BEHAVIOUR & INFORMATION TECHNOLOGY
https://doi.org/10.1080/0144929X.2022.2030798

What really matters?: characterising and predicting user engagement of news postings using multiple platforms, sentiments and topics

Kholoud Khalil Aldous [a], Jisun An[b] and Bernard J. Jansen[c]

[a]College of Science and Engineering, Hamad Bin Khalifa University, Doha, Qatar; [b]School of Information Systems, Singapore Management University, Singapore, Singapore; [c]Qatar Computing Research Institute, Hamad Bin Khalifa University, Doha, Qatar

ABSTRACT

This research characterises user engagement of approximately 3,000,000 news postings of 53 news outlets and 50,000,000 associated user comments during 8 months on 5 social media platforms (i.e. Facebook, Instagram, Twitter, YouTube, and Reddit). We investigate the effect of sentiments and topics on user engagement across four levels of user engagement expressions (i.e. views, likes, comments, cross-platform posting). We find that sentiments and topics differ by both news outlets and social media platforms, and both sentiments and topics by the four levels of user engagement expression. Finally, we predict a volume of four user engagement levels for given news content, with an 83% maximum average F1-score for the external posting of news articles from one platform to another using language and metadata features. Implications are that news outlets can benefit by developing a platform, sentiment and topic, and strategies to best achieve user engagement objectives.

ARTICLE HISTORY
Received 12 October 2020
Accepted 10 January 2022

KEYWORDS
User engagement; cross platforms; news organisation; topical analysis; sentiment analysis; social media

1. Introduction

Due to the revolutionary effects of social media platforms, many news organisations highly depend on them to distribute content and reach larger audiences. Moreover, multiple social media platforms are providing a multitude of social interactions for users to express feelings or opinions (Thonet et al. 2017). Almost 68% of news consumers, about two-thirds of American adults, get part of their news on social media, while 'convenience' is the main positive advantage of doing so Matsa and Shearer (2018).

As a result, news organisations are highly dependent on different platforms as online channels for distributing their postings. Those platforms are considered a revenue source where news organisations are competing for the $109 billion spent on digital advertising by 2018 (Stocking and Khuzam 2019). Thus, for news organisations, social media platforms are essential for revenue generation and content distribution (Stocking and Khuzam 2019). The success of a social media post depends, in part, on the audience engagement it generates. Hence, many challenges are faced by content creators and editors who are dependent on understanding what interests audiences both on individual and across multiple platforms (Aldous, An, and Jansen 2019a).

User engagement on social media platforms can be measured through different web analytic metrics. Higher user engagement for news content indicates better performance, which means higher audience growth, such as followers. Higher user interactions include the number of views, likes, and comments (Balbi, Misuraca, and Scepi 2018). Audience engagement metrics are context-specific; however, each metric is a different type of experience that is unique and impactful in ways different from others on each social media platform (Voorveld et al. 2018; Noguti 2016). A unified approach of categorised engagement metrics is needed to support the analysis of news organisations' data and across multiple social media platforms. Classifying those metrics into levels is one approach, where each level encompasses similar actions, expressions, or impact. Hence, content with more public user engagement expressiveness is more impactful due to the network effect of social media platforms. This premise is grounded in the marketing notion of electronic word-of-mouth (eWOM) concept, in which people promote services and products loudly (Chen, Hong, and Li 2017; Jansen et al. 2009). The network effect concept is based on the fact that each individual has a certain reach. For example, each person shares content with his friends, and each of them can, in turn, share the same content on their friends' networks and so on. This network effect illustrates how content can go viral over social media platforms. News organisations highly benefit from the

CONTACT Kholoud Khalil Aldous kaldous@mail.hbku.edu.qa; khulood.k.2011@gmail.com

eWOM concept, by which they enlarge their audience base in a costless manner. As a result, organising user engagement metrics into levels based on audience interest in expressiveness from more public to more private is an effective way to leverage the network effect. From this theoretical background and building from Noguti (2016), in this research, we derive a framework of user engagement consisting of user engagement levels ordered according to their degree of public expressiveness regarding content.

Building from this model of user engagement, we do a cross-platform topical analysis of content to understand user preferences among different social media services for multiple news organisations. One of the biggest challenges news content editors face is understanding their audiences across multiple platforms (Aldous, An, and Jansen 2019a). For this, we analyse and compare the topical distribution of content across platforms to determine the best platform for each topic. We also do sentiment analysis for both posts and comments across platforms, as well as for individual topics, to help news organisations better understand their audiences. Results show that depending on the platform and topic, the sentiment of users' comments may vary. A higher engagement (e.g. Level-3 comments) of a topic in a given platform would not necessarily mean it is a positive engagement. So a simple counting of comments might give an incorrect impression of user engagement.

In summary, this research focuses on analysing user engagement across multiple social media platforms and many news organisations. It builds on the previous research (Aldous, An, and Jansen 2019d) showing the topical effect on user engagement. The current research extends the prior findings (Aldous, An, and Jansen 2019d) with an in-depth investigation of the topical effect on engagement and the sentiment aspect with prediction experiments of multiple engagement metrics organised into levels. Therefore, the contributions of this research are the following:

- first, analysing the psychological aspect of user engagement using sentiments of both posts and comments;
- second, using comments data in addition to posts data for understanding the audience's opinion for news organisations;
- third, analysing the topical distribution and user engagement differences among the platforms and news organisations;
- fourth, we exploit the correlation between posts' topics and the sentiment of both posts and comments; and
- fifth, we conduct additional prediction experiments to predict four user engagement levels across five social media platforms.

1.1. Engagement framework

There are multiple user engagement metrics across platforms without a clear framework. An effective approach is by organising user engagement metrics into levels based on audience interest expressiveness from more private to more public. Out of this theoretical background of eWOM (Chen, Hong, and Li 2017; Jansen et al. 2009) and building from Noguti (2016), in this research we provide a theoretical description of an engagement framework that has four levels, defined below. The metrics are ordered according to the degree a user is willing to publicly associate with the online content. Hence, a higher engagement level (Level-4) indicates more public association with the content, while a lower engagement level (Level-1) means more private engagement. Figure 1 shows an illustration of the four levels.

- Level-1 (# views): Private engagement by viewing social media posts or videos. The number of views of a given online content is an example of Level-1 engagement, where users only watch the content without expressing any preferences or giving feedback.
- Level-2 (# likes): Users interact with the content more publicly in Level-2 by pressing the like or favorite button. Liked posts, most probably, are shown in the users' feeds on the platform where the like action is taken. That makes the posts more visible to the users' friends list (e.g. Facebook and Twitter).
- Level-3 (# comments and # shares): When a user chooses to share opinions or feelings, that is considered a more public engagement (Level-3), such as adding a comment or sharing a post on the same platform where the content is found. Publicly shared content is made visible in the user's timeline with notifications, if enabled, sent to the user's network, which makes the content more visible on the same platform. Also, Level-3 allows private sharing to selected users through private messaging, where the user may share the content with only friends who are interested.
- Level-4 (# external postings): The act of sharing content from one platform to another is public sharing, which we consider to have the highest engagement level in the framework (Level-4). This is because publicly shared content is exposed to other platforms with potentially wider audiences. The difference between Level-3 and Level-4 is that the sharing of Level-4 is done on another platform rather than the

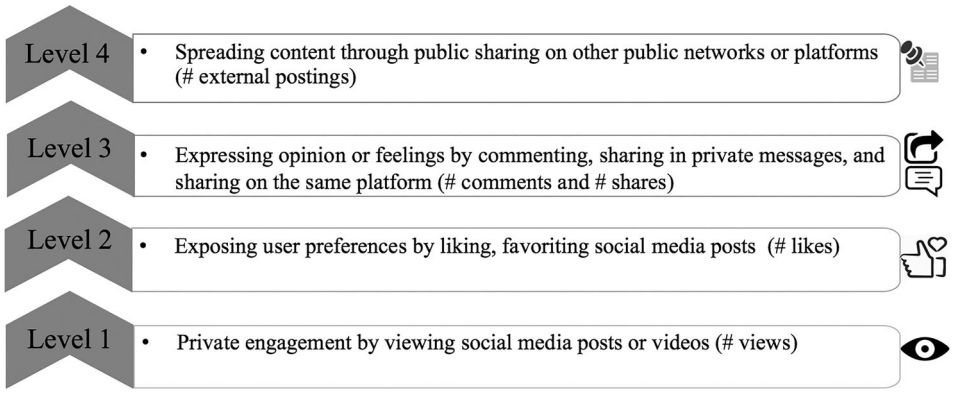

Figure 1. Engagement levels ordered by degree of public expressiveness, from more private (level 1) to more public (level 4).

platform where the content occurs. Level-4 can be measured on almost all social media platforms if the data is afforded, thus Reddit is one of the good networks due to public data availability. On Reddit, one can add a post into different available communities, such as r/worldpolitics and r/Positive_News, where the content reach is expanded into the subscribers of the targeted community (e.g.r/worldpolitics has 894K subscribers). As a result, we use Reddit data to examine the topical distribution of users' externally shared content.

There is a shortage of prior work that simultaneously deals with multiple social media platforms and multiple organisations within a domain focusing on user engagement in a systemic manner. This research addresses these shortcomings.

2. Related work

In this section, we review three different related work sets: user engagement on social media platforms, the topical effect, and the sentiment effect.

2.1. User engagement on social media platforms

One definition of user engagement is 'the emotional, cognitive and behavioural connection that exists, at any point in time and possibly over time, between a user and a resource' (Attfield et al. 2011), where a resource is an online application or content. In this work, we study the interaction of users with online *news* content posted in *social media platforms*. We adopt the analytical approach where user engagement metrics are recorded user behaviours (Jansen 2009; Lalmas, O'Brien, and Yom-Tov 2014; An and Weber 2018). Social media engagement metrics are commonly used in businesses and research to measure performance improvements in which they need to be connected to marketing goals (Peters et al. 2013). We study the sentiment of news posts and the sentiment of users' comments in response to those posts, which can be considered as psychological user engagement (Oh, Bellur, and Shyam Sundar 2018).

User behaviour in social media is an important indicator of news organisation success, where some behaviours are more common than others. For example, social media users are more likely to click on a news post link or like the post than commenting into it as less effort is required to do so Larsson (2018) and Chung (2008). Understanding those engagement behaviours is also essential for news organisations as they provide audience feedback and an indirect user contribution to the journalistic process (Hille and Bakker 2014). Consequently, organisational performance can be enhanced, such as sales growth, cost reduction (Harrigan et al. 2017; Kumar et al. 2010).

There are multiple existing research on social media engagement, but most of them have studied an individual platform such as Facebook (Van Canneyt et al. 2018; Srinivasan et al. 2013; An, Quercia, and Crowcroft 2014), Instagram (Jaakonmäki, Müller, and vom Brocke 2017; Ferrara, Interdonato, and Tagarelli 2014), Twitter (Muñoz-Expósito, Ángeles Oviedo-García, and Castellanos-Verdugo 2017; Van Canneyt et al. 2018; Bandari, Asur, and Huberman 2012; An et al. 2014), YouTube (Ma, Yan, and Chen 2017; Vallet et al. 2015), and Reddit (Stoddard 2015). There is limited cross-platform analysis in which metrics are compared based on their public expressiveness, but without considering the topical or domain aspects (Rajapaksha, Farahbakhsh, and Crespi 2019; Rowe and Alani 2014; Liu et al. 2019). Hence, there is a clear need for a comprehensive analysis of audience engagement across various platforms from content providers of the same domain (e.g. news).

User engagement is a complex construct that is impacted by many factors, including the context, content, and creator (Jaakonmäki, Müller, and vom Brocke 2017). This research tackles one factor, which is content, where the scored amount of a user engagement metric is affected by multiple content aspects such as the format (O'Brien 2017), topic, and sentiment. This work considers both the content topic and sentiment of social media posts in relation to user engagement.

2.2. Topical effect

Content topic is one of the most important features affecting user engagement all over online platforms. The topical interest of users varies across social media platforms (Ferrara, Interdonato, and Tagarelli 2014; Guo et al. 2015). Although some users have a wider range of topical interests (Wang 2017), many others are very specialised with few topics (Ferrara, Interdonato, and Tagarelli 2014). Hence, their preferences have a large effect on their interaction behaviour with the content, such as sharing (Shi et al. 2018) and consumption, which is why topics are used for personalisation (Zhang and Pennacchiotti 2013).

Understanding the users' topical interest helps in boosting the engagement volume on Twitter (Yang and Rim 2014). Thus, many scholars have worked on findings hidden, but popular topics from users' posting for better information retrieval, instead of using trending topics provided by the platforms (Lee et al. 2011; Indira et al. 2019) Other researchers use topical analysis to detect events of breaking news streams (Altınel and Ganiz 2018; Liu et al. 2017) and predict users who will share a given piece of content (Zarrinkalam, Kahani, and Bagheri 2018). Other research found that one might increase user engagement by improving the saliency of topics that interest users (McCay-Peet, Lalmas, and Navalpakkam 2012).

Most of the topicality in previous studies focused on individual platforms (Hong and Davison 2010; Kouloumpis, Wilson, and Moore 2011; Pak and Paroubek 2010). However, there are also a few cross-platform studies, most of which either focused on an event analysis (e.g.Ebola) (Pokharel et al. 2019), social networking (Ting, Wu, and Chang 2009), or users' preferences and behaviour (Lee, Hoang, and Lim 2017; Hu et al. 2017). Generally, it is difficult to conduct a comprehensive cross-platform topical analysis because it is challenging to collect data across multiple social media platforms. Thus, a cross-platform topical analysis can provide in-depth understanding of topics' trends and popularity. In this work, we leverage topic analysis for identifying and comparing the popular topics across social media platforms for news postings while also considering the sentiment and engagement aspects.

The problem of identifying latent topics from text data has been extensively studied through different topic modelling techniques (Cataldi, Di Caro, and Schifanella 2010; Saha and Sindhwani 2012). The most popular topic modelling technique is the Latent Dirichlet Allocation (LDA), which finds hidden topics using unsupervised learning models for clustering text posts (Blei, Ng, and Jordan 2003; Darling 2011). Hence, each topic is a cluster of concurrent terms, and each post is assigned to a topic set according to its probability distribution. However, it is known that the standard LDA does not work well for short texts such as social media posts since each post is mostly related to a single topic due to its length (Jelodar et al. 2017). To tackle this problem, researchers proposed Twitter-LDA (Zhao et al. 2011), which addressed the short text length issue. In comparison to LDA, Twitter-LDA found to capture the topics of short text posts more properly and meaningfully (Fang et al. 2016) with multiple real applications of social media already adopting it Diao et al. (2012), Yang, Chen, and Bao (2016) and Jiang, Qiu, and Zhu. (2013). As a result, this research adopts Twitter-LDA for identifying the latent topics and their persistence across social media.

2.3. Sentiment effect

Sentiment analysis has been integrated into different approaches to studying its effect on user engagement (Nimala and Jebakumar 2019). For example, reordering some comments on news articles is preferable based on the user's gender and the interest level of the news article (Arapakis et al. 2014). Also, sentiment can be used to predict user engagement (Arapakis et al. 2014). Additionally, it can be used to calculate user ratings for non-rated reviews (e.g. reviews in social media platforms) where the rating of the review is dependent on the sentiment (Qiu et al. 2018).

Many recent studies are focusing on predicting sentiment using a different set of features and algorithms (Kaladevi and Thyagarajah 2019; Ceron, Curini, and Iacus 2016; Huang et al. 2020; Kong et al. 2020). In this research, we leverage sentiment analysis to understand the correlation between posts and comments sentiment and the effectiveness of user engagement in a larger scale dataset in cross-platform and multiple news outlet environments with the topical aspect taken into account. We use the Valence Aware Dictionary and sentiment Reasoner (VADER) to extract a sentiment label (i.e. negative, neutral, or positive) which is optimised for short texts (Hutto and Gilbert. 2014).

2.4. Summary

Since topics are affecting user engagement differently, and sentiment of content might affect the overall engagement in social media platforms, there is no comprehensive analysis of both aspects in one study in cross-platforms and for multiple organisations. Meanwhile, understanding the differences of this effect across platforms and for a large set of content producers in one domain is important for content producer decision-making.

3. Research questions

This research aims to provide guidelines and a framework for studying comprehensive user engagement levels of content across multiple platforms and from multiple news organisations. By focusing on the topicality and sentiment of news content, we can discover the effect of various topics published by news organisations, the way users are engaging with online social platforms, and the corresponding sentiment. Typically, news organisations have content editors who work simultaneously on pushing new content on different social media platforms (Aldous, An, and Jansen 2019a). However, as we mention in the literature review, there are a limited number of studies involving the multiple cross platforms factor (Glenski, Weninger, and Volkova 2018; Del Vicario et al. 2017) and multiple similar domain organisational postings (Rieis et al. 2015). This research begins addressing these shortcomings.

Findings from this research work can be adopted into systems and models that magnify social media engagement while the content is generated. This aim is the motivation of our research questions, which are as follows:

- RQ1: What are the topical distribution and user engagement differences among (a) the platforms and (b) news organisations?
- RQ2: (a) Is user engagement affected by the content topic? (b) If so, what topics generate which levels and what volume of user engagement?
- RQ3: How are posts' topics related to the sentiment of (a) posts and (b) audience's comments on those posts?
- RQ4: Can we predict the volume of the different engagement levels for a given content: (a) Level-1, (b) Level-2, (c) Level-3, and (d) Level-4?

The novel contributions of this research have several aspects. One is that the engagement metrics of multiple social media platforms are ordered into one framework based on the degree of public expressiveness of user engagement. Second, the conducted analysis is across multiple social media platforms (Facebook, Instagram, Twitter, and YouTube), a community-based social network (Reddit), and multiple organisations (53 news organisations) from the news domain. Third, the duration of the collection of news postings is 8 months, where multiple topics are covered, which are used to compare and understand user engagement. Lastly, we focus on the organisational postings to understand the topical distribution of content across platforms, whilst the limited previous researches focused on the topicality of users' posts in their social media profiles. Thus, the current research has a potential impact on the general understanding of news content production strategies across platforms, and the topical and sentiment effect on user engagement.

4. Methodology

4.1. Data collection

For answering our research questions, we collect our dataset by first selecting (a) multiple popular news outlets and (b) multiple social media platforms where news outlets post the most. After that, we collect the news posts of all the selected news outlets from the multiple selected social media platforms. Furthermore, we also collect all comments in response to those posts in our data collection to analyse the audience sentiment.

4.1.1. Popular news outlets

A list of 60 English-based outlets is constructed after considering multiple news-ranking websites, including PewResearch[1] and Wallethub.[2] When checking their activity across different social media platforms, we excluded a few news outlets (7) from the list. Thus, 53 outlets are used in the final list of popular news outlets in this research. Most of these are US-based (44), then UK-based (5), and other countries (4). The news outlets are ABC News, Los Angeles Times, The New York Times, AlJazeera, Mail Online, The Verge, BBC News, Mic, The Wall Street Journal, Bleacher Report, MSNBC, The Washington Post, Bloomberg, National Public Radio, The Week, Boston.com, NBC News, The Blaze, Breitbart News Network, NDTV, TIME, Business Insider, New York Post, U.S. News & World Report, Buzz Feed, Newsweek, Upworthy, CBS News, NY Daily News, USA Today, Chicago Tribune, Qatar Tribune, VICE, CNBC, Salon, Vox, CNN, Slate, Xinhua News Agency, CNN Digital Network, The Associated Press, Examiner.com, The Atlantic Magazine, Financial Times, The Boston Globe, Forbes, The Daily Beast,

Fortune, The Economist, Fox News, The Guardian, Huffington Post, and The Hill.

4.1.2. Popular social media platforms

News outlets are mostly publishing their news content across four popular social media platforms: Facebook, Instagram, Twitter, and YouTube (Kallas. 2017). We collect the news content from those four platforms to conduct our research. By using the verified English account of each news outlet, which contains the general news content, we ensure the diversity of topics across each outlet and platform. That is, no specialised topic account (e.g. fashion, sport) is selected. For example, Aljazeera news outlet has different Instagram accounts (e.g.aljazeeradocumentary, aljazeerasports); however, we select the one account where they post their general news (aljazeeraenglish).

For examining Level-4 engagement, we collect news content from the Reddit, which has a collection of different communities called subreddits (e.g.r/TechNewsToday, r/worldnews, etc.). On Reddit, individuals publicly share and discuss news content in different subreddits with URL links to the original articles on the news outlets' official websites.

4.1.3. Data overview

Using the open API of each platform, we collect the news postings of 53 news outlets from January to August 2017 inclusively. The data collection time is more than 30 days after the last post in our dataset is published (31 August 2017). Thus, the engagement metrics of those news postings are already stable given that most of the social media engagement occur within the first few days of publishing.[3]

In total, we collect more than 3M news postings and 50M comments across the five platforms. On YouTube, there are two idle news outlets that have no posting, while two others (the Associated Press and Mic) have deactivated the user comments option. For each platform, Table 1 shows the number of news outlets (# Outlets), total posts (# Posts), comments (# Comments), and likes (# Likes).

Table 1. The total number of news outlets, collected posts, comments and the associated total likes for Facebook, Instagram, Twitter, YouTube and Reddit.

Platform	# Outlets	# Posts	# Comments	# Likes
Facebook	53	27,117	984,266	70,557,281
Instagram	53	35,289	11,732,837	723,493,279
Twitter	53	571,270	14,426,570	13,604,785
YouTube	51	43,103	4,674,630	33,265,610
Reddit	53	2,486,594	18,200,179	147,521,797
Total		3,163,373	50,018,482	988,442,752

Note: For Reddit, we refer to score values as likes to unify the metrics naming across platforms.

Facebook (FB): By creating a web crawler using Facebook API, we manage to crawl the news postings of all the targeted news outlets (see Table 1). Also, for each post, we collect users' comments with their metadata, including their timestamp, which is needed for the temporal analysis in our study. The collected posts are also associated with engagement metrics and timestamp. A Facebook post can be an image, video, text, or a combination of those. As this research focuses on the text data, we only consider the posts with text, which we can use to label the post topic. We note that Facebook collected comments are random samples due to API limitation.

Instagram (IG): Similar to Facebook, we build a crawler specifically for collecting all Instagram profile postings for the 53 news outlets. We then only retain the posts within the eight-month collection time. The collected posts include all their metadata, such as the posting time, number of likes, number of comments, and the actual user comments. In total, we collect more than 35K Instagram posts with their user comments (11M).

Twitter (TW): In a similar way, we collect the tweets of the 53 news organisations through a crawler for their Twitter pages. The crawler specifically returns all the tweets' IDs between a fixed start and end date based on our collection period, which we use later for pulling the actual tweets' text and metadata using Twitter API. In total, we collected more than 500K Twitter posts and their 14M comments.

YouTube (YT): Through adopting the YouTube search function, we are able to collect the list of posts for each news outlet YouTube channel, plus retrieving whatever comments have been added by the channel subscribers to those posts. In total, we have more than 43K posts and their 4M comments.

Reddit (RD): We extract all Reddit posts that include domain name of each news outlet (e.g. bbc.com, edition.cnn.com) from publicly available Reddit dataset.[4] We collect more than 2M Reddit posts during the time from January to August 2017. Reddit opens doors for machine-generated posts (bots), which causes a problem for the accuracy of our study. As a result, it is important to filter out the bot posts from our Reddit dataset. To do so, we followed an easy approach with the simple assumption that highly active users are most probably bots. So, we sorted the users by their posting volume in descending order and then manually checked the user names of the top 100 users to find patterns. We then remove those users with a posting volume of more than 1000 posts with bot keyword indicators: bot, auto, news, or admin. Accordingly, we

removed 796 (out of 128,956) user accounts distributed across the keywords which are 426 'bot', 203 'news', 62 'auto', and 36 'admin'. All posts from those accounts are then excluded from our Reddit dataset. As a result, we have 602,870 posts out of the $2M+$ collected for 128,160 real Reddit users.

4.2. Engagement metrics

For each engagement level of the four levels of framework, we calculate an engagement metric: (a) Level-1 Normalised View (NV), (b) Level-2 Normalised Likes (NL), (c) Level-3 Normalised Comments (NC), and (d) Level-4 Normalised External Posting (NEP). The original values of the engagement metrics in our dataset are highly skewed; as a result, we use the log normalisation function to calculate the four normalised engagement metrics. According to the platforms' differences and data limitations, some metrics are measurable across the five platforms, such as the number of likes and comments, which helps when finding the cross-platform differences of Level-2 and Level-3 engagement. In Reddit, we call the score value of a post, which is a combination of users' up and down ratings, as the number of likes for unifying the metrics naming across platforms. Also, the number of views is only available on the YouTube posts of our dataset, while the Level-4 metric is available on Reddit. For Level-4, we calculate the NEP through log-normalising the count of the number of times a news article is posted on a public network (Reddit), which we call the External Posting Count (EPC). In this work, we use Reddit public network for measuring Level-4 engagement; however, other public platforms can be used, which is subject to data availability.

4.3. Topical analysis

Analysing the content topic across platforms and its effect on user engagement requires defining the common topics from the news posts of the five social media platforms. Using Twitter-LDA (Zhao et al. 2011), we extract the hidden topics from our dataset after employing a few data cleaning steps for the posts' text and determining the best fitting number of topics, as described below.

Cleaning the Posts Text: We first remove the email addresses, hyperlinks, punctuation, and stop words. Also, we omit some domain-specific frequently repeated terms such as news and articles while also omitting the names of the news outlets' (e.g. AlJazeera, Times, etc.). After that, we tokenize the posts' text and apply the Porter stemmer. After all of these steps are completed, we only retain posts with five terms or more.

Finding the Optimum Number of Topics: In general, using LDA requires specifying the number of topics as an input to the model, so finding the optimal number of topics is required to guarantee the effectiveness of the generated topics and minimise the overlapping topics. We test the topics' coherence for determining the best number of topics (Fang et al. 2016), which is calculated using the top (n) most frequent words for individual topics (t). Thus, we pair all words to construct a list of the top-n term pairs and find the semantic similarity of each pair of words using the pointwise mutual information (PMI) measure (Fang et al. 2016). The average PMI values of all word pairs represent the topic's coherence (C), which is calculated using the following formula:

$$C(t) = \frac{1}{\sum_{n-1}^{m=1} m} \sum_{i=1}^{n} \sum_{j=i+1}^{n} PMI(w_i, w_j), \tag{1}$$

where the number of top words is donated by n, and the PMI score of a word pair (w_i, w_j) is computed as follows:

$$PMI(w_i, w_j) = \log \frac{p(w_i, w_j)}{p(w_i) * p(w_j)}, \tag{2}$$

where $p(w_i)$ is the probability of observing word w_i, and $p(w_i, w_j)$ is the probability of the two words $w_i and w_j$ appearing in the same post.

Having better coherence means getting higher PMI average across the topics. This is done by testing 10 different topic numbers (k), starting with 10 and ending with 100, with a gap of 10 between each tested number. Also, the coherence is examined at three different numbers of top topics (@5, @10, and @20). In Table 2, we show the average PMI score results of a total of 29 experiments using different topic numbers and coherence numbers. According to the results in Table 2, the optimal number of topics is when $k = 20$ all over the coherence values @5, @10, and @20, where the highest average PMI is 3.5 for coherence@5. Consequently, in this study, we use a Twitter-LDA model with (k = 20) as the number of topics, where we manually assign a

Table 2. The coherence of the Twitter-LDA model with different number of topics k at three coherence average points.

#topics	coherence@5	coherence@10	coherence@20
10	3.0	2.5	–
20	**3.5**	**3.0**	**2.1**
30	2.4	2.2	1.8
40	2.0	1.7	1.5
50	1.8	1.6	1.4
60	1.5	1.4	1.2
70	1.4	1.3	1.2
80	1.2	1.1	1.0
90	1.2	1.1	1.0
100	1.0	0.9	0.8

label for each topic using its top 20 frequent words. Then, for each social media post, we assign a topic using the predefined topic model.

4.4. Sentiment analysis

For analysing the sentiment effect on engagement and its relationship with topics, we need to be able to measure sentiments of all posts and comments. For that, we use VADER (Valence Aware Dictionary and sentiment Reasoner) (Hutto and Gilbert. 2014) that also gives a compound score (−1 to 1) for a given post or comment. It can be labelled with positive sentiment if the compound score is ($>= 0.05$), neutral sentiment if (> -0.05 and < 0.05), and negative sentiment if the compound score ($<= -0.05$). Accordingly, we labelled each comment and post with a corresponding sentiment polarity (positive, neutral, or negative).

4.5. Level-1 prediction model

Due to the data affordance limitations by platform, we use YouTube posts to predict Level-1 engagement using the number of views metric. However, one can perform similar experiments to other platforms (e.g. Facebook and Instagram) if the view count of each content is obtainable. The prediction experiment is done for each news media separately, and the results are summarised using average, minimum, median, maximum, and standard deviation.

To start, for each individual news media, we normalise the number of views for each YouTube post using the log normalisation after adding one to accommodate zero values. Then, we calculate the quantile value of 0.33 (low-quantile) and the quantile value of 0.66 (high-quantile) for the normalised view counts of the posts within each news source. That means each news source has different quantile values based on its follower volume. We label Class-1 to the news posting with normalised views more than or equal to the high-quantile value, which, most likely, refers to the top 33% of Level-1 engagement from the posts. Meanwhile, we add a Class-0 label to the YouTube posts with normalised views less than or equal to the low-quantile value, which again refers to the bottom 33% of Level-1 engagement. Thus, we conduct a binary classification of experiments of the two labelled portions of the posts, while the middle 33% of the content is saved for future improvement. Plus, we balance the number of posts within each class using under-sampling, and we only keep the news sources with a minimum of 100 posts in each class, which accommodates up to 23 out of 51 news sources.

For evaluating the model's performance, we use Precision, Recall, F1-score, and Area Under the Curve (AUC) with 10-fold cross-validation. However, we report only F1-scores because other values are approximately the same due to the balanced datasets. We compare the results of three different algorithms: AdaBoost, Decision Tree, and Random Forest. There are three different features sets we examine: Language features (L), Metadata features (M), and combining both features (L+M).

Language Features(L): we use the Term Frequency-Inverse Document Frequency (TF-IDF) features vector. We set the TF-IDF parameters with 0.8 and 0.01 for the maximum and minimum IDF, respectively. Also, we fix the value of the maximum number of features to 5,000 words.

Metadata Features(M): For the metadata features, we include a Twitter-LDA topic vector with 20 features. Also included are four general content features: content length by characters, binary features of an emoji, question marks, and exclamation marks, which indicate the existence of a feature in the post. One more feature is the sentiment label of individual posts using VADER, which can be either positive, negative, or neutral.

Language and Metadata (L+M): We combine both feature sets and use them as input to the model.

4.6. Level-2 prediction model

For Level-2 engagement, our dataset facilitates the prediction of the volume of likes for Facebook, Instagram, Twitter, and YouTube. By following the same described prediction experiment for Level-1 engagement, we build the prediction models for the normalised number of likes for each news outlet and each platform. When filtering the news organisations based on posting volume (minimum 100 for each prediction class), we include 52 news organisations on Twitter, 51 on Facebook, 28 on Instagram, and 23 YouTube.

4.7. Level-3 prediction model

Our datasets afford the volume of the comments of each news source posting over four platforms Facebook, Instagram, Twitter, and YouTube; thus, the prediction experiments are conducted over the four platforms. In the filtering process to eliminate news sources with less than 100 posts for each classification class (total 200 posts), we end up with the following number of news sources: 51, 28, 52, and 23 for Facebook, Instagram, Twitter, and YouTube, respectively.

4.8. Level-4 prediction model

In this research, we build two different prediction models to examine Level-4 engagement volume: predicting whether a news article will be posted at least once on the Reddit public network (posted once), and predicting whether a news article will be posted multiple times on Reddit (posted multiple times). Those two models where we use the headlines of news articles as an input of the prediction models can be used by news producers. We test the same three algorithms and three feature sets of the other prediction tasks (levels 1 to 3), and similarly, report F1-score as an evaluation measure.

Posted Once: For each news article, we predict whether it will be posted publicly (or not) by Reddit users. Our Reddit dataset includes the positive cases (Class-1) of this prediction task, which are articles that have been posted on Reddit by users. Consequently, to create the list of news articles that are not posted on Reddit, the negative cases (Class-0), we crawl all RSS feeds of each news outlet during the same eight-month period of our data collection process. As a result, we download 914,671 unique news articles in total, which we then filter to keep only the negative cases 535,841 (58%) by removing articles posted at least once on Reddit. We use random sampling to balance the number of positive and negative cases and keep our model's random baseline (0.5). In this prediction task, we use the same feature sets used to predict the other engagement levels (L, M, and L+M) using the news articles' titles. Also, we conduct a random forest feature selection to reduce the number of L features to 1,000 and prevent over-fitting.

Posted Multiple Times: we predict the volume of Level-4 engagement through building a model that predicts those articles that would be publicly shared multiple times (>= 2) on Reddit (Class-1) and those which will be shared only once (Class-0). We use our collected Reddit dataset to build both prediction classes, which we balance by random sampling; thus, the prediction's random baseline is (0.5). Moreover, we use the same feature sets used for the other prediction tasks using the text of news articles' headlines. This prediction task is important because a news article that is posted multiple times into Reddit indicates a higher exposure of the article to different Reddit sub-communities, which means its public sharing reaches Level-4 engagement.

5. Results

khI do not understand the differences among Tables 5–8. I think that Tables 6 and 7 are more complete. For this reason, I think that the Tables 5 and 8 should be similar.

5.1. Exploratory analysis

For answering the first research question, we first show the topical distribution in our dataset. Figure 2 shows the top 10 topics in our data collection with the distribution ratio across the platforms. The topics' labels are assigned by manually checking the top 10 words in each topic as shown by Twitter-LDA (Zhao et al. 2011). We identify the top topics by summing up the frequency rate over the five platforms for each topic, rank them in descending order, and then select the top 10 ranked topics.

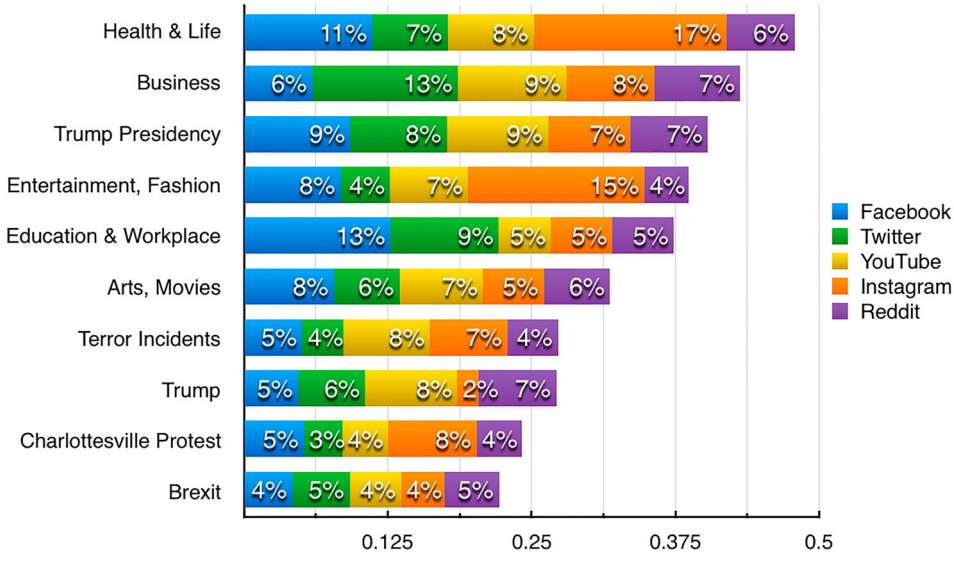

Figure 2. The top 10 topics and the distribution frequencies across the five social media platforms.

Topics have different distribution rates across platforms. For example, the Health & life topic has more posts on Instagram (17%) than on Facebook (11%), YouTube (8%), Twitter(7%), and Reddit (6%). This indicates that news outlets target a specific platform to push more content about a given topic. Content on the Education & workplace topic is more numerous on Facebook, with 13% among the 53 news outlets, making Facebook the ultimate targeted platform for similar content. Moreover, some topics do have almost similar distribution rates across multiple platforms, such as Trump Presidency, which has 9% on both Facebook and YouTube, 8% on Twitter, and 7% on both Instagram and Reddit.

The content-topic with various distribution across platforms strike different possible reasons, including the platform's distinguishable traits or users' platform-specific preferences of topics. Thus, for a given post topic, some platforms are more suitable and popular for that topic, while others have similar distribution rates (see Figure 2).

5.2. RQ1:What are the topical distributions and user engagement differences among (a) the platforms and (b) news organisations?

5.2.1. Top five topics average engagement

Now we start answering part (a) of the first research question RQ1. The average number of likes and comments for each of the top five topics for the five platforms (FB, IG, TW, YT, and RD) is shown in Table 3. The topics are ranked by the number of postings in our dataset for each topic with topic number #1 having the highest number of postings, while topic #5 having the lowest number of postings. From the table, we can highlight two different observations. One is that the higher number of postings of a topic does not guarantee a higher number of likes and comments. For example, topic Health & life is ranked 5 in YT; however, some other higher ranked topics on YT (e,g., Business, and Terror Incidents) have less average comments (C) and likes (L). Also, topic Health & life in YT has higher than average comments (183) and likes (1934); this is a higher average than all other topics: comments (142) and likes (672). The second observation is that some topics appear on the top five most frequently related posts on multiple platforms (e.g. Health & life) while having distinct average engagement ranges across platforms. This might be wired to either difference in the user base or their platform preferences. In Table 3, we show the volume of topic distribution in ranked order with the average engagement (likes and comments) across five platforms, which answers RQ1 (a). In this,

Table 3. The Average Likes (L) and Comments (C) for the top 5 frequent topics of each social media platform and the average (Avg.) of all topics in the last column.

Platform	#1			#2			#3			#4			#5			Avg.		
FB	Education & Workplace			Health & life			Trump Presidency			Entertainment, Fashion			Arts, Movies			All topics		
	#P	C	L	#P	C	L	#P	C	L	#P	C	L	#P	C	L	#P	C	L
	3,372	514	2837	2,949	490	2554	2,441	823	2472	2,229	574	3562	2,063	350	2022	1,323	563	2643
IG	Health & life			Entertainment, Fashion			Sports			Charlottesville Protest			Business			All topics		
	#P	C	L	#P	C	L	#P	C	L	#P	C	L	#P	C	L	#P	C	L
	5,708	152	8889	5,230	283	16517	2,621	1370	95783	2,620	249	9060	2,598	79	5072	1,705	359	16947
TW	Business			Education & Workplace			Trump Presidency			Health & life			Trump			All topics		
	#P	C	L	#P	C	L	#P	C	L	#P	C	L	#P	C	L	#P	C	L
	71,378	8	69	53,148	17	145	47,505	37	205	37,047	18	133	32,747	44	205	28,256	28	276
YT	Business			Trump Presidency			Trump			Terror Incidents			Health & life			All topics		
	#P	C	L	#P	C	L	#P	C	L	#P	C	L	#P	C	L	#P	C	L
	4,056	45	347	3,778	301	904	3,450	161	370	3,231	102	219	3,230	183	1934	2,150	142	672
RD	Business			Trump			Trump Presidency			Health & life			US Policy Immigration			All topics		
	#P	C	L	#P	C	L	#P	C	L	#P	C	L	#P	C	L	#P	C	L
	185,237	7	54	167,886	10	87	166,401	9	78	146,795	7	56	146,252	7	62	123,821	7	58

we use two engagement metrics (likes and comments), which are measurable over the five platforms to facilitate the cross-platform comparison.

5.2.2. Organisation topic similarity across platforms

Part (b) of RQ1 asks, 'What is the topical distribution and user engagement differences among (a) the platforms and (b) news organisations?' In answer to this, we show how the content topics are distributed differently over the five platforms by news outlets. We use the Jensen-Shannon Divergence (JSD) score, which calculates the similarities between two different probability distributions (Lin 1991). We then create pairs of platforms using FB, TW, IG, YT, and RD data to calculate the JSD score. The created 10 platform pairs are: FB-IG, FB-TW, FB-YT, IG-TW, IG-YT, TW-YT, RD-FB, RD-IG, RD-TW, and RD-YT. Since we are comparing news outlets across platforms, we use 51 news outlets for this analysis because the number of available news outlets for YT is 51. For pairs with Reddit (RD) data, the comparison is between to topics of posts published by the news outlets themselves and the topics of posts posted by Reddit users and linked to the news article of the same news outlet.

The JSD scores are not affected by the platform pair order; thus, TW-FB and FB-TW have the same scores, so we only keep one pair. The JSD score of a news outlet (n) for a specific platform pair (p_i and p_j) is calculated using the following equation:

$$JSD(p_i \| p_j | n) = \frac{1}{2} D(p_i \| p_j | n) + \frac{1}{2} D(p_j \| p_i | n), \qquad (3)$$

where JSD is based on Kullback-Leibler divergence $D(p_i \| p_j | n)$, which is calculated using the following equation:

$$D(p_i \| p_j | n) = \sum_k P(k|p_i, n) \log \frac{P(k|p_i, n)}{P(k|p_j, n)}, \qquad (4)$$

where k is number of topics (k=20 in our case) and $p(k|p_i, n)$ is the probability of topic k when news outlet (n) posts on platform p_i.

By calculating the JSD scores over the pair of platforms, we determine the differences of topic distribution across each pair of platforms and for each news outlet. A JSD score of zero indicates that a news outlet is distributing completely different topics across the two platforms, while a JSD score of 1 refers to the identical distribution of topics among the tested pair of platforms. In Figure 3, we show the total number of news organisations for each JSD score (0.1 to 0.8) and platform pair (10 pairs). Among all platform pairs, the FB-TW pair has the lowest similarities between topics for (46) news outlets, which means they share different topics. Of the 510 pairs, 240 (47%) have a JSD score of $<= 0.1$. However, there are (21) news outlets with JSD scores greater than 0.5, meaning they share similar topics between FB-IG, IG-TW, IG-YT, RD-FB, RD-IG, RD-TW, and RD-YT. The overall low similarities of content across social media platforms for individual news organisations, as shown by the JSD scores, indicate that a model predicting user engagement on one platform is likely not to be transferable to another platform. This answers part (b) of RQ1.

5.3. RQ2: (a) is user engagement affected by the content topic? (b) if so, which topics generate which levels and what volume of user engagement?

After answering the first research question, we know that news outlets distribute topics differently across

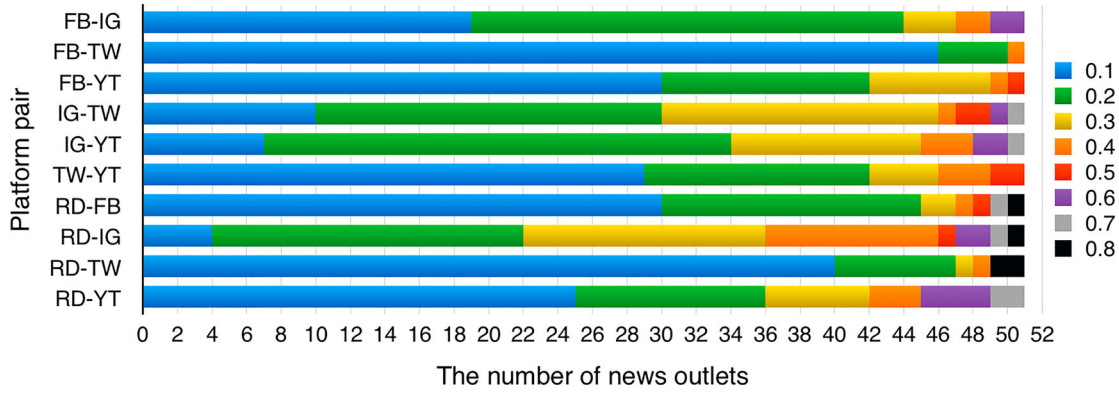

Figure 3. JSD scores for the 10 social media platform pairs.

platforms, and generally, a higher posting volume for one topic does not guarantee having more engagement. However, we want to analyse further individual topics and their effect on user engagement across the five platforms. We move on to answer RQ2: (a) *Is user engagement affected by the content topic?* and RQ2 (b) *If so, which topics generate which levels and what volume of user engagement?*. To find the significant effect of topics on user engagement, we use the chi-square measure. By using the chi-square measure, we can find whether there is a significant difference between the number of high engaging posts and the low engaging posts for each topic. We identify the high engaging posts by considering the top 33% quantile of posts for a given engagement metric (e.g. views, likes, comments). Meanwhile, we consider the posts of the bottom 33% quantile for the same engagement metric as low engaging posts. We do not use the middle 33% of posts to ensure there is a significant difference between the engagement values of high and low engaging posts. Then we calculate the chi-square metric, Table 4, in each platform and for the four engagement levels using their engagement metrics: NV, NL, NC, NEP, and EPC. When using the EPC of Level-4 engagement, we compare the number of posts posted once on Reddit with the number of posts posted multiple times.

In Table 4, we show the significant results of the 20 topics, which reveal the differences in user engagement among topics and across platforms. From this, we observe two types of topics. One is topics that have either high or low engagement across most platform and engagement levels. For example, the topics of Charlottesville, Police Violence, and North Korea have high engagement across most platforms. The other type is topics that have mixed engagement (low and high) varying by platform and engagement metric. For example, the Arts, Movies topic and the Entertainment topic have low engagement on YouTube (Level-1 and Level-3) but high engagement on Twitter (Level-2 and Level-3). Through comparing the chi-square results with the top five topics of each platform (Table 3), we also find that the volume of posting for topics does not always generate higher engagement. As an example, the topic of Education & Workplace is the top-ranked topic on Facebook in terms of posting volume (3372 posts) by news outlets; however, there is no significant correlation with engagement on Facebook.

Additionally, a few other topics have different user engagement volume (high and low) over different engagement levels (levels 1 to 4) in the same platform. On YouTube, posts related to the Immigration topic have low views and likes, but higher comments. With EPC, which analyses if posts on Reddit are made multiple times, it can be noticed that many topics have multiple postings, such as Business and Entertainment. Multiple postings regarding content indicate a higher level of Level-4 engagement. Based on these results, news content producers can target a specific topic and a particular engagement level for a platform to generate higher user engagement. For example, for collecting users' feedback about a movie, one would better focus on Twitter since users of that platform comment more actively on movie-related posts in comparison to other platforms (YouTube, Facebook, and Instagram). Thus, each news outlet has to identify the key performance indicators, which, in turn, helps it adjust the distribution strategies of its posts based on the topic and the targeted platform. The results of the chi-square measures over the 20 topics addressed by RQ2. Our results complement the work advanced by another study (Ferrara, Interdonato, and Tagarelli 2014), which has been done on Instagram only, where many factors might affect the popularity of online content including the topical interests, the platforms structural features, and the content production.

5.4. RQ3: how are posts topics related to sentiment of (a) posts and (b) audience' comments on those posts?

In the previous analysis of the first two research questions (RQ1 and RQ2), we study the behavioural aspect of user engagement through comparing the interaction between users and content topics (view, like, comment, share) across multiple platforms and news outlets. Now, we move to the psychological aspect of engagement by studying the sentiment of news posts and the comments on them. Here, we address RQ3: *How are post topics related to the sentiment of (a) the posts and (b) the audience's comments on those posts?*. In answering this question, one can predict whether the audience of news outlets are interacting positively or negatively with different news topics. Across social media platforms (FB, IG, TW, and YT), the comments are one way to understand the audience's point of view and whether it correlates to the content polarity of actual posts. We expect posts with positive sentiment to trigger more positive comments; similarly, negatively polarised posts are expected to gain more negative comments. Also, since the level of audience engagement differs by platform and topic, we expect that the audience might behave differently across platforms (e.g. having higher scores of negative comments on Twitter rather than on Facebook).

An exploratory box plot of sentiment scores for comments and posts is shown in Figure 4 for the four

Table 4. Chi-square results for the 20 topics and 4 engagement levels (L-1, L-2, L3, and L4).

platform	YouTube			Facebook		Instagram		Twitter		Reddit			
level	L-1	L-2	L-3	L-2	L-3	L-2	L-3	L-2	L-3	L-2	L-3	L-4	L-4
Eng. Metric	NV	NL	NC	NL	NC	NL	NC	NL	NC	NL	NC	EPC	NEP
Arts, Movies	Low*		Low***					High***	High***		High***	High***	
Brexit			High***		High*			High**	High***	High**	High***		
Business	Low***	Low***	Low***					High***	High***		High***	High***	
Charlottesville	High***	High***	High***				High***		High**	High***	High***		
Education and workplace	Low***		Low***					High*		High*	High**	High***	
Entertainment	Low***		Low***					High***	High***	High*	High***	High***	
Health & life	Low***		Low***			High**	Low**		High***	High*	High***	High***	Low*
Immigration	Low**	Low**	High***			Low**				High**			
North Korea	High***	High***	High***					High*		High*	High***		
Police Violence	High***	High***	High***						High***	High***	High***	High*	
Sexual Assault	Low**	Low***	High**			Low*			High**		High***	High***	
Sports			Low**						High***	High***	High***		
Terror incidents	High***	Low***	Low***		Low**				High***		High***		
Trump	High**		High***	Low**				High**	High***		High***	High***	High***
Trump presidency	High***	High***	High***	High*			High***		High***	High***	High***	High**	
US election	Low***	Low***						High***	High**		High***		
US environmental policy			High***						High***	High*	High***		
US Health Care	Low***	Low***	Low***				High*		High***	High**	High***		
US sports			Low***						High***	High***	High***		Low*
War on Terror	High**								High*		High***		

Notes: Low indicate low engagement of topics, while High means high engagement. Using Reddit data, we compare news articles posted once versus those posted multiple times using EPC engagement metric column. Significant level codes: * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

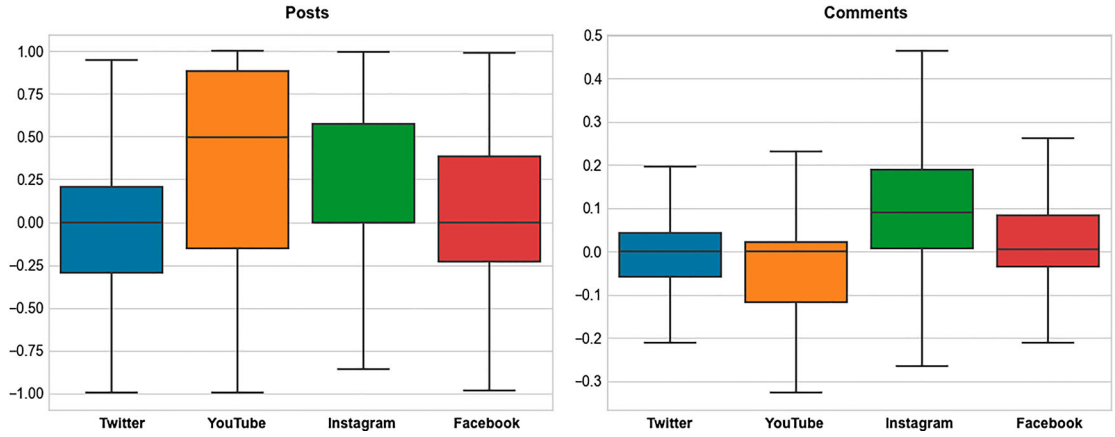

Figure 4. The box plot of average Vader sentiment scores of posts and comments over the four platforms.

platforms and 53 news outlets, which shows the general differences without the topical effect. For example, the sentiment of YouTube posts compared to that of other platforms is skewed towards the positive ($>= 0.5$). However, the sentiment scores of YouTube users' comments are, on average, neutral (< 0.5 and > -0.5). This might not be true if we analyse the sentiment by topics and in relation to post and comment volume.

For understanding the higher level of sentiment effect according to the posting volume, we calculate the Spearman correlation between the normalised number of comments and the percentages of positive, negative, and emotional (positive + negative) comments separately. The number of comments is normalised through dividing by the maximum number of comments of all posts of a given news organisation. The posts of all news organisations are used as one input for each platform to calculate the correlation. We find a significant positive correlation over the four platforms; that is, increasing the number of posts in a platform would generally increase the amount of positive, negative, and emotional comments. For example, 0.36***, 0.24***, 0.56***, and 0.63*** are the correlation values of the normalised number of comments and the percentage of emotional comments for Facebook, Instagram, YouTube, and Twitter respectively, where *** indicates having a p-value $<= 0.001$.

But this correlation is general and does not show the differences across topics. As a result, for each topic, we calculate the same correlation of its posts, which end up having 20 Spearman correlation values–one for each topic and for each platform. Accordingly, on Facebook, Twitter, and YouTube, all the 20 topics have significant positive correlations between the volume of comments and the number of positive, negative, and emotional comments. However, we notice some differences with Instagram topics. The first difference is that the topics 'Trump Presidency' and 'US Policy Immigration' have significant negative correlations with positive sentiment comments, which means that posts with a higher number of comments have fewer positive comments. Another difference is that some topics (e.g. 'Trump,' 'US Environmental Policy,' 'US Health Care Debate,' 'Police Violence,' 'Sexual Assault,' 'Sports,' and 'US Election 2016') do not have significant correlations in terms of positive sentiment comments while other topics (e.g. 'War on Terror,' 'Life, Entertainment, Fashion,' 'Arts, Movies,' 'Terror Incidents, Europe,' and 'Education and Workplace') have significant positive correlations between comments volume and positive sentiment comments volume. In terms of negative sentiment comments, the results show that a higher volume of comments for a post indicates a higher volume of negative sentiment comments for all 20 topics. That means comments tend to be more negative on Instagram.

We now focus on how different topics are covered with different sentiments by news outlets and how this relates to audience emotions. Specifically, we compute the proportions of positive and negative posts for each

Table 5. Level-1 engagement prediction results: the aggregated F1-score (average (avg.), minimum (min), median (med), maximum (max), and standard deviation (std)) of Level-1 views prediction on YouTube.

		YT				
		Avg.	Min	Med	Max	Std
AdaBoost	L	0.612	0.481	0.610	0.810	0.086
	M	0.603	0.449	0.604	0.720	0.057
	L+M	0.653	0.485	0.650	0.808	0.070
Decision Tree	L	0.613	0.469	0.610	0.803	0.084
	M	0.594	0.474	0.597	0.716	0.051
	L+M	0.648	0.498	0.644	0.800	0.069
Random Forest	L	0.633	0.480	0.631	0.814	0.085
	M	0.596	0.445	0.598	0.721	0.058
	L+M	0.651	0.480	0.657	0.810	0.070

Table 6. Level-2 engagement prediction results: the aggregated F1-score (average (avg.), minimum (min), median (med), maximum (max), and standard deviation (std)) for Level-2 prediction results.

		FB					IG					TW					YT				
		Avg.	Min	Med	Max	Std	Avg.	Min	Med	Max	Std	Avg.	Min	Med	Max	Std	Avg.	Min	Med	Max	Std
AdaBoost	L	0.585	0.430	0.565	0.807	0.084	0.652	0.486	0.680	0.752	0.066	0.677	0.500	0.665	0.999	0.086	0.613	0.497	0.599	0.779	0.083
	M	0.543	0.432	0.538	0.701	0.061	0.595	0.479	0.593	0.672	0.046	0.612	0.492	0.589	0.977	0.090	0.606	0.463	0.605	0.725	0.051
	L+M	0.611	0.497	0.608	0.826	0.076	0.667	0.500	0.676	0.758	0.065	0.678	0.497	0.670	0.998	0.085	0.653	0.431	0.656	0.791	0.071
Decision Tree	L	0.584	0.426	0.570	0.806	0.081	0.646	0.506	0.669	0.734	0.062	0.681	0.513	0.670	0.999	0.083	0.612	0.483	0.595	0.772	0.080
	M	0.542	0.441	0.540	0.703	0.057	0.590	0.480	0.595	0.666	0.044	0.606	0.503	0.588	0.982	0.090	0.596	0.474	0.595	0.706	0.047
	L+M	0.611	0.489	0.597	0.861	0.077	0.663	0.520	0.670	0.752	0.061	0.677	0.505	0.665	0.998	0.085	0.647	0.434	0.645	0.778	0.071
Random Forest	L	0.597	0.434	0.587	0.811	0.083	0.668	0.526	0.687	0.776	0.064	0.717	0.562	0.706	0.999	0.075	0.631	0.471	0.628	0.783	0.082
	M	0.537	0.420	0.533	0.717	0.064	0.591	0.481	0.593	0.676	0.052	0.605	0.503	0.587	0.972	0.091	0.599	0.449	0.602	0.720	0.053
	L+M	0.608	0.475	0.605	0.774	0.074	0.670	0.494	0.679	0.764	0.065	0.686	0.501	0.672	0.998	0.088	0.654	0.429	0.657	0.788	0.070

Table 7. Level-3 engagement prediction results: the aggregated F1-score (average (avg.), minimum (min), median (med), maximum (max), and standard deviation (std)) for Level-3 prediction results.

		FB					IG					TW					YT				
		Avg.	Min	Med	Max	Std	Avg.	Min	Med	Max	Std	Avg.	Min	Med	Max	Std	Avg.	Min	Med	Max	Std
AdaBoost	L	0.611	0.422	0.600	0.809	0.089	0.623	0.486	0.612	0.748	0.067	0.661	0.481	0.649	0.799	0.086	0.615	0.429	0.609	0.761	0.086
	M	0.558	0.471	0.551	0.677	0.053	0.565	0.493	0.549	0.700	0.055	0.573	0.478	0.555	0.767	0.065	0.619	0.526	0.629	0.761	0.055
	L+M	0.619	0.499	0.610	0.772	0.067	0.626	0.531	0.619	0.748	0.061	0.674	0.449	0.673	0.843	0.086	0.669	0.539	0.662	0.802	0.065
Decision Tree	L	0.612	0.427	0.605	0.799	0.087	0.620	0.489	0.607	0.735	0.066	0.662	0.486	0.649	0.796	0.084	0.616	0.440	0.606	0.755	0.081
	M	0.556	0.463	0.553	0.660	0.051	0.560	0.492	0.546	0.687	0.051	0.568	0.484	0.554	0.752	0.060	0.608	0.520	0.615	0.747	0.052
	L+M	0.620	0.506	0.599	0.777	0.068	0.622	0.533	0.612	0.747	0.059	0.677	0.476	0.673	0.879	0.087	0.661	0.534	0.650	0.793	0.063
Random Forest	L	0.616	0.400	0.611	0.821	0.092	0.633	0.502	0.626	0.760	0.069	0.687	0.486	0.686	0.814	0.086	0.630	0.395	0.630	0.779	0.089
	M	0.554	0.465	0.553	0.669	0.054	0.560	0.484	0.537	0.698	0.056	0.566	0.468	0.548	0.761	0.064	0.610	0.507	0.622	0.756	0.057
	L+M	0.617	0.487	0.605	0.774	0.071	0.627	0.520	0.611	0.752	0.063	0.668	0.424	0.680	0.810	0.079	0.668	0.520	0.664	0.802	0.067

Table 8. Level-4 engagement prediction results: the aggregated F1-score (average (avg.), minimum (min), median (med), maximum (max), and standard deviation (std)) across 29 news organisations by three models (adaboost, decision tree, and random forest) and using three features sets (languages (L), metadata (M), and (L+M)).

		F1-score				
	Features	Avg.	Min	Med	Max	Std
AdaBoost	L	0.62	0.54	0.61	0.93	0.09
	M	0.63	0.54	0.63	0.78	0.06
	L+M	0.68	0.58	0.67	0.94	0.08
Decision Tree	L	0.68	0.57	0.66	0.93	0.08
	M	0.66	0.56	0.66	0.86	0.07
	L+M	0.71	0.62	0.71	0.93	0.08
Random Forest	L	0.64	0.55	0.62	0.93	0.08
	M	0.67	0.57	0.65	0.80	0.06
	L+M	0.67	0.56	0.66	0.94	0.08

topic. Then, for each topic, we aggregate all comments on the posts on the topic and compute the proportions of positive and negative comments for that topic. Then, we compare the percentage values in each platform. Using the proportions of positive and negative posts and comments, we reveal topics that are more positively (or negatively) covered by news outlets and reacted to by audience.

Facebook: When calculating the percentages of positive and negative posts and comments for individual topics, we observe several different distributions of sentiments across topics. First, there are some topics with higher percentages of positive posts than negative ones (e.g.topic 'Arts, Movies', 'Sports', 'US Sports') with a difference in percentages ranging from 1% to 10% for 13 topics. Second, there are 7 topics with negative posts dominating positive posts with a difference range of 1% to 7.5%. For example, for the topic 'Terror Incidents, Europe,' 20.3% of related posts are negative while 12.78% of posts are positive, making a difference of 7.53%. Third, for 12 topics, the percentage of positive comments is a little higher than the percentage of negative comments (0.39–2.64), where the topic 'Sports' has the highest difference. Meanwhile, the remaining 8 topics have more negative than positive comments with a percentage difference of 0.47–3.84, where the topic 'US election' has the highest number of negative comments (15%). Fourth, when comparing the sentiment of both posts and comments, we observe two topics having higher percentages of negative sentiment for both comments and posts (topic 'Police Violence' and 'US election'), with a minimum difference $> 2\%$. Moreover, five topics are rated higher as positive posts (with $> 6\%$ difference) and have higher positive comments ($> 2\%$), which can indicate positive audience reactions to positive posts on those topics (topic 'Sports', 'Health & life', 'Business', 'Education and Workplace',and 'Life, Entertainment, Fashion'). Another interesting observation is that some topics (e.g.topic 'Arts, Movies', 'Brexit', and 'US Sports') have a higher number of positive posts (with $> 6\%$ difference from negative percentage); however, the comment differences between positive and negative are low $< 2\%$, hence audiences have almost equal opinions over these topics. The last observation in this analysis is that the topic 'Terror Incidents, Europe' has 12.78% and 20.33% positive and negative posts, respectively, while having 13.24% and 12.08% of positive and negative comments, respectively. Thus, a higher number of negative posts on a topic does not necessarily lead to more negative comments.

Instagram: The percentage of news postings with positive sentiments exceeded the negative ones over almost all topics, except 'Sexual Assault,' with a range of 1.33%–23.94%. Similarly, the percentages of comments with positive sentiment are valued more than negatively polarised comments (with a difference range of 7.19–1.37) over all topics except the topic 'US Environmental Policy,' which has a balanced amount of both sentiment types with zero difference. We notice here that the percentage range of posts and comments with positive sentiment are higher on Instagram compared to Facebook, which can indicate that both news organisations and audiences prefer to act more positively over most topics on Instagram.

Twitter: The sentiment of both news postings and users' comments are more negatively oriented across most of the topics where the posts (comments) difference range is 0.41–11.72 (2–5.36). However, for one topic, 'Sports,' the total number of positive posts and comments is greater than the negative ones with 0.95% for posts and 1.94% for comments. Also, the topic 'US Sports' has 12.91% more negative postings than positive postings (11.59%), but its corresponding user comment sentiment is almost similar: 11.67% (positive) and 11.39% (negative).

YouTube: Positivity is pronounced more on news posting of YouTube, with 16 topics having 15.54%–57.59% higher percentages of positive sentiments posts over negative posts. Out of the 16 topics, there are 9 topics (topic 'US Policy Immigration' and 'Charlottesville') with higher positive postings but more negative user comments and 7 topics (topic 'Life, Entertainment, Fashion' and 'Education and Workplace') with higher percentages of positive comments. Whilst, the two topics (Terror Incidents, Europe and Sexual Assault) having a very low difference between positive and negative posts percentages (0.90 and 0.13) where both have ($> 11.24\%$) higher negative comments percentages

than positive percentages, the final two topics (War on Terror and Police Violence) have (> 3%) higher negative posting and (> 12%) higher negative comment percentages over the positive ones.

Depending on the platform and topic, the sentiment of users' comments may vary. A higher engagement (e.g. Level-3 comments) of a topic in a given platform would not necessarily mean it is a positive engagement. For checking the significance of comments' volume with the comments' sentiment over topics, we check the Spearman correlation to know whether increasing the number of comments would lead to more positive or negative comments. Generally, for a given topic, increasing the number of comments will increase the amount of negative sentiment in Facebook and YouTube whilst decrease the amount of positive comments in Twitter and YouTube.

We have, so far, compared the amount of comments in relation to the volume of comments with positive and negative sentiments. For understanding whether the sentiment score of posts is significantly affecting the sentiment score of users' comments for each topic, we calculate the Pearson correlation of the sentiment score (Vader score) of posts and the average sentiment of the comments for each topic. For Facebook, Instagram, Twitter, and YouTube, the results show significant positive correlations between the sentiment score of a post and its comments' average sentiment scores for all 20 topics. This finding is congruent with a previous study showing that the average sentiment of comments on positive posts is more positive than the average sentiment of comments on negative posts (Svetlov and Platonov 2019).

In calculating the percentage of positive (or negative) sentiment posts and the average percentage of positive (or negative) comments per post for each topic, we used the Spearman rank correlation to calculate their correlation across all 20 topics; thus, the input to the correlation here is 20 rows, and each row presents a topic. We found a positive significant correlation for the topics on Facebook and YouTube, but not Twitter and Instagram. On the other hand, over the four platforms (Facebook, Instagram, Twitter, and YouTube), we found a significant positive correlation between the percentage of negative posts and the corresponding average of negative comments across topics. That is, for a given topic, having more negative posts about the topic results in more negative user comments. A previous work showing that the number of comments is higher for the posts with a negative sentiment (Svetlov and Platonov 2019) confirm with our findings, where we add in top of that more negative comments are associated with negative posts.

For further investigation, we compared the top and bottom 5 topics, based on their correlation, across platforms. For the top 5 topics, the correlation coefficient is greater than or equal to 0.236 (FB), 0.336 (IG), 0.211 (TW), and 0.279 (YT). We found some topics have a very high correlation, appearing in the top 5 list, across three platforms, such as the topics 'War on Terror' (FB, TW, IG), 'Police Violence' and 'Terror Incidents, Europe' (TW, IG, YT), and 'US Policy Immigration' (FB, IG, YT). Meanwhile, there is one topic present in the list of bottom 5 correlation of posts-comments sentiment score across the four platforms; this is 'Business' (FB, IG, TW, YT). Also, the topics 'Brexit' and 'Sports' have less correlation across three platforms. A few topics appear in the top correlation in one platform but in the bottom list of another platform. For example, the topic 'Trump' has a high correlation in TW but has less correlation in FB.

We examine different correlation factors using the engagement volume of comments and the sentiment of both posts and comments within 20 different topics. There exists a relationship between the sentiment of posts and the sentiment of comments which vary by platform and topic. These results address RQ3.

5.5. RQ4:Can we predict the volume of the different engagement levels for a given content: (a) Level-1, (b) Level-2, (c) Level-3, and (d) Level-4?

The Tables 5–8, show the engagement prediction results for level-1 (views), level-2 (likes), level-3 (comments), and level-4 (External posting), respectively. Due to data limitation of level-1 engagement, we experiment using only YouTube data. Also, for level-4 engagement we have the external posting metric on Reddit, while for other platforms the data is not available due to data privacy reasons.

5.5.1. Predicting Level-1 engagement

Now, we show the results of the first prediction experiment of Level-1 engagement, which answers part (a) of RQ4. Due to data limitations, we use this result to predict the volume of views for YouTube posts.

The prediction summary of the 23 news sources is shown in Table 5, where the average F1-score over the three algorithms is 0.65 using the (L+M) feature set. The maximum F1-score is 0.813 using (L) features and Random Forest Classifier for (MSNBC) a news source where the total of posts is 1,714 (857 for each class). Based on this result, we can conclude that one can predict the first level of engagement with content before posting with an average F1-score of 0.65, which addresses part (a) of RQ4.

5.5.2. Predicting Level-2 engagement

Table 6 shows the results grouped into average, minimum, medium, maximum, and standard deviation. We observe from the minimum and maximum values that some news organisations have high F1-score s, such as USA Today, which has a 0.861 F1-score using Decision Tree and L+M feature set for the Facebook platform. Other news sources have a low F1-score, even though the classes are balanced. For example, on Facebook, using Decision Tree and L+M features, the minimum F1-score is 0.489 for The Atlantic news source. This partially indicates that some news media have a more predictable Level-2 engagement than others; however, in the second case, the content features are not enough, and the extended features set (e.g. audience interest and location) need to be investigated in future studies. However, the overall average of prediction over Facebook news sources is 0.611 using the same setting as the best-performing news source, 'USA Today.'

That answers part (b) of RQ4, where the engagement of Level-2 can be predicted for a given news outlet with different average F1-scores over platforms.

5.5.3. Predicting Level-3 engagement

Moving into the third level of engagement, we use the number of comments to present the prediction power of this level. In this section, we answer part (c) of RQ4. Our dataset affords the comment volume of each news source posting over four platforms: Facebook, Instagram, Twitter, and YouTube. Thus, the prediction experiments are conducted over the four platforms. In the filtering process to eliminate news sources with fewer than 100 posts for each classification class (a total of 200 posts), we end up with the following number of news sources: 51, 28, 52, and 23 for Facebook, Instagram, Twitter, and YouTube, respectively.

In Table 7 we show the results for the comment volume prediction experiments in aggregated values over the news sources. The average approximate F1-score over the three algorithms is better using both a language and metadata (L+M) feature set. In Facebook, the average F1-score (0.611) is the highest and equal for both Ada Boost and Decision Tree, while Instagram Random Forest quietly performed better, with an average F1 of (0.627). Moreover, for Twitter, the Decision Tree average F1 is higher (0.677). Meanwhile, for YouTube, the Ada Boost model performed a little better with an average F1-score of 0.669. As a result, we can predict the comment volume of Level-3 engagement with an average F1-score of > 0.61, which answers RQ4 part (c).

5.5.4. Predicting Level-4 engagement

We now move to Level-4 engagement prediction of RQ4 part (d). Here, we predict if a news article will receive the highest engagement level and volume using two different prediction models: (1) predicting whether a news article will receive Level-4 engagement if users post it at least once on Reddit (posted once) and (2) predicting whether a news article will receive a higher volume of Level-4 engagement if users post it multiple times on Reddit (posted multiple times).

Posted Once: The engagement metric we predict in this model is Level-4 external posting for each news organisation separately. More precisely, for a given news article, we predict whether Reddit users would post it at least once on Reddit or not. We assemble 4000 random samples of news articles, divided equally to the two prediction classes (2000 each). Class-1 contains positive samples of news articles posted at least once on Reddit, while Class-0 consists of the negative samples of articles not posted on Reddit. In this prediction experiment, we use 29 news organisations and excluded the rest, which contained less than 4000 sample cases.

We use the same feature sets of the previous prediction experiments (L, M, and L+M) and the same three algorithms (i.e. AdaBoost, Decision Tree, and Random Forest). In Table 8, we report the aggregated F1-score results (average, minimum, median, maximum, and standard deviation) of the 29 news organisations. When using only L features, the best-performing model is a decision tree, with a 0.68 average F1-score; when using only M features, random forest is the best with a 0.67 average F1-score. However, the decision tree model with L+M features performs better than other features with a 0.71 average F1-score. Also, we noticed a significant difference between results from different news organisations when comparing the maximum and minimum F1-scores. For instance, Forbes news media has a maximum performance (0.93) compared to BBC News, which has the lowest performance (0.62) when using a decision tree with L +M features.

Overall, the prediction results of articles that will be posted once on Reddit show that Level-4 user engagement can be predicted with 70% average precision. Thus, news content creators can use this model to build advanced knowledge about articles that will receive higher engagement or external sharing before posting the article into other platforms. Even more, by using simple metadata features (topic, textual, and sentiment), content creators can foretell Level-4 user engagement with an average precision of 68% when using our model.

Posted Multiple Times: By looking into the network structure of Reddit and its sub-community, it is obvious that news articles get more exposure if they are posted multiple times on Reddit. In measuring the volume of Level-4 engagement for this prediction task, we distinguish news articles that will be posted multiple times by users on Reddit from those articles that will be posted only once. For this task, we also consider a sample of 2000 positive cases of news articles posted multiple times on Reddit (Class-1) and 2000 news articles that are posted only once (Class-0). Thus, we use 29 news organisations that have enough cases for both classes. We use the best performing algorithm (Decision Tree) from the posted-once prediction experiment and the L+M feature set. The average, minimum, median, maximum, and standard deviation values of the aggregated F1-score of 29 news organisations are 0.83, 0.63, 0.85, 0.94, and 0.08, respectively.

On average, our model predicts news articles that will receive multiple Reddit postings by 0.83 F1-score. Compared to the Level-4 posted-once prediction task, predicting articles that will receive multiple postings by users is easier than predicting if it will be posted once. Through combining both prediction models, content creators can predict articles that will receive a higher engagement level (Level-4), which answers part (d) of RQ4.

6. Discussion and implications

In this research, we provide both theoretical and practical frameworks for user engagement, divided into four levels. Through adopting the eWOM concept (Chen, Hong, and Li 2017; Jansen et al. 2009) as the foundation outcome of the network effects of social media, we provide the theoretical framework of multiple levels of user engagement metrics that measure the extent to which a user would publicly engage with social media content. The more popularity a social media news content gains, the higher the user engagement and outreach; thus, the ultimate goal of news outlets is to widen their user reach.

It is clear, from our analysis, that news outlets push content into social media platforms differently. That is, for selective topics, they highly push content to one platform but more (or less) for other platforms. Similarly, for the content sentiment, they try to be more positive on some platforms (e.g. Instagram), relative to others. The difference can be related to the fact that news outlets are actually adapting to some degree to the various user bases across platforms. Researchers in Al-Rawi (2019) who studied two platforms (Twitter and YouTube) confirmed the differences in the news consumption patterns of online audiences across platforms.

Moreover, there are significant differences between high and low user engagements across platforms and across the four engagement levels. That means, for each topic, the user interactions for a topic for a specific platform are less likely to be perceived on another platform, with few exceptions, which show similar user engagement patterns.

There are few topics (e.g. Police Violence, North Korea, and Charlottesville) that have significantly high engagement on YouTube for three engagement levels (Level-1, Level-2, and Level-3). Meanwhile, some other topics (e.g. Business, US Health Care) have low engagement on YouTube for the three similar engagement levels. On Facebook, most of the topics (16 out of 20) do not have a significant difference between low and high engagement volume. The exceptional four topics are Brexit (Terror Incidents), which has high (low) Level-3 engagement, and Trump (Trump Presidency), which has low (high) Level-2 engagement. Interestingly, some platform pairs show fewer differences among topics and engagement levels. For example, Twitter and Reddit have high Level-3 engagement across 17 topics, while other platform pairs (e.g. YouTube and Facebook) have more pronounced differences in engagement volumes across topics, which indicate users' willingness to publicly show their engagement based on the topic.

In practice, content producers can specify their targeted platform and engagement level for a given content and check the effectiveness through employing similar analysis as this research. More specifically, in this work, we examined Level-4 engagement (public sharing), which indicates the highest volume of user engagement in our framework, thus reaching more users (directly or indirectly) on other platforms than the existing user base of an organisation.

The volume of engagement is not always a good indicator, since the sentiment may vary positively or negatively. A higher engagement (e.g.Level-3 comments) of a topic in a given platform would not necessarily mean it is a positive engagement. As such, in this research, we also analyse how topics of posts are related to the sentiment of posts and the audience's comments across platforms. Interestingly, on Instagram, both news organisations and audiences prefer to act more positively over most topics. On Facebook, we observe that higher negative postings of a topic do not necessarily lead to more negative comments. While on Twitter, the sentiment of both news postings and users' comments are more negatively oriented across most of the topics; however, positivity is pronounced more in the

news postings on YouTube. That confirm with the previous work (Al-Rawi 2019) which found that YouTube news consumers prefer to read and share positive news posts. Moreover, the Pearson correlation of a post's Vader sentiment score and the average sentiment score of comments for each topic across Facebook, Instagram, Twitter, and YouTube shows a significant positive correlation for each topic.

Previously, researchers in Castillo et al. (2014) found that the user engagement of social media news content can help predict future visitation patterns of news articles. Furthermore, we examine the prediction power for the four levels of user engagement, which varies among organisations. The variation we found confirm with the previous work (Aldous, An, and Jansen 2019b, 2019c), where news organisations use stylistic features differently across social media platforms. We relate this variation to the differences in the origin countries of the 53 news organisations, which possibly affects the prediction models. Although the majority are US-based, a few others are based in other countries (e.g. UK). Another possible reason for the prediction variety is the global platform adoption by users; individuals from various nationalities express engagement differently across the platforms. The findings suggest there are other factors affecting engagement such as technological affordances, cultural, political, and social factors (Kalogeropoulos et al. 2017; Schlozman, Verba, and Brady 2010). For the external or public sharing engagement metric (Level-4), we use data from Reddit only, which is a US-based platform. Reddit users share content from news organisations based in the US, which makes it difficult for non-US organisations to make predictions using only information from Reddit. Thus, using publicly shared information from different platforms (e.g. YouTube and Twitter) is an important extension for Level-4 prediction tasks in the future, which are subject to data availability.

6.1. Theoretical contribution

This research adopts the theory of the eWOM concept, derived from the network effects of social media, from which we provide the theoretical framework of multiple levels of user engagement metrics. The levels of user engagement measure the extent to which a user publicly engages with social media content. Given the wide range and spectrum of social media metrics, a unified framework is of theoretical value to the field.

The novelty aspects of this research are multiple organisations and multiple platforms. This research is one of few cross-platform studies with many organisations within the same domain, which opens doors for other aspects of studies in the future. These include studies such as demographics and user analysis. The specific theoretical contributions of this work are the following:

- An engagement framework consists of four levels from private (or silent) (e.g. viewing) to public (or vocal) (e.g. public sharing) expressiveness of user engagement.
- User engagement to news content differs across the platforms, implying that there may be audience population differences at work. So, all platforms can not be treated the same, even in the same domain.
- User sentiment does not always correlate to content sentiment, implying that there are topics that engender specific sentiment in certain segments.
- There are content and user sentiment trends on certain platforms, which imply that audiences may tailor their social media engagement sentiment based on the platform.

6.2. Practical contributions

In terms of practicality, this research offers many contributions. By using multiple organisations (53) from one domain (news) over 8 months and analysing multiple platforms (5) simultaneously in a multiple-platform setting, we extract common topics distribution behaviour by multiple organisations over the platforms. Further to this, we analyse the interplay of engagement across platforms over topics and highlight the most engaging topics in each platform. Moreover, we analyse the effect of topics and sentiment on engagement by platform and topic. Finally, we provide the prediction models of four engagement levels across platforms. Using these models, content editors or creators can build advanced knowledge of the expected reach for a content before posting it online on social media platforms.

The specific practical contributions of this work are as follows:

- The engagement framework consisting of four levels provides some structure to the array of social media metrics that news organisations can use to specifically target the level of engagement they seek.
- Organisations can use our platform difference findings to target content more appropriately to the social media platform that is most receptive.
- Knowing that certain platform audiences are more receptive to certain content or sentiments, content-publishing organisations can more specifically target those platforms.

6.3. Future work

There are many future directions that can be derived from this research. For example, in this work, we only analyse the textual content, while video and images are other content types that need further analysis. Also, we consider the content features, while other context features, such as location-based features, need further studies. Moreover, there is a need to analyse how some content creator factors (e.g. age and number of followers) affect user engagement. Another future direction is to improve prediction results using deep neural network models. Also, for Level-4 prediction, since we use only Reddit information in this work, we would like to use public sharing information from different platforms (e.g.YouTube and Twitter), which is subject to data availability.

7. Conclusions

This research revealed a framework for user engagement publicity in four levels, which can guide future researchers in the cross-platform analysis. We analysed the topical distribution over five different social media platforms for a collection of 53 news outlets. We studied the topical and sentiment aspect of effecting user engagement across platforms. Our results show that user engagement is highly affected by the content's topic, with some topics being more engaging in a particular platform. This means that the volume of user engagement varies by both platform and topic. Also, the sentiment of the posts influences the volume of comments, and their sentiments depend on different levels, topics, and the platform. News outlets can benefit from findings to overcome some of the challenges they are facing, including finding the best platform to post about a topic and understanding the effect of topics and sentiment on engagement across platforms. Thus, this research contributes to the content creation process for making better management decisions.

Notes

1. http://www.journalism.org/2011/05/09/top-25/
2. https://wallethub.com/blog/best-news-sites/21699/
3. https://cappuccinoandfashion.com/lifespan-of-social-media-posts/
4. https://files.pushshift.io/reddit

Acknowledgments

Open Access funding provided by the Qatar National Library.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Kholoud Khalil Aldous http://orcid.org/0000-0003-1188-5724

References

Al-Rawi, Ahmed. 2019. "Viral News on Social Media." *Digital Journalism* 7 (1): 63–79.

Aldous, Kholoud Khalil, Jisun An, and Bernard J. Jansen. 2019a. "The Challenges of Creating Engaging Content: Results from a Focus Group Study of a Popular News Media Organization." In *CHI*.

Aldous, Kholoud Khalil, Jisun An, and Bernard J. Jansen. 2019b. "Predicting Audience Engagement Across Social Media Platforms in the News Domain." In *Social Informatics*, edited by Ingmar Weber, Kareem M. Darwish, Claudia Wagner, Emilio Zagheni, Laura Nelson, Samin Aref, and Fabian Flöck, 173–187. Springer International Publishing.

Aldous, Kholoud Khalil, Jisun An, and Bernard J. Jansen. 2019c. "Stylistic Features Usage: Similarities and Differences Using Multiple Social Networks." In *Social Informatics*, edited by Ingmar Weber, Kareem M. Darwish, Claudia Wagner, Emilio Zagheni, Laura Nelson, Samin Aref, and Fabian Flöck, 309–318. Cham: Springer International Publishing.

Aldous, Kholoud Khalil, Jisun An, and Bernard J. Jansen. 2019d. "View, Like, Comment, Post: Analyzing User Engagement by Topic at 4 Levels across 5 Social Media Platforms for 53 News Organizations." In *ICWSM*, Vol. 13. 47–57.

Altınel, Berna, and Murat Can Ganiz. 2018. "Semantic Text Classification: A Survey of Past and Recent Advances." *Information Processing & Management* 54 (6): 1129–1153.

An, Jisun, and Ingmar Weber. 2018. "Diversity in Online Advertising: A Case Study of 69 Brands on Social Media." In *SocInfo*.

An, Jisun, Daniele Quercia, Meeyoung Cha, Krishna Gummadi, and Jon Crowcroft. 2014. "Sharing Political News: the Balancing Act of Intimacy and Socialization in Selective Exposure." *EPJ Data Science* 3 (1): 12.

An, Jisun, Daniele Quercia, and Jon Crowcroft. 2014. "Partisan Sharing: Facebook Evidence and Societal Consequences." In *COSN*.

Arapakis, Ioannis, Mounia Lalmas, B. Barla Cambazoglu, Mari-Carmen Marcos, and Joemon M. Jose. 2014. "User Engagement in Online N Ews: Under the Scope of Sentiment, Interest, Affect, and Gaze." *Journal of the Association for Information Science and Technology* 65 (10): 1988–2005.

Attfield, Simon, Gabriella Kazai, Mounia Lalmas, and Benjamin Piwowarski. 2011. "Towards a Science of User Engagement (Position Paper)." In *WSDM Workshop on User Modelling for Web Applications*, 9–12.

Balbi, Simona, Michelangelo Misuraca, and Germana Scepi. 2018. "Combining Different Evaluation Systems on Social

Media for Measuring User Satisfaction." *Information Processing & Management* 54 (4): 674–685.

Bandari, Roja, Sitaram Asur, and Bernardo A. Huberman. 2012. "The Pulse of News in Social Media: Forecasting Popularity." In *ICWSM*.

Blei, David M., Andrew Y. Ng, and Michael I. Jordan. 2003. "Latent Dirichlet Allocation." *Journal of Machine Learning Research* 3: 993–1022.

Castillo, Carlos, Mohammed El-Haddad, Jürgen Pfeffer, and Matt Stempeck. 2014. "Characterizing the Life Cycle of Online News Stories Using Social Media Reactions." In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing* (Baltimore, Maryland, USA) *(CSCW '14)*, 211–223. New York, NY, USA: Association for Computing Machinery. doi:10.1145/2531602.2531623 .

Cataldi, Mario, Luigi Di Caro, and Claudio Schifanella. 2010. "Emerging Topic Detection on Twitter Based on Temporal and Social Terms Evaluation." In *Proceedings of the Tenth International Workshop on Multimedia Data Mining*, 4.

Ceron, Andrea, Luigi Curini, and Stefano Maria Iacus. 2016. "iSA: A Fast, Scalable and Accurate Algorithm for Sentiment Analysis of Social Media Content." *Information Sciences* 367: 105–124.

Chen, Zifei Fay, Cheng Hong, and Cong Li. 2017. "The Joint Effect of Association-based Corporate Posting Strategy and EWOM Comment Valence on Social Media." *Internet Research* 27 (5): 1039–1057.

Chung, Deborah S. 2008. "Interactive Features of Online Newspapers: Identifying Patterns and Predicting Use of Engaged Readers." *Journal of Computer-Mediated Communication* 13 (3): 658–679.

Darling, William M. 2011. "A Theoretical and Practical Implementation Tutorial on Topic Modeling and Gibbs Sampling." In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, 642–647.

Diao, Qiming, Jing Jiang, Feida Zhu, and Ee-Peng Lim. 2012. "Finding Bursty Topics from Microblogs." In *ACL*.

Fang, Anjie, Craig Macdonald, Iadh Ounis, and Philip Habel. 2016. "Examining the Coherence of the Top Ranked Tweet Topics." In *SIGIR*.

Ferrara, Emilio, Roberto Interdonato, and Andrea Tagarelli. 2014. "Online Popularity and Topical Interests through the Lens of Instagram." In *HT*.

Glenski, Maria, Tim Weninger, and Svitlana Volkova. 2018. "Identifying and Understanding User Reactions to Deceptive and Trusted Social News Sources." *CoRR*.

Guo, Weiyu, Shu Wu, Liang Wang, and Tieniu Tan. 2015. "Social-Relational Topic Model for Social Networks." In *CIKM*.

Harrigan, Paul, Uwana Evers, Morgan Miles, and Timothy Daly. 2017. "Customer Engagement with Tourism Social Media Brands." *Tourism Management* 59 (3): 597–609.

Hille, Sanne, and Piet Bakker. 2014. "Engaging the Social News User: Comments on News Sites and Facebook." *Journalism Practice* 8 (5): 563–572.

Hong, Liangjie, and Brian D. Davison. 2010. "Empirical Study of Topic Modeling in Twitter." In *Proceedings of the First Workshop on Social Media Analytics*, 80–88.

Hu, Gang, Jie Shao, Fumin Shen, Zi Huang, and Heng Tao Shen. 2017. "Unifying Multi-Source Social Media Data for Alized Travel Route Planning." In *SIGIR*, 893–896.

Huang, Minghui, Haoran Xie, Yanghui Rao, Jingrong Feng, and Fu Lee Wang. 2020. "Sentiment Strength Detection with a Context-dependent Lexicon-based Convolutional Neural Network." *Information Sciences* 520: 389–399.

Hutto, Clayton J., and Eric Gilbert. 2014. "VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text." In *ICWSM*.

Indira, D. N. V. S. L. S., R. Kiran Kumar, G. V. S. N. R. V. Prasad, and R. Usha Rani. 2019. "Detection and Classification of Trendy Topics for Recommendation Based on Twitter Data on Different Genre." In *Smart Intelligent Computing and Applications*. 143–153.

Jaakonmäki, Roope, Oliver Müller, and Jan vom Brocke. 2017. "The Impact of Content, Context, and Creator on User Engagement in Social Media Marketing." In *HICSS*.

Jansen, Bernard. 2009. "Understanding User-Web Interactions Via Web Analytics." *Synthesis Lectures on Information Concepts, Retrieval, and Services* 1 (1): 1–102.

Jansen, Bernard, Mimi Zhang, Kate Sobel, and Abdur Chowdury. 2009. "Twitter Power: Tweets As Electronic Word of Mouth." *Journal of the Association for Information Science and Technology* 60: 2169–2188.

Jelodar, Hamed, Yongli Wang, Chi Yuan, and Xia Feng. 2017. "Latent Dirichlet Allocation (LDA) and Topic Modeling: Models, Applications, a Survey." *Computing Research Repository* 78: 15169–15211.

Jiang, Jing, Minghui Qiu, and Feida Zhu. 2013. "It Is Not Just What We Say, But How We Say Them: LDA-based Behavior-Topic Model." In *SDM*.

Kaladevi, P., and K. Thyagarajah. 2019. "Integrated CNN- and LSTM-DNN-based Sentiment Analysis Over Big Social Data for Opinion Mining." *Behaviour & Information Technology* 40: 1–9. doi:10.1080/0144929X.2019.1699960

Kallas, Pritt. 2017. "Top 15 Most Popular Social Networking Sites." Accessed 4 January 2019. http://www.ebizmba.com/articles/social-networking-websites.

Kalogeropoulos, Antonis, Samuel Negredo, Ike Picone, and Rasmus Kleis Nielsen. 2017. "Who Shares and Comments on News?: A Cross-national Comparative Analysis of Online and Social Media Participation." *Social Media+ Society* 3 (4): 2056305117735754.

Kong, Li., Chuanyi Li, Jidong Ge, FeiFei Zhang, Yi Feng, Zhongjin Li, and Bin Luo. 2020. "Leveraging Multiple Features for Document Sentiment Classification." *Information Sciences* 518 (1): 39–55.

Kouloumpis, Efthymios, Theresa Wilson, and Johanna Moore. 2011. "Twitter Sentiment Analysis: The Good the Bad and the Omg!." In *ICWSM*.

Kumar, V, Lerzan Aksoy, Bas Donkers, Rajkumar Venkatesan, Thorsten Wiesel, and Sebastian Tillmanns. 2010. "Undervalued Or Overvalued Customers: Capturing Total Customer Engagement Value." *Journal of Service Research* 13 (3): 297–310.

Lalmas, Mounia, Heather O'Brien, and Elad Yom-Tov. 2014. "Measuring User Engagement." *Synthesis Lectures on Information Concepts, Retrieval, and Services* 6 (4): 1–132.

Larsson, Anders Olof. 2018. "Diversifying Likes: Relating Reactions to Commenting and Sharing on Newspaper Facebook Pages." *Journalism Practice* 12 (3): 326–343.

Lee, Kathy, Diana Palsetia, Ramanathan Narayanan, Md. Mostofa Ali Patwary, Ankit Agrawal, and Alok Choudhary. 2011. "Twitter Trending Topic Classification." In *IEEE 11th International Conference on Data Mining Workshops*, 251–258.

Lee, Roy Ka-Wei, Tuan-Anh Hoang, and Ee-Peng Lim. 2017. "On Analyzing User Topic-Specific Platform Preferences Across Multiple Social Media Sites." In *WWW*.

Lin, Jianhua. 1991. "Divergence Measures Based on the Shannon Entropy." *IEEE Transactions on Information Theory* 37 (1): 145–151.

Liu, Bang, Di Niu, Kunfeng Lai, Linglong Kong, and Yu Xu. 2017. "Growing Story Forest Online from Massive Breaking News." In *CIKM*.

Liu, Jinwei, Wingyan Chung, Yifan Huang, and Cagri Toraman. 2019. "CrossSimON: A Novel Probabilistic Approach to Cross-Platform Online Social Network Simulation." In *ISI*.

Ma, Changsha, Zhisheng Yan, and Chang Wen Chen. 2017. "LARM: A Lifetime Aware Regression Model for Predicting YouTube Video Popularity." In *CIKM*.

Matsa, Katerina, and Elisa Shearer. 2018. "News Use Across Social Media Platforms 2017." *Pew Research Center*. https://www.pewresearch.org/journalism/2017/09/07/news-use-across-social-media-platforms-2017/

McCay-Peet, Lori, Mounia Lalmas, and Vidhya Navalpakkam. 2012. "On Saliency, Affect and Focused Attention." In *SIGCHI*. 541–550.

Muñoz-Expósito, Miriam, M. Ángeles Oviedo-García, and Mario Castellanos-Verdugo. 2017. "How to Measure Engagement in Twitter: Advancing a Metric." *Internet Research* 27 (5): 1122–1148.

Nimala, K., and R. Jebakumar. 2019. "Sentiment Topic Emotion Model on Students Feedback for Educational Benefits and Practices." *Behaviour & Information Technology* 40: 311–319. doi:10.1080/0144929X.2019.1687756

Noguti, Valeria. 2016. "Post Language and User Engagement in Online Content Communities." *European Journal of Marketing* 50 (5/6): 695–723.

O'Brien, Heather L. 2017. "Antecedents and Learning Outcomes of Online News Engagement." *Journal of the Association for Information Science and Technology* 68 (12): 2809–2820.

Oh, Jeeyun, Saraswathi Bellur, and S. Shyam Sundar. 2018. "Clicking, Assessing, Immersing, and Sharing: An Empirical Model of User Engagement with Interactive Media." *Communication Research* 45 (5): 737–763.

Pak, Alexander, and Patrick Paroubek. 2010. "Twitter as a Corpus for Sentiment Analysis and Opinion Mining." In *LREc*, Vol. 10., 1320–1326.

Peters, Kay, Yubo Chen, Andreas M. Kaplan, Björn Ognibeni, and Koen Pauwels. 2013. "Social Media Metrics–A Framework and Guidelines for Managing Social Media." *Journal of Interactive Marketing* 27 (4): 281–298.

Pokharel, Rashmi, Pari Delir Haghighi, Prem Prakash Jayaraman, and Dimitrios Georgakopoulos. 2019. "Analysing Emerging Topics Across Multiple Social Media Platforms." In *Proceedings of the Australasian Computer Science Week Multiconference (ACSW 2019)*, 16:1–16:9.

Qiu, Jiangtao, Chuanhui Liu, Yinghong Li, and Zhangxi Lin. 2018. "Leveraging Sentiment Analysis At the Aspects Level to Predict Ratings of Reviews." *Information Sciences* 451: 295–309.

Rajapaksha, P., R. Farahbakhsh, and N. Crespi. 2019. "Scrutinizing News Media Cooperation in Facebook and Twitter." *IEEE Access* 4: 1–14.

Rieis, Julio Cesar, Fabrício Benevenuto, Pedro Raquel Oliveira Prates, Haewoon Kwak, and Jisun An. 2015. "Breaking the News: First Impressions Matter on Online News." In *ICWSM*.

Rowe, Matthew, and Harith Alani. 2014. "Mining and Comparing Engagement Dynamics Across Multiple Social Media Platforms." In *Proceedings of the 2014 ACM conference on Web science*, 229–238.

Saha, Ankan, and Vikas Sindhwani. 2012. "Learning Evolving and Emerging Topics in Social Media: A Dynamic NMF Approach With Temporal Regularization." In *WSDM*, 693–702.

Schlozman, Kay Lehman, Sidney Verba, and Henry E. Brady. 2010. "Weapon of the Strong? Participatory Inequality and the Internet." *Perspectives on Politics* 8 (2): 487–509.

Shi, Juan, Ping Hu, Kin Keung Lai, and Gang Chen. 2018. "Determinants of Users' Information Dissemination Behavior on Social Networking Sites: An Elaboration Likelihood Model Perspective." *Internet Research* 28 (2): 393–418.

Srinivasan, Balaji Vasan, Anandhavelu Natarajan, Ritwik Sinha, Vineet Gupta, Shriram Revankar, and Balaraman Ravindran. 2013. "Will Your Facebook Post Be Engaging?." In *UEO*.

Stocking, Galen, and Maya Khuzam. 2019. "State of the News Media".

Stoddard, Greg. 2015. "Popularity and Quality in Social News Aggregators: A Study of Reddit and Hacker News." In *WWW*.

Svetlov, Kirill, and Konstantin Platonov. 2019. "Sentiment Analysis of Posts and Comments in the Accounts of Russian Politicians on the Social Network." In *2019 25th Conference of Open Innovations Association (FRUCT)*, 299–305. IEEE.

Thonet, Thibaut, Guillaume Cabanac, Mohand Boughanem, and Karen Pinel-Sauvagnat. 2017. "Users Are Known by the Company They Keep: Topic Models for Viewpoint Discovery in Social Networks." In *CIKM*.

Ting, I-Hsien, Hui-Ju Wu, and Pei-Shan Chang. 2009. "Analyzing Multi-Source Social Data for Extracting and Mining Social Networks." In *2009 International Conference on Computational Science and Engineering*, Vol. 4., 815–820.

Vallet, David, Shlomo Berkovsky, Sebastien Ardon, Anirban Mahanti, and Mohamed Ali Kafaar. 2015. "Characterizing and Predicting Viral-and-Popular Video Content." In *CIKM*.

Van Canneyt, Steven, Philip Leroux, Bart Dhoedt, and Thomas Demeester. 2018. "Modeling and Predicting the Popularity of Online News Based on Temporal and Content-related Features." *Multimedia Tools and Applications* 77 (1): 1409–1436.

Del Vicario, Michela, Sabrina Gaito, Walter Quattrociocchi, Matteo Zignani, and Fabiana Zollo. 2017. "Public Discourse and News Consumption on Online Social

Media: A Quantitative, Cross-Platform Analysis of the Italian Referendum." *CoRR*.

Voorveld, Hilde A. M., Guda van Noort, Daniël G. Muntinga, and Fred Bronner. 2018. "Engagement with Social Media and Social Media Advertising: The Differentiating Role of Platform Type." *Journal of Advertising* 47 (1): 38–54.

Wang, Wen-Chia. 2017. "Understanding User Experience of News Applications by Taxonomy of Experience (ToE)." *Behaviour & Information Technology* 36 (11): 1137–1147. doi:10.1080/0144929X.2017.1359337

Yang, Min-Chul, and Hae-Chang Rim. 2014. "Identifying Interesting Twitter Contents Using Topical Analysis." *Expert Systems with Applications* 41 (9): 4330–4336.

Yang, Yinfei, Cen Chen, and Forrest Sheng Bao. 2016. "Aspect-Based Helpfulness Prediction for Online Product Reviews." In *ICTAI*.

Zarrinkalam, Fattane, Mohsen Kahani, and Ebrahim Bagheri. 2018. "Mining User Interests Over Active Topics on Social Networks." *Information Processing & Management* 54 (2): 339–357.

Zhang, Yongzheng, and Marco Pennacchiotti. 2013. "Predicting Purchase Behaviors from Social Media." In *WWW*.

Zhao, Wayne Xin, Jing Jiang, Jianshu Weng, Jing He, Ee-Peng Lim, Hongfei Yan, and Xiaoming Li. 2011. "Comparing Twitter and Traditional Media Using Topic Models." In *ECIR*.