

Political Tweets and Mainstream News Impact in India: A Mixed Methods Investigation into Political Outreach

Sunandan Chakraborty
Indiana University
Indianapolis, Indiana
sunchak@iu.edu

Priyank Chandra
University of Michigan Ann Arbor
Ann Arbor, Michigan
prc@umich.edu

Joyojeet Pal
University of Michigan Ann Arbor
Ann Arbor, Michigan
joyojeet@umich.edu

Daniel M. Romero
University of Michigan Ann Arbor
Ann Arbor, Michigan
drom@umich.edu

ABSTRACT

Citizens' perception of politicians and political issues is increasingly influenced by social media. However, little is known about the potential of second order effects of social media in parts of the world where the majority of voting citizens are not online. In this paper, we examine whether a politician can move to communicating through social media as their primary means of outreach, and still present their message to the mainstream population through traditional media. By studying of the use of Twitter by Indian Prime Minister Narendra Modi between 2009 and 2015, the second-most followed elected official in the world, we present evidence of the impact of social media on print news. We use computational text mining techniques to automatically identify print news reports that use Modi's tweets as a source, alongside manual qualitative coding of tweets to analyze the role of tweet themes in print news coverage. We conclude that while Modi's social media messaging does get coverage in the print news, it is his more "newsworthy" tweets, such as references to celebrities, other politicians, or major events such as holidays that have a greater likelihood of coverage.

ACM Reference Format:

Sunandan Chakraborty, Joyojeet Pal, Priyank Chandra, and Daniel M. Romero. 1997. Political Tweets and Mainstream News Impact in India: A Mixed Methods Investigation into Political Outreach. In *Proceedings of ACM COMPASS (COMPASS'18)*, Jennifer B. Sartor, Theo D'Hondt, and Wolfgang De Meuter (Eds.). ACM, New York, NY, USA, Article 4, 11 pages. https://doi.org/10.475/123_4

1 INTRODUCTION

In the last few years, there has been much discussion on political actors who turn away from the mainstream media and convert social media into their main form of output. Indian Prime Minister Narendra Modi and US President Donald Trump have been among the high-profile elected leaders whose direct outreach to citizens has been an important campaign strategy. These leaders did not

divest from mainstream media, but limited or eliminated direct contact with political journalists, instead using social media as an outreach vehicle to be consumed by the press and people alike.

In this study, we examine the use of Twitter in the 2014 Indian general elections by Minister Narendra Modi, who challenged the incumbent Indian National Congress (INC) party and ran a populist campaign on the issues of governance, anti-corruption, and economic development. Modi's party, the Bharatiya Janata Party (BJP) places at the right of political spectrum, and is part of the Hindu chauvinist organization called the Rashtriya Swayamsewak Sangh (RSS), which has traditionally been guided by Hindutva, a philosophy of polity and social life based on ideals from Hinduism.

The 2014 campaign is already perhaps the most widely researched in India's postcolonial history because of the landslide results giving India its first non-coalition government in over three decades that saw the consolidation of certain political alignments [7] [6] and reframing of others [52]. It was the first time in India's history that a party on the Hindu right has been able to form a government on absolute majority. The election has also been an object of much media research because of the outreach strategies used in the election including a heavy reliance on personality politics [15], charismatic imagery [56], and marketing-style political brand management [16].

Modi's use of social media, specifically Twitter, has been widely documented through a series of studies focusing on the network effects [17], the role of language [41], imagery [4]. Two motivating factors drove Mr. Modi's investment into social media. First, he had an image as a Hindu hardliner in the political spectrum, which has traditionally been difficult in coalition politics because they tend to favor centrist/compromise candidates. Second, he had a problematic relationship with large sections of the mainstream media, following his alleged role in Hindu-Muslim riots in his state, which led to him being banned from entering the United States and subject to embargoes on interaction with EU officials [52]. This drove the need for an independent channel of communication that could help reshape public opinion [42].

Much qualitative research in recent years has focused on Modi's avoidance of mainstream news commentators and investment into social media channels to speak directly through a controlled medium to present an alternative narrative of his political positioning [42]. However, little work has attempted to get at the question of whether second-order effects of social media can enable a politician to

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
COMPASS'18, June 2018, San Francisco, CA USA
© 2018 Copyright held by the owner/author(s).
ACM ISBN 123-4567-24-567/08/06.
https://doi.org/10.475/123_4

reach citizens who are not online. To this end, we present a mixed-methods study of the second-order effects of social media, by using computational methods to mine the text of three mainstream English-language media sources - The Times of India[54], The Hindu[53], and NDTV[40] for tweets from Narendra Modi. We use qualitative thematic coding methods to analyze the relationship between tweet content and inclusion in the media.

Our text mining tool is capable of retrieving mentions and coverage of tweets from news articles in various forms, from direct quotes to cursory mentions. Comparison between short, unstructured texts of the tweets, with no inherent grammar and larger, more formal texts of news articles is a challenging task. This tool has been developed with careful deliberation and deep understanding of these underlying challenges. The final version of this tool is capable of extracting articles, which quoted or used tweets, with a precision of 0.894 and a recall of 0.813.

This tool was used to retrieve all media articles that have reportedly used Modi's tweets between February 2009 and October 2015. These *tweet to article* pairs were used to understand typical themes and features of the tweets that are more likely to be covered by the mainstream media. Specifically, this paper addresses the following:

- Detecting salient features of Modi's tweets
- Thematic conditions that influence mainstream media to use Modi's tweets in their reports
- Effects of media coverage on the popularity of the tweets

Our goal is not to generalize across politicians, but rather to present a case study of a politician whose campaign provides important insight into an increasingly global movement of undermining traditional mainstream media by politicians speaking mainly through social media [43]. We propose that this case has important implications for the future of democratic discourse and informed citizenry, particularly in parts of the world susceptible to media domination by politicians. While ICTD studies in the past have traditionally taken instrumentalist views of technologies built for developmental outcomes, our study highlights the need for cross-discipline collaboration on studying media environments *situated within* development contexts.

2 BACKGROUND

The official Twitter account of Narendra Modi, @narendramodi, was started in February 2009. Narendra Modi started interacting with citizens relatively early, with his own website starting in 2007, well before many of his political colleagues or rivals had dedicated online sites. In 2009, Modi started accounts on both Twitter and Facebook, although Twitter would eventually become the main location where he posted messages, which were then mirrored to other online social media sites. Modi tweets almost exclusively in English. While it is unclear who composes the messages, they are all written in the first person. The Twitter account is also the main point of outreach, tweets are frequently mirrored as status updates in other social media accounts, including Facebook.

Until the end of 2010, the Twitter account was used somewhat sporadically. Starting in 2011, the account has been used consistently, with some messaging almost daily. As we can see in figure 1, initially his messages had relatively weak and sporadic reach in terms of retweets, but in 2012 the account @narendramodi started

to have a consistent growth in following, as well as having messages retweeted. This finding is consistent with reports of Modi hiring brand managers and social media experts to work on his social media team, charged with making his social media presence the primary means of output [46]. After his election, Modi's Twitter page had a dramatic rise in following, consequently the retweet rates of his messages have also risen. Modi currently has more followers than the next several media outlets in India combined.

In this work, we look at the second order effects of the tweets reaching mainstream print and news television audiences through the three largest English-language sources – the two top national newspapers, The Times of India [54], and The Hindu [53] and the largest English-language television news channel, NDTV[40], and their online article archive.

3 RELATED WORK

There have been multiple computational examinations of social media as a tool for news consumption and dissemination [19]. Lerman et al. showed that the network structure plays a crucial role in the spread of news on social media [21]. Others have developed systems that predict and recommend news that users are likely to be interested in [1, 44]. Researchers have also developed techniques for early detection of trending topics and breaking news that can then be disseminated on the network [23, 45]. Most of these studies use topic modeling or other text mining techniques to cluster content into topics automatically [47], or use the hashtags as a way of classifying content into topics [49]. Since this paper focuses on tweets that are related to political outreach and elections that need contextual knowledge to appropriately code, we employ expert classification of tweets rather than automatic techniques.

Social media and Twitter, in particular, has been shown to be used by politicians to engage with journalists and the public in various countries. Australian politicians were found to have more active on Twitter than non-politicians and that they cluster by party [12], Dutch politicians and journalists were found to be highly co-dependent [55], Norwegian politicians differentiated by medium based on whether they were self-marketing (Facebook) or engaging in dialog (Twitter) [10], while Swiss politicians online tended to be younger, and more dependent on journalists than the other way around [48].

Work on the interplay between social and traditional news media have traditionally been driven by communications studies and highlighted the importance of strengthening rather than weakening a professional journalist corps. Research has shown that journalists using microblogging sites express their opinions more freely than in traditional media outlets [20] and that social media tends to disseminate entity-oriented topics that have low coverage in traditional news outlets [57]. Little research however tackles the question of how politicians may benefit from the structural changes in mainstream print news, and how circumventing journalist contact can be turned into a feasible strategy for politicians with antagonistic relationships with sections of the media [24]. Modi has largely managed communications through advertising agencies and brand managers [14], leading to a "rebranding" in the mainstream media through a careful use of the affordances of social media such as

Table 1: Top themes with description and example

| Theme | Classification rule | Example Tweet |
|-------------------|---|---|
| political callout | A tweet that calls out or mentions the name of a major political figure, domestic or foreign. | Met Tamil Nadu Chief Minister Jayalalithaa ji.âĀĀ |
| general callout | A reference to an individual citizen or group of people (who is not a public figure) | Congratulating a beneficiary who has been allotted a house in the draw todayâĀĀ |
| foreign affairs | Tweets about meetings with foreign leaders or emissaries, summit meetings, state visits etc., or references to issues of international concern. | I thanked PM Sheikh Hasina for her good wishes and emphasised on the strong historical and cultural ties between India & Bangladesh. |
| elections | Tweets about domestic elections âĀĀ both general and state-level | I congratulate Assam BJP for the commendable performance of the Party in the Municipal Polls held across the State |
| greetings | Tweets featuring a greeting to a specific figure or population | Har Har Mahadev! Greetings on the occasion of Maha Shivratri. May Lord Shiva bless our people with strength & wisdom |
| gratitude | Tweets expressing appreciation to an individual or group | PutinRF_Eng I would like to thank you for your good wishes on our Republic Day. |
| science and tech | Mentions of science and technology, including scientific innovations, computers, mobile devices, and science programs, or figures associated with these | From making electronic goods that are globally competitive to enhancing cyber security, India must show the way. |
| function | Mentions of public functions or commemorative events in India, or abroad, excluding religious holidays | My dear engineer friends, greetings to you on EngineersDay I salute your hardwork and innovation, which has benefitted India immensely. |
| development | References to economic or social development, or related initiatives or policies | Guj is among India's fastest growing states and has a unique growth model, says a CLSA report "Gujarat:Growth to Sustain |
| international | Tweet on any international issue | I thank President of Nepal Shri Ram Baran Yadav, Prime Minister Shri Sushil Koirala & former PM Shri Baburam Bhattarai for their wishes |
| tribute | Paying homage to an individual such as a historical or contemporary figure, or a group (i.e. Teachers, on Teachers Day). | Tributes to Ram Prasad Bismil, Ashfaqulla Khan & Roshan Singh, who were martyred on this day. Every Indian remembers their brave sacrifice |
| BJP | Tweets with a mention of or a direct reference to the BJP (Bharatiya Janata Party) or a figure associated with the party | Best wishes to Madhya Pradesh CM Shri Shivraj Singh ji Chouhan on his birthday. May God bless him with a long & healthy life |
| celebrity callout | A tweet in which a popular celebrity is referred to | Met Amitabh Bachchan ji. Glad to know he is very happy with the love people of Gujarat have given him bit.ly/y7sra0 @SrBachchan |
| youth | References to students or youths/young adults | Dear @PoweredByATP Very soon u will find Chief Minister's internship programme for students & sabbatical prog for working professional |
| criticism | Attacks and criticism against political opponents | Highlighted how Cong kept playing with emotions of our armed Defense on 'One Rank, One Pension' issue for decades without giving a solution |
| Vivekananda | Tweets about historical Hindu and nationalist figure Swami Vivekananda or commemoration events or initiatives in his honor. | A few heart-whole, sincere, and energetic men and women can do more in a year than a mob in a century nm4.in/ybFnfJ vivekananda150 |
| Initiative | Any reference to a state or national initiative launched by the government | Launched new schemes for Skill Development including i-KVK, Kaushalya Rath, e-Skilling in the presence of youngsters from all over Gujarat. |
| Congress | Anything that refers to the Congress Party, whether it be the party itself, its policies, its leaders | Anything that refers to the Congress Party, whether it be the party itself, its policies, its leaders |
| Hinduism | References to the ideas, institutions, artifacts, religious figures, or other relevant aspects of the Hindu religion | I consider myself blessed to have closely interacted with Pujya Pramukh Swami Maharaj for years. He has deeply inspired me |
| Birthday | Tweets referring to birthdays | Wishing President Nursultan Nazarbayev of Kazakhstan on his 75th birthday. I pray for his long life & good health |

selfies and hashtags [4, 42]. He rarely offers unstaged interviews with the press [46], and indeed in his entire 2014 campaign, did not do a single interview with print media, journalists had to rely on Twitter to get sound-bites from Modi [52]. This in itself offers an important case for the future of electoral campaigns and democratic deliberation in the Global South, since the traditions of head of government press briefings or other forms of direct interaction with a professional media corps may be less common. Indeed, as we see in various parts of the world, leaders and representatives who have minimal direct interaction with professional journalists use social media as a means of communicating with citizenry [43]. Specifically in India, as recent research has suggested, media practices surrounding the coverage of Modi's campaign have created a heteronomic effect on journalistic practices, as political hegemony has strengthened their grip on mainstream media discourse [22].

4 METHODOLOGY

We use a case study method [9] to do a deep dive into two aspects of Narendra Modi's social media output - his tweets, and their coverage in the news. Similar approaches have studied various aspects of social media presence of major political leaders such as Barack Obama [2, 5], Hugo Chavez [13], David Cameron, [3], and although such studies do not lend themselves easily to replication in other contexts, they allow a nuanced understanding of the construction of various political movements online.

We use three methods to analyze the relationship between the tweets and their coverage in the news.

First, we coded 9040 tweets from @narendramodi to add meta information to the content of individual messages. Second, we build a tool that mines the text of the three news sources to identify articles that have covered Modi's tweets. Finally, we conducted a discursive reading of specific tweets to add qualitative depth in explaining our analysis.

4.1 Qualitative Coding

To capture granular and nuanced information on the messaging, we used manual coding to analyze tweets. We mined 9040 tweets from the account @narendramodi starting with the first tweet sent out on Feb. 2, 2009, until Oct. 2, 2015. Each tweet was hand-coded with up to four themes by a team of experts, each familiar with contemporary Indian politics. Every tweet was coded by at least two primary coders and one arbitrator, such as in the example below:

Tweet text: "Anna says the battle is only half won. Efforts to bring back black money, which has held the nation back, should be accelerated." [27]

Themes: Economics, Corruption, Political Callout

Date: 27/08/2011

The coding scheme was freely generated at the first stage from a sampled set of 100 tweets each separately open-coded by three separate coders. The first set of themes was used to populate the initial codebook. Each theme (code) was a subjective descriptor, and we had a mix of generic themes such as youth, elections, foreign affairs, as well as India-specific themes such as Congress, BJP, Vivekananda. These themes were created through discussions among all the team

members and added to a list during weekly team meetings early in the coding process.

In order to establish the quality and reliability of their coding, we initially asked the two primary coders to label 994 out of the 9040 tweets. Due to the nature of exploratory multi-label coding wherein a tweet can have up to 4 codes of equal weight chosen out of both existing and new themes, we instead measured inter-coder reliability using percentage agreements measures.

We chose to not use Cohen's kappa (κ) [8] for two reasons - 1) the regular kappa measure does not account for exploratory multilabel coding, and 2) if we consider the multilabel coding as vectors and take a weighted kappa that accounts for partial agreements between vectors, then the imbalanced (and symmetrical) marginals due to most existing themes coded as '0' leads to significantly low kappa values [11].

The average percentage agreements for all 994 tweets was 0.58, with 84.5% of the tweets having at least 1 theme in common and 43.9% tweets having at least 2 themes in common. The researchers thought this was high enough for this nature of qualitative work to justify the coding scheme which was subsequently used to code the rest of the tweets. The coding itself was an iterative process where the primary coders independently coded the tweets and met with an arbitrator at regular intervals to discuss and reconcile codes with low intercoder agreement. New themes were added to the codebook when new topics were covered in the tweets. Some themes were merged in the codebook if needed, and in such cases we recoded the lapsed themes. By the end of the coding process, we had 129 themes, with the entire process taking about 1200 worker-hours.

4.2 Tweet-Article Similarity

In this section, we present the methods used to search and retrieve articles that have mentioned Narendra Modi's tweets. While we focus on the media mention of Modi's tweets the methods are general and can be used to find media mention of any other tweets.

There are numerous ways an article can use or mention a tweet.

- **Direct quote:** The article has mentioned the entire content of the tweet.
- **Extract:** The article has only mentioned a smaller part of the tweet instead of the entire text.
- **Abstract:** The article has mentioned the content of the tweet without quoting any part of the tweet. The entire or part of the content is reproduced by changing the words and the phrases.
- **Cursory mention:** In this case the article has mentioned about a particular tweet without giving much detail about the content.

This task can be considered an Information Retrieval (IR) problem, where each of Modi's tweet can be used as a query to search the news corpus and find the top-ranked articles. In other words, if \mathcal{D} represents a news corpus and \mathcal{M} represents a set of tweets, then for a given tweet $m_i \in \mathcal{M}$, the task is to find a set of articles $\mathcal{D}^{m_i} \subset \mathcal{D}$, where \mathcal{D}^{m_i} represents the news articles that have quoted or mentioned the tweet m_i . Here, whether an article $d_j \in \mathcal{D}$ is in \mathcal{D}^{m_i} depends on a distance metric ($dist(m_i, d_j)$) that computes the similarity between the query (i.e. the tweet in our case) and the target documents (the news corpus). In a standard IR task the

resulting set is always a list of documents ranked inversely based on the distance metric.

There are many tweets that are not reported or mentioned by the media. Hence, unlike a standard IR task – where, the resulting documents (\mathcal{D}^{m_i}) is a set of ranked of documents and usually it is not an empty set – in this case \mathcal{D}^{m_i} can be an empty set. For a tweet, if there are no media articles about it, then a non-empty result set containing irrelevant articles can introduce inaccuracies in the subsequent analyses. For a tweet m_i and a news article d_j , we define a function $\mathcal{F}(m_i, d_j)$, such that if m_i is mentioned or quoted in d_j then $\mathcal{F}(m_i, d_j) = 1$, otherwise $\mathcal{F}(m_i, d_j) = 0$. The value of \mathcal{F} is determined based on the distance $dist(m_i, d_j)$. As \mathcal{F} is designed to have value 0 we define \mathcal{F} as follows,

$$\mathcal{F}(x, y) = \begin{cases} 1 & \text{if } dist(x, y) \leq \theta \\ 0 & \text{otherwise} \end{cases}$$

where, θ is threshold and any article whose distance is greater than θ will be discarded.

Measuring the distance between documents has been extensively studied in field of Information Retrieval and Extraction. However, our task requires computing the distance between a very short sample of text (i.e. the tweets) with a comparatively large text (i.e. the news articles). Measuring the distance of these vastly different types of text differs from the standard IR task [25]. Our approach is based on a few assumptions:

Assumption 1: Any article that does not have the word “Narendra”, “Modi”, “Prime Minister”, “Chief Minister”¹ and an inflected form of the words “Twitter” and “Tweet” will have an arbitrarily large distance. We identified an exhaustive list of possible inflections of these words after carefully examining the news corpus. They are – *tweet, tweeted, tweetering, tweeting, tweets, twits, twitted, twittering, twitter, twitting*.

Assumption 2: The mention or coverage of a tweet in an article will be in the proximity of the above words. More precisely, coverage of a tweet in an article is done in the same paragraph where either “Narendra”, “Modi” is mentioned or some form of the words “tweet” or “Twitter” is mentioned. This assumption significantly reduces the search space and makes the comparison simpler. This assumption ensures that the compared strings have similar lengths.

However, an article can be about some other person whose first name is “Narendra” or last name is “Modi” and even if the politician Narendra Modi is mentioned, the words “tweet” or “twitter” can be related to other person’s tweet. The distance metric needs to consider these corner cases to identify whether an article is covering Modi’s tweet and which particular tweet of his. We examined a range of similarity measures to compute the distance between the texts. We also experimented with different vector representations of text to evaluate the similarity measures.

4.2.1 Similarity Measures. We experimented with two different word weight representations – Term Frequency-Inverse Document

Frequency (TF-IDF) and Word2Vec with two different similarity measures – Cosine Similarity and Word Mover’s Distance.

TFIDF weight is a popular method to represent and compare text [51]. This metric is composed of two different components – Term Frequency, which signifies the popularity of a particular word in the given context and Inverse Document Frequency that signifies the rarity of the term, where a rare term carries more information. TFIDF weight vector is often used to compute *cosine similarity* between two pieces of text. This similarity metric computes the distance as the angle between the vectors representing the two documents.

Recent advances in Neural language modeling techniques saw the emergence of Word2Vec [26]. Word2Vec maps a word into a high-dimensional vector space, such that words appearing in similar contexts are placed nearby within the space. Any similarity measure, such as *cosine similarity* that can operate on vectors can be used to find the distance between two samples of text. The advantage of Word2Vec over TFIDF is that it measures meaningful similarity in the presence of polysemous and synonymous words.

Word Mover’s Distance (WMD) [18] is a text similarity metric recently proposed by Kusner *et al.* This metric, unlike others, actually considers each word in a document separately and the net distance is measured by the aggregated distance between the individual word pairs. WMD can be computed on top of any vector representation of words, including Word2Vec. The choice of the final word representation scheme and the distance metric was decided based on the experimental results performed on a training set.

The methods used were evaluated using a portion of the larger datasets. We have two different sets of data. The twitter dataset contains all the tweets posted by Narendra Modi, starting from his first tweet on Feb 2, 2009, until Oct 2, 2015. To complement this dataset, we collected all the news articles posted online by the media during the same time period. The news corpus from February 2009 till October 2015 has around 1,500,000 articles with a vocabulary size of 700,000. To evaluate our methods, we used the first part of the datasets [2009-2012] to build the training and cross-validation sets. For all the tweets in this time period, matching articles were found through manual inspection. This task was performed by one human annotator and verified and corrected by a second annotator. The labeled dataset was represented as pairs of tweet and sentences within news articles. There were two types of pairs – **positive pair**, where the tweet was found mentioned in the article sentence and **negative pair**, where the tweet was combined with a sentence from a news article which had no connection with the tweet. A randomly selected 80% of this labeled set was used as the training set and the rest 20% was kept for evaluation. The objective of the training was to understand the typical distance between the positive pairs and negative pairs (based on the three methods described above), which is required to determine the threshold θ . Assuming that both the positive distances and the negative distances are normally distributed, we took the threshold θ as the average of the data point value at the 95th percentile of the negative distance distribution and point at the 5th percentile of the positive distribution. If $qnorm(p, \mu, \sigma)$ represents the quantile function of a normal distribution with mean μ and standard deviation σ then,

¹Narendra Modi became the Prime Minister of India in May, 2014. Before that he was the Chief Minister of the Indian state of Gujarat

$$\theta = \frac{1}{2} \sum_{k \in \{pos, neg\}} qnorm(p_k, \mu_k, \sigma_k)$$

where $\mu_{pos}, \mu_{neg}, \sigma_{pos}, \sigma_{neg}$ are the means and the standard deviations of the positive and negative pair distance distributions respectively. p_k is the the percentile for each category, such as $p_{pos} = 0.05$ and $p_{neg} = 0.95$. The training objective for each of the methods used is to determine the value of this threshold to understand the distribution of the distance metric when a pair is *positive* or *negative*. Based on the learned value of θ , we evaluated the method on the evaluation set.

4.2.2 Evaluation. Selection of the best method from the ones described in the previous section was based on their individual performance on a gold standard test set. The purpose of this test set is to have a manually verified ground truth data. This test set consists of numerous *true* tweet-article pairs, where the article has either a direct quote or a cursory mention of the corresponding tweet. The test set was constructed using all of Modi’s tweets between 2013 and 2014, and the articles published in the same time period. The manual identification of matching news articles was done by two different human coders. The final matching pairs that were retained in the test set were based on the unanimous agreement of the two coders. The manual inspection of the articles was done on a reduced set of articles based on the Assumptions 1&2 stated in the previous section. This significantly reduced the search space to a set of 6,652 articles to find articles that quoted or mentioned Modi’s tweets.

In the experiments involving the different distance metrics, each metric was used to identify the articles corresponding to every tweet within the test time period (2013-2014). The articles retrieved in this process were matched with the test set articles, and the corresponding precision-recall values were used to evaluate the method. The performance of the three experimented distance metrics is summarized in Table 2.

Table 2: Performance of the methods

| Word weight | Distance metric | Precision | Recall | F-value |
|-----------------|------------------------------|--------------|--------------|--------------|
| tfidf | cosine similarity | 0.729 | 0.476 | 0.576 |
| word2vec | cosine similarity | 0.802 | 0.655 | 0.721 |
| word2vec | word mover’s distance | 0.894 | 0.813 | 0.851 |

5 RESULTS

5.1 Modi’s Tweets in Media

Based on the results shown in Table 2, we chose to use *WMD* distance metric with *word2vec* weights of the words for rest of the analysis presented in this paper. To preserve the effect of newsworthiness of the tweets, we considered a tweet mentioned in an article to be a true match only if the article was published within

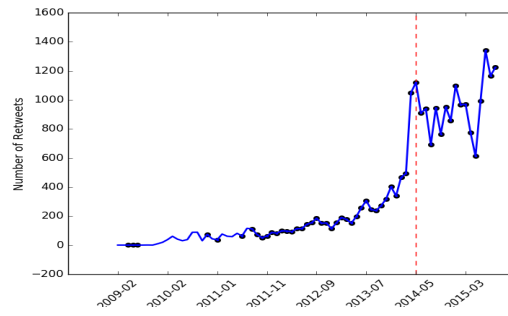


Figure 1: Monthly Median of Retweet Rate of Modi’s Tweet. The dots in the line depict the months where at least one tweet was mentioned by media . The vertical dashed line denotes the 2014 Parliamentary Election time.

seven days of the tweet’s post date. This is to ensure that the articles have used the tweets not to report in retrospect but about events or incidents that were contemporary. With this restriction, out of 9,040 tweets we had in our Twitter dataset, **204** were covered by *The Times of India*, **365** by *The Hindu*, and **732** by *NDTV*. Among all the tweets in our dataset, **20** tweets were found to be mentioned by all the three sources, **322** tweets were mentioned by exactly two sources, **597** tweets were found in exactly one source and the remaining **8,101** tweets had no mention in any of the three sources. There were **939** unique tweets that were mentioned by at least one media source.

The first tweet that was mentioned by any media was on Apr 27, 2009, by *NDTV*, two and half months after his first tweet. Media reports that used Modi’s tweets gained in maximum momentum during his campaign for the 2014 Parliamentary elections and his victory thereafter. Figure 1 shows a time-series of the frequency of Modi’s tweets that were covered by the media . Two trends are seen in the visualization. First that there is a gradual increase in the median monthly retweet rate from 2011 to 2015, with a leap in retweets around the time of his election in May 2014. We also see the concentration of dots on the line increases after the election, at which point Modi’s tweets are relatively more covered by the media. So while a number of his tweets were actively retweeted, only 60 tweets from @narendramodi made it to the news between 2009-2012, another 279 tweets were covered till the day of the election results, and the remaining 600 were covered between May 2014 and December 2015. In other words, our data shows that Modi gradually "became news" as first the mainstream press started to cover his tweets as a front-running campaign, but more importantly, thereafter, a selection of his tweets in and of themselves becoming newsworthy.

Thus, from a mediatization perspective, the data shows that Modi was successful in both getting more throughput in terms of retweets from the general Twitter public, as well as coverage from the mainstream press.

Among the tweets covered by the media , the predominant themes (Table 1) were – *Political callout* [10.09%], *Other collouts* [11.88%], *Foreign Affairs* [5.97%], *International* [3.67%] *Elections* [3.48%], *BJP* [2.78%] and *Events* [2.39%]. The average retweet rate for these tweets were 1497.093, whereas the favorite rate was 2981.526.

Table 3: Logit Regression Results to analyze the effects of the top 20 themes on twitter being mentioned in the media. The dependent variable is the number of article mention for each tweet

| Variables | Coefficient | std err | z | P > z |
|-------------------|-------------|---------|--------|--------|
| Const | -3.2028 | 0.142 | -22.54 | 0.000* |
| Political callout | 0.4917 | 0.107 | 4.599 | 0.000* |
| General callout | 0.3477 | 0.116 | 2.986 | 0.003* |
| Foreign affairs | -0.0731 | 0.127 | -0.577 | 0.564 |
| Elections | 0.0763 | 0.161 | 0.474 | 0.636 |
| Events | -0.667 | 0.163 | -4.082 | 0.000* |
| Development | -0.4908 | 0.188 | -2.614 | 0.009* |
| Greetings | 0.6644 | 0.123 | 5.407 | 0.000* |
| Gratitude | 0.0729 | 0.14 | 0.519 | 0.604 |
| Science and Tech | -0.1971 | 0.193 | -1.02 | 0.308 |
| International | 0.1866 | 0.143 | 1.304 | 0.192 |
| Tribute | 0.267 | 0.145 | 1.845 | 0.065 |
| BJP | 0.1933 | 0.161 | 1.2 | 0.230 |
| Celebrity callout | 0.5174 | 0.145 | 3.577 | 0.000* |
| Public address | -0.7585 | 0.273 | -2.782 | 0.005* |
| Initiative | -0.4792 | 0.25 | -1.917 | 0.055* |
| Criticism | -0.2874 | 0.233 | -1.235 | 0.217 |
| Youth | -0.256 | 0.209 | -1.226 | 0.220 |
| Vivekananda | -0.6639 | 0.297 | -2.234 | 0.025* |
| Congress | 0.2774 | 0.219 | 1.264 | 0.206 |
| Hinduism | -0.2593 | 0.197 | -1.319 | 0.187 |

* represents statistically significant coefficients

5.2 Thematic Coverage in the Media

We selected the 20 themes represented most frequently throughout the sample (Table 1), and analyzed their occurrence in the 939 Tweets that made it to the news. We see in Table 3, that the themes "Political Callouts", "General Callouts", "Greetings", and "Celebrity Callouts" were significantly more likely to be covered by the media outlets, whereas themes "Events", "Development", "Public Address", "Initiatives", and "Vivekananda" were significantly less likely to be covered.

These findings are intuitive, in that callouts are potentially newsworthy in and of themselves. First, for celebrity callouts, we find that the majority of the tweets captured by the news were one of two categories – 'regards' or 'requests'. Regards were either greetings or condolences, whereas requests were typically associated with some initiative of the Prime Minister's for which he sought the celebrity's support. An example is a tweet mentioning Sachin Tendulkar, India's cricketing star, who is himself widely followed on social media:

RT @sachin_rt: Shared his experience, interesting anecdotes and sound advice.... #SwachhBharat.. Swast Bharat.. Sach Bharat !! [50]

In the tweet above and in the corresponding news article, we find that the interaction between the prime minister and the country's best-known sportsman is news in and of itself. The wording of the tweet indicates how the news story is converted to a political advantage - by using an alliteration of Swachh (clean), Swast (healthy), and Sach (true) Bharat (India), a wordplay on the prime minister's flagship sanitation initiative, Swachh Bharat.

Similarly, "Political Callouts" are relatively more covered by the news media since these refer to an interaction between Modi and another politician. Callouts to foreign politicians are more likely to make the news, and their wording is frequently informal and affiliative. We see this in both the tweets below to the British and Canadian prime ministers respectively. In this tweet below, for instance, Modi congratulates David Cameron for his election victory, but in doing so uses clever wordplay on his own widely recognized

campaign slogan "Abki Baar Modi Sarkar", meaning, "This Time, Modi's Government" to a pun "One more time, Cameron's Government." The playful tone of the tweet suggests an affable relationship between the two leaders, suggesting warm relations between the two governments.

Congratulations @David_Cameron. As you rightly pointed out-its "Phir Ek Baar, Cameron Sarkar!" My best wishes. May 8, 2015[36]

The frequency of tweets relating to foreign leaders making it to the news is attributable to these being casual talking points for members of the press. Whenever Modi visits a foreign leader or has one come to him, the tone and content of the messaging is essentially a foreign relations artifact. The example Tweet to Canadian Prime Minister Stephen Harper also uses an informal, friendly tone, suggesting not only a congenial relationship between the two but a willingness on Modi's part to publicly embrace western leaders.

A special thanks to @pmharper- a fine host, a wonderful human being and a very dear friend. April 17, 2015 [37]

Such use of carefully crafted, often emotive tweeting has been a hallmark of Modi's social media approach [41]. We can see though that such framing, when used in the context of an exchange with global leaders, finds newsworthiness in the press since nested in the tweet is a larger story about international affairs. With Modi having a minimal media presence during his foreign trips and unwilling to give press conferences, tweets such as the above become the primary source of information for journalists looking to cover the event.

Besides specific celebrity and political callouts, we also find that general callouts have a higher likelihood of being covered in the news. The three most common sub-themes within the general callout-themed tweets that were covered in the news were "Condolences", "Greetings" and "Gratitude." For instance, the expression of condolences to citizens who have been in the news for the past for their work, but are not necessarily celebrities, serve as news vignettes for obituaries. Such as here:

Shri Baleshwar Aggarwal will be remembered for his efforts to keep Indian diaspora close to Motherland & in building Hindustan Samachar. RIP[28]

The tweets about Baleshwar Aggarwal, as well as Dada (J.P) Vaswani below are interesting for their making it to the news. To the untrained eye, they are casual citizens, there would be no reason for their birthdays to make it to the news. But both are in fact figures related to the Hindu Right. Vaswani is a spiritual leader with a strong international following, whereas Baleshwar Agarwal was an early figure in the international Hindutva movement who ran some of the Rashtriya Swayamsevak Sangh's (RSS)² publications. While obituaries of moderately known persons may make the news, the case of a birthday greeting highlights the ability of the prime minister to turn a private event into a mainstream event using social media.

Warm birthday greetings to Dada Vaswani, who has devoted his life to service & spirituality. I pray for his good health & long life.[38]

A feature of campaign tweets that make it to news is the use of a nested appeal. An example of a tweet that is an appeal to his followers is the following highly retweeted elections-themed tweet from October 2014 by Modi which was featured in the Times of India.

²https://en.wikipedia.org/wiki/Rashtriya_Swayamsevak_Sangh

It is an explicit urge for voter turnout, preceding the state elections and is worded to get Twitter users to propagate the message. The tweet makes it to the mainstream news, but also as a call to action galvanizes supporters to retweet aggressively.

Urging the people of Haryana & Maharashtra to go out & cast their votes. Youngsters must show the way & ensure record turnout. 14 Oct 14[30]

However, it is interesting to see what happens when such appeals are used in conjunction with a celebrity mention. There are two broad sets of appeals that Modi makes to celebrities - one is on voter turnout, and another is on state initiatives. 23 of the 30 tweets that included some kind of appeal to citizens that made the news were explicitly about voter turnout. Here below, he tweets to popular actor and Hindi film star Priyanka Chopra, asking her to urge voter registration, in a message sent out to various celebrities:

Dear @priyankachopra lets encourage voter registration among 18-24 yr olds. Large number of them unregistered. EC drive presently underway. September 20, 2013[29]

Since Priyanka Chopra is a widely covered celebrity herself, the conversation is deemed newsworthy, even though there is little about a message of asking young Indians to register online that is newsworthy, being a cyclical issue that is raised around elections. Since it appears to be a non-partisan public service appeal, it is within the realm of a reasonable request rather than something that would require the star to explicitly align themselves politically with Modi. The real news item is that Narendra Modi appears to have a one-on-one conversation with the leading film star, even when he actually isn't.

There are other similar cases of a tweet primarily intending to emphasize an affiliative relationship. An example is a tweet from early in his 2014 campaign drops the information that Salman Khan, one of the biggest stars of Bollywood, was dropping by for lunch. The newsworthiness of the story is the innuendo of friendship.

Enjoying Uttarayan in Ahmedabad. Salman Khan will join for lunch. @BeingSalmanKhan Jan 13, 2014 [31]

Interestingly, neither Salman Khan nor Priyanka Chopra responded to the tweets at the point, suggesting that the tweets were intended for public consumption, rather than as direct messages to the celebrities. Greetings-related tweets, which are disproportionately more likely to be covered by the news, are typically aimed at the population at large, such as for festivals, national commemorations etc. Two examples with greetings themes that were featured in the news highlight this

As the 2015 Cricket World Cup begins, my best wishes to the Indian Cricket Team Feb 2, 2015[39]

Its wonderful how all sections of Malayali society come together to celebrate Onam, making it a wonderful occasion of national integration. September 6, 2014 [32]

In both the cases, the greetings have some newsworthiness, since the prime minister calling out probably the biggest sports event in India. The second tweet from September 2014 comes at a time of public debates on secularism, his framing of the Kerala festival of Onam is worded in terms of national integration, alluding to the participation of Hindus, Muslims, and Christians in a festival in the South. The greetings thus are at once greetings related to

the specific events - i.e. World Cup, Onam, but also callouts to his followers.

Besides references to individual celebrity influencers, we also find tribute-themed tweets that on topics or figures that evoke emotions. Examples include tributes to Australian cricketer Phil Hughes, Hindu nationalist figure Vinayak Savarkar, and to the terror attack victims in Mumbai.

Heart-rending funeral in Australia. Phil Hughes, we will miss you. Your game & exuberance won you fans all over! RIP. Dec 2, 2014[33]

We remember the horrific terror attacks in Mumbai on this day in 2008 & pay homage to the innocent men & women who lost their lives. Nov 25, 2014 [34]

Veer Savarkar is remembered as a prolific writer, thinker, poet & a social reformer. Leaving for Parliament to pay tributes to Veer Savarkar May 27, 2014 [35]

Three tweet themes are significantly less likely to make the news - these are "events", "development", "public address" and "Vivekananda". Of these, events and public addresses are announcements - in other words, while the prime minister (or candidate before that) being present at an event may be newsworthy, the specific text of that message announcing it is arguably not. 'Event' tweets, structured as mini press releases, are part of Modi's media strategy that circumvents the mainstream media and directly addresses his Twitter followers referring to specific functions or gatherings that he was part of. However, devoid of wordplay and unattached to callouts or initiatives they are not deemed newsworthy enough to be used by the media.

On "Vivekananda", while it was used primarily early in his tweeting when he posted quotes from spiritual guru Swami Vivekananda as brief parables. These too are thus not newsworthy in and of themselves, since they are typically brief aphorisms, that are repeated as a sort of casual advice from a politician seeking to add a social message to his political tweeting.

Development, on the other hand, has different characteristics that may explain why the tweets with this theme are not newsworthy. The nature of political tweeting in India requires that the country's development be an important topic that the candidate engage with, consequently there is a very large volume of development-themed tweets, sixth in overall mention among all the themes. These tweets are typically not newsworthy for what is referred to as the 'banalities' of political tweeting, which are casual daily political performances of the politician calling for changes and improvement [41].

In summary, we see here that the themes that make the news appear to make the news because of a combination of factors, an important part of which is that the presence of a newsworthy element other than just Modi tweeting it out.

5.3 Relationship between Media Coverage and Retweeting

Media coverage of a tweet can potentially have effects on many factors including the retweet rate. We designed an experiment to investigate this effect.

For this experiment, we computed the daily mean retweet rate of Modi’s tweets and split it into groups controlled by whether a tweet has been mentioned by at least one media article. The goal of this experiment was to check if there is a significant difference between trends in the retweet rate of tweets that were mentioned in the media compared to those ones, which did not make to the media. Table 5 shows the results of this experiment performed on the retweet rate of tweets belonging to these two groups. The retweet rate for the group of tweets that were mentioned in media has an average value of 1360.03 which is significantly higher than the mean of the other group (617.06). Following a paired t-test to compare the mean difference between these two groups, we found that the difference is significant with the $p - value = 0.003$. The results strongly indicate that retweet rate is likely to increase for a tweet that is mentioned by the media. This observation can be justified by the fact that media coverage advertises a tweet to a larger audience, hence more people get a chance to read the tweet, which increases the chance of it being retweeted. It is also likely that tweets that are more likely to be retweeted because of their content and associations are likely to be covered by the media for the same reasons. For example, a tweet that is a callout to a celebrity benefits from its association to the celebrity in terms of both retweets (from both followers of the celebrity and of Modi) and in being deemed newsworthy by the media.

5.4 Effect of Becoming the Prime Minister

Becoming the Prime Minister after the 2014 general election has been a milestone for Narendra Modi’s political career. To understand whether his social media activities and effects also took a turn after this event, we split our data into two parts – before and after he became the Prime Minister. The aim of this experiment was to understand the effect of Modi’s new appointment on the daily percentage of tweets that are covered by the media. We computed the percentage of daily tweets that were covered by the media and split the tweets into two groups, based on whether it was posted before or after May 16, 2014³, controlling for the increase in following. Comparing the mean difference between these two groups using a paired t-test, we found that the daily percentage of media coverage of his tweets increased significantly ($p - val : 0.008$) after he became the Prime Minister. The mean percentage of media coverage of his tweets was 6.57% before he was the Prime Minister and this value rose to 14.42% after he won the elections. The detailed results of this comparison are presented in Table 4. This observation is intuitive as becoming the Prime Minister made him more newsworthy and his actions statements caught more media attention.

6 DISCUSSION

In this case study of Narendra Modi’s tweeting, we find evidence that a direct social media strategy has sizable mainstream media effects. Furthermore, with a media-savvy curation of subjects, language, and timing, we see how Modi is able to find resonance in traditional media. For the untrained observer, news reports of Modi tweeting to film stars, cricketers, and other celebrities can easily be interpreted as a two-way conversation.

³the day results of the election were announced

Table 4: Comparison of Tweets covered in media before and after Modi became the Prime Minister using paired t-test. Compares the percentage of tweets covered by the media for each group

| | Before PM | After PM |
|---------|-----------|----------|
| Mean | 0.0657 | 0.1442 |
| SD | 0.2477 | 0.3513 |
| SEM | 0.0036 | 0.0053 |
| p-value | 0.008 | |

Table 5: Comparison of Retweet Rate: mentioned and not mentioned in the media

| | Tweet not mentioned in media | Tweet mentioned in media |
|---------|------------------------------|--------------------------|
| Mean | 617.0625 | 1360.0302 |
| SD | 748.6947 | 1731.3625 |
| SEM | 7.9469 | 88.2002 |
| p-value | 0.003 | |

The boundaries between traditional and social news media are further blurred when the politician is able to piggyback upon the topics that are newsworthy in the mainstream media without directly interacting with journalists. Much of the computationally oriented research that studies this phenomenon has focused on understanding the mechanism behind the spread of news and information on social media or how news organizations and journalists rely on social media to disseminate their content. However, given the rising centrality (and in many cases exclusivity, in terms of primary output) of social media platforms like Twitter to political outreach, many politicians such as Narendra Modi – and to a large extent Donald Trump – heavily use social media as their means of connecting to both the general public and by extension, to the mainstream media, when they start to sanction the access that print journalists have to them. Narendra Modi’s use of social media is a valuable case study because the journalists were not only excluded from direct access to him during the campaign trail but frequently continue to be left out of foreign visits and major events. What we find the politician able to do, despite the presence of a relatively free press, is insert themselves into the discourse on the terms best suited to the politician.

Undoubtedly, there is dissent to Modi’s narrative that our work is unable to capture, but we see here that tweets act as public relations missives. If it hits the ground running by means of followers catching on or the subjects in it being newsworthy enough, then the exact text and framing of the Tweet, as the politician intends it, remains in the public. For Modi, this reduces the likelihood of being misquoted, misconstrued, or cornered by a member of the press. A large curation of Modi’s feed of images is on Instagram, he is a LinkedIn "influencer", every single speech and radio message of Modi’s is now on his YouTube channel, and he even has a host of videos on Yoga exercises for various physical ailments. Without owning a single integrated news source, social media nonetheless offers a significant, unified footprint to the political actor.

There are contextual editorial decisions that cause certain social media messages to end up on the mainstream news. The wording and consequent themes in a tweet are created by a person, likewise, the decision to include or exclude a tweet in the news is driven by an editorial decision. This makes the study of social media and its impact on mainstream media a candidate for careful interdisciplinary collaboration. It also makes a case study approach the most viable means of deep, contextual dive into a politician's strategy. While this may mean that there are problems with replicating in other settings, or for that matter, even temporally as new political and technical environments crop up, it nonetheless serves as an important lesson on what worked and what did not. That members of Modi's social media team have swiftly been courted by other politicians is already an indicator of professionals in the world seeking to emulate his experiences.

In Modi's case, we see two patterns in what gets covered in the news. The first pattern is that tweets are not necessarily covered because of something thematic that is in and of itself newsworthy, but rather there is a trigger event such as a birth or death, a sports event, etc. Here, the news item is the event itself, and the politician's tweet is one among several public figures' messages to people on the event. However, the political actor, once significant enough as an online entity, can be a newsmaker by virtue of them tapping into an existing event. Thus birthday greetings or condolences on relatively unknown figures make the corresponding tweets worthy of coverage by the news media.

The second broad pattern is the inclusion of tweets which have more intentionality – i.e. the political actor does something to make it newsworthy. These tweets often include the use of celebrities or other political actors who are also newsworthy to increase the possibility of it being covered in the news.

This research has implications for questions around the mediation, i.e. the tension between media logic, in which politicians trying to appeal to the mainstream media through what the media sees as newsworthy versus political logic, in which the politicians define the agenda for what is considered worth covering in the news. Importantly, Modi's tweets have by and large been in the news not for their rhetorical framing, as compared to US President Donald Trump, for instance, who has come to be known for his incendiary messaging, but rather for their subjective content.

Here, the development question is particularly salient. While the second order outreach effects are true for US President Donald Trump, there is a sense that the level and spread of media development means that it offers a counterpoint to the politician's narrative. Traditional news media relying on politicians' social media feed reverses the roles of social media and traditional news media. Rather than social media serving as a tool to disseminate content from news organizations, news now originates in social media and news organizations cover it. While India has a sufficiently large and diverse media environment (and indeed the largest being in regional language media, which we do not study), this trend can pose significant challenges in countries with a smaller mainstream media footprint.

6.1 Limitations

We limited the scope of this study solely to news media effects of the tweets. While there were a large number of important trends and analyses related to the thematic coding, retweets etc. that is an entirely separate set of issues that require an in-depth examination that would be a distraction in this piece. We chose to study the incidence of tweet text in the text of articles instead of other measures such as news related to tweets sent out by Narendra Modi, which may have other sets of effects, but are also difficult to isolate. Our study also does not capture the follow-up effects of a tweet appearing in the news such as subsequent stories about the subject – i.e. if Modi's tweet puts a topic on the map. Finally, while Modi's is a case of a single politician, we do not see this as a concern since he is currently the most followed elected sitting politician in the world and has changed the media landscape for political communication in the Global South, making him an important case study on social media and political outreach.

7 CONCLUSION

The ICTD universe needs to look beyond traditional notions of development to examine media democracy. Well before Donald Trump turned the political tweet into a regular news spectacle, Narendra Modi was able to make the news despite staying low on sensationalist topics. Even when he has piggy-backed upon newsworthy elements such as celebrities, he has done so in benign and careful ways. Modi's is a remarkable case of how social media can change the public outreach landscape for a politician, even in a country where a relatively small proportion of the population uses these technologies.

The case brings to light the urgent need for more mixed-methods studies that bring together the nuanced perspectives that trained political scientists or communications scholars can bring, with the powerful computational data mining techniques that can help identify the nature and spread of arguments online. As political communications cover new languages, devices, and communications technologies, the need of the day is for continued collaboration across disciplinary boundaries to track the future of democracy as it plays out online.

REFERENCES

- [1] Fabian Abel, Qi Gao, Geert-Jan Houben, and Ke Tao. 2011. Analyzing user modeling on twitter for personalized news recommendations. In *International Conference on User Modeling, Adaptation, and Personalization*.
- [2] Lorien C Abroms and R Craig Lefebvre. 2009. Obama's wired campaign: Lessons for public health communication. *Journal of health communication* 14, 5 (2009), 415–423.
- [3] Noa Aharony. 2012. Twitter use by three political leaders: an exploratory analysis. *Online Information Review* 36, 4 (2012), 587–603.
- [4] Anirban Baishya. 2015. Selfies# NaMo: The Political Work of the Selfie in the 2014 Indian General Elections. *International Journal of Communication* 9 (2015), 15.
- [5] Jody C Baumgartner, Jenn Burleson Mackay, Jonathan S Morris, Eric E Otenyo, Larry Powell, Melissa M Smith, Nancy Snow, Frederic I Solop, and Brandon C Waite. 2010. *Communicator-in-chief: How Barack Obama used new media technology to win the White House*. Lexington Books.
- [6] Pradeep Chhibber and Rahul Verma. 2014. The BJP's 2014 Modi Wave: An Ideological Consolidation of the Right. *Economic and Political Weekly* 49, 39 (2014), 50–56.
- [7] Pradeep K Chhibber and Susan L Ostermann. 2014. The BJP's fragile mandate: Modi and vote mobilizers in the 2014 general elections. *Studies in Indian Politics* 2, 2 (2014), 137–151.

- [8] J. Cohen. 1960. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement* 20, 1 (1960), 37.
- [9] Kathleen M Eisenhardt. 1989. Building theories from case study research. *Academy of management review* 14, 4 (1989), 532–550.
- [10] Gunn Sara Enli and Eli Skogerbø. 2013. Personalized campaigns in party-centred politics: Twitter and Facebook as arenas for political communication. *Information, Communication & Society* 16, 5 (2013), 757–774.
- [11] Alvan R Feinstein and Domenic V Cicchetti. 1990. High agreement but low kappa: I. The problems of two paradoxes. *Journal of clinical epidemiology* 43, 6 (1990), 543–549.
- [12] Will J Grant, Brenda Moon, and Janie Busby Grant. 2010. Digital dialogue? Australian politicians’ use of the social network tool Twitter. *Australian Journal of Political Science* 45, 4 (2010), 579–604.
- [13] Craig Hayden. 2013. Engaging technologies: A comparative study of US and Venezuelan strategies of influence and public diplomacy. *International Journal of Communication* 7 (2013), 25.
- [14] Christophe Jaffrelot. 2013. Gujarat elections: The sub-text of Modi’s high-tech populism and the neo-middle class. *Studies in Indian Politics* 1, 1 (2013), 79–95.
- [15] Christophe Jaffrelot. 2015. The Modi-centric BJP 2014 election campaign: New techniques and old tactics. *Contemporary South Asia* 23, 2 (2015), 151–166.
- [16] Varsha Jain, Christopher Pich, BE Ganesh, and Guja Armannsdottir. 2017. Structured Abstract: Comprehending Political Branding and Brand Image of the BJP in India. (2017), 1307–1313.
- [17] Kawaljeet Kaur Kapoor and Yogesh K Dwivedi. 2015. Metamorphosis of Indian electoral campaigns: Modi’s social media experiment. *International Journal of Indian Culture and Business Management* 11, 4 (2015), 496–516.
- [18] Matt J Kusner, Yu Sun, Nicholas I Kolkun, and Kilian Q Weinberger. 2015. From word embeddings to document distances. In *Proceedings of the 32nd International Conference on Machine Learning (ICML 2015)*. 957–966.
- [19] Haewoon Kwak, Changhyun Lee, Hosung Park, and Sue Moon. 2010. What is Twitter, a social network or a news media?. In *WWW*.
- [20] Dominic L Lasorsa, Seth C Lewis, and Avery E Holton. 2012. Normalizing Twitter: Journalism practice in an emerging communication space. *Journalism studies* 13, 1 (2012), 19–36.
- [21] Kristina Lerman and Rumi Ghosh. 2010. Information contagion: An empirical study of the spread of news on Digg and Twitter social networks. *ICWSM 10* (2010).
- [22] Swati Maheshwari and Colin Sparks. 2018. Political elites and journalistic practices in India: A case of institutionalized heteronomy. *Journalism* (2018), 1464884918761630.
- [23] Michael Mathioudakis and Nick Koudas. 2010. Twittermonitor: trend detection over the twitter stream. In *SIGMOD*. 1155–1158.
- [24] Nalin Mehta. 2010. Ashis Nandy vs. the state of Gujarat: authoritarian developmentalism, democracy and the politics of Narendra Modi. *South Asian History and Culture* 1, 4 (2010), 577–596.
- [25] Donald Metzler, Susan Dumais, and Christopher Meek. 2007. Similarity measures for short segments of text. In *ECIR*.
- [26] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013. Distributed representations of words and phrases and their compositionality. In *NIPS*.
- [27] @narendramodi. 2011. . <https://twitter.com/narendramodi/status/107696770540253184>.
- [28] @narendramodi. 2013. . <https://twitter.com/narendramodi/status/337897720981377024>.
- [29] @narendramodi. 2013. . <https://twitter.com/narendramodi/status/380978803129847808>.
- [30] @narendramodi. 2014. . <https://twitter.com/narendramodi/status/522207898114330624>.
- [31] @narendramodi. 2014. . <https://twitter.com/narendramodi/status/422993619918675968>.
- [32] @narendramodi. 2014. . <https://twitter.com/narendramodi/status/508445371136110592>.
- [33] @narendramodi. 2014. . <https://twitter.com/narendramodi/status/539997879205502976>.
- [34] @narendramodi. 2014. . <https://twitter.com/narendramodi/status/537413515607158784>.
- [35] @narendramodi. 2014. . <https://twitter.com/narendramodi/status/471512748119957504>.
- [36] @narendramodi. 2015. . <http://t.co/xf5tJfW0SE>.
- [37] @narendramodi. 2015. . <http://t.co/HoMyB1hDdI>.
- [38] @narendramodi. 2015. . <https://twitter.com/narendramodi/status/627700149171523584>.
- [39] @narendramodi. 2015. . <https://twitter.com/narendramodi/status/565850675397939201?lang=en>.
- [40] NDTV_Archive. [n. d.]. . <http://www.ndtv.com/>.
- [41] Joyojeet Pal. 2015. Banalities Turned Viral Narendra Modi and the Political Tweet. *Television & New Media* (2015).
- [42] Joyojeet Pal, Priyank Chandra, and VG Vinod Vydiswaran. 2016. Twitter and the Rebranding of Narendra Modi. *Economic & Political Weekly* 51, 8 (2016), 53.
- [43] Joyojeet Pal and Andre Gonawela. 2016. Political Social Media in the Global South. In *Conference on e-Business, e-Services and e-Society*. Springer, 587–593.
- [44] Owen Phelan, Kevin McCarthy, and Barry Smyth. 2009. Using twitter to recommend real-time topical news. In *ACM conference on Recommender systems*.
- [45] Swit Phuvipadawat and Tsuyoshi Murata. 2010. Breaking news detection and tracking in Twitter. In *Web Intelligence and Intelligent Agent Technology (WI-IAT)*, Vol. 3. 120–123.
- [46] Lance Price. 2015. *The Modi effect: inside Narendra Modi’s campaign to transform India*. Hachette UK.
- [47] Daniel Ramage, Susan T Dumais, and Daniel J Liebling. 2010. Characterizing Microblogs with Topic Models. *ICWSM 10* (2010).
- [48] Adrian Rauchfleisch and Julia Metag. 2015. The special case of Switzerland: Swiss politicians on Twitter. *new media & society* (2015).
- [49] Daniel M Romero, Brendan Meeder, and Jon Kleinberg. 2011. Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on twitter. In *WWW*. ACM.
- [50] @sachin_rt. 2014. . https://twitter.com/sachin_rt/status/522725083928743936.
- [51] Gerard Salton and Michael J McGill. 1986. Introduction to modern information retrieval. (1986).
- [52] Rajdeep Sardesai. 2014. *the election that changed India*. New Delhi: Penguin.
- [53] The_Hindu_Archive. [n. d.]. . <http://www.thehindu.com/archive/>.
- [54] TOI_Archive. [n. d.]. . <http://timesofindia.indiatimes.com/archive.cms>.
- [55] Peter Verweij. 2012. Twitter links between politicians and journalists. *Journalism Practice* 6, 5-6 (2012), 680–691.
- [56] Lars Willnat, Roshni Verghese, and Rashad Mammadov. 2017. Symbols, Slogans, and Charisma: Political Posters in India’s 2014 National Election. (2017), 187–209.
- [57] Wayne Xin Zhao, 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*.