Using Social Media for Research, Monitoring and Evaluation in the MENA Region: World Food Programme Case Study

Middle East and North Africa Division

Department for International Development

December 2016

Contents

Tables

Figures

Executive Summary

- This report presents an evaluation of the utility of social media for conducting research, monitoring and evaluation in the MENA region, using a case study approach. The Twitter based reaction to the reported cancellation/reduction of food provision by World Food Programme (WFP) for Syrian Refuges was taken as the case study. The selection of this case study was based on convenience, given the timing of the announcement and the high-profile nature of the WFP. This report makes no comment on or about the working or policies of the WFP. This report builds upon the DFID Practice Note: Using Social Media Data in International Development Research, Monitoring and Evaluation by providing an extended demonstration of the methods covered in this note.

- Current estimates put the social media world population at near 2.5 billion non-unique users, with Facebook, Twitter, Instagram and YouTube making up a sizable portion of this total. In the Middle East and North Africa region, on average 85% of the population use Facebook and 41% of population use Twitter (Dennis et al. 2015)

- Access to 'big data' sources of this type provides unprecedented opportunities for researchers to gain insights into the social world in near-real-time, often at low cost as compared to conventional methods, such as social surveys and interviews.

- The following questions were addressed via an analysis of the tweet dataset: Do the spikes in frequency of tweets collected with the key term set correspond to WFP press releases and Google Trends? Which social media users are most influential in spreading information about the WFP cuts? Can we identify clusters of geo-located data in different regions? How does sentiment change over time? Are social media posts containing negative sentiment in relation to the WFP cuts more likely to spread compared to posts containing positive sentiment?

- Frequency Analysis: The application of frequency analysis to the case study showed that information regarding ebb and flow of the *popularity* or *awareness* of the cuts to the WFP could be obtained from Twitter. Spikes in popularity or awareness on Twitter related to key reports in the media relating to the cuts. These spikes were subjected to further analysis using topic classification.

- Topic Classification: Simple topic classification was performed on the spikes identified in the frequency analysis. Wordclouds are simple but effective methods for viewing the high-level discussion on social media. Words are sized in the wordcloud based on their frequency – the bigger the word, the more frequently it is used. The accounts @60Minutes and @bbcworld were dominant in the visualizations, confirming the broadcast reaction on Twitter. To further refine the analysis these Twitter accounts were visualised in retweet networks.

- Social Network Analysis: Social Network graphs were used to visualise interactions between accounts posting about cuts to the WFP on Twitter. The size of nodes (accounts) in the graphs corresponded to amount of interactions with other nodes (accounts), indicating their relative importance in propagating information about the cuts to the WFP on Twitter. The graphs also identified nodes (accounts) that acted as bridges for information to travel from one Twitter community to another (the removal of these accounts would have the impact of stemming the flow of information between these communities).

- Geospatial Analysis: Approximately 1% of all tweets are geotagged globally. Only a very small proportion of the original tweets (40, or 0.48%) generated by our WFP key word search query included GPS coordinates. However, it is important to note that retweets, representing 65% of the dataset for this study, do not contain GPS information by default. Given the low number of geotagged tweets in the dataset, we were unable to make any conclusions on the location of posts in relation to the Twitter reaction to the cuts to the WFP. This limitation is likely to present itself in other Twitter datasets on similar topics.

- Sentiment Analysis: Sentiment analysis provided insights into the emotive response of Twitter users to the cuts to the WFP. The WFP Twitter dataset showed that there were more extreme negative views than extreme positive views, and that peaks in negative views corresponded to the press releases during the study window. However, care should be taken during interpretation as sentiment analysis is prone to producing false results due to sarcasm, irony and the use of multiple emotions towards different entities.

- Information Propagation Modelling: Information propagation statistical modelling was used to successfully estimate the effect of various factors (such as sentiment and type of Twitter user) on the *size* (how many retweets) and *survival* (how long it is retweeted for) of retweets about the cuts to the WFP. Information propagation modelling can be used to identify the inhibiting and enabling factors of the spread of Twitter messages during and following an event of interest. A focus on these factors can assist decision making with respect to communication strategy around an event. For example, if it is identified that certain types of accounts are important to the spread of information (e.g. local MENA agents) then it would be possible to seek endorsement of key MENAD messages by these agents, via retweeting.

- Ethics has emerged as a contentious issue in the use of social media data in social and government research. Key issues include informed consent, confidentiality and potential harm to social media users. The emerging consensus from government and academia, is that while it is acceptable to collect and analyse social media posts without consent from users, informed consent should be sought (opt-in or opt-out) if posts are to be directly quoted in research reports.

This material has been funded by UK aid from the UK government; however the views expressed do not necessarily reflect the UK government's official policies.

BACKGROUND TO THE STUDY

Current estimates put the social media world population at near 2.5 billion non-unique users, with Facebook, Twitter, Instagram and YouTube making up a sizable portion of this total. In the Middle East and North Africa (MENA) region, on average 85% of the population use Facebook and 41% of population use Twitter (Dennis et al. 2015). Access to 'big data' sources of this type provides unprecedented opportunities for researchers to gain insights into the social world in near-real-time, often at low cost as compared to conventional methods, such as social surveys and interviews. MENA is a focus of a considerable amount of the Department for International Development's (DFID) effort. To help ensure that DFID's programme response is appropriate the Middle East and North Africa's Division (MENAD) continually seeks out opportunities to better inform itself about the situation within the region. Unfortunately, across MENA there is real scarcity of timely information and accessible data sources. The capacity of its governments to produce official statistics is limited[1] and conditions within the region make social and market research often very difficult to undertake. Some MENA countries do produce some economic and vital statistics as well as an occasional household survey or census. However aside from these official statistics there is a general lack in the region of any broader public perception or public opinion survey work. This project sought to understand how social media data can be used to provide a more informed picture of public social reactions to events in the MENA region. Via an extended case study this project aimed to examine the reaction on social media in the MENA region and beyond to the announcement of the cancellation/reduction of food provision by the World Food Programme (WFP) for Syrian Refuges. The project objectives were to identify the extent to which social media analysis can contribute to:

- The Middle East and North Africa Division's understanding of the public mood in the region and their reaction to events;
- Informing the Middle East and North Africa Division's programme response to events in the region to help its programmes flex to meet changing situations; and
- Extending the possible evidence base available to the Middle East and North Africa Division so that it is better informed.

This report provides a social media analysis of the global reaction to the cuts to the World Food Programme, using the freely available COSMOS software[2]. Multiple social media analysis tools were applied to the dataset (including frequency, topic, social network, sentiment and geospatial analysis) and an evaluation of their potential use within MENAD is presented. The results from information propagation modelling are presented to show how statistical analysis can be applied to social media datasets to inform decision making relating to shaping the flow of information following events. At the end of the report an evaluation of social media data for research on the MENA region is presented, followed by a thorough review of the ethical dimensions of using these data and a requirements outline for a MENAD social media analytics solution.

[1] See World Bank indicators on statistical capacity
[2] www.socialdatalab.net/software

Methodological Context

Currently, some of the social interactions produced on social media platforms are free to collect for research purposes. In particular, data from Twitter is free of charge up to a limit of between 3-5 million interactions per day. Access to 'big data' sources of this type provides unprecedented opportunities for researchers to gain insights into the social world in near-real-time, often at low cost as compared to conventional methods, such as social surveys and interviews. Within government social research the advent of 'big data' and all the problems associated with using it monitoring and evaluation are feeding methodological innovation. The key challenges of harnessing these data (known as the 6Vs: volume, velocity, veracity, variety, virtue and value) are not easily overcome (see page 34). Indeed, the term 'big social data' in itself is less than precise as it can refer to a vast range of transactional and open communications sources. Despite these challenges, recent advances have been made in the ability of social and government researchers to marshal these data to address pressing research questions. Notably, transactional data generated during internet searches has been used to track the spread of flu in the US (Ginsberg et al. 2009) and to build psychological constructs of nations linked to GDP (Noguchi et al. 2014). Open source social media communications have been used to investigate the spread of hate speech following terrorist attacks (Williams & Burnap 2015) and to estimate offline crime patterns (Williams et al. 2016). Indeed, increasing attention is being given to the role of big data in understanding collective behaviour in the 'offline' world (see Moat et al. 2014).

Conducting government social research using these new forms of data requires a reconsideration of what methods, approaches and research designs are appropriate. For academic and government social research, the sheer volume of data and its constant, locomotive, longitudinal nature provides an opportunity to take the 'pulse of nations' every minute of the day rather than relying on cross-sectional, punctiform and time-consuming terrestrial methods such as social surveys. But despite the reported success of studies claiming to have used social media data to predict 'real-world' outcomes, these new forms of data are not without their limitations. For example, unlike conventional research methods, it is difficult to ascertain with exact precision the demographic characteristics of all social media users (in particular Twitter users). Therefore, it is unlikely that social media data could act as a *surrogate* to conventional methods, such as surveys, that are designed to generate more complete and reliable datasets on populations under study. Therefore, researchers prefer to consider how these new data of might *augment* traditional social research designs. The addition of social media data to conventional designs provides an opportunity to include a near-real-time dimension to social research that captures data from populations that are often hard-to-reach with conventional methods. Table 1 provides a diagrammatic representation of the distinctive contribution of social media data to social research strategy and design. Unlike conventional methods, such as interviews and surveys, social media provides insights into both the extensive (large samples) and locomotive (in process, as opposed to snap-shot) dimensions of behaviours and opinions. This new ability allows social researchers to gain insights into key events as they unfold over time. For example, the ways in which different populations produce content that relates to major incidents of civil unrest, foreign policy interventions, global sporting events and disasters and states of emergency.

		Research Design/Data	
		Locomotive	Punctiform
Research Strategy	Intensive	E.g. Participant Observation	E.g. Qualitative interviewing
	Extensive	Social Media Analysis	E.g. Surveys and experiments

Table 1: The distinctiveness of social media analysis

Data Collection

The data collection period for this extended case study spanned 13 months (01/11/14 to 01/12/15) during which announcements were made of the cuts to the World Food Programme. The study period was initially guided by a series of WFP press releases (see Table 2).

WFP Forced To Make Deeper Cuts In Food Assistance For Syrian Refugees Due To Lack Of Funding	1st July 2015
WFP Forced To Suspend Syrian Refugee Food Assistance, Warns Of Terrible Impact As Winter Nears	1st December 2014
Funding Shortfall Forces WFP To Announce Cutbacks To Syrian Food Assistance Operation	18th September 2014

Table 2: Press releases guiding period of study

In addition, an examination of Web search trends using the ''World Food Programme'' keyword set to query the Google Trends service indicated that an issue attention cycle around this event (the duration within which public attention to this event rises and falls away) spanned roughly 13 months from Nov 2014 (see Figure 1). This became the analysis sampling time frame for this study.

Figure 1: Google trends search query for "World Food Programme"

Data were purchased from the Twitter social media network via their data-reseller GNIP. The Twitter social media network differs from others such as Facebook, in that it is largely used as a public medium and the data are easily accessible by researchers. Twitter also has an open friendship network (non- reciprocal linking between users means that the followed are not required to follow their followers) resulting in a digital public space that promotes the free exchange of opinions and ideas. As a result, Twitter has become the primary space for online citizens to publicly express their reaction to events of international significance. A hashtag convention has emerged amongst Twitter users that allows tweets to be tagged to a topic which is searchable. The term 'trending' is used to describe hashtags that become popular within the tweet-stream, indicating a peak or pulse in discussion usually surrounding an event. Data for this project were generated via a key term query related to the WFP submitted via the GNIP Historical Power Track service (https://gnip.com/historical/):

English: (World Food Programme OR WFP OR Syrian refugee OR Syrian Refugees OR Syria) AND Food OR cuts OR reduction OR reduce OR suspend

Arabic:

انهيار غذائي (roughly translated as nutritional collapse)
أزمة الغذاء (food crisis)
العالمي الأغذية برنامج (World Food Programme)

Key term queries produce robust social media data samples due to the interactive nature of keywords and hashtags, where followers of events on Twitter actively seek out the most popular, or trending topics/hashtags in order to identify relevant information and subsequently add to the flow by replicating the keyword or hashtag in their posts. This selection procedure generates a census of tweets containing the most common keywords around a topic, and hence a large sample of all tweets about the event in question. The key term query generated 23,693 tweets. The sample was subject to data pre-processing and recoding prior to analysis. Given our sampling technique ensured the collection of all tweets containing the most popular terms surrounding the event for 13 months, we are confident that the sample is representative of *non-trivial* information flows on Twitter surrounding this event.

COSMOS OPEN DATA ANALYTICS PLATFORM

A series of analyses were conducted using the Cardiff Online Social Media Observatory (COSMOS) Social Media Analytics Platform. COSMOS is an open data (inc. social media) collection, analysis and fusion platform. It programmatically collects data from a number of sources using publicly accessible Application Programme Interfaces (APIs).

COSMOS has been collecting a random 1% sample from the Twitter API (commonly referred to as the 'spritzer') since 2012 and the database currently holds over 7 billion tweets. It is also possible to collect 10% ('gardenhose'), 20% (decahose) or the full 100% of tweets ('firehose') with permission from Twitter. However, the data storage and retrieval requirements for these larger data streams are impractical for many social research and government projects.

COSMOS also has a persistent connection to the Office for National Statistics API allowing access to all national curated datasets with the capacity for linking (data fusion) these with social media data geographically. Data import also allows for the loading of CSV files and RSS feeds, meaning almost any open quantitative and qualitative data source from the MENA region can be subject to the analytical tools within COSMOS.

COSMOS provides a single interface to a number of tools, with no data collection overhead and automated translation of input files from one format to another. Below we demonstrate a range of COSMOS tools by interrogating the WFP social media dataset collected for this project. To direct the analysis we posed the following questions:

- Do the spikes in frequency of tweets collected with the key term set correspond to WFP press releases and Google Trends?
- Which social media users are most influential in spreading information about the WFP cuts?
- Can we identify clusters of geo-located data in different regions?
- How does sentiment change over time?
- Are social media posts containing negative sentiment in relation to the WFP cuts more likely to spread compared to posts containing positive sentiment?

Each of these questions can be interrogated by using the various big data inspection tools within COSMOS. With Twitter data it is currently possible to conduct, at the individual tweet level, language, geospatial, sentiment (positive, neutral, negative), topic, gender, age, and occupation/social class classification. At a corpus (dataset) level it is possible to conduct keyword/hashtag/tweet frequency analysis, topic frequency (often visualised via a wordcloud or cluster) social network analysis, and information flow analysis (indicating what features lead to virality). These forms of analysis can be combined into workflows to create models that allow for the visualisation and prediction of an array of phenomena.

USING COSMOS TOOLS ON WFP DATA

Frequency Analysis and Subsample Selection

COSMOS allows for tweet corpora to be analysed by the occurrence of specified keywords (or tweet frequency if no keywords are specified) over time. This allows a researcher to visually identify points of high and low activity in relation to an event or topic. Figure 2 shows tweets between 01/11/14 to 01/12/15 obtained from the GNIP historic power tracker using the keyword set described previously. The panels (from top to bottom) display day, hour and minute frequency of the Twitter corpora. Several peaks in frequency can be observed in the time period under study, the first of which corresponds with the Google Trends visualisation in Figure 1 and the press release on cuts to the WFP on 1st Dec 2014. The second spike, on 15th February 2015, relates to a story about footballer Zlatan Ibrahimović tattooing 50 names of starving people in support of the UN WFP.

In Figure 3 the first peak in frequency is selected (highlighted in orange), revealing the hour and minute distribution, allowing for a more fine-grained visualisation. This selection also produces a summary frequency of tweets in the spike in the lower left hand corner of the display. The first spike on Monday 1st December, 2014 consists of 3,658 tweets. Figure 4 shows this subsample re-visualised in tabular form, displaying individual tweet properties, including Timestamp, Text (including if the tweet is a Retweet or original tweet), estimated Gender of Tweeter, Account Name, GPS coordinates (if available), Sentiment Scores based on the posted text (positive and negative), and Language of the tweet text. From an initial inspection of the Text field, it is clear that at least half of the tweets in view are retweets relating to the 60 Minutes US TV programme that is covering the cuts to the WFP. For a fuller inspection of the content of tweets, we can use topic classification and the COSMOS wordcloud tool (see page 18). We can also examine the other tweet properties using COSMOS tools.

Figure 2: Frequency analysis

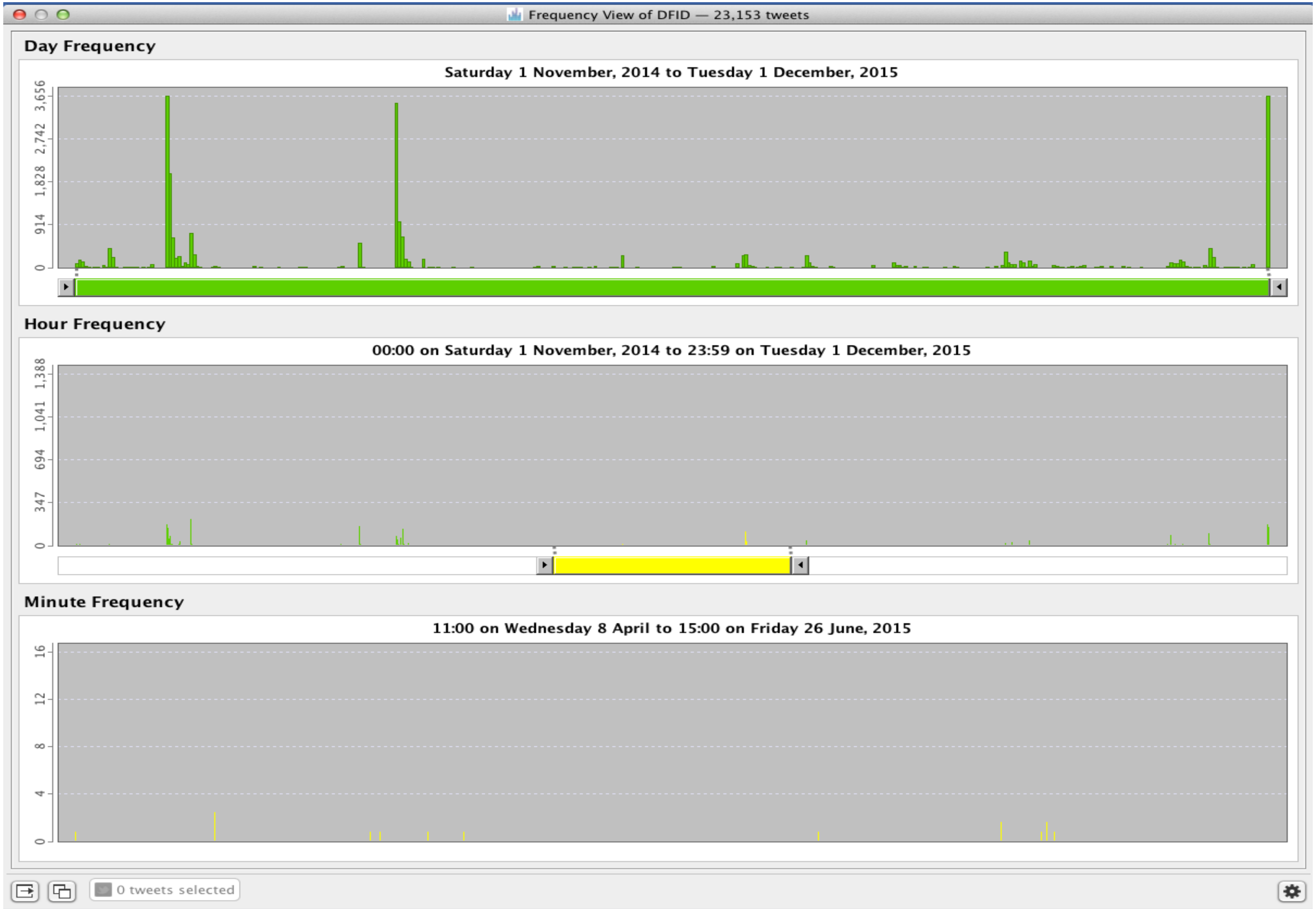


Figure 3: Subsample Selection

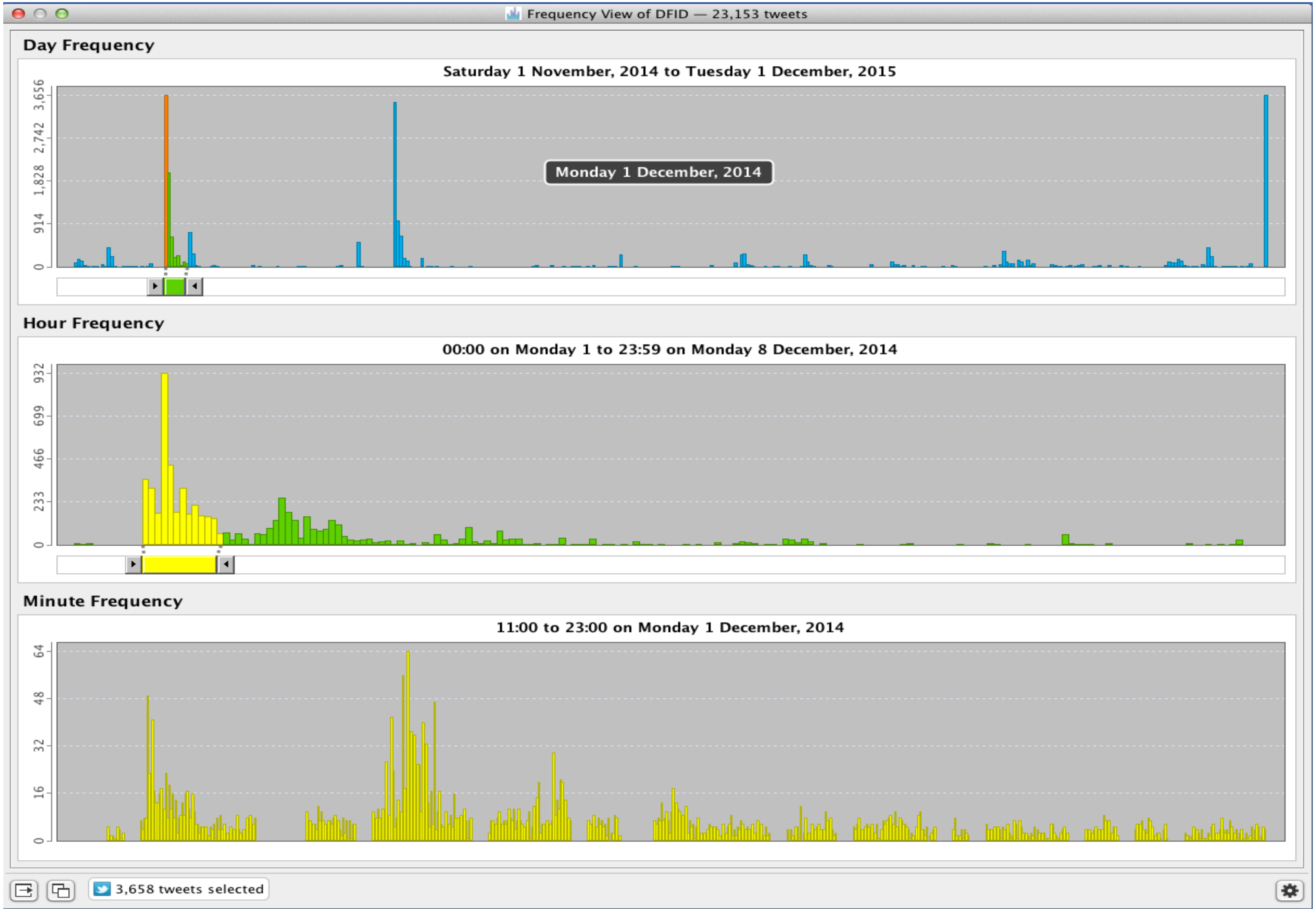


Figure 4: Tweet sub-sample from first spike on 1st Dec, 2014

Table View of Spike 1 — 3,658 tweets

Timestamp	Tweet Text	Gender	Account	Latitude	Longitude	+ve Senti...	-ve Senti...	Language
2014-12-01 00:5...	RT @FooteWashington: World Food Programme faces unusual quadruple food emergency: ...	FEMALE	najatd	0.0	0.0	1.0	-2.0	EN
2014-12-01 00:5...	Really appreciate 60Minutes piece tonight on food crisis for Syrian refugees and work of ...	FEMALE	tracym...	0.0	0.0	3.0	-3.0	EN
2014-12-01 00:5...	@WFP is the World Food Program they feed refugees in Syria. There are millions folks and t...	MALE	oldlady...	42.9990...	-89.45...	1.0	-2.0	EN
2014-12-01 00:5...	Watching report on @WFP in Syria/jordan on @60Minutes then my husband switches over ...	FEMALE	docabw	0.0	0.0	1.0	-1.0	EN
2014-12-01 00:5...	BBC News World Food Programme: 'Without enough food people can't fight Ebola' http://...	FEMALE	abronx...	0.0	0.0	1.0	-1.0	EN
2014-12-01 00:5...	The diff btwn cooking your own food and eating something off a truck is dignity. 60Min...	FEMALE	noranoh8	0.0	0.0	2.0	-2.0	EN
2014-12-01 00:5...	Give to WFP The World Food Programme SavingALife savinghope	FEMALE	lisamatto	0.0	0.0	1.0	-1.0	EN
2014-12-01 00:5...	RT @LisaMatto: Give to WFP The World Food Programme SavingALife savinghope	MALE	reswift...	0.0	0.0	1.0	-1.0	EN
2014-12-01 00:5...	I just donated to the World Food Programme to end hunger. You can help too. Donate no...	FEMALE	quellel...	0.0	0.0	1.0	-2.0	EN
2014-12-01 00:5...	The U.S. government pays the biggest part of the bill as the World Food Programme feeds ...	MALE	carlos...	0.0	0.0	1.0	-1.0	EN
2014-12-01 00:5...	RT @WFP: WFP is struggling to meet food needs of 6M displaced Syrians help us help the...	UNKNO...	celtic a...	0.0	0.0	1.0	-2.0	EN
2014-12-01 00:5...	I just donated to the World Food Programme to end hunger. You can help too. Donate no...	UNKNO...	camp629	0.0	0.0	1.0	-2.0	EN
2014-12-01 01:0...	I donated to Syria Crisis: Help deliver food to families amp children this winter. http://t....	UNKNO...	c bosher	0.0	0.0	1.0	-3.0	EN
2014-12-01 01:0...	60minutes on world food programme in Syria is heartbreaking. Hunger not just food acce...	MALE	eipark	0.0	0.0	1.0	-4.0	EN
2014-12-01 01:0...	I just donated to the World Food Programme to end hunger. You can help too. Donate no...	FEMALE	tammy...	0.0	0.0	1.0	-2.0	EN
2014-12-01 01:0...	This story is heart wrenching I will add this organization to those I regularly donate to h...	FEMALE	shireen...	0.0	0.0	1.0	-1.0	EN
2014-12-01 01:0...	Sad to watch 100's in Syria starving given that 40 of food produced in U.S. goes uneaten ...	MALE	upfront...	0.0	0.0	1.0	-4.0	EN
2014-12-01 01:0...	World Food Programme in Kenya Cuts Refugee Food Rations by Half BloombergNews htt...	FEMALE	abronx...	0.0	0.0	1.0	-2.0	EN
2014-12-01 02:1...	RT @CBSNews: @60Minutes Scott Pelley reports on efforts of the World Food Programme t...	MALE	alirawaf	0.0	0.0	2.0	-2.0	EN
2014-12-01 02:1...	RT @CBSNews: @60Minutes Scott Pelley reports on efforts of the World Food Programme t...	UNKNO...	gracely...	0.0	0.0	2.0	-2.0	EN
2014-12-01 02:1...	RT @A22523: 60 Minutes Program: War and hunger http://t.co/UBuSZT9u9y CBS WFP Sy...	UNKNO...	turtlew...	0.0	0.0	1.0	-3.0	EN
2014-12-01 02:1...	RT @CBSNews: @60Minutes Scott Pelley reports on efforts of the World Food Programme t...	MALE	irwincb...	0.0	0.0	2.0	-2.0	EN
2014-12-01 02:2...	RT @CBSNews: @60Minutes Scott Pelley reports on efforts of the World Food Programme t...	FEMALE	emac5	0.0	0.0	2.0	-2.0	EN
2014-12-01 02:2...	RT @CBSNews: @60Minutes Scott Pelley reports on efforts of the World Food Programme t...	UNKNO...	60minu...	0.0	0.0	2.0	-2.0	EN
2014-12-01 02:2...	RT @CBSNews: @60Minutes Scott Pelley reports on efforts of the World Food Programme t...	MALE	steveta...	0.0	0.0	2.0	-2.0	EN
2014-12-01 02:2...	UN World Food Programme Yemen: Food insecurity by governorate as of July 2014 31 J...	FEMALE	rinkui	0.0	0.0	1.0	-2.0	EN
2014-12-01 02:2...	RT @CBSNews: @60Minutes Scott Pelley reports on efforts of the World Food Programme t...	MALE	mhend...	0.0	0.0	2.0	-2.0	EN
2014-12-01 02:2...	RT @WFP: Make this the video you share today: A 360 view of a camp for refugees from S...	MALE	tipzip	0.0	0.0	1.0	-2.0	EN
2014-12-01 02:2...	RT @WFP: Today:@USAIDFFP gives WFP additional 125M and the ability to feed hungry fa...	MALE	tipzip	0.0	0.0	1.0	-2.0	EN
2014-12-01 02:2...	RT @CBSNews: @60Minutes Scott Pelley reports on efforts of the World Food Programme t...	MALE	mikieb...	0.0	0.0	2.0	-2.0	EN
2014-12-01 07:3...	RT @CBSNews: @60Minutes Scott Pelley reports on efforts of the World Food Programme t...	MALE	wfpgiulio	0.0	0.0	2.0	-2.0	EN
2014-12-01 07:3...	RT @CBSNews: @60Minutes Scott Pelley reports on efforts of the World Food Programme t...	UNKNO...	metalra...	0.0	0.0	2.0	-2.0	EN
2014-12-01 11:2...	RT @WFP: Syria: 'Are we willing to lose a generation of children to hunger ' via @60Minut...	FEMALE	dykhalil	0.0	0.0	1.0	-2.0	EN
2014-12-01 11:2...	RT @WFP: Syria: 'Are we willing to lose a generation of children to hunger ' via @60Minut...	FEMALE	lalitape...	0.0	0.0	1.0	-2.0	EN
2014-12-01 11:2...	RT @WFP: Syria: 'Are we willing to lose a generation of children to hunger ' via @60Minut...	UNKNO...	anexas8	0.0	0.0	1.0	-2.0	EN
2014-12-01 11:2...	RT @WFP: Syria: 'Are we willing to lose a generation of children to hunger ' via @60Minut...	MALE	janerid...	0.0	0.0	1.0	-2.0	EN
2014-12-01 11:2...	RT @WFP: Syria: 'Are we willing to lose a generation of children to hunger ' via @60Minut...	MALE	faresdjak	0.0	0.0	1.0	-2.0	EN
2014-12-01 11:2...	RT @WFP: Syria: 'Are we willing to lose a generation of children to hunger ' via @60Minut...	MALE	karand...	0.0	0.0	1.0	-2.0	EN
2014-12-01 11:2...	RT @WFP: Syria: 'Are we willing to lose a generation of children to hunger ' via @60Minut...	FEMALE	celeste...	0.0	0.0	1.0	-2.0	EN
2014-12-01 11:2...	RT @WFP: Syria: 'Are we willing to lose a generation of children to hunger ' via @60Minut...	UNKNO...	greenvi...	0.0	0.0	1.0	-2.0	EN
2014-12-01 11:2...	RT @WFP: Syria: 'Are we willing to lose a generation of children to hunger ' via @60Minut...	FEMALE	nilimaj...	0.0	0.0	1.0	-2.0	EN
2014-12-01 11:2...	RT @WFP: Syria: 'Are we willing to lose a generation of children to hunger ' via @60Minut...	UNKNO...	widesc...	0.0	0.0	1.0	-2.0	EN
2014-12-01 11:2...	Grim day for Syria refugees as @WFP forced to suspend aid as funds dry up: http://t.co/5...	MALE	gregory...	0.0	0.0	1.0	-3.0	EN
2014-12-01 11:2...	RT @WFP: Syria: 'Are we willing to lose a generation of children to hunger ' via @60Minut...	FEMALE	inekep...	0.0	0.0	1.0	-2.0	EN
2014-12-01 11:2...	RT @WFP: Syria: 'Are we willing to lose a generation of children to hunger ' via @60Minut...	UNKNO...	abuaay...	0.0	0.0	1.0	-2.0	EN
2014-12-01 11:2...	RT @WFP: Syria: 'Are we willing to lose a generation of children to hunger ' via @60Minut...	UNKNO...	tatom2k	0.0	0.0	1.0	-2.0	EN
2014-12-01 11:2...	RT @WFP: Syria: 'Are we willing to lose a generation of children to hunger ' via @60Minut...	MALE	reehlte	0.0	0.0	1.0	-2.0	EN
2014-12-01 11:2...	RT @WFP: Syria: 'Are we willing to lose a generation of children to hunger ' via @60Minut...	UNISEX	ng engi...	0.0	0.0	1.0	-2.0	EN
2014-12-01 11:2...	RT @WFP: Syria: 'Are we willing to lose a generation of children to hunger ' via @60Minut...	MALE	magnu...	0.0	0.0	1.0	-2.0	EN
2014-12-01 11:2...	RT @WFP: Syria: 'Are we willing to lose a generation of children to hunger ' via @60Minut...	UNKNO...	dipenduc	0.0	0.0	1.0	-2.0	EN
2014-12-01 11:2...	RT @WFP: Syria: 'Are we willing to lose a generation of children to hunger ' via @60Minut...	UNKNO...	yoursel...	0.0	0.0	1.0	-2.0	EN
2014-12-01 11:2...	RT @WFP: Syria: 'Are we willing to lose a generation of children to hunger ' via @60Minut...	UNKNO...	jhrconu	0.0	0.0	1.0	-2.0	EN
2014-12-01 11:2...	RT @WFP: Syria: 'Are we willing to lose a generation of children to hunger ' via @60Minut...	MALE	mecena...	0.0	0.0	1.0	-2.0	EN
2014-12-01 11:2...	RT @WFP: Syria: 'Are we willing to lose a generation of children to hunger ' via @60Minut...	MALE	martina...	0.0	0.0	1.0	-2.0	EN
2014-12-01 11:2...	RT @WFP: Syria: 'Are we willing to lose a generation of children to hunger ' via @60Minut...	UNKNO...	drelnora	0.0	0.0	1.0	-2.0	EN
2014-12-01 11:2...	1.7	UNKNO...	lbci news	0.0	0.0	1.0	-1.0	EN
2014-12-01 11:2...	RT @WFP: Syria: 'Are we willing to lose a generation of children to hunger ' via @60Minut...	UNKNO...	aster is	0.0	0.0	1.0	-2.0	EN
2014-12-01 11:2...	RT @WFP: Syria: 'Are we willing to lose a generation of children to hunger ' via @60Minut...	UNKNO...	hendsa	0.0	0.0	1.0	-2.0	EN

0 tweets selected

Highlights

- Frequency analysis can provide an additional layer of information regarding the *popularity* or *awareness* of an event or topic on social media
- Spikes in popularity or awareness may relate to events of public interest on the topic being monitored via social media
- Spikes should be used as a pointer for a data subset for fine-grained analysis – "what is going on in this spike?"
- COSMOS allows spikes to be easily pulled into a study dataset for further analysis

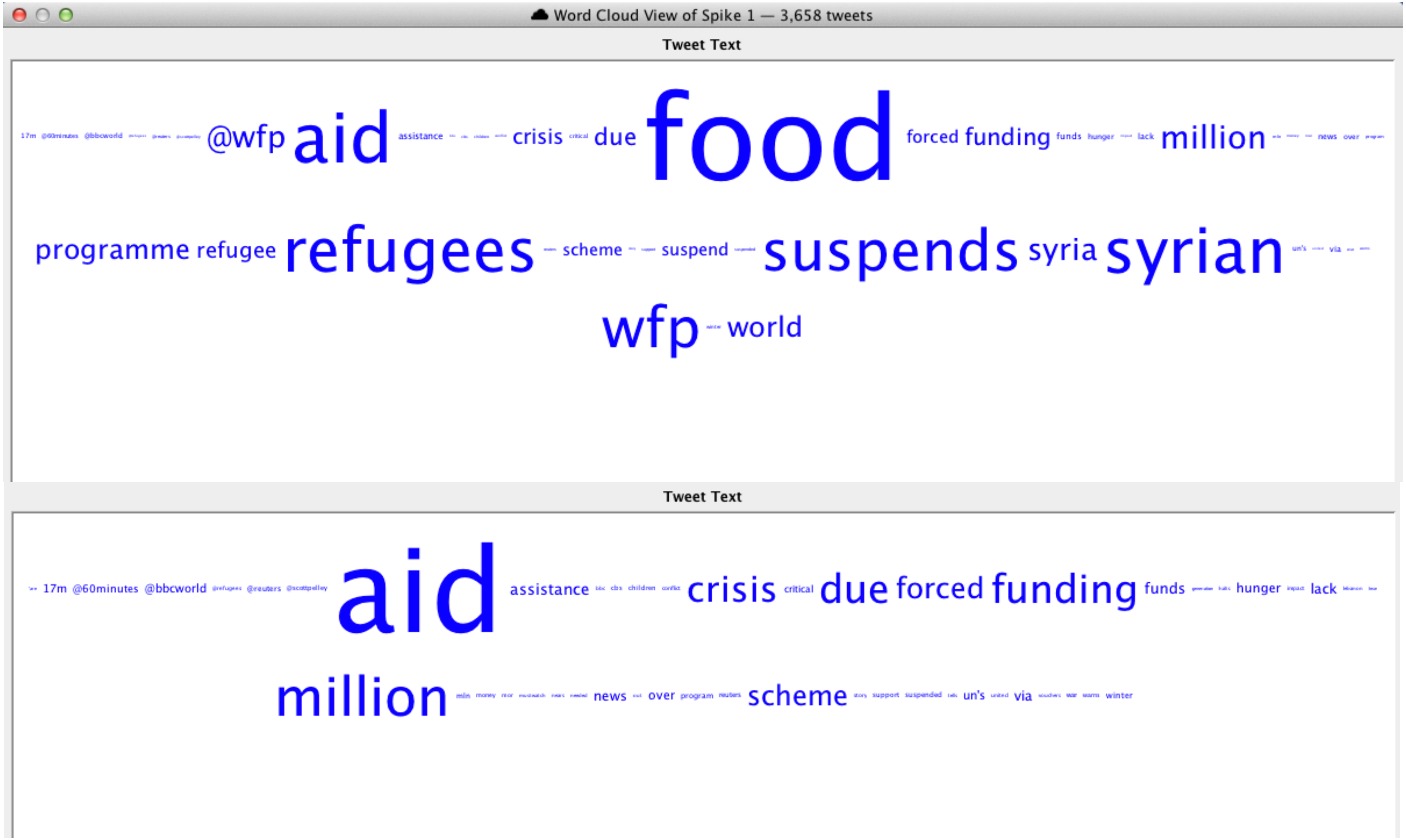
Topic Classification

Many social media analysis platforms provide topic classification tools. Topic classification can utilise an array of algorithms to reduce vast amounts of textual information into summaries of the main themes of the tweet corpus. In its most simple form, the frequency of terms is visualised in a wordcloud, with more frequent terms appearing larger in the visualisation. In platforms such as COSMOS, words can be removed from view with a right mouse click, allowing the researcher to refine the visualisation by removing irrelevant content.

Figure 5 shows wordcloud displays of the tweets in the first spike. The top wordcloud includes the key search terms, hence their large size. In the second wordcloud the keywords of been removed from view, to reveal underlying topics. The words ‘aid’, ‘million’, ‘crisis’, and ‘forced’ are dominant, and were not part of the keyword search query. These terms can be isolated using the COSMOS software for closer inspection, allowing the researcher to identify the context in which they are being used.

In both wordcloud visualisations, the @60Minutes Twitter handle is visible, confirming the broadcast and the reaction to it on Twitter was partly responsible for the first observed spike. The Twitter handles @bbcworld and @wfp are also visible, indicating these accounts have been retweeted and mentioned many times on this day. To confirm this, and the possible centrality of these accounts in the spreading of information on Twitter during this period, we conducted network analysis and present the results next.

Figure 5: Wordclouds of first spike (with and without search terms)



Highlights:

- Wordclouds are simple but effective methods for viewing the high-level discussion on social media
- Words are sized in the wordcloud based on their frequency – the bigger the word, the more frequently it is used
- Remember the word cloud will also include the terms used to collect data from social media so their size should be controlled for
- COSMOS allows the removal of data collection terms from the wordcloud using a right mouse click

Social Network Analysis

Social network analysis is useful for visualising and describing the interactions of social media users. The purpose of the network is to highlight prominent user accounts in the dataset. Each tweet is analysed to determine whether it contains any interaction with another Twitter user. This can include a *direct mention* or a *retweet,* with the latter representing information propagation (often but not always used to endorse content), and the former representing public messages between accounts. Both can be visualised in a network to quantify the number of interactions between users – indicating level of influence. Networks also help to identify important connecting users – those with high *Eigenvector centrality* who provide a link or *bridge* between communities in the graph.

Figures 6 and 7 show retweet networks[3] for the first and second spikes in the Twitter corpora. The first network confirms our assumption based on the wordcloud, that media accounts and the @wfp account were central to the spread of information around the topic on Dec 1st 2014. It is clear from the network that the @wfp account was the most retweeted on the day (334 retweets), indicated by its size and central position in the network. Media accounts, including @bbcworld, @reuters, dominate other parts of the network. Also central to the network, but less dominant in terms of retweets, are the UNHRC associated accounts @Refugees and @Melissarfleming. The second network shows the accounts @bbcsporf and @transfersite were primarily resendable for the spread of information on the story regarding the support of premier league football player to the WFP.

[3] Mention networks not provided as they show very similar patterns in the graph.

Figure 6: Retweet network analysis of the first spike

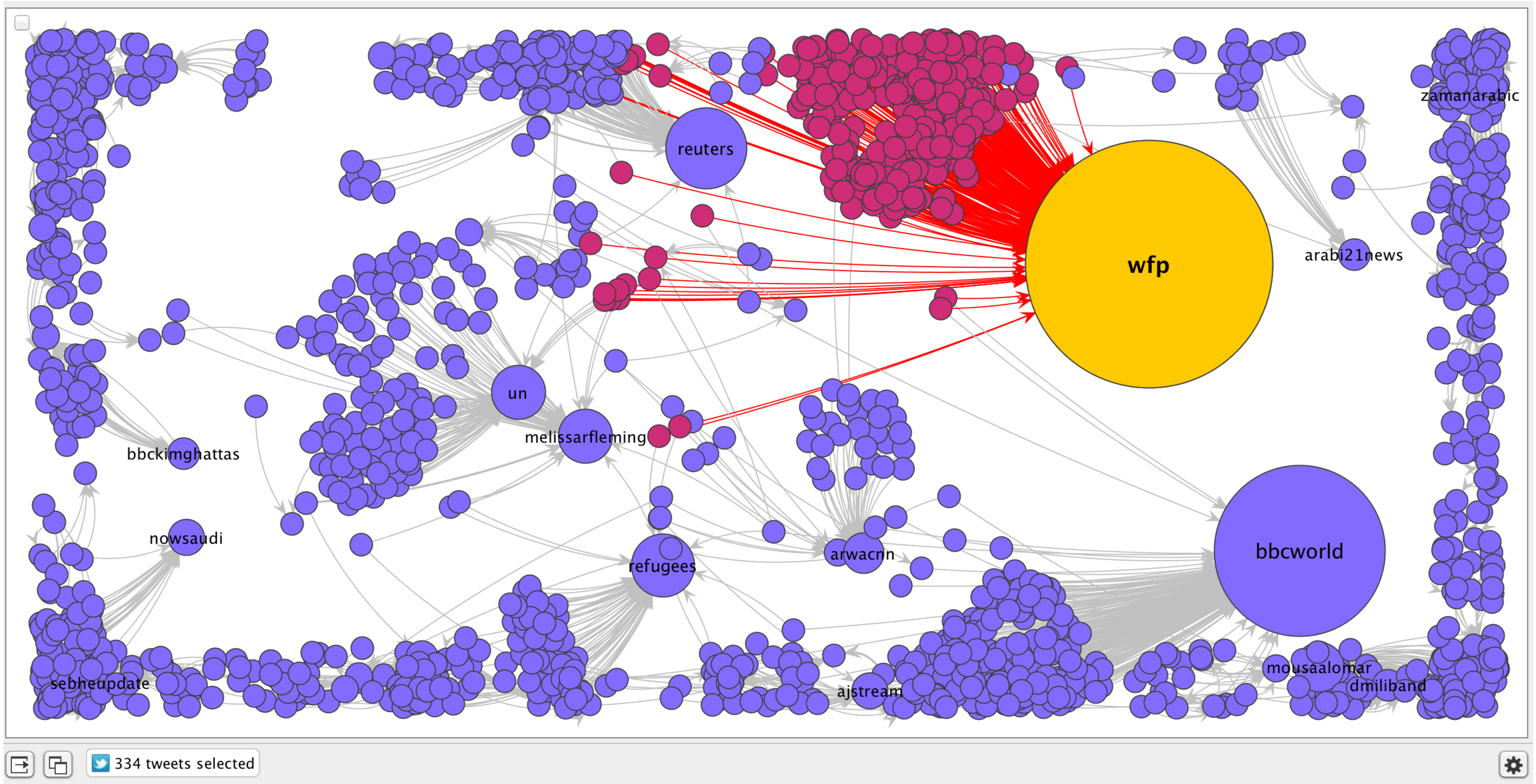
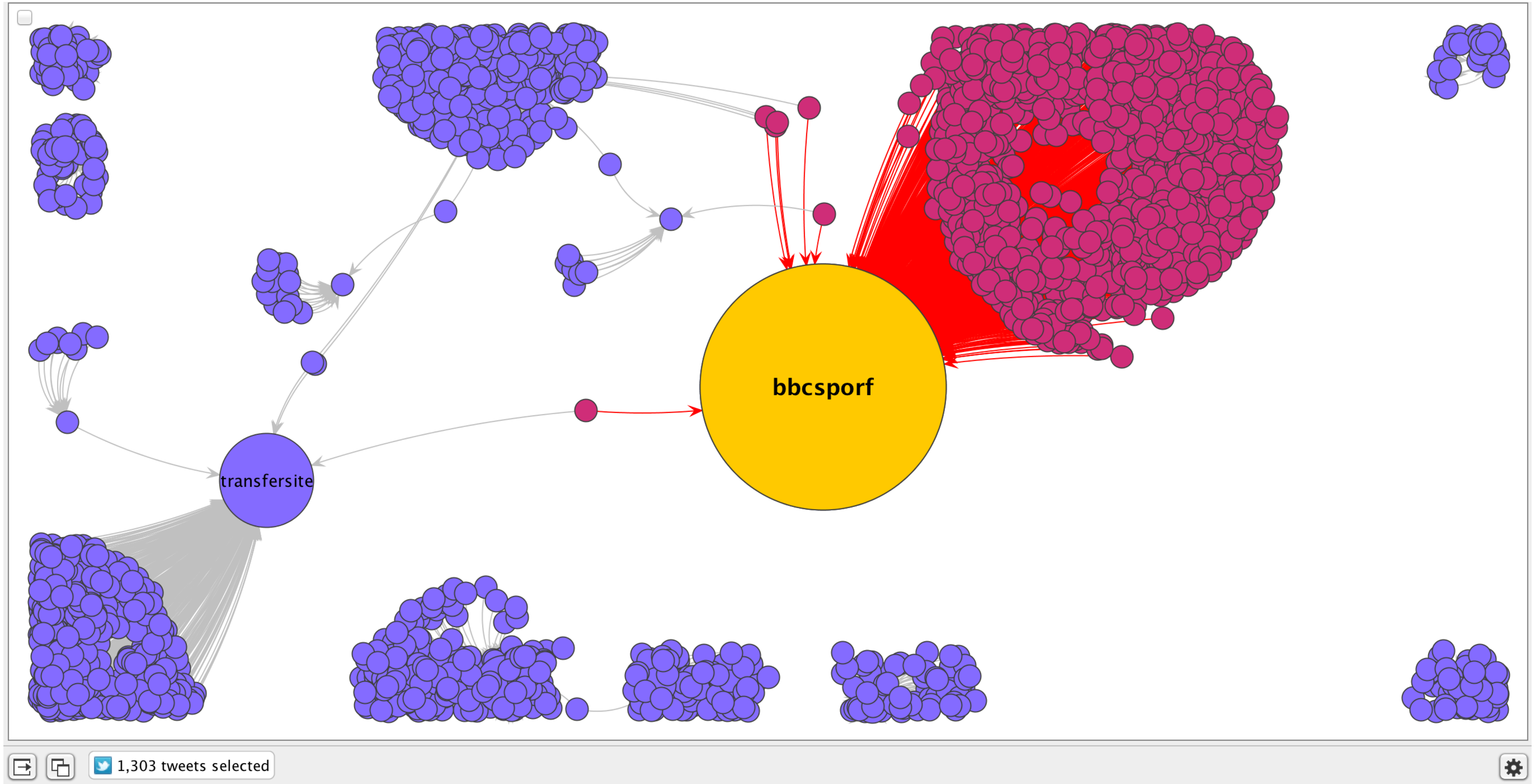


Figure 7: Retweet networks of second spike



Highlights:

- Social Network graphs visualise interactions between accounts on social media
- Interactions can include @mentions and @retweets, with the latter representing information propagation (often but not always used to endorse content), and the former representing public messages between accounts
- Figures 6 and 7 show larger and smaller circles or nodes. The larger the node, the more interactions it has had with the nodes attached to it by lines or edges.
- Figure 6 shows larger nodes for ‘wfp’ and ‘bbcworld’. These are account names. The nodes are larger because they have been retweeted more than others. This shows they are the main influencers in this social media dataset
- In figure 6 ‘wfp’ is yellow because it has been selected within COSMOS. This highlights the node and all the nodes (in red) that have interacted with it via a retweet
- Size of node is not the only important factor. The node to the immediate left of ‘bbcsporf’ in Figure 7 shows a ‘bridge’ between ‘bbcsporf’ and ‘transfersite’, who are another influencer in the network. Therefore, the ‘bridge’ is an account that links two active communities and could be key for spreading information between communities.

Geospatial Classification

From a social and government research perspective location data is incredibly valuable as it enables researchers to establish the geographic context in which the tweeter is immersed at the point of data creation. Having a geo-spatial point enables researchers to position tweets within existing geographies to which demographic and contextual data can be linked, thus mitigating criticisms of social media sources being ‘data-light’. For example, Figure 8 shows a subsample of the WFP corpora that contains geolocation information for tweets. These tweets may be located within regions for which administrative data is recorded on reductions in food parcels, potentially illustrating whether there is a link between changes in opinions of the WFP with actual cuts.

Twitter provides three opportunities for collecting geographical information about users: from the user profile, from geo-tagged tweets, and from the content of tweets. The first opportunity comes from the location field in the Twitter user’s profile. This is the location where Twitter users say they live. When analysing user-supplied locations there are some challenges: users can write nothing at all in the field, users can write sequences of nonsensical letters, punctuation and other symbols, and users can write actual place names but customise them with creative use of spelling, word boundaries, symbols and punctuation. Another challenge is that some users lie about where they live, which is particularly obvious when they make humorous or wishful references to real or imaginary places such as 'on the beach' or 'in hell'.

Despite the challenges of deriving geographic information from the location profile field, it is possible to identify the country for over 50% of Twitter users. The Yahoo! PlaceFinder geographic database can be used to extract location information. Yahoo! PlaceFinder returns rich, hierarchical location descriptions that include the city, county and country in which a location is found as well as the latitude and longitude co-ordinates of the location. Yahoo! PlaceFinder can return accurate location information from text, such text found in the location field of Twitter user profiles. For example, given the partial postcode SW19, Yahoo! PlaceFinder can identify this postcode as belonging to the Wimbledon district in London. Using this service, it is possible to locate the country for 52% of Twitter users, the state for 43% of users, the county for 36% of users, the city for 40% of users and the post/zip code for 10% of users worldwide.

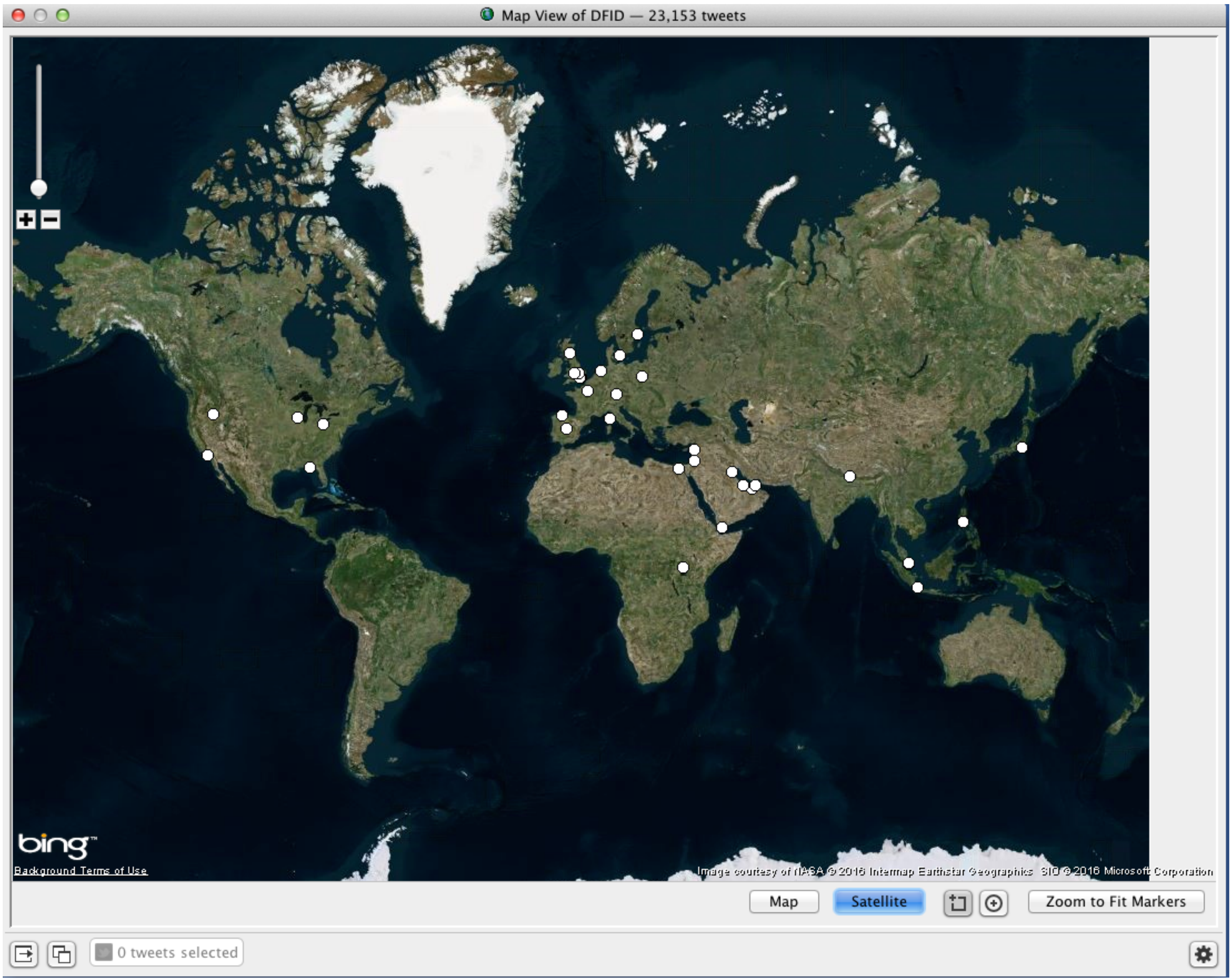
Beyond including geolocation information in the profile, Twitter users have the option to enable location services on their account. This feature is *off* by default and requires users to opt in, but once it is enabled users can geotag their tweets with precise location data in the form of latitude and longitude. Only tweets with original content can be geotagged. Retweets generated by invoking the retweet command in the Twitter user interface are not classed by Twitter as original content and are never geotagged. However, retweets generated by copying and pasting the content of a tweet into the tweet-composition box are classed as original content and can be geocoded (if the user chooses). Approximately 1% of all tweets are geotagged, meaning that the exact position of where the tweeter was when the tweet was posted is recorded using longitude and latitude measurements (although this percentage can vary depending on the tweeter's activity).[4] Figure 8 demonstrates that only a very small proportion of the original tweets (40, or 0.48%) generated by our WFP key word search query included GPS coordinates. However, it is important to note that retweets, representing 65% of the Twitter corpora for this study, do not contain GPS information by default. Table 3 provides data on the typical volume of tweets containing geotags by language use.

There is a conceptual difference between geo-tagging and profile-based locations. Geotagged data tell researchers where a person is when they publish the tweet, whilst profile data could tell researchers any number of things including where people were born, lived, employed, are passing through or simply identify with. For all these reasons, geotagged tweets have become the gold standard. They contain the most information in the most useful and accurate format for social and government research.

Deriving geographical information about Twitter user locations, such as the city, county and latitude and longitude co-ordinates is important for linking traditional curated and administrative data sources with social media data. For example, when we know the district in which a Twitter user lives we can look up a wide range of official statistics about their area. Linking traditional and social media data in this way leads to new opportunities for analysing and visualising data about populations. For example, we could identify tweets that contain comments indicating rising tensions between groups and by locating these within existing geographies we can show which areas can be characterised by proportionally high levels of social tension. This can then be cross-referenced with official statistics on the religious and ethnic population composition and the relationship between social tension in tweets and contextual area factors can be explored.

[4] For example, Twitter users tend to geolocate content more frequently when at events or on vacation.

Figure 8: Geocoded Tweets in full sample



The representativeness of geo-tagged tweets

The low number of tweets that contain GPS location data in the WFP corpora is not unusual. There remains significant concern over who geotags their tweets. Fundamentally this is a question about representativeness, not in relation to the Twitter population as a subset of the general population, but whether this group is representative of other Twitter users. While males are more likely to geotag their tweets than females (by 0.1%), the difference in proportion is quite small. The differences in age and occupation also very small.[5] The biggest differences in geolocation-based behaviour are related to language. Table 3 shows the least likely groups to use geotagging are those who tweet in Korean (0.4%), followed by Japanese (0.8%), Arabic (0.9%), Russian and German (both 2.0%). Turkish tweeters are the most likely to geotag (8.3%), then Indonesian (7.0%), Portuguese (5.9%) and Thai (5.6%). There is a language effect in play that shapes geotagging behaviour. User language is not a proxy for location so these cannot be dubbed as country level effects, but perhaps there are cultural

[5] As the class-based analysis can only be applied to users in the UK (using timezone as a proxy), these results cannot be generalised beyond the UK Twitter population.

differences in attitudes towards Twitter use and privacy for which language acts as a proxy. Whether these differences are cultural or whether user tweet language can be considered a proxy for location are interesting questions for social and government researchers. In the case of location proxy researchers may make some tentative observations about technological infrastructure and levels of smartphone use and it may be the case that decisions about behaviour on Twitter are primarily cultural for some groups but a function of technological necessities for others, or even a mix of both. Regardless of the cause, clear differences exist based on language that demonstrates inconsistent adoption of geotagging Twitter content.

Language:	Code:	Not Geotagged:	Geotagged:	Total:
English	en	96.7% (n = 8083523)	3.3% (n = 275321)	835884
Japanese	ja	99.2% (n = 4762822)	0.8% (n = 38231)	480105
Spanish	es	95.6% (n = 2343493)	4.4% (n = 107758)	245125
Arabic	ar	99.1% (n = 1352119)	0.9% (n = 12707)	136482
Portuguese	pt	94.1% (n = 931212)	5.9% (n = 58225)	98943
Indonesian	in	93.0% (n = 812496)	7.0% (n = 61579)	87407
Russian	ru	98.0% (n = 719338)	2.0% (n = 14816)	73415
Turkish	tr	91.7% (n = 567254)	8.3% (n = 51249)	61850
French	fr	97.1% (n = 427340)	2.9% (n = 12901)	44024
Tagalog	tl	95.7% (n = 298373)	4.3% (n = 13415)	31178
Korean	ko	99.6% (n = 256037)	0.4% (n = 1117)	25715
Thai	th	94.4% (n = 184890)	5.6% (n = 10899)	19578
Italian	it	96.3% (n = 159880)	3.7% (n = 6074)	16595
German	de	98.0% (n = 114658)	2.0% (n = 2331)	11698
Dutch	nl	96.6% (n = 98897)	3.4% (n = 3484)	10238
Polish	pl	97.8% (n = 42613)	2.2% (n = 978)	4359
Swedish	sv	95.5% (n = 40047)	4.5% (n = 1887)	4193
Haitian	ht	96.1% (n = 39975)	3.9% (n = 1621)	4159
Estonian	et	94.5% (N = 36023)	5.5% (n = 2094)	3811
Slovenian	sl	96.3% (n = 26653)	3.7% (n = 1031)	2768
Total		96.9% (n = 21297643)	3.1% (n = 677718)	2197536

Table 3: Percentage of geo-tagged tweets by language use on Twitter

Highlights:

- Using Yahoo! Placefinder with Twitter data it is possible to identify the geolocation of the country for 52% of Twitter users, the state for 43% of users, the county for 36% of users, the city for 40% of users and the post/zip code for 10% of users worldwide
- Approximately 1% of all tweets are geotagged, meaning that the exact position of where the tweeter was when the tweet was posted is obtained
- The least likely groups to use geotagging are those who tweet in Korean (0.4%), followed by Japanese (0.8%), Arabic (0.9%), Russian and German (both 2.0%). Turkish tweeters are the most likely to geotag (8.3%), then Indonesian (7.0%), Portuguese (5.9%) and Thai (5.6%)
- Only a very small proportion of the original tweets (40, or 0.48%) generated by our WFP key word search query included GPS coordinates. However, it is important to note that retweets, representing 65% of the Twitter corpora for this study, do not contain GPS information by default
- This information is useful to track where people are in relation to their beliefs and opinions around topics such as food cuts

Sentiment Analysis

Sentiment analysis is a form of opinion mining that generally requires the identification of an entity on which the opinion is focused (e.g. a person); attributes of the entity (e.g. the person's political perspective); views, attitudes or feelings towards the entity and its attributes (commonly defined as sentiment); an opinion holder and a time at which the sentiment was expressed. The outcome of sentiment analysis is often subjective and based on the existence of a list of keywords in a message. Therefore, many social media software platforms have validated the 'ground truth' of their sentiment analysis tools with human users and compared the human annotated sentiment score with results returned from the sentiment algorithm. This is a semi-automated approach where human input is required to tailor the machine's interpretation of what is positive or negative, and can dramatically increase the speed at which a general opinion on a topic can be obtained.

The COSMOS platform has integrated the SentiStrength sentiment analysis tool, which provides a positive and negative score for each English language text string it examines[6] (see Figure 9). The average + ve and − ve sentiment scores are plotted on a line chart with time as the horizontal axis and a range of − 5 to +5 on the vertical axis, representing variation in sentiment over time, and making peaks and troughs obvious to the user. Figure 9 shows sentiment scores plotted as a line chart over the full period of this study (top=positive; bottom=negative). High density in the line graph (visualised as blocks of colour) indicate high frequency of particular sentiment. What is apparent is that the first spike observed in the frequency chart corresponds to high density and extreme spikes in negative and positive sentiment. This indicates that highly emotive language was being used in the information being spread on Twitter on Dec 1st 2014. Further spikes can be observed throughout the year, but none of them display the same amount of density and sustained extremity as the first spike.

The analysis in this study was conducted on historic data. In a real-time setting a social or government researcher might use a similar keyword search on an unfolding news story related to a policy change in a MENA region using the Twitter Streaming API. As tweets are produced sentiment scores are computed and displayed in a line chart in real-time in an aggregate fashion, allowing for the identification of peaks in positive or negative content. Repeated peaks in negative content (shown as blocks of colour) might indicate a pattern is forming, which may prompt the researcher to visually inspect the content of these highly negative tweets, and maybe invoke other tools, such as network analysis to identify key thought leaders in the negative social media discourse.

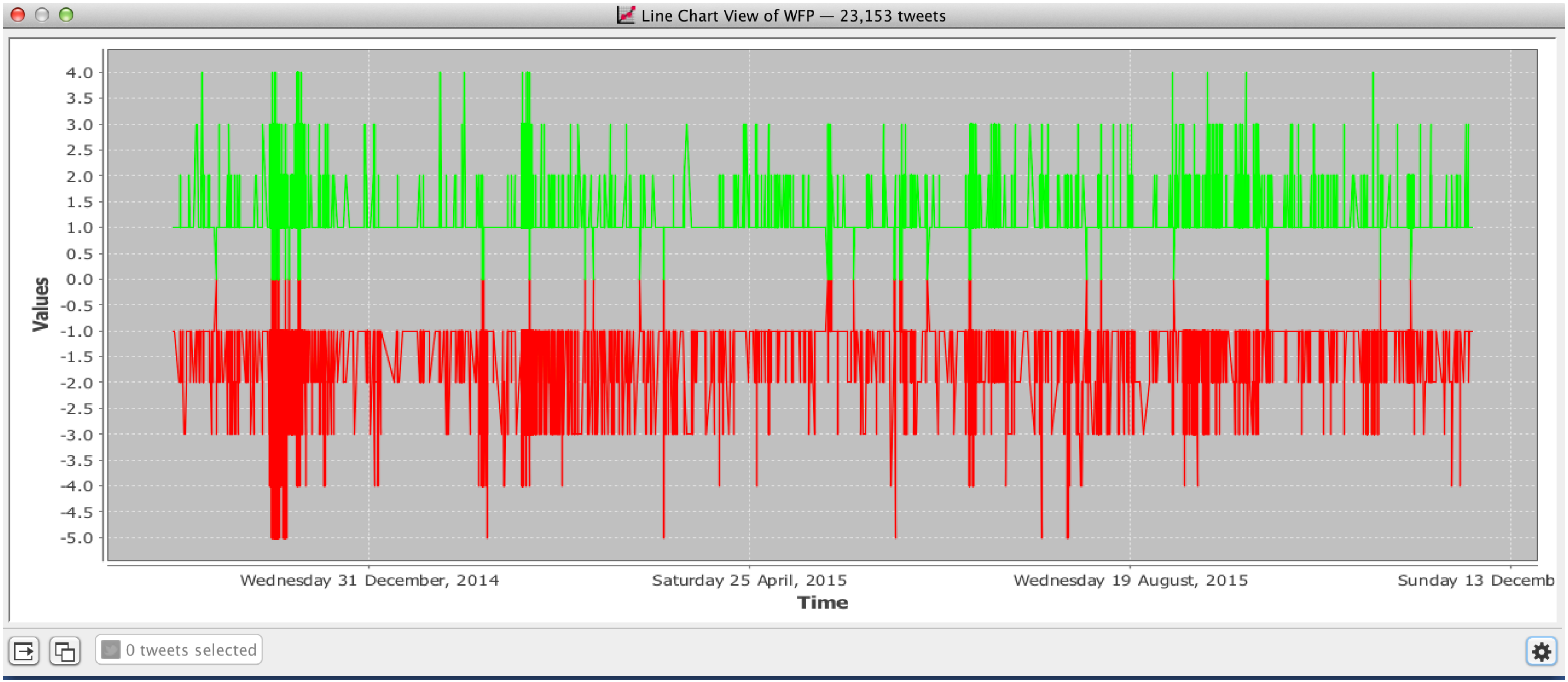
[6] Arabic sentiment analysis tools are also available

It is important for social and government researchers to be aware of the problems associated with employing generic sentiment analysis tools to specific contexts. Generic tools are trained and tested using terms that are classified as either positive or negative. This approach means sentiment analysis is not sensitive to certain forms of communication, such as sarcasm (which may be responsible for the spike in positive sentiment on 1st Dec 2014). As a result, many tweets can be misclassified resulting in erroneous results and visualisations. If the sentiment of content is of importance to a research project, adequate resources should be set aside to develop bespoke algorithms designed around the research problem.

Highlights:

- Sentiment analysis maps words onto pre-defined emotive labels. The examples provided use positive and negative, but this can be extended to more fine grained emotions
- Sentiment provides insight into the way social media users are reacting to events such as news of cuts to the WFP
- Sentiment analysis is prone to producing false results due to sarcasm, irony and the use of multiple emotions towards different entities (e.g. “I *hate* the decision because I *love* this country)
- Figure 9 shows a sentiment timeline with positive emotion on the top and negative on the bottom. The scale on the left shows the lines going from 0 (neutral) to +4 (positive) and -5 (negative). Polar extremes on this chart suggest extreme differing opinions on a topic within the dataset

Figure 9: Sentiment analysis over full sample



INFORMATION PROPAGATION MODELLING

Outside of COSMOS more sophisticated forms of social media analysis are possible. Using conventional statistical software, such as STATA, it is possible to build statistical models of the flow of information on social media networks to gain better insights into the dynamics of propagation. Information propagation modelling allows for an understanding of the various factors that lead to the spread of social media messages. There are two *Dependent* measures in information propagation modelling: the *Size* of information (measured by counting number of retweets) and *Survival* of information (measured by counting the seconds between the first and last retweet). In terms of size, the number of retweets is a measure of the volume of public interest and endorsement of the information, while survival (or duration) is a measure of persistence of interest over time.

To model the two Dependent measures of information propagation we used two modelling techniques: Poisson regression and Cox's Proportional Hazards regression. Poisson regression was used to model the Size measure as this is best described as a *count* of retweets. Count variables represent types of events that are largely not experienced by the majority of the sample (in this case retweets where the majority of tweets are not retweeted with a minority being retweeted). Linear regression models are not appropriate for count variables given the nonlinear distribution of the data. Poisson related models are suited to this kind of data as they are built on assumptions about error distributions that are consistent with the nature of event counts. The second dependent - survival - was a measure of the lifetime of an information flow. Our interest here is to model the factors that affect the survival of information flows following the WFP media announcements. This can be framed as hazards to survival, thus we adopted Cox's proportional hazards model (Cox 1972).

Three sets of variables were entered as independent predictors of information flow size and survival in the models: Content factors; Social factors and Control factors. Content factors relate to the text of the tweet, and include the sentiment expressed in text and whether or not the text included a URL pointing to an external source (such as a news item). Previous research has indicated that sentiment and the presence of URLs can influence the likelihood of information spreading on social media (Burnap et al. 2014). Social factors relate to the user of the Twitter account, and includes the 'agent type': Media agent (local, national and international press and broadcasters); Official agent (government departments, non-governmental organizations, non-state actors etc.); MENA individual agent (influential social media figures from the Middle East and North Africa region); and Other agent (all other Twitter users not included in these categories). Previous research has indicated agent type has a significant influence over the likelihood of information propagating on Twitter (Williams & Burnap 2015). Multiple control factors were included that have been shown to influence the flow of information in social media networks (Zarrella 2009). These include time of day (entered as categories) and day of week.

Table 4 presents the results of the Size model. It is important to note that the model only includes original tweets as cases (n=8,383), with the number of retweets entered as the dependent. Many of the control factors emerged as statistically significant, justifying their inclusion. Both the coefficients and the incidence-rate (IRR) ratio are included to ease interpretation. An IRR is a univariate transformation of the estimated coefficient for the Poisson model. It is a relative difference measure used to compare the incidence rates of events (retweets) occurring at any given point in time. A score above 1 indicates an increased incidence rate ratio and below 1 a reduced incidence rate ratio for retweets. The Content and Social factors of interest emerged as significant in the model. Holding all other factors constant, both sentiment and the inclusion of a URL in tweet text significantly predicted the likelihood of retweeting content related to the cuts to the WFP. Tweets containing negative sentiment related to the WFP cuts were statistically more likely to be retweeted compared to tweets containing positive content, confirming our initial interpretation of Figure 9. Using the IRR we can say that tweets containing negative sentiment had a retweet rate 1.33 times higher than those tweets containing positive sentiment. Tweets about cuts to the WFP containing URLs were also significantly more likely to be retweeted. Tweets with links to external sources had a retweet rate 1.63 times higher than those without links, confirming the interpretation of the network visualization (Figure 6), that media outlets were central to spreading information around this topic. Holding all other factors constant, social factors were also statistically significant in predicting the rate of retweets. Compared to tweets produced by Other Agents (general Tweeters) tweets produced by Media Agents had a retweet rate near 45 times greater, Official Agents had a retweet rate 20 times greater, and MENA Agents had a retweet rate near 74 times greater over the 13 month study window. The MENA tweet included a picture featuring individuals feasting on excessive amounts of food, purported to be taken on the last day of the conference on the global food crisis (see Figure 10). In interpreting these findings it is important to note the small number of identified MENA Agents in the dataset, meaning this correlation may be spurious, and further research is needed to confirm any association in future observations. Overall the model for size of retweets performs well, being significant (p=0.00) and explaining near 40% (Pseudo R square=0.3905) of the variance in the dependent variable (rate of retweets).

Figure 10: Image from the tweet from MENA Agent

Table 4: Poisson Regression Predicting Counts of Retweets (Size Model)

	Coef.	SE	IRR
Content factors			
Sentiment	-0.28***	0.09	0.75
URL	0.49***	0.15	1.63
Social Factors			
Media Agent	3.81***	0.12	45.21
Official Agent	3.03***	0.21	20.69
MENA individual Agent	6.56***	0.16	73.53
Ref: Other Agent			
Control factors			
Commute Morning (GMT)	-0.29	0.17	0.75
Commute Evening (GMT)	0.28*	0.14	1.32
Evening (GMT)	0.20	0.13	1.22
Night (GMT)	0.39**	0.14	1.47
Ref: Work (GMT)			
Sunday	-0.22	0.19	0.80
Monday	-1.09***	0.16	0.34
Tuesday	-1.22***	0.16	0.29
Wednesday	-1.18***	0.19	0.31
Thursday	-2.01***	0.25	0.13
Friday	-2.61***	0.33	0.07
Ref: Saturday			
Model Fit			
Log Likelihood			-1680.9549
Chi-Square			2153.58
Sig.			*p*=0.00
Pseudo R sqaure			0.3905
N[1]			8,383

**p*<.05 ***p*<.01

[1] Reduction due to removal of retweets, leaving only original tweets.

Table 5 presents the results of the Cox proportional hazards model predicting tweet survival over the study period. The first point to make is that the model performs poorly being just statistically significant (*p*=0.047). This is likely because of the length of the study period being modelled. It is highly unlikely many tweets would last this amount of time in the study window. The longest lasting tweet was 27 days, far shorter than the 13 month study period, and on average, tweets in the dataset lasted just over 1 hour. The poor performing model is confirmed when examining the independent variables. Only sentiment emerges as statistically significant, with positive sentiment more likely to die out more rapidly as compared to negative sentiment. Figure 11 is a Kaplan Meier graph displaying the survival estimates of positive and negative content.

Information propagation modelling can be used to identify the inhibiting and enabling factors of the spread of Twitter messages during and following an event of interest. A focus on these factors can assist decision making with respect to communication strategy around an event. For example, if it is identified that certain types of accounts are important to the spread of information (e.g. local MENA agents) then it would be possible to seek endorsement of key MENAD messages by these agents, via retweeting.

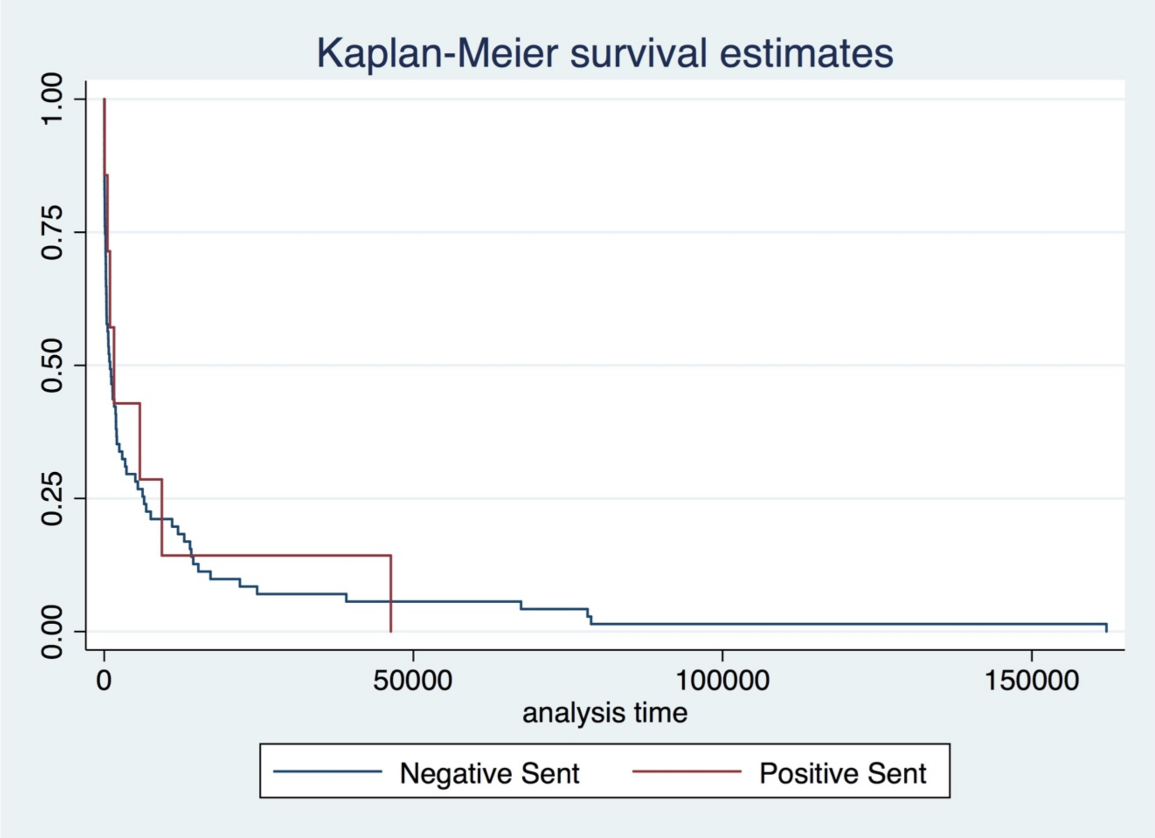
Table 5: Cox Regression Predicting Hazards to Tweet Survival (Survival Model)

	Coef.	SE	Hazard Ratio
Content factors			
Sentiment	0.60**	0.11	1.81
URL	-0.39	0.19	0.68
Social Factors			
Media Agent	0.02	0.25	1.02
Official Agent	0.52	0.72	1.68
MENA Individual Agent	-0.44	0.22	0.64
Ref: Other Agent			
Control factors			
Commute Morning (GMT)	-0.03	0.30	0.97
Commute Evening (GMT)	-0.06	0.27	0.94
Evening (GMT)	0.00	0.30	1.00
Night (GMT)	-0.31	0.30	0.73
Ref: Work (GMT)			
Sunday	0.49	0.79	1.63
Monday	0.14	0.45	1.15
Tuesday	0.38	0.58	1.46
Wednesday	-0.10	0.39	0.90
Thursday	0.60	0.83	1.82
Friday	0.32	0.70	1.38
Ref: Saturday			
Model Fit			
Log Likelihood			-567.45853
Chi-Square			25.19
Sig.			*p*=0.047
N[1]			8,383

**p*<.05 ***p*<.01

[1] Reduction due to removal of retweets, leaving only original tweets.

Figure 11: Sentiment survival



Highlights:

- Information propagation modelling can be used to identify the inhibiting and enabling factors of the spread of Twitter messages during and following an event of interest.
- Information propagation on social media can be measured in terms of size (how many retweets) and survival (how long it is retweeted for).
- For the study of reactions to WFP food cuts, sentiment and the inclusion of a URL (additional evidence and longer blogs) in the tweet text significantly predicted the likelihood of content being retweeted.
- Tweets with –ve sentiment were more likely to be retweeted by a factor of 1.33 compared to +ve tweets.
- Tweets containing URLs were more likely to be retweeted by a factor of 1.63 compared to those without.
- During the study window, Media tweeters were more likely to be retweeted by a factor of 45, official tweeters by a factor of 20, and MENA tweeters by a factor of 75 compared to general tweeters.
- Content containing positive sentiment was shown to die out more rapidly than content containing negative sentiment.
- A focus on these factors can assist decision making with respect to communication strategy around an event.

EVALUATION OF USING SOCIAL MEDIA DATA TO RESEARCH THE MENA REGION

Social researchers in academia and government are being challenged by new forms of socially relevant data produced at volume and speed on social media networks. The exponential growth of social media uptake and the availability of vast amounts of information from these networks has created fundamental methodological and technical challenges and benefits for social research. As stated at the beginning of this document, these can be summarised as the 6 Vs: volume, variety, velocity, veracity, virtue and value.

The challenge of the 6 Vs

Volume refers to the vast amount of socially relevant information uploaded on computer networks globally every second. Ninety per cent of the world's data was created in the two years prior to 2013 (BIS 2013). This is partly due to the global adoption of social media over the past half a decade. Of these online social interactions, a sizable portion are relevant to social and government research problems. A comparison with conventional curated and administrative data shows how these new datasets far exceed the size of those that social and government researchers have come to rely on to gain their scientific insight. For example, the whole UK Data Archive currently holds between 2.2 and 15 terabytes of data. These sizes are dwarfed by the volume of social media data being produced daily (Facebook produces 500 terabytes a day). This volume creates technical and methodological challenges and benefits. It is challenging as the technology to collect, store, search and retrieve such vast amounts of data is rarely available to social and government researchers, meaning the insights contained within massive datasets often remain undiscovered. Furthermore, our existing modes of analysis (qualitative and quantitative) may not be appropriate for such sizable datasets. For example, manual qualitative analysis of millions of tweets is simply too time consuming, while statistical modelling on such large datasets remains largely uncharted[7]. Yet despite these limitations, the volume of data allows researchers to potentially reach 'new populations' that are inaccessible using conventional methods. We are beginning to gain an understanding of how these online populations map on to geographic regions, including the MENA region, but further work is needed (see Sloan et al. 2013, 2015, Dennis et al. 2015).

Velocity refers to the speed at which these new forms of data are generated and propagated by users. The rapid and continual production of these naturally occurring data means social and government researchers can observe events as they unfold, as opposed to retrospectively gathering data months or even years after the event. Recent social unrest illustrates how social media information can spread over large distances in very short periods of time as evidenced by the Tunisian and Egyptian Revolutions (Choudhary 2012; Lotan 2011). However, dealing with the speed of these data is a technical and human challenge. Gaining insights from data in real-time is an attractive offering, however, being able to confirm the veracity of posts rapidly, and to analyse data as they are produced, are significant challenges that we are yet to overcome.

[7] For example, big data approaches tend to produce models and algorithms that are over fit to the idiosyncrasies of a particular data set, leading to spurious results that often do no not reflect reality.

Variety relates to the heterogeneous nature of these data, with users able to upload text, images, audio and video. This multi-modal mixed dataset is rich in meaning that can be harnessed by researchers. However, unlike qualitative and quantitative data that are often labelled, coded and structured within matrices and ordered transcripts, social media data are messy, noisy, complex and unstructured making them difficult to manage and analyse. Currently we are able to handle large amounts of text data with a degree of confidence, but the technology relating to dealing with and interpreting massive amounts of data gained from images and is in its infancy.

Veracity relates to the quality, authenticity and accuracy of these messy data. At the most benign end of the spectrum, tweeters may use sarcasm that may be misinterpreted by other tweeters or machine algorithms designed to measure sentiment. At the more malicious end of the spectrum, some tweeters spread misinformation and rumours, in an attempt to mislead other users and the authorities (e.g. fake news). Bots, automated accounts designed to spread information as widely as possible, also interrupt the flow of information in social media networks, sometimes skewing results. To establish the veracity of social media posts, researchers suggest triangulating information flows with more conventional sources, such as curated and administrative data, and on the ground observations. Instead of social media data acting as a surrogate for established sources, they should instead augment them, adding a longitudinal extensive dimension to existing research strategies and designs.

Virtue relates to the ethics of using this new form of data in social research. The ESRC Framework for Research Ethics highlights the two key principles of informed consent and harm to participants. But it is practically difficult to seek informed consent from Twitter users in research, and Twitter's Terms of Service require users to consent for Twitter to share any content posted with third parties. We may argue therefore that researchers in this field must accept that consent has been provided, as long as researchers adhere to basic principles of social science ethics, while ensuing results are presented at an aggregate level. Additional individual level consent should be sought if researchers wish to directly quote online communications (see page 37).

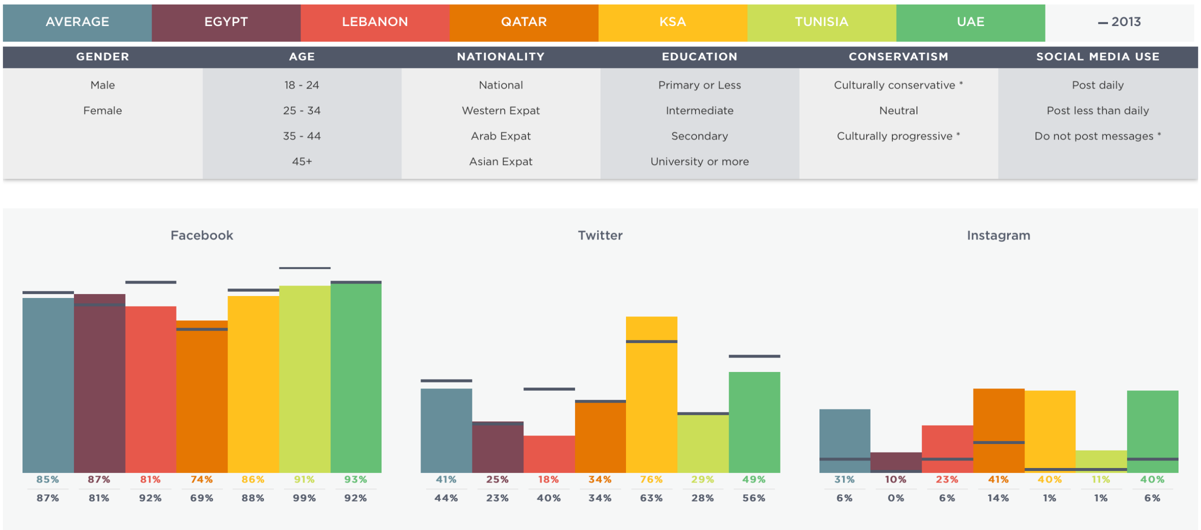
Finally, **value** links the preceding five Vs – only when the volume, velocity and variety of these data can be computationally handled, and the veracity and virtue established, can social researchers in academia and government begin to marshal them and extract meaningful information. However, to date, few academic or government projects have collected and analysed social media data. This is primarily due to the lack of existing computational infrastructure to support researchers in gaining access to and analysing these data, and the lack of interdisciplinary working practices within government.

Accessing Twitter Data on the MENA Region

Twitter use in the MENA Region

There is limited data on social media penetration in the MENA region, but efforts are underway to build a picture of use in some countries. Dennis et al. (2015) conducted surveys of social media use in several counties in the MENA region over the past 3 years. The latest survey in 2015 was based on N=6,093 face-to-face and telephone interviews with the general population 18 years and older across six countries: Egypt, Lebanon, Qatar, Saudi Arabia, Tunisia, and the UAE. Figure 12 provides a comparison between countries and years.

Figure 12: Social Media Use in MENA Countries in 2015 (compared to 2013)



Source: Dennis, E. E., Martin, J. D., & Wood, R. (2015). *Media use in the Middle East*, Northwestern University in Qatar. Retrieved from www.mideastmedia.org

Real-time data

The Twitter Streaming API is more open and accessible compared to other social media platforms. Twitter provides three levels of data access (the lowest of which is free) and the data can be obtained using an online query or dedicated software. The free random 1% of the Twitter stream is dubbed the 'spritzer'. Both the 'garden hose/deca hose' (providing access to a random 10%) and the 'fire hose' (providing 100% access) can be obtained via agreement with Twitter for research or can be purchased via their data-resellers e.g. Gnip (https://gnip.com). For most social research projects the 1% feed is usually sufficient and provides access to approximately 3.5-5 million tweets per day free of charge. Several social media data collection and analysis software platforms make accessing the 1% stream of data straightforward.

Historic data

The Twitter Search API provides free access to historical data up to 7 days into the past from the day of the query. It is important to know that the Search API is focused on relevance and not completeness. This means that some Tweets and users may be missing from search results. For research where match for completeness is important researchers should consider using the Streaming API (above) instead. For social media posts over 7 days old data can be purchased via Twitter's chosen data-resellers e.g. Gnip (https://gnip.com). The cost of data is dependent upon volume of returned posts and the length of the query. For example, a query spanning 1 week that returns 1M posts may cost the equivalent of a query spanning 1 year that returns 10K posts.

Sharing Data

Twitter Terms of Service for the use of its API restricts researchers from programmatically sharing data over the Internet. However, researchers are able to share datasets manually, in batches of no more that 50,000 tweets. This means that government departments can share data without breaching Twitter's ToS.

ETHICS

Government standards

Ethics has emerged as a contentious area of debate in the use of social media data in social and government research. Recently the Cabinet Office Data Science Ethical Framework (Cabinet Office 2016) was developed and applies to all government research. It includes the following ethical guidelines that are relevant to research and evaluation and in development contexts:

- The Data Protection Act requires researchers to have an understanding of how people would reasonably expect their personal data to be used and therefore need to be aware of shifting public perceptions of these data;
- Social media data, commercial data and data scraped from the web allow us to understand more about the world, but come with different terms and conditions and levels of consent that need to be fully understood by researchers;
- Researchers should always use the minimum data necessary to achieve the public benefit, and the use sensitive personal data should be kept to a minimum;
- Safeguarding people's privacy is paramount in any research write-up. This can be achieved by de-identifying or aggregating data to higher levels, and using synthetic data.

Learned Society Standards

Several learned societies have introduced ethical principles for research in digital settings, including the British Psychological Society (BPS), the British Society of Criminology (BSC), the British Educational Research Association (BERA), the European Society for Opinion and Market Research (ESOMAR) and the Association and the Association of Internet Researchers (AoIR). Broadly, most guidelines adopt the 'situational ethics' principle: that each research situation is unique and it is not possible simply to apply a standard template in order to guarantee ethical practice. The most through set of guidelines are those developed by AoIR. AoIR was one of the first learned societies to introduce a set of guidelines, which are now in their second iteration. These guidelines highlight three key areas of tension: the question of human subjects online; data/text and personhood; and the public/private divide (AoIR 2012). The guidelines advance that the notion of the 'human subject' is complicated when applied to online environments. The Internet complicates the conventional construction of 'personhood' and the 'self', questioning the presence of the human subject in online interactions. In some cases this may be clear-cut: emails, instant message chat, newsgroup posts are easily attributable to the persons that produced them. However, when dealing with aggregate information in social media repositories, such as collective sentiment scores for sub-groups of Twitter users, the connection between the object of research and the person who produced it is more indistinct. Attribute data on very large groups of anonymised Twitter users could be said to constitute non-personalised information, more removed from the human subjects that produced the interactions as compared to, say, an online interview. In these cases, the AoIR (2012: 7) guidelines state 'it is possible to forget that there was ever a person somewhere in the process that could be directly or indirectly impacted by the research'.

In relation to informed consent BERA specifically state that social networking and other on-line activities, present challenges for consideration of consent issues and the participants must be clearly informed that their participation and interactions are being monitored and analysed for research. On anonymity the guidelines state one way to protect participants is through narrative and creative means, which might require the fictionalising of aspects of the research or the creation of composite accounts, such as in vignettes, providing generalized features based on a number of specific accounts. In relation to consent ESOMAR state that if it has not been obtained researchers must ensure that they report only depersonalised data from social media sources. If researchers are using automated data collection services, they are recommended to use filters and controls to remove personal identifiers such as user names, photos and links to the user's profile. In relation to anonymity the guidelines state where consent is not possible their analysis must only be conducted upon depersonalised data and if researchers wish to quote publicly made comments they must first check if the user's identity can be easily discoverable using online search services. If it can, they must make reasonable efforts to either seek permission from the user to quote them or mask the comment.

Legal Considerations

Data extracted from the Twitter APIs contain personal data meaning they are subject to the UK Data Protection Act (DPA), and as such it must be processed fairly and lawfully. In cases where informed consent cannot be sought from users (likely to be the majority of cases if thousands of posts are being subject to analysis), a social researcher should establish the fair and lawful basis for collecting personal information. A researcher can accept that social media networks terms of service provide adequate provision to cover this aspect of the DPA. However, if the data has been collected using a service that provides additional meta data on users, such as sensitive personal characteristics (e.g. ethnicity and sexual orientation) based on algorithms that make estimations, the legal issue of privacy may be compounded. Under the DPA sensitive personal information can only be processed under certain circumstances. Deriving insights and making conclusions about a person or persons' views or philosophy and publicly releasing this information could lead to stigmatization or actual bodily harm (in this case of extremist views for example), should the location of the social media persona be established. Furthermore, it is possible that legal proceedings could follow. The DPA allows cases to be brought on a personal basis, so it is possible that the researcher and not the institution could be liable.

Within the EU the 1995 Data Protection Directive and the EU Court of Justice of the European Union decision in Google Spain SL, Google Inc. v Agencia Española de Protección de Datos, Mario Costeja González 2014 (application of the Right to be Forgotten to search engines) have implications for researchers as social media data controllers. Under these protections citizens in EU countries have the right, under certain conditions, to ask for online information to be removed where the information is inaccurate, inadequate, irrelevant or excessive for the purposes of data processing. The proposed Data Protection Regulation allows for the erasure of personal data, the abstention from further dissemination and a duty to obtain from third parties (potentially including researchers) a guarantee of the erasure to any links or replication of the data (where the data are no longer necessary for the purposes under which they were collected, where the data subject withdraws consent on which the processing is based or the data subject objects to the processing of personal data). However, the proposed Regulations are also specific as to the reasons of public interest that would justify keeping data online – the limitations of the right to be forgotten. These include the exercise of the right of freedom of expression, the interests of public health as well as cases in which data is processed for historical, statistical and scientific purposes. To what extent these proposals will impact upon social media research data controllers is yet to be fully realised.

Public Attitudes

The Eurobarometer Survey 359 *Attitudes on Data Protection and Electronic Identity in the EU* (2011) found that 58 percent of European Internet users read online privacy policies. Over 70 per cent were aware of the purposes for which social media networks can and may collect, use, and share personal data of users. Around 70 percent of European citizens were concerned about how companies use their data. Beninger *et al.* (2014) conducted a series of interviews with social media users in London to develop recommendations for researchers as part of the National Centre for Research Methods funded New Social Media New Social Science (NSMNSS) methods network. The study found that Internet users differ in how familiar and comfortable they are with the privacy and security settings that are provided by social media networks, and recommended researchers should not assume all users have read and understood terms and conditions that govern issues such as consent and privacy. Users also expressed concern over their photos, Twitter handles (screen names) and personal and sensitive posts being published in research papers. There was an expectation from users that they be approached for consent if there was intent to publish these kinds of data. More recently Evans *et al.* (2015) conducted a survey of users' attitudes towards social media research in government and commercial settings. Three in five respondents reported knowing that their social media data could be shared with third parties under the terms of service they sign up to. However, 60 per cent felt that social media data should not be shared with third parties for research purposes. These views softened when users were offered anonymity and where only public data were to be used in the research. The majority of users rejected the position that accepting the terms of service was enough to establish consent, preferring instead opt-in consent for each individual research project.

Publishing Twitter Content

The publication of quantitative findings from social media research is largely ethically unproblematic as the data are aggregated. However, it is likely that the publication of qualitative findings will be enhanced by the inclusion of full tweet text, yet problems arise here as this makes users identifiable. While Twitter do not provide guidance for social researchers, they do provide a set of Best Practices for Media (static uses and publication):

- Show name, @username, unmodified Tweet text and the Twitter bird nearby, as well as a timestamp
- If displaying Tweets, make sure they are real, from legitimate accounts and that you have permission from the author when necessary
- Display the associated Tweet and attribution with images or media
- If showing screenshots, only show your own profile page, the @twitter page, the Twitter “About” page or a page you have permission from the author to show

Twitter provide additional information for developers on the maintenance of their products. Violation of these guidelines can result in Twitter taking action, such as preventing access to their data.

Maintain the integrity of Twitter’s products:

- @username must always be displayed (and name if possible) with tweet text
- Respond to content changes such as deletions or public/private status of tweets
- Do not modify, translate or delete a portion of the content

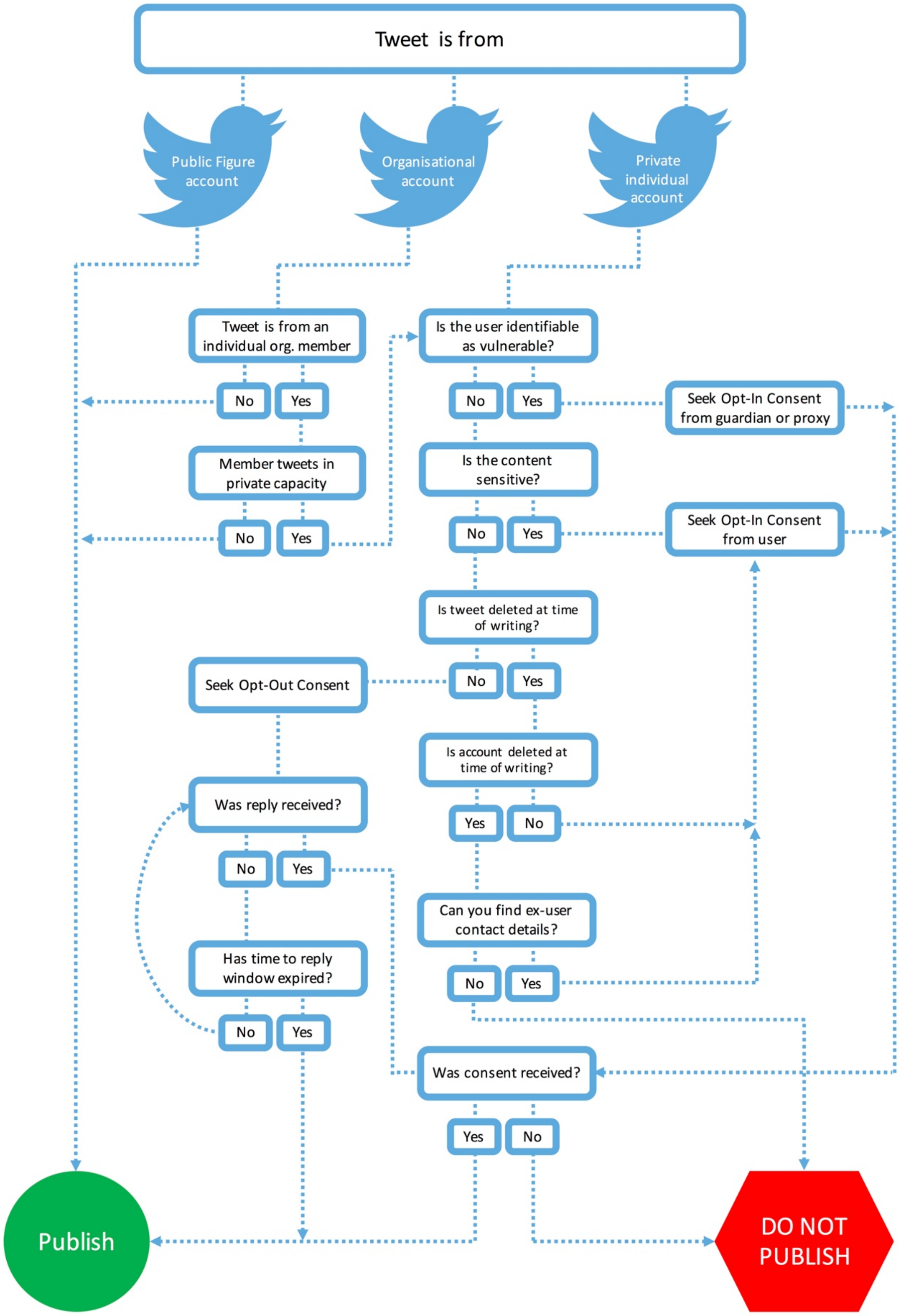
Respect Users’ Privacy and get the user’s express consent before you do any of the following:

- Take any actions on a user’s behalf, including posting Content and modifying profile information
- Store non-public Content such as direct messages or other private or confidential information
- Share or publish protected Content, private or confidential information

Twitter's Terms of Service require users to provide their consent for Twitter to share any content posted with third parties. While it might be acceptable for social researchers to accept users have provided informed consent for their data to be shared with them under these ToS, they should not accept that this provides them with informed consent to publish the content of individual tweets (anonymised or not). Doing so could put users at risk of harm, including reputational, personal and/or physical.

The decision flow chart below has been informed by Twitter guidelines and assists researchers in choosing whether or not to publish the original content of tweets.

Figure 13: Publishing tweet content decision flow chart



Deciding on the status of tweeters (e.g. public, organisational, private, vulnerable) and their tweets (e.g. organisational, private, sensitive) is at the discretion of the researcher and/or the ethics review board. In seeking to reach these decisions researchers should consult existing ethical guidelines (see Section 6.4) that provide definitions of public figures (e.g. politicians and celebrities who aim to communicate to a wide audience), vulnerable individuals (e.g. children, learning disabled and those suffering from an illness) and sensitive content (e.g. posts about criminal activity, financial problems, mental health issues and feelings of suicide, extramarital sexual activity, controversial political opinions and activism). As social media accounts can lack personal details, and it is difficult to find additional identifying details, researchers and ethics review boards may be satisfied with the use of the information presented on the profile and in posts alone to reach decisions on the status of users.

REFERENCES

AOIR (2012) Ethical Decision-Making and Internet Research: Version 2.0 – Recommendations from the Association of Internet Researchers Working Committee. Available at: <http://aoir.org/reports/ethics2.pdf>

Beninger, K, Fry, A., Jago, N., Lepps, H., Nass, L. and Silverster, H. (2014) Research Using Social Media: Users Views, London: NatCen

BIS (2013) *Seizing the data opportunity: A strategy for UK data capability*, London: Department of Business, Innovation and Skills

Burnap, P, Williams, M. L. & Sloan, L. (2014) 'Tweeting the terror: modelling the social media reaction to the Woolwich terrorist attack', *Social Network Analysis and Mining*, 4: 206

Cabinet Office (2016) *Data Science Ethical Framework*, London: Cabinet Office

Choudhary, A., Hendrix, W., Lee, K., Palsetia, D. and Liao, W. (2012) 'Social Media Evolution of the Egyptian Revolution', *Communications of the ACM*, 55: 5, pp 74-80

Cox D. (1972), 'Regression Models and Life Tables', *Journal of Royal Statistical Society B*, 34: 187–220

Dennis, E. E., Martin, J. D., & Wood, R. (2015). *Media use in the Middle East*, Northwestern University in Qatar. Retrieved from www.mideastmedia.org

Ginsberg, J.; Mohebbi, M. H.; Patel, R. S.; Brammer, L.; Smolinski, M. S.; and Brilliant, L. 2009, 'Detecting influenza epidemics using search engine query data', Nature 457 (7232)

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) Feature: 1375-1405

Moat H, Preis T, Olivola C, Liu C, Chater N (2014) Using big data to predict collective behaviour in the real world. Behavioral and Brain Sciences 37:1 92–93

NatCen (2014) *Research Using Social Media: Users' View*, London: Natcen

Noguchi T, Stewart N, Olivola C, Moat H, Preis T (2014) Characteristing the Time-Perspective of Nations with Search Engine Query Data. PLoS One 9:4

Sloan, L. and Morgan, J. (2015) 'Who Tweets with Their Location? Understanding the Relationship Between Demographic Characteristics and the Use of Geoservices and Geotagging on Twitter', *PLOS ONE*

Sloan, L., Morgan, J., Burnap, P. and Williams, M. L. (2015) 'Who Tweets? Deriving the Demographic Characteristics of Age, Occupation and Social Class from Twitter User Meta-Data', *PLOS ONE*

Sloan, L., Morgan, J., Williams, M. L., Housley, W., Edwards, A., Burnap, P., Rana, O. (2013), 'Knowing the Tweeters: Deriving sociologically relevant demographics from Twitter', *Sociological Research Online* 18:7

Williams, M. L. and Burnap, P. (2015) 'Crime Sensing with Big Data: The Affordances and Limitations of using Open Source Communications to Estimate Crime Patterns', *British Journal of Criminology.* Online Advance Access.

Williams, M. L. and Burnap, P. (2015) 'Cyberhate on social media in the aftermath of Woolwich: A case study in computational criminology and big data', *British Journal of Criminology.* 56(2), pp. 211-238

Zarrella D (2009) The science of retweets, http://danzarrella.com/science-of-retweets.pdf