Using news media and Twitter to provide a quantitative overview of factors related to the outbreak, successes and failures of the "Arab Spring"

Kenneth Joseph *[1], Matthew Benigni †[1], Wei Wei ‡[1], Kathleen M. Carley§[1], and Mia Bloom¶[2]

[1]Computation, Organization and Society Program, Carnegie Mellon University, 5000 Forbes Ave. , Pittsburgh, PA, USA
[2]Center for Terrorism and Security Studies, University of Massachusetts, Lowell, Lowell, MA, USA

May 13, 2015

Abstract

asdfsadf

1 Introduction

On December 17th, 2010, Mohamed Bouazizi immolated himself in Sidi Bouzid, Tunisia in response to harassment from both a local policewoman and local municipality officers. Though Bouazizi was not the first to engage in this form of protest, for some reason his case resounded with others who themselves took to the streets in protest of constant harassment and victimization by a corrupt government. Although early protests were relatively small and were met with violence from government forces, social media sites like Twitter, Facebook and YouTube were used to record these events and display them to the broader public. These events are widely considered to be the beginning of what has come to be known, for better or worse (Gelvin, 2015), as the Arab Spring.

There is little doubt that social media, and new media (Baym, 2010) more generally, played an important role in Arab Spring. However, it is almost common knowledge at this point that popular emphasis on social media as *the* cause of the revolutions is overblown (Bruns et al., 2013; Comunello and Anzera, 2012; Goldstone, 2013). Recent research has thus instead focused on how social media may have aided certain aspects of the revolutions in important ways for different people (Gallé et al., 2013; Lotan et al., 2011; Starbird and Palen, 2012; Tufekci and Wilson, 2012), and in how social processes that were carried out via new media are reflective of those that occurred "offline" (Comunello and Anzera, 2012). Similarly, recent work has suggested that data from newspaper

*kjoseph@cs.cmu.edu
†mbenigni@cs.cmu.edu
‡weiwei@cs.cmu.edu
§kathleen.carley@cs.cmu.edu
¶mia_bloom@uml.edu

articles written during the time of the Arab Spring also may be of use in better understanding these processes (Joseph et al., 2014a; Pfeffer and Carley, 2012). Thus, social media and the coverage of news media should be seen as both pieces and reflections of a complex system of causal structures that were at play.

To date, however, prior work has largely considered how the news media or Twitter are useful in understanding the social processes at play during particular events across a small set of nations () or a series of events in a particular nation, most often Egypt (). Further, few studies have considered news and Twitter data side-by-side, leaving questions as to the similar and different ways in which these media responded to different long and short term social processes during the Arab Spring.

In the present work, we use a corpora of around 70 M tweets and around 700K newspaper articles to provide an overview of the change in topical focus over time in sixteen nations relevant to the Arab Spring. We take a breadth-over-depth approach, attempting to reconcile patterns in Twitter usage and news media coverage over a wide range of countries and time periods. Additionally, we utilize the same methodology for both datasets, allowing us to compare results across media. Specifically, we here focus on the following three research questions:

- **RQ1:** How did the topical focii of our news and Twitter data differ over time and across different nations?
- **RQ2:** How did (un)successful government overthrows change the topical focii of news and Twitter data?
- **RQ3:** Can we develop new hypotheses based on our data for the relationship between topical focii in news and Twitter data and (a lack of) social change? a holistic view of the topical focii of

In order to study these three research questions, we begin by developing a set of N human-curated topical themes of interest based on prior literature and identifying terms that, when mentioned, are relevant to these themes. We then search for these terms across all of our Twitter and newspaper data. Where we find a term used in a particular tweet or news article, we determine the time at which the content was produced and the particular nation(s) in the Arab World that the content is relevant to. We thus are left with a set of counts, over time, of the discussion of our different thematic content in different nations in the Arab world.

Naïvely, we could then use this count data, or rates directly calculated from this count data, to address our research questions. We could compare normalized rates of usage across news and Twitter data to address RQ1, look at changes in these rates before and after government overthrows to address RQ2 and develop a case study to explore the changes in these rates in RQ3. As detailed in recent work by Eisenstein et al. (2014), and as we will show in this article, however, the direct utilization of count data, or rate data based off these counts, is a methodologically unsound decision.

As Eisenstein et al. (2014) discuss for Twitter, term count may be biased by unknown irregularities in the way Twitter provides tweets through its API (Morstatter et al., 2013) or unique properties of the keyword or spatial queries researchers construct to obtain data from the API (Joseph et al., 2014b). Similarly, superfluous coverage by news media on particular nations may lead to artificial increases in counts or, if focused on themes not of interest in the present work, superfluous decreases in rates. In the present work, we thus adapt the statistical model developed and employed by Eisenstein et al. (2014). This model mediates the effect of these biases by controlling for spatial and temporal patterns in the rate at which data is obtained and by smoothing rate estimates via an autoregressive model. While a plethora of issues still must be considered when analyzing social media (Tufekci, 2014) and news media () data, Eisenstein et al.'s (2014) model allows us to move beyond these statistical irregularities in the data and thus gives us more freedom to draw inferences about the relationship between these adjusted rates of keyword usage and actual events occurring during the Arab Spring.

After explaining in more detail the methods utilized by Eisenstein et al. (2014) and how they were adapted for the present work, we consider three analyses that address each of our three research questions. With respect to RQ1, we find that **todo Kenny: [[??]]**. With respect to RQ2, we find that **todo Kenny: [[??]]**. With respect to RQ3, we find that **todo Matt: [[??]]**.

2 Related work

The immense number of complexities and historical artifacts that played a role in the beginnings, successes and failures of the Arab Spring prohibit a full exploration here. Instead, we provide only a summary of some of the more widely accepted factors, and ones that readers should be aware of and keep in mind throughout our analysis. For a more detailed overview, we direct the reader to Gelvin's (2015) recent book. In this section, we also discuss recent work utilizing social and/or news media during the Arab Spring, detailing how the methods and data used here are both similar and different.

2.1 Causal factors of the Arab Spring

Two broad classes of causal factors can be considered - those that prompted revolution, and those which affected the success or failure of revolutions. We briefly review each set of factors here.

2.1.1 Causes of protests and their spread

A host of historical factors led to the conditions in 2011 that made many nations in the Arab world ripe for protest Gelvin (2015). One long-standing issue was the increasing extent of economic problems caused by ineffective, corrupt and state-run economies. These issues led to high levels of unemployment and inflation Dewey et al. (2012) as well as to both food shortages and huge hikes in food prices, all of which contributed to high levels of civil unrest Comunello and Anzera (2012); Goldstone (2011). The effect of high unemployment rates was particularly a problem because its effect was particularly strong on well-educated youth populations, individuals who had often been promised that their education efforts would be rewarded with jobs Dewey et al. (2012); Gelvin (2015). This, combined with a "youth buldge" in which a disproportionate percentage of the population was between the ages of 15-29 in many of the MENA region countries, provided a fodder of civil unrest that required only a spark to ignite and a gust of wind to spread throughout the region.

This spark came, as noted above, in the form of Mohamed Bouazizi. While Bouazizi may have provided the spark, it is generally aggreed upon that social media, or more aptly, the existence of communication infrastructures that supported all forms of new media(Tufekci and Wilson, 2012; Wolfsfeld et al., 2013), served as the wind to spread the flame to Egypt, to Libya and on to several other nations in the Arab world. Even before the protests, however, the existence of internet connectivity allowed individuals within the Arab region to observe the democratic processes existent in other regions of the world, stoking their desire to live in that type of environment Hussain and Howard (2013). During the early moments of the revolutions, these tools allowed both deliberate diffusion processes, defined as those "carried out via the conscious sharing of tactics and frames by activists who are linked by networks that may be transnational" and demonstration diffusion effects, " ' the power of precedent' " to occur, rapidly engulfing an increasing number of actors in an increasing number of nations in protest (). Through both logics, a further cause of spread was the strengthening of an underlying Arabic identity, which united protesters in their unified goal of empowerment and the ending of corruption.

2.1.2 Causes of Varying Outcomes

Goldstone (2013) suggests that several additional predictors were important in determining whether or not a particular government was ultimately overthrown. Chief amongst these predictors, he argued, was the structure of the overarching regime had a strong impact on whether or not the government was ultimately overthrown. This view is a piece of a consistent take on the properties of the ruling regime as the primary factor in the success of a particular revolutionBellin (2012); Comunello and Anzera (2012); Goldstone (2013). Goldstone defines a *personalist* regime as one in which a single individual – who may have begun as an elected leader, or head of a military or even party regime – takes total or nearly total control of the national government. He then provides qualitative evidence for his believe that "the single best key to where regimes in MENA have been overturned or faced massive rebellions is where personalist regimes have arisen".

Goldstone's (2013) work argues that personalist regimes were the most susceptible because their power was tied to their ability to provide the necessary economic and political incentives to their constituency, particularly in nations that depended on oil production. While Comunello and Anzera (2012) notes that personalist, or as he refers to them, neopatrimonial states, had controlling arms that made it difficult to organize any sort of formal protest, three other factors led to conditions in which such formal protests could arise.

One such factor that has already been discussed was the economic conditions under which the revolution occurred. While these economics played a role in bringing about revolution, they also prevented personalist regimes from being about to "buy their way out" of the protests, and thus also had a role in the revolutions' successes. The second factor that played into the success of personalist regimes in the face of revolution was the relationship of the regime to the military Battera (2014); Comunello and Anzera (2012). This relationship can be boiled down to one decision the military made - whether or not to shoot at the protestors Bellin (2012)[1]. In countries where the military made the decision to quell protests with violent force, protesters were able to flood the streets without fear of the full wrath of the state. This lack of impunity led to stronger protests that eventually led to the downfall of the Tunisian and Egyptian regimes. In contrast, in Saudi Arabia, for example, where the military opted to side with the regime and use force to quell protests, regimes managed to maintain power.

In making the decision of whether or not to fire on protestors, one important factor was the extent to which previously disparate social groups formed a cross-class coalition in their protests and revolutionary efforts Goldstone (2011). A unified coalition of protestors made it more difficult for the military to justify the use of force in their response to the protest for two reasons. First, the combination of various social groups lessened the military's ability to claim that violent actions were a response to a particular out-group in the interest of protecting the "nation". Second, the sheer size of such a coalition would relegate military action to being viewed as "illigitimate slaughter" Bellin (2012).

The development of these unified coalitions across various social groups also provided an opportunity for the news media to characterize protestors under a national identity, rather than as protests via one specific subgroup. As we discuss shortly, the portrayals of the revolutions by news media thus had an important impact on revolutionary outcomes in that their coverage ultimately

[1] While regimes certainly had other forces willing to deal with the protestors, the extent and breadth of the protests eventually came to a point at which it "is sufficient to look at the character of the military and its capacity and will to repress in order to reckon the immediate chances of regime survival." Bellin (2012)(pg. xx).

influenced a final factor in revolutionary success, which was the extent and type of involvment in the revolutions in the various nations from international powers Comunello and Anzera (2012); Goldstone (2013). This applies both to the actions of the West during the revolutions as well as the actions taken by Arab nations themselves.

Social media, of course, had much the same effect, increasing both the level of information spread and the extent to which individuals felt compelled to participate in protests Bellin (2012); Tufekci and Wilson (2012); Wolfsfeld et al. (2013). As insinuated above, it is generally agreed upon that social media, and new media more generally, played at least some role in the spread and success of certain revolutions that occurred during the Arab Spring. There do exist, however, many scholars who feel this effect has been overstated, or that no such effect exists at all. Comunello and Anzera (2012) give an overview of a significant amount of work focusing on the relationship between social media and the Arab Spring. The authors consider how, across 8 different properties of socail media, technological determinists (people who believed social media played *the* causal factor) and the techno-realists (people who believed social media played *no* causal role) differ on their opinions regarding the effect of social media. Additionally, Hussain and Howard (2013) provide a unique analysis of the role that communication infrastructures played in coordination with other effects on the revolutions.

2.2 Using new media to study the Arab Spring

Having given our stance on the question of whether or not social media played a role in the revolution, we now turn to how data from social media has been used to better understand the processes inherent to the revolution as a whole. In particular, we focus on Twitter, as this is the data available to us in the present study. The earliest work we are aware of to look at the relationship between new media and the Arab Spring was that of Lotan et al. (2011). The authors stress the strong interplay between Twitter and news media, and showed that the relationships between mainstream media outlets, activists, journalists and bloggers differed in datasets collected from Tunisia and Egypt. Lotan et al. (2011) considered data that used the hashtags #egypt or #libya, finding that Egypt and Libya displayed differing amounts of tweets in the two major languages of interest, Arabic and English. Egypt having significantly more tweets in Arabic than English, and Libya the other way around. Later work by Bruns et al. (2013) observed that the English-speaking world lost interest in the events relevant to the Arab Spring much earlier than those tweeting in Arabic. As we focus largely on English-language keywords, such language effects must be kept in mind.

Perhaps most relevant to the present work, however, are the efforts ofBorge-Holthoefer et al. (2014). These authors use Twitter data to test the extent to which individuals switched between Secularist and Islamist and pro and anti-military "camps" in Egypt during the Arab Spring. While the authors focus on Arabic tweets, their efforts show that Twitter data provides a unique lens through which important and interesting social processes relevant to the revolutions can be studied. We extend their efforts to new questions, new data and new methods in the present work.

2.3 The role of the news media

As argued by Hussain and Howard (2013), international news media organizations were important in the Arab Spring in that they brought world-wide attention to the events in the region, helping to "stave off overtly violent reactions from security forces". We would thus expect that the degree of coverage for particular topics, especially those relating to violence and revolution, should have interesting relationships with the success and/or failure of uprisings across the Arab world. Goldstone (2011) argues in a similar fashion that the extent to which the news media portrayed "protestors as representative of the whole society, rather than as one particular group seeking partisan advantages for itself" was likely to impact the success of the revolution. We thus consider the extent to which this is the case in our data.

Other work by Joseph et al. (2014a) has considered quantitative approaches to using news media data to better understand the Arab Spring. Specifically, the authors use news media to instantiate a dynamic network agent-based model and use model output to predict successful revolutions. While their use of newspaper data chiefly incorporates off-the-shelf statistical tools, it is an important indicator that news media data can be used in large quantities to provide accurate models of underlying social processes.

3 Data

The present work focuses on discussions of N human-curated topical categories in large corpora of news and Twitter data. We focus on 15 countries in the Arab World, plus data relevant to Iran. The countries of interest are given in Figure **??**. Below, we describe the Twitter and news datasets utilized. For more information on the Twitter data, we refer the reader to . For more information on the news data, we refer the reader to . Note that we use a subset of the Twitter data described in previous work here, and an expanded news dataset collected in the same fashion as the data used in previous work.

3.1 Twitter data

The present work utilizes a corpora of approximately 70M tweets collected between February, 2011 and February, 2013. The data was collected from a set of two sources. The first utilized a variety of approaches, including geospatial bounding boxes, keyword searches and searches for specific Twitter users, to capture data that data collectors deemed relevant to the Arab Spring. As the data was obtained from an outside source, an exact listing of terms and spatial regions from which this set of data was collected is unavailable to us. However, most of the collection is known to have focused on events relevant to Egypt, Libya, Syria, Tunisia and Yemen. The second source of data is obtained from a 10% sample of all Twitter data made available to members of our university. From this sample, we extract all tweets that were geo-tagged from a location inside the nations of interest. Any geo-tagged tweet from a particular country is assumed to be relevant to that country. In addition, any tweet that uses a country's name, in Arabic or in English, or a tweet that uses the name of any of the country's five most populous cities in English or in Arabic, is assumed to be relevant to that country. Note, then, that the same tweet may be considered relevant to multiple countries.

3.2 News data

4 Methodology

There are three major parts to the methodology completed in this paper. We first cover the manner in which we determined the themes, or categories, of interest to us and how we selected sets of terms used as a proxy for the discussion of each of these categories. We then describe the methodology we utilize to determine the extent to which each category was being discussed in different countries over time. Finally, we detail methods used to analyze the output of this model in order to better understand our three research questions.

4.1 Categories

As the data we use only covers periods during and after the first revolts, the categories we considered in our analysis were chiefly focused on identifying discussions on topics relevant to the spread of the Arab Spring and the extent to which revolutions in different nations succeeded. After iterating on possible categories, we first settled on twenty categories that we were interested in better understanding. Following a decision on the categories of interest, our next task was to determine a set of terms[2] that, if mentioned in a tweet or news article, the authors agreed pointed to a discussion pertinent to that topic. This list was refined by the authors until all agreed on the set of terms relevant to each category. Finally, we searched through every tweet and news article in our dataset for these terms. Due to the relatively data-hungry nature of the statistical model described in the following section, we eventually chose to ignore any category that was found in fewer than 5,000 news articles or 5,000 tweets. This left us with a final set of twelve categories that we considered in our analysis.

Category	**Reason for Use**	**Num. Terms**	**Representative Terms**
Terrorist Orgs (terrorist_org)	Indicator of instability	3619	Al-qaeda, terrorism
Adaptation	Indicator of change	130	adjustments, amendment
Protest	Indicator of unrest	19	protests, demonstrations
Violence	Indicator of unrest	39	violence, violent
War	Indicator of unrest	46	war, wartime
National identities (nationality)	Goldstone (2011)	245	Algerian, Syrian
Youth	Goldstone (2011)	10	young person, youngster
Non-national identities (tribe)	Goldstone (2011)	1489	Shi'a, Sunni
Profession	Goldstone (2011)	395	teacher, textile worker
Noise 1 (stop4)	Check for spuriousness	14	good, person, problem
Noise 2 (stop5)	Check for spuriousness	16	good, person, problem
Noise 3 (stop6)	Check for spuriousness	14	good, person, problem

Table 1: Categories Used

Table 1 presents the set of categories used in the present work, along with a description of why they were chosen, the number of terms that map to the category and a set of search terms representative of the category. With respect to reasons for use, there are three general reasons why the categories were selected. First, we desired a set of general indicators for social processes involved in revolution, in particular indicators for varying levels of unrest (protest, violence, and war) and indicators for processes of change (adaptation). In addition, we included a large number of terrorist organizations. The existence of terrorist organizations in a nation-state is known to be a general indicator of instability, as highly unstable countries are potential places where terrorist organizations can gain a foothold.

Second, we included several concepts related to Goldstone's (2011) theoretical model of cross-class revolution in order to better understand the extent to which his predictions are reflected in the

[2]Note that a term may comprise multiple words, e.g. "Michael Jordan" might be a single term for the category "sport"

language of social media and news media. Goldstone's (2011) recent assessment of the drivers of instability and revolutionary success during the Arab Spring suggests that a recurrent causal story played a role in the early revolutionary successes in Egypt, Tunisia and Libya. Goldstone states that "virtually all successful revolutions were forged by cross-class coalitions...pitting society as whole against the regime". He then provides qualitative evidence that the revolutions in three countries were no different. In Egypt, Tunisia and Libya, previously disparate social groups combined under a united "cross-class coalition" of protesters and revolutionaries, which made the revolutions significantly more likely to succeed.

Finally, we use a set of three "noise categories" to better understand the existence of spurious correlations that the methodology we utilize here may induce. These noise categories were constructed by taking the top 25 most frequently used nouns and adjectives, as provided by Wikipedia[3] and randomly splitting them into three categories. Of these 50 terms, we removed five nouns (man, woman, child, life, government) and one verb (young) that were relevant to categories of interest, leaving us with 44 terms split between the three noise categories.

With respect to the number of terms used, it is clear that there is large variation across categories. This imbalance is only to be expected- while certain categories, in particular the Terrorist Organization and Non-national identity ones, can contain a virtually limitless number of terms, many of the other categories (e.g. national identity) are restricted to a very limited number of possible terms. However, there are two reasons why one should not expect these differences to affect our analyses. First, it is important to realize that the vast majority of terms in these two categories were found in only a few, if any, tweets and news articles. Second, as we show in the next sections, categories are analyzed independently and only compared based on rates of change.

4.2 Statistical Approach

The statistical model we utilize is drawn from the work of Eisenstein et al. (2014). Eisenstein and his colleagues were interested in understanding the extent to which new words diffused on Twitter across metro areas in the United States. To capture the extent to which words in one metro area r_1 spread to a different area r_2, they formulated an autoregressive model that captured the extent to which the level of "activation" for the words in their dataset at time $t-1$ in metro area r_1 predicted the level in r_2 at time t. As noted above, a naïve use of count data, or proportions based on count data, would provide a biased analysis. Briefly, this bias can be broken into two general, exogenous factors that influence fluctuations in both counts and proportions beyond the desired influence relationship. First, as Eisenstein et al. (2014) notes, sampling rates in the data may differ *over time* due to fluctuations in the rate at which Twitter's API provided data over time. These temporal patterns may also lead to random spikes in the data, spikes which may require smoothing. Second, as our sample differentially focused on different countries (or in Eisenstein et al. (2014), because people tweeted more regularly from larger metro areas), there also exists and exogenous *spatial* factor on count data that one does not actually wish to model[4].

Our interest in the present work is not to understand diffusion, but rather to capture the extent to which different categories we are interested in were being discussed in different countries at different times. Thus, large portions of the work performed by Eisenstein et al. (2014) still apply, in particular the way in which they partial out temporal and spatial biases in their data. While the autoregressive portion of their model is not used here to estimate a diffusion network, we find that it still allows for a desirable smoothing of estimates over time that provides a more realistic picture of long-term change in discussion topics.

[3]`http://en.wikipedia.org/wiki/Most_common_words_in_English`

[4]For more detail on these biases, we refer the reader to the original article

Parameter	Explanation
$c_{w,r,t,m}$	The *c*ount, or number, of users (newspaper articles) that mentioned category w related to country r at time t in media m
$s_{r,t,m}$	The total number of users (newspaper articles) related to country r at time t in media m
$\eta_{w,r,t,m}$	The *activation rate* of category w related to country r at time t in media m
$v_{w,t,m}$	The activation rate of w across all countries at time t in media m
$\mu_{r,t,m}$	The base activation rate of all users interested in country r at time t in media m
$A_{w,m}$	The autoregressive component for each w across all countries and times in media m
$\sigma_{w,r,m}$	The standard deviation of η draws for a particular w and r in media m

Table 2: table

Table 2 presents an overview of the variables used in our description of the statistical model. Note that each model variable is differentiated by media via the subscript m. In all cases, results for the two media are estimated independently. Thus, to ease notation in this section, we will focus on Twitter, and consequently drop the m subscript in our discussion. The analog to a Twitter user is a newspaper article. All data is aggregated by month and by country. From this aggregated data, we extract two primary values. First, the variable $c_{w,r,t}$ gives the number of Twitter users (newspaper articles) who used category w in at least one of their tweets related to the country r during month t. Second, $s_{r,t}$ represents the total number of users who sent one or more tweets about *anything* in country r during month t. Thus, it is straightforward to represent $c_{w,r,t}$ as being distributed binomially, $c_{w,r,t} \sim \text{Bin}(s_{r,t}, p)$.

It is obvious to see that the MLE for p, $\hat{p} = \frac{c_{w,r,t}}{s_{r,t}}$ in the naive model. To address the biases above, Eisenstein et al. (2014) introduce a logistic model for the p parameter that allows an additive structure for parameters to estimate these biases independently, along with the "true" indicator of the popularity of category w in country r at time t. Equation 1 defines their basic model, where the logistic function Logistic(x) is $\frac{1}{1+\exp(x)}$.

$$c_{w,r,t} \sim \text{Binomial}(s_{r,t}, \text{Logistic}(\eta_{w,r,t} + v_{w,t} + \mu_{r,t})) \tag{1}$$

In Equation 1, the three parameters in the logistic function affect the log-odds of an increase in $c_{w,r,t}$. An increase in a parameter represents an increase in these log-odds, and thus the higher the value of each parameter, the more likely a user is to use the term. Eisenstein et al. (2014) uses the term "activation" to describe these increases, presumably in the context of the logistic model frequently used in cognitive activation theory . We follow this terminology here. In Equation 1, the parameter $v_{w,t}$ is the overall activation of category w at month t, and $\mu_{r,t}$ is the activation of the country r at time t. These terms control for temporal and spatial biases, respectively. The parameter of interest to our analysis is thus $\eta_{w,r,t}$, which represented the unbiased activation for word w at time t in region r.

As noted above, we also would like to smooth our estimates of word activations over time to better understand longer term trends in the data. As $\eta_{w,r,t}$ is our parameter of interest, the smoothing is enacted as an autoregressive model on *eta*, as described in Equation 2. Here, the parameter $A_{w,r}$ is the lagged influence variable, estimated for each category w and each country r.

$$\eta_{w,r,t} \sim \text{N}(A_{w,r}\eta_{w,r',t-1}, \sigma^2_{w,r}) \tag{2}$$

The full autoregression model can be specified as a Markov model with an *observation model* (on c) and a *dynamics* model on η. Overall, the model we thus wish to estimate is:

$$P(\eta, c|s; A, \sigma^2, \mu, v) = P(c|\eta, s; \mu, v)P(\eta; A) \tag{3}$$

In order to do so, we adopt the same estimation process as was used by Eisenstein et al. (2014). We first consider the estimation of v and μ for each word assuming η is 0. To do so, we utilize a stepwise procedure. First, we obtain a simplified $\bar{v_w}$ as the inverse logistic function ($\log(\frac{x}{1-x})$) of the total proportion of users that utilized word w across all time steps. Using this, we now would like to compute the maximum likelihood estimate of each $\mu_{r,t}$ using $\bar{v_w}$ as the value for each $v_{w,t}$ (and again, setting all $\eta = 0$). Below we derive the MLE for a particular μ_r at a single time point t (where the t subscript is implicit), a step absent (presumably due to space) from the original article. We first derive the form for the MLE:

$$\begin{aligned}
\hat{\mu_r} &= \text{argmax}_\mu \prod_w P(c_w|s; \mu, v_w) \\
&= \text{argmax}_\mu \prod_w \tbinom{c_w}{s} \frac{1}{1+\exp(-(v_w+\mu))}^{c_w} (1 - \frac{1}{1+\exp(-(v_w+\mu))})^{s-c_w} \\
&= \text{argmax}_\mu \sum_w \log\tbinom{c_w}{s} + \log(\frac{1}{1+\exp(-(v_w+\mu))}^{c_w}) + \log((1 - \frac{1}{1+\exp(-(v_w+\mu))})^{s-c_w}) \\
&= \text{argmax}_\mu \sum_w \log\tbinom{c_w}{s} - c_w \log(1+\exp(-(v_w+\mu)) + (s-c_w)\log(1 - \frac{1}{1+\exp(-(v_w+\mu))})
\end{aligned} \tag{4}$$

As there is no closed-form solution to the MLE, we compute the derivative of the above expression and use it for a gradient descent approach to maximizing the function above. The gradient can be derived as follows:

$$\begin{aligned}
&\frac{\partial}{\partial \mu} \sum_w \log\tbinom{c_w}{s} - c_w \log(1+\exp(-(v_w+\mu))) + (s-c_w)\log(1 - \frac{1}{1+\exp(-(v_w+\mu))}) \\
&\qquad = \sum_w \frac{(c_w - s)\exp(v_w+\mu)}{1+\exp(v_w+\mu)} + \frac{c_w}{1+\exp(v_w+\mu)}
\end{aligned} \tag{5}$$

An analogous derivation can be constructed for each v, and we thus exclude it from the present work. Now that we have our estimates for μ and v, we will estimate values for $\eta_{w,r,t}$ and $\sigma^2_{w,r}$. Because the observation model is non-Gaussian, the traditional Kalman Filter algorithm to infer parameter values cannot be used. Instead, we resort to Bayesian methods to perform approximate draws from the distribution of the η s over time. We can then use maximum likelihood estimation to update the diagonal of $\tilde{A_w}$ and $\sigma^2_{w,r}$. We then update expectations for $\tilde{A_w}$ and $\sigma^2_{w,r}$ and iterate again to generate a draw from an updated version of the distribution for $\eta_{w,r,t}$.

This process results in a Monte-Carlo EM algorithm to estimate A_w, $\sigma^2_{w,r}$ and $\eta_{w,r,t}$. In the E step, we get an expectation for η using Forwards-Filtering Backwards Sampling (FFBS) . Eisenstein et al. (2014) do not specify an initial distribution for η_0, so we draw from N($\hat{eta_0}$, 1), where $\hat{eta_0}$ is the MLE for eta_0, given in Equation 6, where $p = \frac{c}{s}$:

$$\begin{aligned}
\hat{\eta_0} &= \text{argmax}_{\eta_0} p(\eta_0|\sigma^2, c_{t=0}, s_{t=0}, \mu_{t=0}, v_{t=0}) \\
&= \log(\frac{p}{1-p}) - v_{t=0} - \mu_{t=0}
\end{aligned} \tag{6}$$

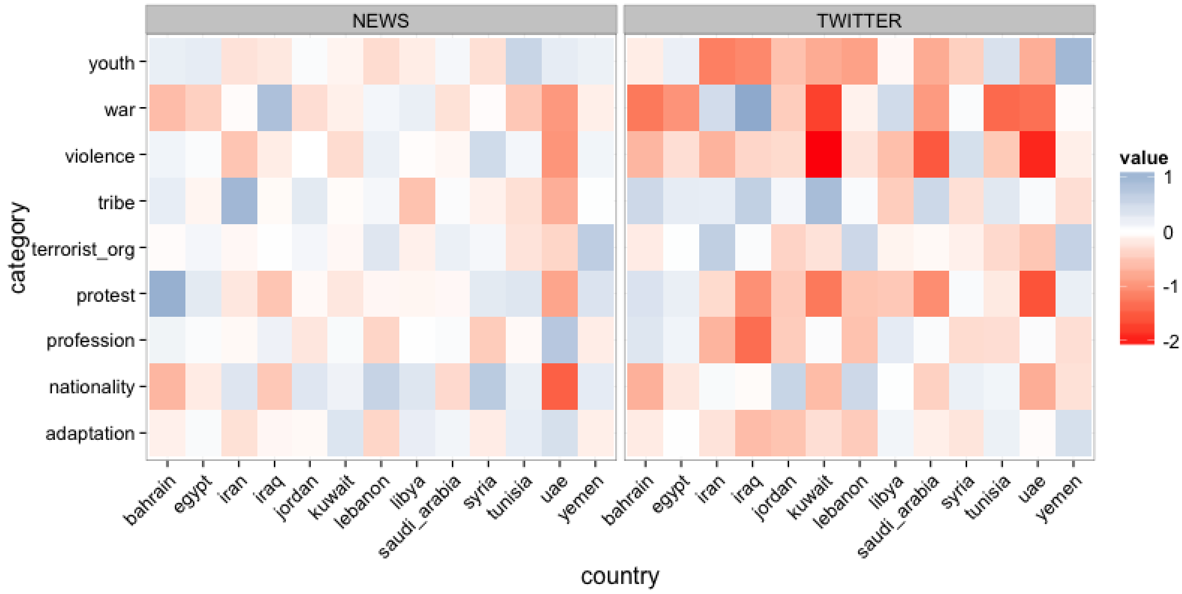

Figure 1

After obtaining a value for η at time 0, we proceed with FFBS with the proposal distribution $Q(x)$ equal to the transition distribution. In contrast to Eisenstein et al. (2014), who use a simple particle filter in the forward-filtering step, we found our estimates were much more stable when we used an additional resampling step during the filtering process. Thus, for each time point, we construct weights for each sample $\eta_{w,r,t}^{(m)}$ as:

$$\omega_{w,r,t}^{(m)} = \omega_{w,r,t-1}^{(m)} * P(c_{w,r,t}|\eta_{w,r,t}, s_{w,t}; \mu_{r,t}, v_{w,t}) \tag{7}$$

We then resample from this distribution, producing a set of samples for $\eta_{w,r,t}$ from a discrete approximation of the true distribution. This process, when run on each η, completes the "filtering" step on the forward pass. On the backwards pass, we draw our samples for $eta_{w,r,t}$ using these weights. This completes the E step, or in other words, provides our samples for $\eta_{w,r,t}$. The M step updates the MLEs for A_w and $\sigma^2_{w,r}$. The update for $\tilde{A_w}$ is simply determined via least-squares estimation. The MLE for $\sigma^2_{w,r}$ can be solved in closed form, $\frac{1}{T}\sum_t^T(\eta_{w,r,t} - \tilde{a_w}, r\eta_{w,r,t-1})$.

The code for this estimation process is available at (REMOVED FOR BLIND REVIEW). Note that the primary differences between our work and Eisenstein et al. (2014) are two-fold. First, Eisenstein et al. (2014) are interested in constructing a diffusion network, and thus continue with a further estimation step to approximate a full transition matrix A across all regions. Second, while Eisenstein et al. (2014) focus on specific terms, we focus on collections of terms. However, the generalization is trivial, as the sum of a set of independent Binomial random variables is still binomially distributed.

4.3 Analytical methods

5 Results

Figure 1 shows the average level of activation across countries, per category of interest. A few things are of interest. First, Twitter looks more variable. Second, similarities across categories.

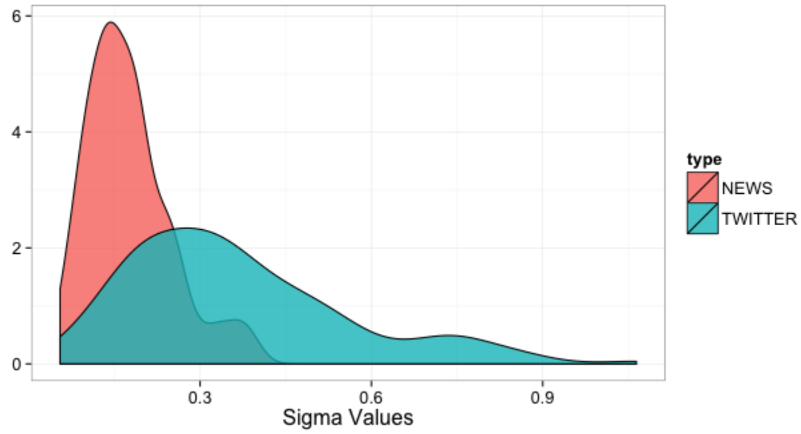

Figure 2: Distribution of Sigmas

Figure 2 shows that the Twitter data is more variable than the news data. This is what we would expect because ... This tells us also that changes in distributions on Twitter may not necessarily be as interesting as changes in the news data.

What about correlations? That is, were news and Twitter correlated? We cant use our data as is, BC non-IID. But because estimates from the model are, bydefinition, AR(1), we can subtract off mean of the last month and then treat the result as IID. Now we ask, given a country and a category, are these two time series independent?

What about categories? Which were most correlated?

What about countries?

Where did the biggest jumps happen? Was that in accordance with anything?

	category	country	type	date	val	sigma
1	war	syria	NEWS	2011-03-01	-0.72	0.18
2	protest	syria	NEWS	2011-03-01	1.02	0.27
3	tribe	uae	TWITTER	2012-10-01	1.55	0.42
4	protest	tunisia	TWITTER	2012-08-01	1.83	0.50
5	war	lebanon	NEWS	2011-03-01	-0.64	0.18
6	adaptation	libya	TWITTER	2012-08-01	1.43	0.41
7	adaptation	tunisia	TWITTER	2011-10-01	-1.42	0.41
8	terrorist_org	libya	NEWS	2012-09-01	0.47	0.14
9	terrorist_org	egypt	TWITTER	2012-11-01	0.72	0.21
10	profession	libya	TWITTER	2012-09-01	-1.16	0.34

5.1 RQ1: News vs. Twitter

5.2 Case Study

6 Conclusion

Our analysis is cowith the acknowledgement of the rash of recent claims over the methodological trapdoors that exist within this type of data Joseph et al. (2014b); Morstatter et al. (2013); Ruths

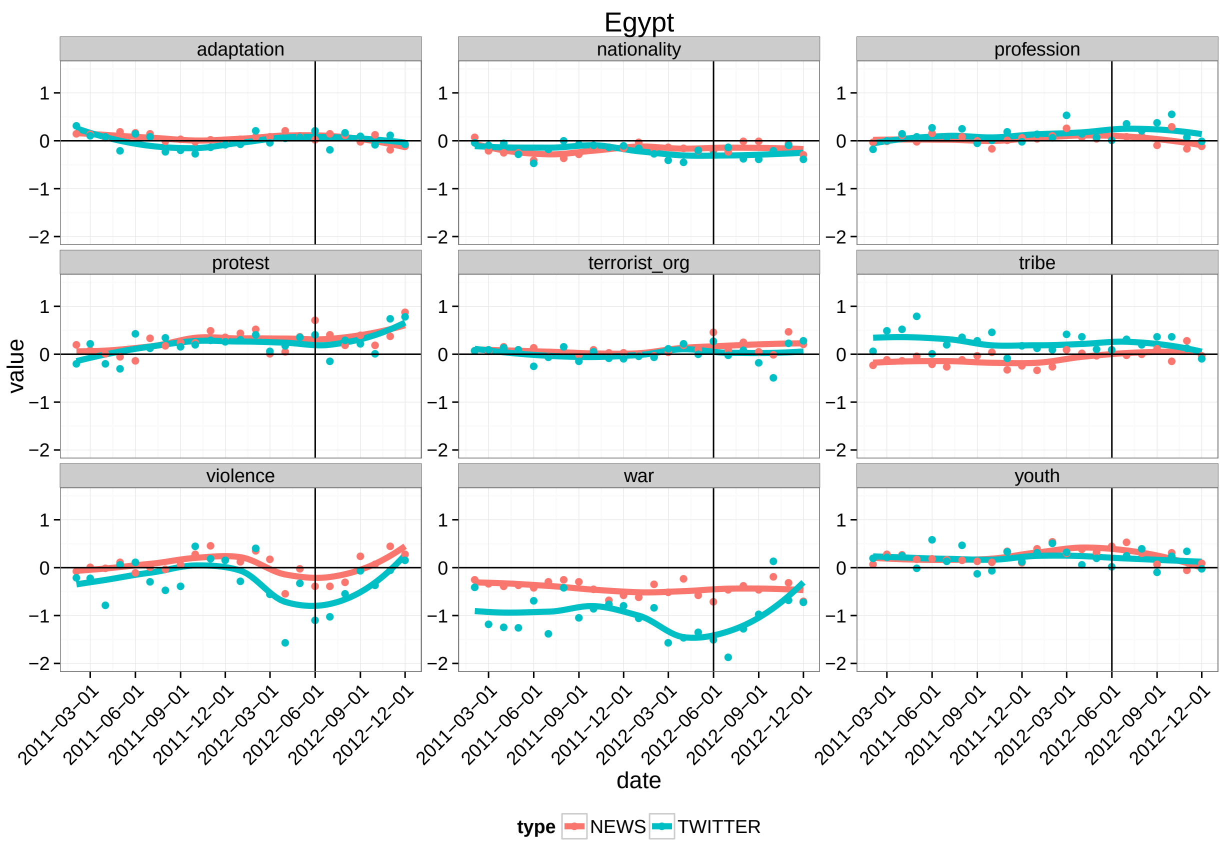

Figure 3: Egypt: Black line is when Morsi is elected

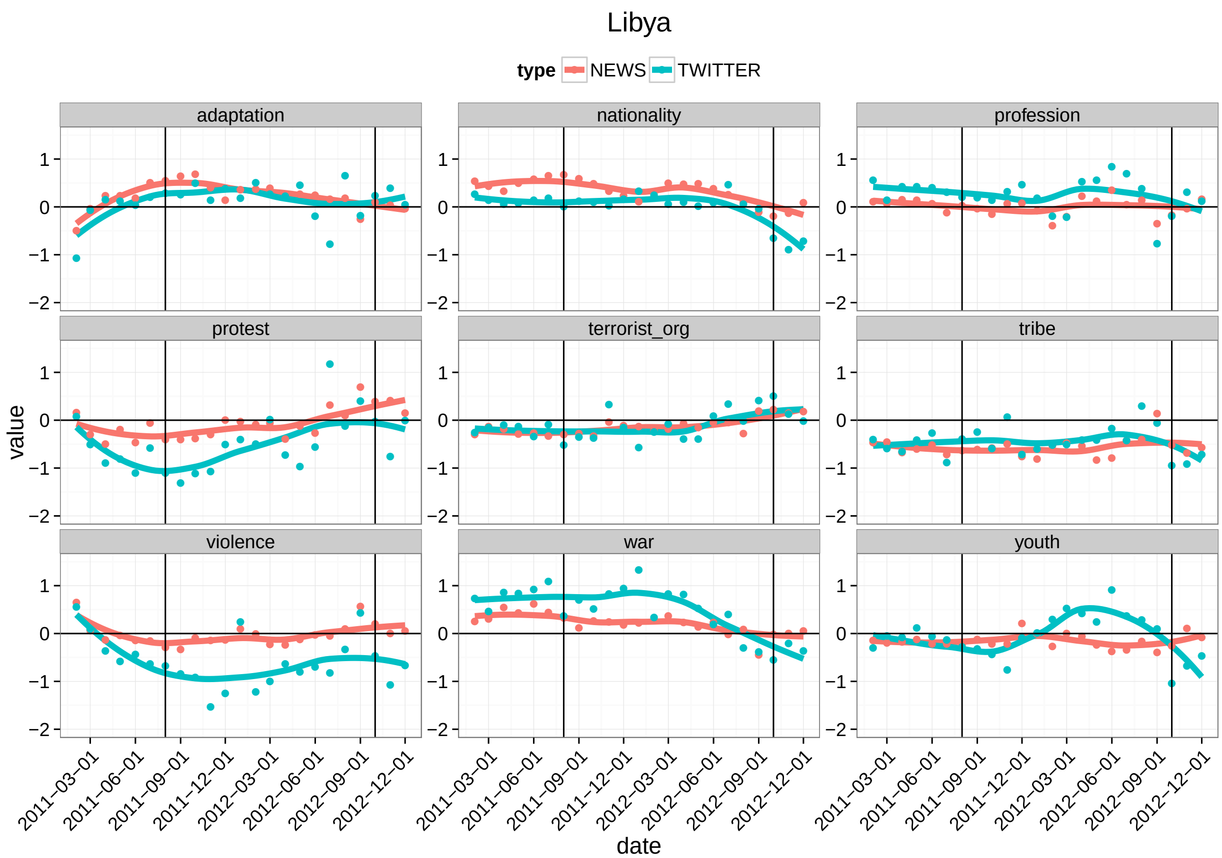

Figure 4: Libya: Black lines are death of Gaddafi and Benghazi

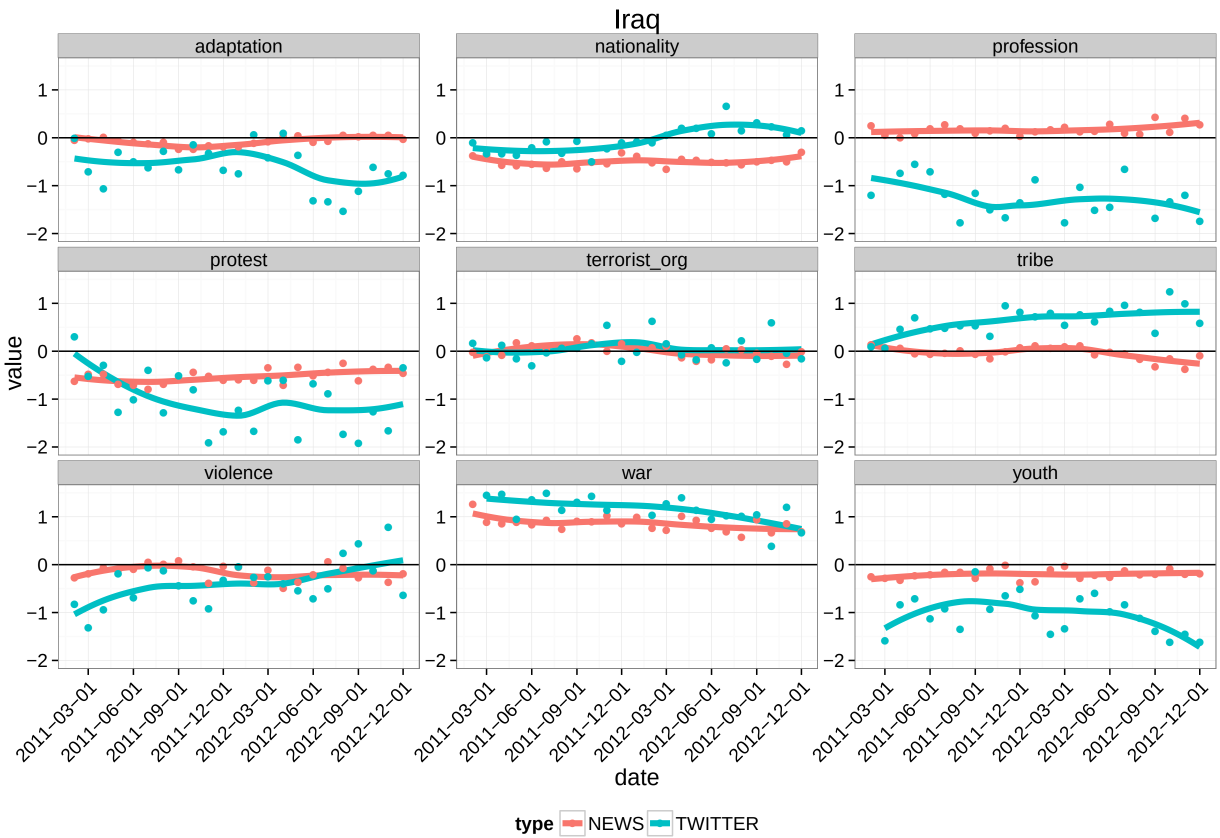

Figure 5: Iraq

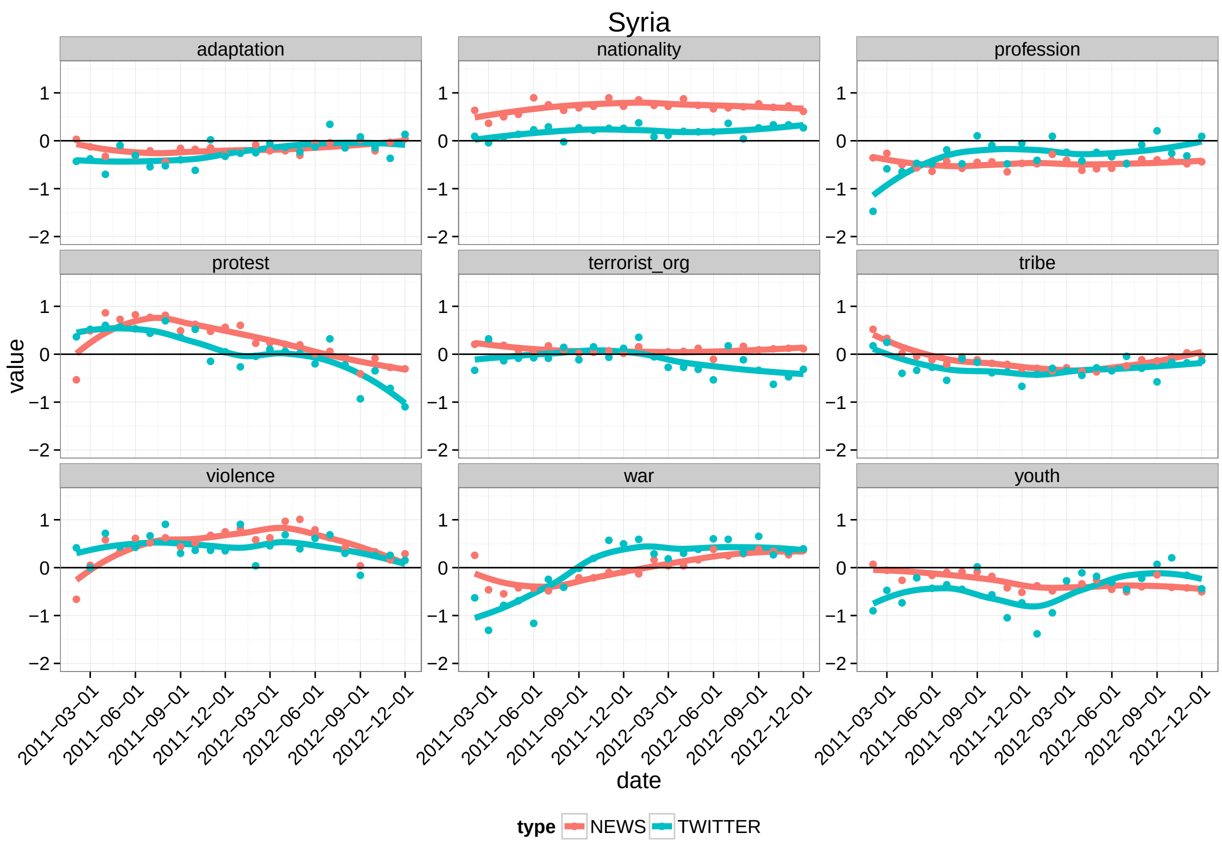

Figure 6: Syria

and Pfeffer (2014); Tufekci (2014). While these issues are not to be ignored, the present work utilizes cautious, nuanced analysis techniques which account for, or at least admit, these possible biases.

References

Battera, F. (2014). Perspectives for change in tunisia, egypt and syria: the military factor and implications of previous authoritarian regimes. *Contemporary Arab Affairs*, 7(4):544–564.

Baym, N. K. (2010). *Personal Connections in the Digital Age*. Polity.

Bellin, E. (2012). Reconsidering the robustness of authoritarianism in the middle east: Lessons from the arab spring. *Comparative Politics*, 44(2):127–149.

Borge-Holthoefer, J., Magdy, W., Darwish, K., and Weber, I. (2014). Content and network dynamics behind egyptian political polarization on twitter. *arXiv:1410.3097 [physics]*. arXiv: 1410.3097.

Bruns, A., Highfield, T., and Burgess, J. (2013). The arab spring and social media audiences english and arabic twitter users and their networks. *American Behavioral Scientist*, 57(7):871–898.

Comunello, F. and Anzera, G. (2012). Will the revolution be tweeted? a conceptual framework for understanding the social media and the arab spring. *Islam and Christian–Muslim Relations*, 23(4):453–470.

Dewey, T., Kaden, J., Marks, M., Matsushima, S., and Zhu, B. (2012). The impact of social media on social unrest in the arab spring. *International Policy Program*.

Eisenstein, J., O'Connor, B., Smith, N. A., and Xing, E. P. (2014). Diffusion of Lexical Change in Social Media. *PloS one*, 9(11):e113114.

Gallé, M., Renders, J.-M., and Karstens, E. (2013). Who broke the news?: an analysis on first reports of news events. In *Proceedings of the 22nd international conference on World Wide Web companion*, pages 855–862. International World Wide Web Conferences Steering Committee.

Gelvin, J. (2015). *The Arab uprisings: what everyone needs to know*. Oxford University Press, 2nd edition.

Goldstone, J. A. (2011). Cross-class coalitions and the making of the arab revolts of 2011. *Swiss Political Science Review*, 17(4):457–462.

Goldstone, J. A. (2013). Bringing regimes back in–explaining success and failure in the middle east revolts of 2011. *Available at SSRN 2283655*.

Hussain, M. M. and Howard, P. N. (2013). What best explains successful protest cascades? ICTs and the fuzzy causes of the arab spring. *International Studies Review*, 15(1):48–66.

Joseph, K., Carley, K. M., Filonuk, D., Morgan, G. P., and Pfeffer, J. (2014a). Arab spring: from newspaper data to forecasting. *Social Network Analysis and Mining*, 4(1):1–17.

Joseph, K., Landwehr, P. L., and Carley, K. M. (2014b). An approach to selecting keywords to track on twitter during a disaster. In *Proceedings of the 11th International Conference on Information Systems for Crisis Response and Management*, State College, PA.

Lotan, G., Graeff, E., Ananny, M., Gaffney, D., Pearce, I., and Boyd, D. (2011). The revolutions were tweeted: Information flows during the 2011 tunisian and egyptian revolutions. *International Journal of Communication*, 5:1375–1405.

Morstatter, F., Pfeffer, J., Liu, H., and Carley, K. M. (2013). Is the sample good enough? comparing data from twitter's streaming api with twitter's firehose. *Proceedings of ICWSM.*

Pfeffer, J. and Carley, K. M. (2012). Rapid modeling and analyzing networks extracted from pre-structured news articles. *Computational and Mathematical Organization Theory*, 18(3):280–299.

Ruths, D. and Pfeffer, J. (2014). Social media for large studies of behavior. *Science*, 346(6213):1063–1064.

Starbird, K. and Palen, L. (2012). (how) will the revolution be retweeted?: information diffusion and the 2011 egyptian uprising. In *Proceedings of the acm 2012 conference on computer supported cooperative work*, pages 7–16. ACM.

Tufekci, Z. (2014). Big questions for social media big data: Representativeness, validity and other methodological pitfalls. In *ICWSM '14: Proceedings of the 8th International AAAI Conference on Weblogs and Social Media.*

Tufekci, Z. and Wilson, C. (2012). Social media and the decision to participate in political protest: Observations from tahrir square. *Journal of Communication*, 62(2):363–379.

Wolfsfeld, G., Segev, E., and Sheafer, T. (2013). Social media and the arab spring politics comes first. *The International Journal of Press/Politics*, 18(2):115–137.