM protests were associated with 16 different emotional and moral changes. We quantify the short-term and long-term percent change in the corresponding collective affect before and after the event for four of the most impactful events (Fig. 4). Consistent with our intuition, Kobe Bryant's Death was associated with a short-term increase in *pessimism* and *sadness* and a decrease in *joy*, as well as a short-term rise in moral language related to *care* and *harm*. In contrast, Valentine's Day brought a short-term increase in *love* and a decrease in *anger* and *disgust*. No long-term changes were seen with these events. On the other hand, the BLM protests was associated with complex short- and long-term changes in affect. We observe increases in negative emotions and decreases in positive emotions. In addition, compared to other three events, we see greater increases in moral sentiments. The moral concerns about *fairness* and *betrayal* had especially increased, expressing a deep sense of the injustice and betrayal in George Floyd's death.

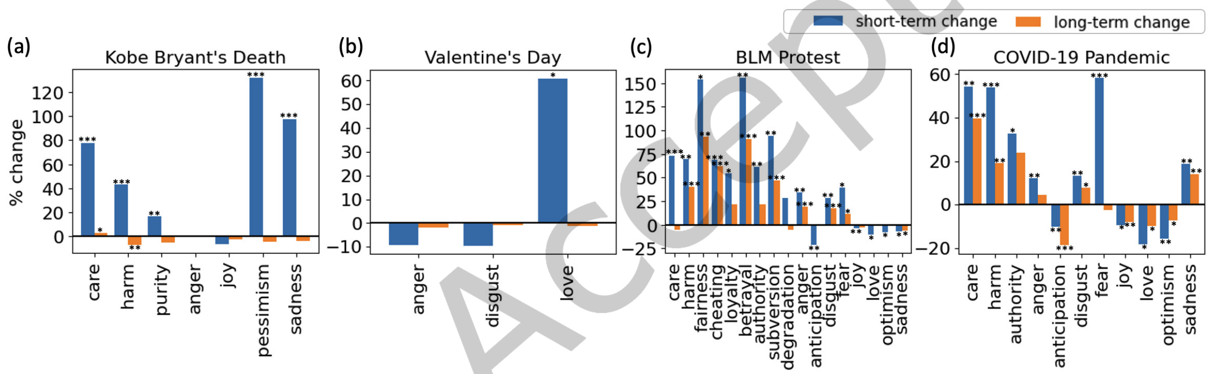

Fig. 4. Short-term and long-term changes of emotions and moral sentiments around four events in 2020 Los Angeles data. Asterisks indicate significance values: * (p-value < 0.05), ** (p-value < 0.01), *** (p-value < 0.001), and no asterisk indicates p-value ≥ 0.05.

The COVID-19 outbreak triggered a cascade of events aimed at mitigating the pandemic that were associated with complex short-term and long-term changes in affect (Fig. 4d). People expressed more *anger*, *disgust*, *sadness*, and more significantly, *fear*. Positive emotions like *joy* and *love* simultaneously decreased, both in the short-term and the long-term. People also expressed more moral sentiments like *care* such as in "Stay safe. We thank you", as well as more *harm* blaming the virus. Interestingly, the moral language around *authority* also increased, possibly due to new policies such as lockdowns to mitigate the pandemic (e.g. "I think governor Newsom is doing a great job..."), and some were critical of government's response, e.g., "we need leadership not a politician". Because Twitter users are predominantly liberal, and this dataset is collected in Los Angeles with 91% users to be liberal and 9% conservative, we found the emotional and moral reactions to these events reflecting liberal perspectives.

Next, we take a deep dive in the COVID-19 emotion analysis, and show the benefit of disentangling emotional reactions by disaggregating topics. We select four top categories discussed and group related topics into these categories: directly covid-related topics, grocery panics, leisure activities and school and education. We study emotions and moral expressions aggregated in all the tweets, as well as in these topic categories (Fig. 5). We

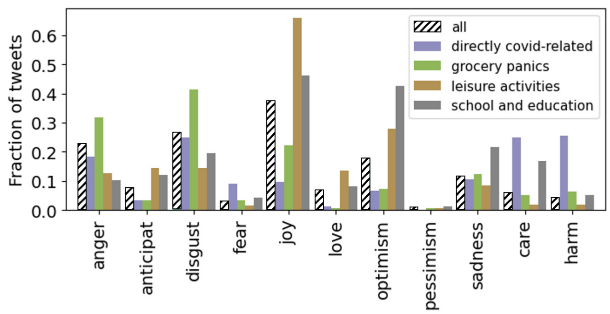

Fig. 5. Emotions and moral sentiments expressed in COVID-related topics during the two weeks after WHO announcement of the pandemic on 2020-03-11. The topics are COVID ("coronavirus, corona, virus"); grocery panics ("grocery, groceries, shelves", "water, dasani, hydro", and "toilet, paper, rolls"); leisure activities ("episode, episodes, show", "cook, cooking, cookout","tickets, ticket, selling"); and education ("teachers, students, learning", "schools, lausd, classes", "schools, lausd, closed").

find that aggregating emotions from all tweets can give misleading impressions. Positive emotions like *joy* were mostly expressed in all tweets (aggregated), but in fact they were mostly dominated by people talking about leisure activities. In COVID-related tweets, few positive emotions were expressed. *Anger* and *disgust* were higher in topics about grocery panics than in topics directly related to COVID. Another example is the expression of *care* and *harm* moral sentiments. Their expression was diluted by other topics in aggregate tweets. After disaggregating, they were largely expressed in directly COVID-related tweets. These results suggest that during times of maximal crisis and uncertainty, people find outlets for positive emotions. They also demonstrate the importance of disaggregating by topics when we study specific issues like COVID-19 that cover a multitude of fine-grained sub-topics.

5 Unfolding the Evolution of Abortion Rights Discussion

Abortion is one of the most politically charged issues in the U.S. The debate was especially intense in 2022, when the Supreme Court of the United States (SCOTUS) struck down federal protections for abortion rights in its Dobbs v. Jackson Women's Health Organization ruling on June 24, 2022. This decision overturned the nearly 50 year old precedent set by the Roe vs Wade decision, which guaranteed women in U.S. access to abortion. In the 2022 Abortion Tweets data, we have identified a total of 24 different events, unfolding the evolution of abortion discussions on Twitter (Figure 6). See a full list of events detected in Table 8. We detected major events, such as the leak of SCOTUS ruling on May 3rd, to which people expressed over 1000% more *surprise* (Fig. 7b), and the overturning of Roe v. Wade on June 24th, followed by a surge of strong emotional reactions, including increasing *anger* and *disgust*, and a dip in *care* and *optimism* (Fig. 7c). This finding is consistent with the previous work by [49]. Along with the SCOTUS ruling, multiple states had issued abortion bans, which further provoked online debates. We detected the issuing of abortion bans in Florida and Oklahoma on March 4th and April 5th, respectively, and observed surging negative moral sentiments, such as *harm*, *cheating* and *subversion*. In the contrast, when Kansas voted to keep abortion legal on August 3rd, *anticipation* and *trust* was expressed.

This historical SCOTUS ruling did not only elicit temporary emotional reactions, but also had long term effects. We detected multiple protests as well as police arrests of protesters, accompanied by online discussions expressing *betrayal*, *subversion*, *fear* and *sadness*. There were also viral stories of individuals experiencing health and legal crises because of the ruling. One example is a woman in Texas who was denied an abortion even with diagnosis of fatal fetal abnormalities. Increasing *surprise*, *sadness* and *subversion* were detected in these discussions (Figure 6 and Table 8).

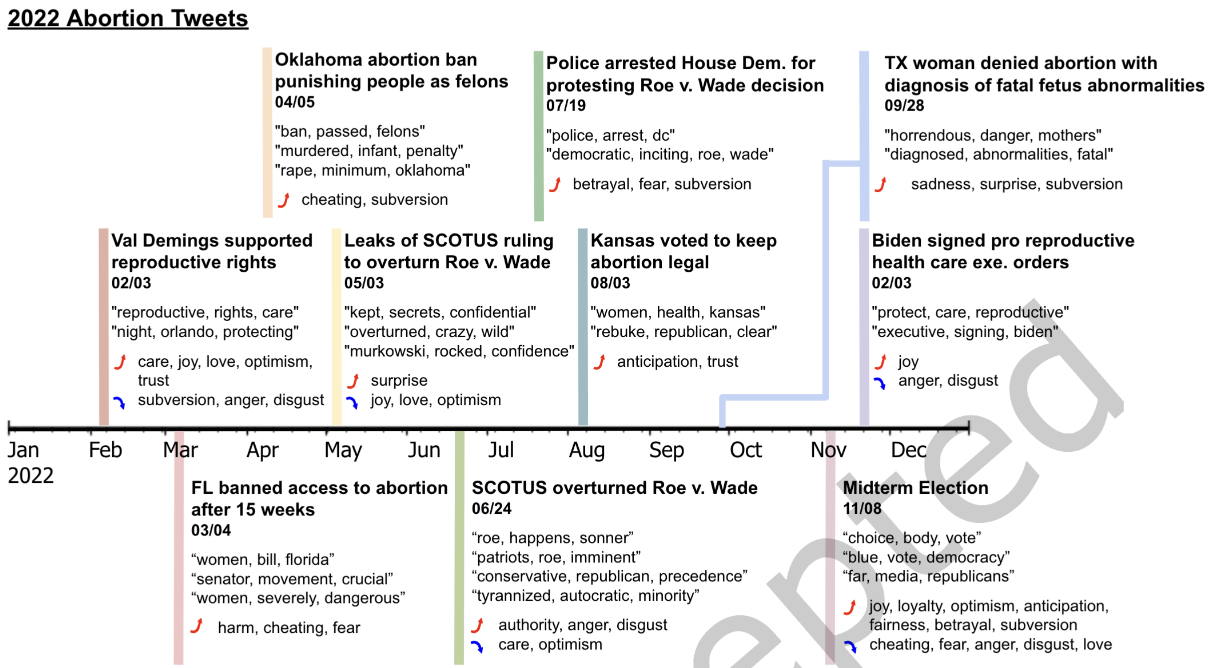

Fig. 6. Top events, their associated topics and corresponding emotional and moral reactions detected in 2022 Abortion data. See the full list of events in Table 8 in Appendices.

Another important and long-term impact of this ruling was in the 2022 US Midterm Elections. Many congressmen pushed out agendas related to abortion rights. Some Democrats such as Biden and Val Demings stated their support for abortion rights (Figure 6). In response, we detected increasing *care*, *love*, and *trust* in the dataset. On the other hand, some Republicans implied a pro-life stance (e.g. event 3 and 11 in Table 8). When the Midterm Election happened on November 8th, we observed complex emotional and moral reactions (Fig. 7d). In general, positive emotions and moral sentiments such as *loyalty*, *anticipation* and *optimism* increased, whereas negative emotions decreased, indicating people were hoping for their favored election results. The 24 events we detected in this dataset include major and momentous events, small events such as people re-sharing a viral tweet, as well as events that took place closely in time and even on the same day (Table 8). This demonstrates again the effectiveness of our proposed method to automatically detect emotional reactions.

However, we observe that the emotional and moral reactions to most of these events are predominantly aligned with left-leaning beliefs. The dataset is largely composed of liberal users: using validated political ideology labels from [49], we find that 72% of users are liberal, while 28% are conservative. To obtain a less biased understanding of emotional reactions, we disaggregate the data into liberal and conservative groups and apply our method separately to each. This approach reveals distinct emotional dynamics and event detection between the two populations. Figure 8a shows the change points detected in each emotion or moral category. Liberals and conservatives have different emotional and moral reactions toward different events. Table 5 provides examples of events detected exclusively within each group. These examples clearly depict the overall pro-abortion perspectives among liberals and the pro-life perspectives among conservatives. Furthermore, even in response to the same event, the two groups express contrasting emotions and moral language. For instance, after the SCOTUS ruling on June 24, conservatives expressed significantly more authority, anticipation, and less fear, while

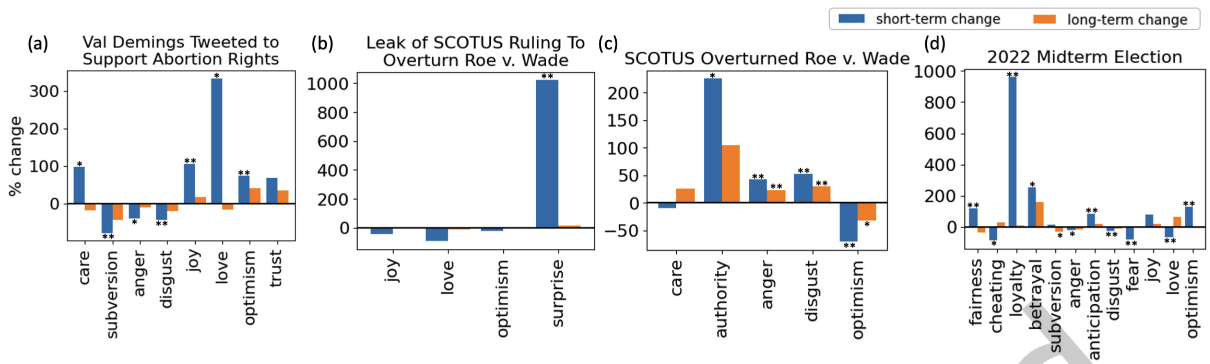

Fig. 7. Short-term and long-term changes of emotions and moral sentiments of some top events in 2022 Abortion Tweets. Asterisks indicate significance values: * (p-value < 0.05), ** (p-value < 0.01), *** (p-value < 0.001), and no asterisk indicates p-value ≥ 0.05.

liberals exhibited higher levels of anger, disgust, and less optimism (Figure 8b). In another example (Figure 8c), liberals and conservatives framed the news "Nebraska cops used Facebook messages to investigate an alleged illegal abortion" in starkly different ways. Conservatives focused on details of the abortion and expressed more fear, whereas the liberals focused on the breaches of privacy by Facebook and expressed more cheating and betrayal. These divergent reactions highlight the potential biases hidden in heterogeneous data and underscore the benefits of data disaggregation [34].

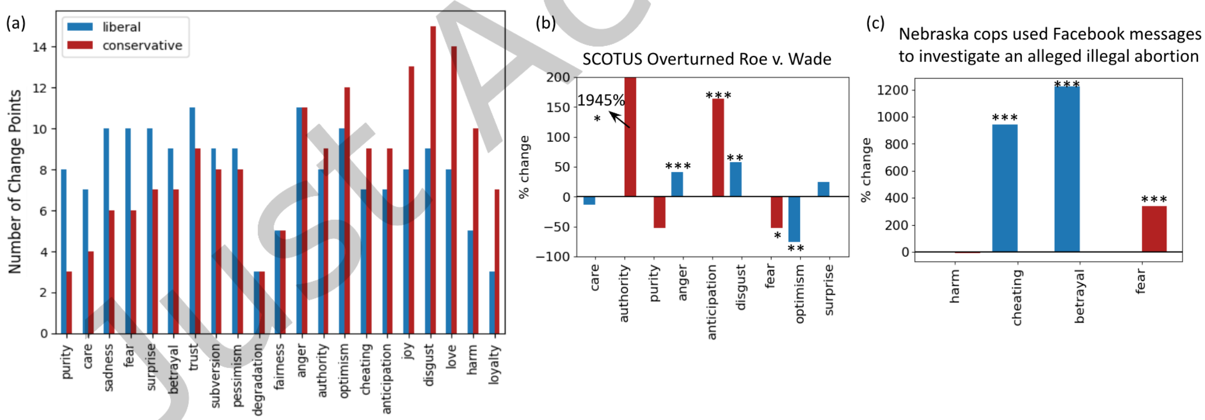

Fig. 8. (a) Number of change points detected in each emotion or moral category in liberal and conservative populations are different. The two populations also have different short-term emotional and moral reactions to (b) SCOTUS overturning Roe v. Wade on June 24 and (c) Nebraska cops using Facebook messages to investigate an alleged illegal abortion on August 9, 2022. Asterisks indicate significance values: * (p-value < 0.05), ** (p-value < 0.01), *** (p-value < 0.001), and no asterisk indicates p-value ≥ 0.05.

Table 5. Examples of Emotional and Moral Reactions Detected Only in Liberals and Conservatives.

Examples of Reactions in Liberals	Examples of Reactions in Conservatives
01/18 Taliban crushed Afghan women's rights protest.	02/23 An Illinois woman's uterus was torn in a botched abortion.
02/03 Val Demings tweeted to support abortion rights.	03/11 Texas Supreme Court ruled unanimously against abortion businesses.
06/14 Florida synagogue sued new abortion ban.	07/09 A pregnant woman in Dallas ticketed for driving alone in the HOV lane argued that her unborn child should count.
09/28 Texas woman denied an abortion after diagnosis of fatal fetal abnormalities.	09/24 FBI raided home of a pro-life speaker in PA with guns drawn in front of kids.
11/19 Biden signed 2 orders to protect access to reproductive health care.	

6 Understanding Emotion Dynamics in the 2022 French Election

Our third case study is the 2022 French Election, mainly among Emmanuel Macron, Marine Le Pen and Jean-Luc Mélenchon. This election took place close in time to the Russia-Ukraine War, which started in February 2022, and the G7 Summit in June, adding complications to the study of emotional reactions in the online population. Previous reports have shown that the Russia-Ukraine war had significant implications for the election campaigns, especially a positive effect on Macron's polling [54]. On the Twitter dataset, we are able to detect both the first and second rounds of presidential elections on April 10th and April 24th, as well as the two rounds of legislative elections on June 12th and June 19th. In addition, we have also detected the G7 summit and various events related to the Russia-Ukraine war, including the initial invasion, the sinking of a Russian warship, and Russia cutting off natural gas to East Europe (Figure 9 and Table 9). The online discussions about the election also involved topics about the Russia-Ukraine war, such as "ukraine, conflict, economics".

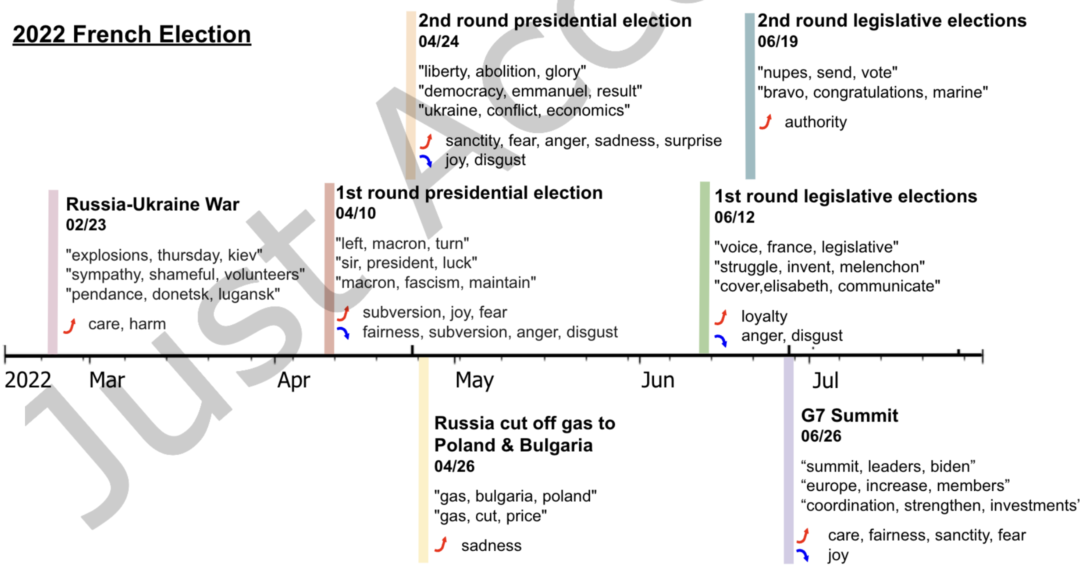

Fig. 9. Top events, their associated topics and corresponding emotional and moral reactions detected in 2022 French Election data. See the full list of events in Table 9 in Appendices

Next, we analyze the dynamics of emotions and moral sentiments during the presidential election cycle. The change points in different emotion and moral categories are detected in a relatively wide range of dates, from

04/07 to 04/15 for the first round, which actually happened on 04/10, and from 04/21 to 04/26 for the second round, which took place on 04/24. This indicates the convoluted influence of the voting rounds. We found an interesting pattern, that *positive emotions and moral sentiments increased before each voting round, but negative sentiments surged up right after the voting round.* This is reflected in the time series with the zigzag shapes in Figure 10a. For example, *anger* and *subversion* first surged then dipped between two rounds, whereas *joy* moved in the opposite way, dipping first and then surging between the two rounds. This indicates that people were showing hope and support before each voting round, but started to reflect and criticize after the voting finished and results came out. Fig. 10b also reveals similar patterns. The most interesting example is the changes in subversion. There were two change points detected in subversion. It decreased by 52.77% on 04/08 before the first voting round, but significantly increased by 107.12% on 04/15 after, and then dropped again right before the second round election started. These change points are consistent with the time series pattern shown in Fig. 10a. In addition, we also observe a significant increase in fear after each voting round. These again show people's hope before voting, and negativity and fear after voting results revealed.

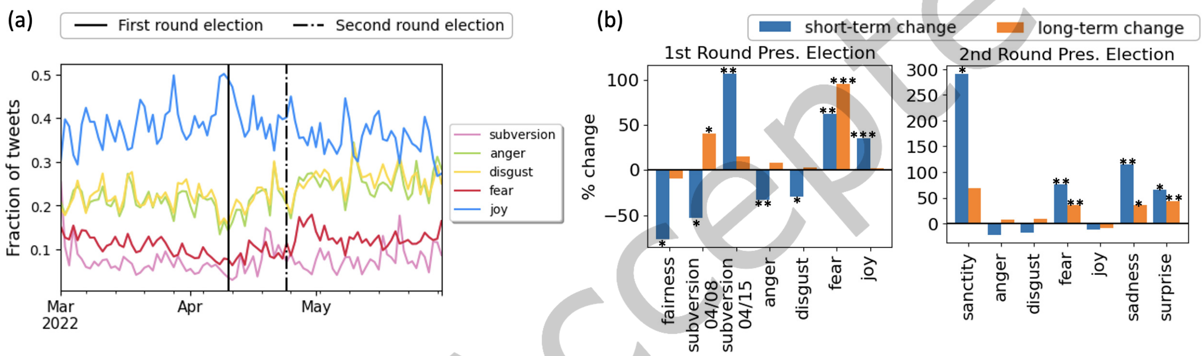

Fig. 10. (a) Time series of emotions and moral sentiments with significant changes during the 2022 French Election. (b) Short-term and long-term changes of emotions and moral sentiments around first and second rounds of 2022 French Election. Asterisks indicate significance values: * (p-value < 0.05), ** (p-value < 0.01), *** (p-value < 0.001), and no asterisk indicates p-value ≥ 0.05.

Lastly, to further examine how emotional reactions varied across the different presidential candidates, we disaggregated the dataset using relevant hashtags into discussions about the top four candidates — Emmanuel Macron, Marine Le Pen, Jean-Luc Mélenchon and Éric Zemmour (see Appendix B for detailed method). Figure 11a shows the short-term emotional shifts in discussions following the first round of voting for each candidate. To better understand the nature of these discussions, we also present word clouds built from all tweets in each subset within one week after the first round (Figure 11b). In the first round, Macron led with 27.9% of the vote, followed by Le Pen with 23.2%, Mélenchon with 22%, and Zemmour with 7.1%. This result meant that Macron and Le Pen advanced to the second round. In discussions about Macron, people expressed increased loyalty, sanctity and joy implying the supportive emotions, but also increase negative emotions like fear and anger, showing people's mixed sentiments. Words like "glory" and "pure" verify the significantly increased sanctity. In discussions about Marine Le Pen, there were heightened expressions of care/harm, sanctity/degradation, anger, disgust and surprise, suggesting that some were surprised or dissatisfied with her second-place finish. For Mélenchon, the increase in loyalty, sanctity, and joy indicated that his supporters were content despite his third-place outcome. Meanwhile, for Zemmour, most emotional categories, as well as the total number of relevant tweets, decreased after the first round, as it became clear he would not advance to the second round.

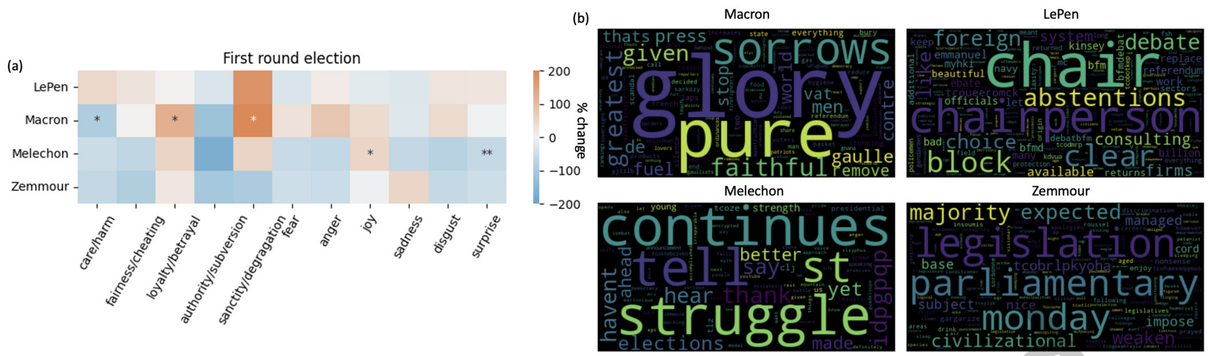

Fig. 11. (a) Short-term changes of emotions and moral sentiments around first round election in subsets relevant to four candidates. Asterisks indicate significance values: * (p-value < 0.05), ** (p-value < 0.01), *** (p-value < 0.001), and no asterisk indicates p-value ≥ 0.05. To have enough data, we combine the virtue and vice of each moral foundation (e.g. care and harm). (b) Word clouds generated using tweets within one week after the first round election in corresponding subsets.

In the French Election dataset, we have successfully detected all events related to the election and those related to the war. We are able to disentangle the complex dynamics of emotional and moral changes, benefiting from the proposed change point detection method. Furthermore, we have also shown the different reactions to different presidential candidates after the first round election by disaggregating the dataset.

7 Conclusion

In this work, we have demonstrated the effectiveness of an unsupervised method to detect and measure public reactions to newsworthy events. We applied our method to three large Twitter datasets. We have disentangled the dynamics of online emotions during a time period punctuated by complex social, health, and political events in the 2020 Los Angeles data, studied the evolution of abortion rights discussion along with the overturning of Roe v. Wade in 2022, and revealed the interesting dynamics of emotions during the 2022 French election. We showed that our method can discover major and subtle events, including those that occur closely in time, and effectively measure emotional and moral reactions to these events. Additionally, we highlighted the importance of disaggregating data by topics and heterogeneous populations to gain a deeper understanding of the complex impacts of significant events. Together, these results suggest the potential of using social media data for sensing and tracking of public reactions to events, as well as discovering significant events that may have been missed by traditional news sources.

Limitations and Future Works. First, Twitter users are predominantly liberal. Hence the emotional reaction analysis we present inevitably include more left-leaning perspectives, especially in the U.S. tweet datasets. We have shown that splitting data by demographic groups and analyze the heterogeneity in different user groups help to mitigate potential biases. However, biases related to political leaning are not the only concerns; other biases may also exist within the studied populations. For instance, the emotional responses measured could be skewed toward a small subset of highly active users. Future research that focuses on user-level representation will therefore be crucial. Second, as discussed in section § 3.5, it is impractical to annotate every tweet in the dataset to create an exhaustive list of events and compute recall. Consequently, obtaining a reliable proxy for the share of events captured by our proposed method is challenging. We demonstrate the retrievability using the total number of events captured following prior works. A thorough validation — such as tallying all news articles recognized as prominent events during the study period and comparing these with the events detected by our method — can be conducted in future studies, though it would require significant labor. Third, our emotional reaction detection method encounters limitations when addressing change points that indicate a dip. In such cases, topic modeling cannot explain the dip, as a decrease in emotional or moral sentiment signifies a reduction in discussion related to an event. However, the decline of certain emotions is often accompanied by an increase in others, allowing us to analyze tweets tagged with the surging emotions to gain insights into the underlying topics.

In future research, we also plan to advance towards causal analysis. In principle, offline events *cause* online emotional and moral reactions, but numerous confounding factors complicate this causal relationship. We intend to expand this work by conducting causal analyses to further disentangle the relationships between offline events and online emotions and to measure the heterogeneous effects on different online groups.

Acknowledgments

This work is supported in part by AFOSR under grants FA9550-22-1-0380 & FA9550-20-1-0224, and DARPA under contract HR001121C0168. The authors thank Eugene Jang and Yuanfeixue Nan from University of Southern California for helping with emotion and moral foundation annotations.

References

[1] Ryan Prescott Adams et al. 2007. Bayesian online changepoint detection. *arXiv preprint arXiv:0710.3742* (2007).
[2] Kholoud Khalil Aldous et al. 2022. Measuring 9 emotions of news posts from 8 news organizations across 4 social media platforms for 8 months. *ACM Transactions on Social Computing (TSC)* 4, 4 (2022), 1–31.
[3] Hassan Alhuzali et al. 2021. SpanEmo: Casting Multi-label Emotion Classification as Span-prediction. In *ECACL*. ACL, 1573–1584.
[4] Sarah A Alkhodair, Steven HH Ding, Benjamin CM Fung, and Junqiang Liu. 2020. Detecting breaking news rumors of emerging topics in social media. *Information Processing & Management* 57, 2 (2020), 102018.
[5] James Allan et al. 1998. Topic Detection and Tracking Pilot Study Final Report. (1998). https://doi.org/10.1184/R1/6626252.v1
[6] Marina Bagić Babac. 2022. Emotion analysis of user reactions to online news. *Information Discovery and Delivery* ahead-of-print (2022).
[7] Francesco Barbieri, Luis Espinosa Anke, and Jose Camacho-Collados. 2022. XLM-T: Multilingual language models in twitter for sentiment analysis and beyond. In *LREC*. 258–266.
[8] David M Blei et al. 2003. Latent dirichlet allocation. *JMLR* 3, Jan (2003), 993–1022.
[9] Johan Bollen et al. 2011. Twitter mood predicts the stock market. *Journal of computational science* 2, 1 (2011), 1–8.
[10] Gerlof Bouma. 2009. Normalized (pointwise) mutual information in collocation extraction. (2009).
[11] Yuwei Cao et al. 2021. Knowledge-preserving incremental social event detection via heterogeneous gnns. In *In WWW*. 3383–3395.
[12] Rong-Ching Chang, Ashwin Rao, Qiankun Zhong, Magdalena Wojcieszak, and Kristina Lerman. 2023. # RoeOverturned: Twitter Dataset on the Abortion Rights Controversy. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 17. 997–1005.
[13] Kai Chen, Zihao He, Keith Burghardt, Jingxin Zhang, and Kristina Lerman. 2024. IsamasRed: A Public Dataset Tracking Reddit Discussions on Israel-Hamas Conflict. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 18. 1900–1912.
[14] Emily M Cody et al. 2015. Climate change sentiment on Twitter: An unsolicited public opinion poll. *PloS one* 10, 8 (2015), e0136092.
[15] Michael D Conover, Emilio Ferrara, Filippo Menczer, and Alessandro Flammini. 2013. The digital evolution of occupy wall street. *PloS one* 8, 5 (2013), e64679.
[16] Adji B. Dieng et al. 2020. Topic Modeling in Embedding Spaces. *TACL* 8 (2020), 439–453. https://doi.org/10.1162/tacl_a_00325
[17] Dodds et al. 2022. Fame and Ultrafame: Measuring and comparing daily levels of 'being talked about' for United States' presidents, their rivals, God, countries, and K-pop. *JQD:DM* 2 (Feb. 2022). https://doi.org/10.51685/jqd.2022.004
[18] Peter Dodds et al. 2011. Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter. *PloS one* 6, 12 (2011), e26752.
[19] Justin Garten et al. 2018. Dictionaries and distributions: Combining expert knowledge and large scale textual data content analysis: Distributed dictionary representation. *Behavior Research Methods, Instruments, and Computers* 50, 1 (1 Feb. 2018), 344–361. https://doi.org/10.3758/s13428-017-0875-9
[20] Dominique Geissler, Dominik Bär, Nicolas Pröllochs, and Stefan Feuerriegel. 2023. Russian propaganda on social media during the 2022 invasion of Ukraine. *EPJ Data Science* 12, 1 (2023), 35.
[21] Scott A Golder et al. 2011. Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. *Science* 333, 6051 (2011), 1878–1881.

[22] Maarten Grootendorst. 2022. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv preprint arXiv:2203.05794* (2022).
[23] Siyi Guo et al. 2023. A Data Fusion Framework for Multi-Domain Morality Learning. In *In ICWSM-2023*, Vol. 17. 281–291.
[24] Jonathan Haidt et al. 2007. The moral mind: How five sets of innate intuitions guide the development of many culture-specific virtues, and perhaps even modules. *The innate mind* 3 (2007), 367–391.
[25] Eva Hauthal et al. 2019. Analyzing and visualizing emotional reactions expressed by emojis in location-based social media. *ISPRS International Journal of Geo-Information* 8, 3 (2019), 113.
[26] Zihao He, Negar Mokhberian, and Kristina Lerman. 2022. Infusing Knowledge from Wikipedia to Enhance Stance Detection. In *Proceedings of the 12th Workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis*. 71–77.
[27] Stig Hebbelstrup Rye Rasmussen and Michael Bang Petersen. 2023. The event-driven nature of online political hostility: How offline political events make online interactions more hostile. *PNAS nexus* 2, 11 (2023), pgad382.
[28] Brian Heredia, Joseph D Prusa, and Taghi M Khoshgoftaar. 2018. Social media for polling and predicting United States election outcome. *Social Network Analysis and Mining* 8 (2018), 1–16.
[29] David V Hinkley. 1971. Inference about the change-point from cumulative sum tests. *Biometrika* 58, 3 (1971), 509–523.
[30] Joe Hoover et al. 2020. Moral Foundations Twitter Corpus: A Collection of 35k Tweets Annotated for Moral Sentiment. *Social Psychological and Personality Science* 11, 8 (2020), 1057–1071. https://doi.org/10.1177/1948550619876629
[31] Kokil Jaidka et al. 2020. Estimating geographic subjective well-being from Twitter: A comparison of dictionary and data-driven language methods. *PNAS* 117, 19 (2020), 10165–10171.
[32] Julie Jiang et al. 2020. Political Polarization Drives Online Conversations About COVID-19 in the United States. *Human behavior and emerging technologies* (June 2020).
[33] Marko Klašnja et al. 2018. Measuring Public Opinion with Social Media Data. Oxford University Press.
[34] Kristina Lerman. 2018. Computational social scientist beware: Simpson's paradox in behavioral data. *Journal of Computational Social Science* 1, 1 (2018), 49–58.
[35] Jure Leskovec et al. 2009. Meme-tracking and the dynamics of the news cycle. In *Proceedings of the 15th ACM SIGKDD*. 497–506.
[36] Chenliang Li et al. 2012. Twevent: Segment-Based Event Detection from Tweets *(CIKM '12)*. ACM, New York, NY, USA, 155–164. https://doi.org/10.1145/2396761.2396785
[37] Muzamil Malik et al. 2022. A Performance Comparison of Unsupervised Techniques for Event Detection from Oscar Tweets. *Computational Intelligence and Neuroscience* 2022 (2022).
[38] Louis Martin, Benjamin Muller, Pedro Javier Ortiz Suárez, Yoann Dupont, Laurent Romary, Éric Villemonte de la Clergerie, Djamé Seddah, and Benoît Sagot. 2020. CamemBERT: a Tasty French Language Mode. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (2020).
[39] Andrew J. McMinn, Yashar Moshfeghi, and Joemon M. Jose. 2013. Building a large-scale corpus for evaluating event detection on twitter. In *Proceedings of ACM CIKM* (San Francisco, California, USA) *(CIKM '13)*. Association for Computing Machinery, New York, NY, USA, 409–418. https://doi.org/10.1145/2505515.2505695
[40] Seyed Amin Mirlohi Falavarjani, Jelena Jovanovic, Hossein Fani, Ali A Ghorbani, Zeinab Noorian, and Ebrahim Bagheri. 2021. On the causal relation between real world activities and emotional expressions of social media users. *Journal of the Association for Information Science and Technology* 72, 6 (2021), 723–743.
[41] Lewis Mitchell et al. 2013. The geography of happiness: Connecting twitter sentiment and expression, demographics, and objective characteristics of place. *PloS one* 8, 5 (2013), e64417.
[42] Saif Mohammad et al. 2018. SemEval-2018 Task 1: Affect in Tweets. In *Proc. 12th Int. Workshop on Semantic Evaluation*. 1–17.
[43] Saif Mohammad and Peter Turney. 2010. Emotions Evoked by Common Words and Phrases: Using Mechanical Turk to Create an Emotion Lexicon. In *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*. Association for Computational Linguistics, Los Angeles, CA, 26–34. https://aclanthology.org/W10-0204
[44] Keval Morabia et al. 2019. SEDTWik: Segmentation-based Event Detection from Tweets Using Wikipedia. In *NACCL workshop*. 77–85.
[45] Liqiang Niu et al. 2015. Topic2Vec: Learning distributed representations of topics. In *2015 IALP*. IEEE, 193–196.
[46] Max Pellert et al. 2022. Validating daily social media macroscopes of emotions. *Scientific Reports* 12, 1 (2022), 11236.
[47] Robert B Penfold and Fang Zhang. 2013. Use of interrupted time series analysis in evaluating health care quality improvements. *Academic pediatrics* 13, 6 (2013), S38–S44.
[48] Alexandru Petrescu, Ciprian-Octavian Truică, Elena-Simona Apostol, and Adrian Paschke. 2023. EDSA-Ensemble: An Event Detection Sentiment Analysis Ensemble Architecture. *arXiv preprint arXiv:2301.12805* (2023).
[49] Ashwin Rao, Rong-Ching Chang, Qiankun Zhong, Kristina Lerman, and Magdalena Wojcieszak. 2023. Tracking a Year of Polarized Twitter Discourse on Abortion. *arXiv preprint arXiv:2311.16831* (2023).
[50] Nils Reimers et al. 2019. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In *In EMNLP-2019*. ACM.
[51] Zahra Rezaei et al. 2022. Event detection in twitter by deep learning classification and multi label clustering virtual backbone formation. *Evolutionary Intelligence* (2022), 1–15.

[52] Rodrigo Sandoval-Almazan et al. 2020. Sentiment analysis of facebook users reacting to political campaign posts. *Digital Government: Research and Practice* 1, 2 (2020), 1–13.
[53] Andrea L Schaffer, Timothy A Dobbins, and Sallie-Anne Pearson. 2021. Interrupted time series analysis using autoregressive integrated moving average (ARIMA) models: a guide for evaluating large-scale health interventions. *BMC medical research methodology* 21, 1 (2021), 1–12.
[54] Hugh Schofield. 2022. French elections: Putin's war gives Macron boost in presidential race. https://www.bbc.com/news/world-europe-60793320
[55] Fei Shen, Erkun Zhang, Wujiong Ren, Yuan He, Quanxin Jia, and Hongzhong Zhang. 2023. Examining the differences between human and bot social media accounts: A case study of the Russia-Ukraine War. *First Monday* 28, 2 (2023).
[56] Jackson Trager et al. 2022. The Moral Foundations Reddit Corpus. *arXiv preprint arXiv:2208.05545* (2022).
[57] Andranik Tumasjan et al. 2010. Predicting elections with twitter: What 140 characters reveal about political sentiment. In *In ICWSM-2010*, Vol. 4. 178–185.
[58] Gerben A. vanKleef et al. 2016. Editorial: The Social Nature of Emotions. *Frontiers in Psychology* 7 (2016), 896.
[59] Jianshu Weng et al. 2011. Event detection in twitter. In *In ICWSM-2011*, Vol. 5. 401–408.
[60] Shouzhi Xu, Xiaodi Liu, Kai Ma, Fangmin Dong, Basheer Riskhan, Shunzhi Xiang, and Changsong Bing. 2023. Rumor detection on social media using hierarchically aggregated feature via graph neural networks. *Applied Intelligence* 53, 3 (2023), 3136–3149.

A Detecting emerging topics after change point

To determine the new topics relevant to each change point, we compare the top 10 topics before the change point with those after the change point. Table 6 shows some examples. The newly emerged topics are highlighted in bold. For most reactions, regardless of how small or impactful, the identified topics clearly relate to an offline event (e.g. row 1). For the second example (row 2), the newly emerged topics point us to several different events. However, by examining the tweets belonging to these topics, we found the Black Lives Matter protests was the most predominant event, and other emerging topics such as "america, vote, trump" and "covid, coronavirus, tested" were related to the protests. Finally, there are also some change points for which we cannot identify meaningful emerging topics (e.g. row 3). We decide whether it is a false positive through manual verification.

Table 6. Examples of topics detected before and after several change points.

	Change Point Date	Emotion	Event	Topics Before Change Point	Topics After Change Point
1	2020-05-26	Betrayal	Black Lives Matter protests	"president", "wearing, masks, mask"	"**looting, starts**", "**people, black, white**", "**cops, police**", "president, leader", "**fox, news, stand**", "**minneapolis, floyd, police**"
2	2020-05-29	Care	Black Lives Matter protests	"love, sending, thank", "california, beverly, hills", "care, feeling, kindness", "donate, families, wines", "prayers, pray, praying", "home, miss, back", "mask, wearing, wear", "dog, leash, sorry", "black", "god, bless, blessing", "police, officers, cops"	"love, sending, much", "donate, donated, black", "**america, vote, trump**", "**protest, protesting, safe**", "**covid, coronavirus, tested**", "praying, pray, prayers", "**peaceful, peach, fighting**", "**black, lives, matter**", "california, los, angeles", "**safe, stay, careful**", "**city, business, santa, monica**"
3	2020-06-01	Pessimism	Unknown	"sad, heart, people", "need, rn, scared", "miss, going, back"	"sad", "heart, feel, sleep", "miss, friend, das", "heartbreaking, child, right", "never, wrong, life"

B Building Subsets of French Dataset For Four Candidates

We first gather all hashtags used in the dataset and look at the top 100 frequent ones. For each candidate, we manually select a set of hashtags relevant to them in the top 100, and gather all tweets containing these hashtags as their subset. The hashtags used are:

Macron (55404 tweets): "Macron", "MacronArena", "TousContreMacron", "StopMacron", "MacronDehors-DesLePremierTour".

Lepen (52608 tweets): “MarinePrésidente”, “DimancheJeVoteMarine”, “LePen”.

Mélenchon (63102 tweets): “Melenchon2022”, “MelenchonBFMTV”, “MelenchonToulouse”, “Melenchon1erMinistre”, “DimancheJeVoteMelenchon”, “MelenchonMarseille”, “MelenchonTF1”, “MelenchonHologrammes”, “AlloMelenchon”, “MelenchonLyon”, ‘Melenchon2eTour”, “Melenchon”, “MelenchonChiffrage”, “MelenchonMatignon”.

Zemmour (94038 tweets): “ZemmourTrocadero”, “JeVoteZemmourLe10avril”, “JeVoteZemmour”, “ZemmourCollineDuCrack”, “ZemmourPresident”, “ZemmourMetz”, “ZemmourEurope1”, “ZemmourToulon”, “JeVoteEricZemmour”, “ZemmourMontSaintMichel”, “ZemmourAgen”, “ZemmourBFMTV”, “Zemmour”, “ZemmourAgriculture”, “ZemmourProgramme”, “ZemmourCroissance”.

C All Events Detected in Different Datasets

The following tables show all the events detected for each dataset.

Table 7. Events and the emotional and moral reactions detected in 2020 Los Angeles Data.

	Event	Date	Time Window	Peaking Emotion/Morality	Declining Emotion/Morality	Relevant Topics
1	MLK Jr. Day	01-20	01-17 to 01-20	joy, loyalty, fairness		“mlk, king, martin”, “rights, king, justice”
2	Trump’s impeachment trial	01-20	01-20 to 01-21	betrayal, subversion		“senate, vote, will”, “treason, traitors, moscow”
3	Earthquake	01-20	01-20	fear		“earthquake, usgs, km”
4	Kobe Bryant’s death	01-26	01-26	anger, pessimism, sadness, care, harm, purity	joy	“kobe, rest, peace”, “prayers, praying, family”
5	Trump’s State of the Union address	02-03	02-03	anger, disgust, authority		“pelosi, nancy, speech”, “pelosi, trump, tear”
6	Valentine’s Day	02-14	02-14	love	anger, disgust	“valentine, valentines, happy”
7	CA Pres. primary	03-03	03-02 to 03-03	fairness, subversion		“vote, california”“biden”
8	COVID-19 pandemic	03-10	03-09 to 03-11	anger, disgust, fear, sadness, care, harm, authority	anticipation, joy, love, optimism	“coronavirus, corona, virus”, “pandemic, virus, administration”, “safe, stay, everyone”
9	Earthquake	04-03	04-03	fear		“usgs, reports, quake”
10	DOJ Drops Flynn Case	05-07	05-07	cheating		“flynn, cheated, cheating”
11	Trump tweeted: ”He got caught, OBAMAGATE!”	05-11	05-11	subversion		“obama, trump, president”
12	BLM protests	05-26	05-25 to 05-30	anger, disgust, fear, care, harm, fairness, cheating, loyalty, betrayal, authority, subversion, degradation	anticipation, joy, love, optimism	“racist, racism, white”, “murder, george, floyd”, “black, injustices, oppresion”
13	Biden criticizes Trump on Russian bounties	06-27	06-27	betrayal		“trump, traitor, must”, “biden, president, treason”
14	Dodgers beat Giants	07-23	07-23	joy		“dodgers, mlb, giants”

Received 30 December 2023; revised 26 October 2024; accepted 11 November 2024

Table 8. Events and the emotional and moral reactions detected in 2022 Abortion data.

	Event	Date	Time Window	Peaking Emotion/Morality	Declining Emotion/Morality	Relevant Topics
1	Justice Gorsuch refused to wear mask	01-18	01-18	betrayal, subversion		"gorsuch, justice, mask", "refuses, gorsuch, prolife"
2	Taliban crushes Afghan women's rights protest	01-18	01-18	subversion	love, joy	"taliban, women, rights"
3	Jason Miyares filed a motion wanting to overturn Roe v. Wade	01-20	01-20	authority	optimism	"virginia, motion, reversed", "jason, position, miyares"
4	disturbing video from SCOTUS on Roe v. Wade memorial	01-24	01-24 to 01-25	disgust, surprise		"wade, roe, memorial", "shocking, today, video"
5	Val Demings tweeted to support abortion rights	02-03	02-03 to 02-04	care, joy, love, optimism, trust	subversion, anger, disgust	"reproductive, rights, care", "night, orlando, protecting"
6	Florida banned access to abortions after 15 weeks	03-04	03-04	harm, cheating, fear		"women, bill, florida", "senator, movement, crucial", "women, severly, dangerous"
7	Memorial of Death of Breonna Taylor	03-08	03-08 to 03-10	degradation, care, betrayal		"breonna, tangible, remembering", "praying, seek, care"
8	Marsha Blackburn accused Ketanji Brown Jackson of supporting Griswold v. Connecticut	03-21	03-21	betrayal		"attacked, griswold, brown", "ketanji, jackson, prosecuting"
9	Anti-abortionist Mark Robinson says he paid for abortion after impregnating a woman	03-23	03-23	surprise		"abortion, paid, impregnating", "surprised, rights, robinson"
10	Oklahoma passed abortion ban punishing people as felons	04-05	04-05	cheating, subversion		"ban, passed, felons", "murdered, infant, penalty", "rape, minimum, oklahoma"
11	Republicans forced vote against TitleX	04-27	02-26 to 02-28	care, degradation, subversion		"titlex, screenings, preventative", "defend, backsliding, titlex", "access, reproductive, program"
12	Leak of SCOTUS ruling to overturn Roe v. Wade	05-03	05-02 to 05-03	surprise	joy, love, optimism	"kept, secrets, confidential", "overturned, crazy, wild", "murkowski, rocked, confidence"
13	Memorial of Uvalde school shooting	05-24	05-23 to 05-25	care, trust, fear	subversion, anticipation	"value, life, kids", "gun, shot, past"
14	Florida synagogue sued new abortion ban	06-14	06-14	purity		"sues, intrusion, prohibits", "suing, judaism, bible", "abortion, god, wrong"
15	COVID vaccine protests	06-17	06-17	subversion		"vaccine, choice, body", "biden, executive, action", "combust, protesters, believe"
16	SCOTUS overturned Roe v. Wade	06-24	06-21 to 06-24	authority, anger, disgust	care, optimism	"roe, happens, sooner", "patriots, roe, imminent", "conservative, republican, precedence", "tyrannized, autocratic, minority"
17	Police arrested House Democrats for protesting Roe v. Wade decision	07-19	07-19 to 07-21	betrayal, fear, subversion		"police, arrest, dc", "democratic, inciting, roe, wade"
18	Kansas voted to keep abortion legal	08-03	08-02 to 08-03	anticipation, trust		"women, health, kansas", "rebuke, republican, clear"
19	Facebook gave police a teenager's private chats about her abortion	08-08	08-08 to 08-09	fear, betrayal		"warrants, chats, abortion", "seize, facebook, phone", "apps, tracking, surveillance"
20	Texas woman denied an abortion after diagnosis of fatal fetus abnormalities	09-28	09-28	sadness, surprise, subversion		"horrendous, danger, mothers", "diagnosed, abnormalities, fatal"
21	2022 Midterm election	11-08	11-03 to 11-10	joy, loyalty, optimism, anticipation, fairness, betrayal, subversion	cheating, fear, anger, disgust, love	"choice, body, vote", "blue, vote, democracy", "far, media, republicans"
22	Biden Signed 2 orders to protect access to reproductive health care	11-19	11-20	joy	anger, disgust	"protect, care, reproductive", "executive, signing, biden"
23	Senate stayted blue after election	12-07	12-07	loyalty, anticipation, optimism	anger, disgust	"vote, senate, codify", "majority, progress, settled"
24	FBI arrested 2 pro-life protesters	12-16	12-16	harm, fear, sadness, trust		"arrested, violence, facism", "federally, jesus, individuals", "advocates, prolifers, attacks"

Table 9. Events and the emotional and moral reactions detected in 2022 French election data.

	Event	Date	Time Window	Peaking Emotion/Morality	Declining Emotion/Morality	Relevant Topics
1	Russian-Ukraine war	02-23	02-22	care, harm		"explosions, thursday, kiev", "sympathy, shameful, volunteers", "pendance, donetsk, lugansk", "emmanuel, macron, condemned"
2	1st round presidential election	04-10	04-07 to 04-15	subversion, joy, fear	fairness, subversion, anger, disgust	"left, macron, turn", "sir, president, luck", "macron, facism, maintain", "will go up, sanctions, moscow"
3	Russian missile cruiser Moskva sank	04-14	04-14	surprise		"moskva, cruiser, sea"
4	2nd round presidential election	04-24	04-21 to 04-26	sanctity, fear, sadness, surprise	joy, anger, disgust	"liberty, abolition, glory", "rally, national, respect", "democracy, emmanuel, result", "ukraine, conflict, economic", "macron, price, commodity"
5	Russia cut off natural gas to Poland and Bulgaria	04-26	04-26	sadness		"gas, bulgaria, poland", "gas, cut, price"
6	Explosions in Moldova	04-26	04-26	fear		"transnistria, moldavia, explosions"
7	1st round legislative election	06-12	06-12	loyalty	anger, disgust	"voice, france, legislative", "struggle, invent, Mélenchon", "cover, elisabeth, communicate"
8	2nd round legislative election	06-19	06-18	authority		"nupes, send, vote", "bravo, congratulations, marine"
9	US supreme court overturned Roe v. Wade	06-24	06-24 to 06-25	care, harm, cheating, degradation, fear		"abortion, right, your", "unborn, about, rights"
10	G7 summit	06-26	06-24 to 06-25	care, fairness, sanctity, fear	joy	"summit, leaders, biden", "europe, increase, members", "coordination, strengthen, investments"