

Visual Social Media and Big Data

Interpreting Instagram Images Posted on Twitter

Dhiraj Murthy, Alexander Gross, Marisa McGarry

Abstract

Social media such as Twitter and Instagram are fast, free, and multi-cast. These attributes make them particularly useful for crisis communication. However, the speed and volume also make them challenging to study. Historically, journalists controlled what/how images represented crises. Large volumes of social media can change the politics of representing disasters. However, methodologically, it is challenging to study visual social media data. Specifically, the process is usually labour-intensive, using human coding of images to discern themes and subjects. For this reason, Studies investigating social media during crises tend to examine text. In addition, application programming interfaces (APIs) for visual social media services such as Instagram and Snapchat are restrictive or even non-existent. Our work uses images posted by Instagram users on Twitter during Hurricane Sandy as a case study. This particular case is unique as it is perhaps the first US disaster where Instagram played a key role in how victims experienced Sandy. It is also the last major US disaster to take place before Instagram images were removed from Twitter feeds. Our sample consists of 11,964 Instagram images embedded into tweets during a two-week timeline surrounding Hurricane Sandy. We found that the production and consumption of selfies, food/drink, pets, and humorous macro images highlight possible changes in the politics of representing disasters – a potential turn from top-down understandings of disasters to bottom-up, citizen informed views. Ultimately, we argue that image data produced during crises has potential value in helping us understand the social experience of disasters, but studying these types of data presents theoretical and methodological challenges.

Keywords: social media; Instagram; Twitter; image posting; crisis communication; humour; big data.

"I expect that the number of photos uploaded to Facebook daily is larger than all artifacts stored in all the world's museums"

LEV MANOVICH (2012: 461)

Introduction

Social media has become an important medium of communication during crises – from natural disasters to violent events such as shootings and terrorist events. In the case of the former, users producing and consuming content on social media help tell the story of what is happening on the ground from the vantage point of victims. These content can also serve to alert emergency services as to what victims are asking for. Social media can be used to both gauge public reactions as well as help track down perpetrators via citizen-based surveillance. In all these cases, big data methods are deployed, but they tend to leverage the analysis of text. Though the posting of images and video during crises has become an important part of how crises are socially experienced, methods to study not only the large volume of data being produced during crises, but the speed and complexity of them has perhaps discouraged some social researchers from studying visual social media.

Indeed, despite social media moving towards a state of including visual content as a norm, scholarship has been slow to grasp the methodological challenges of this change. This is partially attributable to the fact that 'social media are a moving target', creating challenges for identifying consistencies across platforms (Hogan/Quan-Haase 2010). In addition, visual social media platforms such as Snapchat and Instagram have restrictive or inaccessible application programming interfaces (APIs), the systems by which external users can directly access platform data. Closed or semi-closed APIs make it difficult to collect many types of visual social media data. Though these data can be immensely constructive to studies ranging from health (Seltzer et al. 2015) to disasters (Yates/Paquette 2011), a disproportionate amount of work is focused on criminal contexts, such as being able to identify people and places (Zhenguo et al. 2015) and deploy police or security services accordingly, what has sometimes been called 'predictive policing' (Lupton 2015, 100). In addition, the majority of work is largely quantitative and tends to rely on computerized visual recognition (e.g. Xinpeng et al. 2015, whose work seeks to automatically identify hackers through their social media profiles). Images produced and consumed during disasters likely perform multiple functions and such methods are generally unidimensional by definition.

To help bridge this gap, our study uses an empirical case study of 11,964 geolocated images taken by users of the popular mobile photo sharing app Instagram and posted on Twitter during Hurricane Sandy. Not only was the October 29, 2012 storm the most destructive in recent memory, but it is perhaps the first major US disaster where Instagram played a major role in how it was socially experienced. In addition, Sandy was the first and last major US disaster to take place when Insta-

gram images were natively embedded within Twitter feeds, providing a unique case where images are networked across two major social media platforms.

Instagram pictures posted to Twitter are an uncommon case for several reasons. First, Twitter data are readily collectible and are well-known for having an open API that enables large volumes of data to be harvested (Murthy and Bowman 2014). Second, most social media platforms are walled off for commercial reasons. Instagram was still in its infancy and saw convergence with Twitter as a positive force. At the time of our study, an image taken with the Instagram application could be directly embedded into a user's tweet, making it fully accessible via Twitter's relatively open API. In December 2012, Instagram disabled this service and as of June, 2016, the platform shut off most third-party direct application access (Cook 2016). Our study of Hurricane Sandy is distinct both in its study of Instagram images posted on Twitter and its method of human coding nearly 12,000 images. The latter allows us to better understand Sandy through the eyes of its victims. Our study also seeks to understand the collective, and perhaps alternative, narratives that can be discerned through analysing this large volume of image data.

Background

Big data and visual social media

Big data has traditionally been defined through the 3Vs – variety, velocity, and volume (Kaisler et al. 2013). This initial framing was extended by Kitchin (2014: 1–2), who argued that big data are also “exhaustive in scope [...] fine-grained in resolution [...] relational in nature [...] and] flexible”. In other words, big data is not just about being big, but also signifies detail, precision, and unique abilities to correlate across very diverse variables. Simply put: “Bigger data are not always better data” (boyd/Crawford 2012: 668). And how we think about big data research methods also matters. Ruppert (2015) argues that Kitchin's extensions are useful, but big data also encompasses changing data practices and new forms of sociality. This vantage point is critical as it not only emphasizes the social life of these data, but also that big data encompasses the methods used to study them as well. In addition, placing primacy on classification via machine learning or other data science methods raises questions of epistemology and ontology that should not be glossed over (Murthy 2016).

Visual social media kicks up particular challenges in terms of big data. As Vis (2013) argues, visual social media are difficult types of content to text classified by machine learning, presenting challenges to deriving fine resolution in large data sets. In most cases, advanced, custom-tailored recognition algorithms or some form of hand coding would have to be used. Therefore, in commercial contexts, images with easily mineable metadata (tagged by users or by platform providers)

are valued more (Vis 2013). Montgomery (2015) adds that “marketplace forces” are driving technology that can ‘identify and target individuals through the photos they have posted of themselves, as well as images in pictures’. In cases where images cannot be automatically identified, the micro-work platform Amazon Mechanical Turk (AMT) has been used to achieve higher usable sample counts (Rashtchian et al. 2010). Others have opted for smaller random samples through hand coding of images by a research team (Morgan, Snelson/Elison-Bowers 2010).

More creative methods to study large image corpora have also been employed. For example, Hochman and Schwartz (2012), use aggregate big data methods to compare Instagram usage between Tokyo and New York City by creating montages of all posted images to discern an identifiable “local color”. Clearly, big data methods vary from quantitative to more mixed or artistic methods, highlighting how difficult it is to frame not only what big data means, but also what the term means in the specific context of visual social media.

This is potentially compounded by an overabundance of visual social media material. Social researchers clearly want to interpret these data as visually-oriented social media activities may be transforming how people understand and experience events (both in crises and every day). In the case of occupy Wall Street, Milner (2013: 2357) argues that interpreting images (in his case, mimetic ones) is important to understanding “conversation between diverse positions”. The proliferation of images in social media platforms is part of economic and social changes that are a product of the information age. Shifman (2012: 199) argues: “Whereas the old economic system focused on ‘things’, the most valuable resource in the information era is not information but the attention people pay to it.” And images are an important part of garnering attention on social media platforms as Ringrose and Harvey’s (2015) work on ‘sexting images’ reveals. Existing images need to be reposted or new images posted to keep or gain visibility. Feng et al. (2015) argue that “[b]ecause messages come to a user every day, new messages appearing above old messages, an information overload means any message, however popular, will rapidly lose its visibility.” These changes in computer-mediated-communication have implications for social communication more broadly, including our interactions face-to-face (e.g. discussing the newest viral image or video with friends). Therefore, not only are large volumes of images important to everyday social media user practices, but big data methods need to be engaged with meaningfully in the digital humanities and social sciences. Otherwise, we risk not understanding changes in human communication, or letting purely computational methods dominate this field.

Instagram

There has been a proliferation of online socially-augmented applications that attempt to create social communities based on the sharing of various types of content with ‘friends’/followers on those platforms or the public-at-large. Among

the popular types of content shared in such communities are updates/messages (Twitter/Facebook), photo/visual media (Instagram/Snapchat/YouTube), and location (Foursquare). Instagram is a popular mobile photo sharing application that debuted in 2010. In April of 2012, Facebook bought the service for \$1 billion. Instagram has over 500 million monthly active users with 80 % of them outside of the US (Instagram 2016). It is compatible with a range of mobile platforms and devices. The application allows users to take pictures using their mobile device. They can apply filters to these photographs to customize their color, tone, and content amongst other things. These filter-applied images can then be shared on a variety of social media that have included Facebook, Twitter, Foursquare, and Weibo over the years. Users can also follow each other via Instagram.

The app's initial uniqueness, as Lister (2013: 11) argues, is its ability to "turn 'run-of-the-mill' snapshots into retro images". However, the platform quickly broadened from an app to allow people to apply image filters (particularly vintage 35mm film chemistries) and evolved into a mainstream photo sharing service. Instagram posts also tend to include an array of hashtags allowing for the image to become a 'network image' (Lister 2013: 4), highlighting the interconnectedness of posted image content. Images were easily networked across several social media platforms as well. As a result, a social network emerged to allow users to share pictures with each other. Instagram is one of the most popular visual media sharing applications and now generates a significant portion of Facebook's revenue (Verhage 2016). Historically, photos shared on Instagram are original photos taken by the user on their smartphone and are usually posted real time (Zappavigna 2016). Many Instagram images are also tagged with location information. In this sense, studying Instagram image streams can give one insights into what people are doing, and when and where – sometimes to the latitude and longitude – they are doing it.

Social Media and Disasters

A rich literature has emerged both around Instagram and Twitter use during disasters. Though work around Instagram and disasters is highly emergent, Twitter use during disasters has been studied in a variety of contexts. For example, Mendoza et al. (2010) studied tweets posted during the 2010 Chilean earthquake and argue that rumours are more likely to be identified and questioned on Twitter than truths during disasters. Following this thread of authenticity, Gupta et al. (2013) examine the spread of inaccurate/deceptive/misleading images on Twitter during Hurricane Sandy, focusing specifically on the height of storm activity (the night of October 29, 2012). They find that these images mostly spread through retweets and that 90 per cent of all retweets can be attributed to the top 30 users. Importantly, they find that users are often comfortable circulating images from sources they are not currently following. This is in distinction to the authority over images previously held by traditional broadcast and print media.

Mills et al. (2009) argue that Twitter has real utility “in time-pressured and information-critical situations.” However, they note that, as the medium’s user base increases, inaccurate information is more likely to appear and spread quickly. They find that Twitter is the leading informational site in the first hour of a situation and other media catch-up after that. Mills et al. argue that Twitter is particularly valuable for ‘second or third waves of emergency response rather than disaster victims and first wave responders, depending on how dialed-in those people are to the network’. Indeed, part of the utility of any social medium during disasters is its usability as a disaster unfolds. Veinott et al. (2009) found that Twitter has utility as a ‘lightweight, low bandwidth’ broadcast tool, but there is a learning curve for best use practices during an emergency.

Similarly, Vieweg et al. (2010) compare two natural disasters – flooding of the Red River and a major grassfire in Oklahoma. They found that users had very different tweeting behaviours. Most people in the Red River area of North Dakota/Winnipeg had experienced a flood before. Additionally, the location of a river does not have the same unpredictability as a wildfire. They find that geolocated tweets and situational update tweets are more likely to be retweeted. Vieweg et al. categorize disaster-related communications into Preparation, Warning, Response to Warning, Hazard Locations, Advice, and Other Environmental Conditions. These studies not only point to discernible patterns in social media production and consumption during disasters, but emphasize that social media content can be coded into meaningful topical categories.

Methods

Data collection

We utilized Twitter’s ‘Streaming API’ to collect tweets from 50 major US cities. We chose the most populous cities according to the 2011 Census Prediction, which estimates population increase based on the last full census (Wilson/United States Bureau of the Census 2012). Twitter allows for the collection of tweets from specified geographical ‘bounding boxes’, specified by a series of latitude and longitude coordinates. The Streaming API has three major functions: users, locations, or keywords. Following Vieweg et al.’s (2010) argument that geolocated tweets during a disaster are more likely to be retweeted, our study focuses on the Places object within the Twitter API, which allows for real-time delivery of tweets from these major cities to our database. Tweets were then filtered for three storm related terms: “hurricane”, “storm” and “sandy”. This search returned 142,768 tweets. With this geolocated sample, we isolated tweets that contained links to outside media. From there, we followed any links to Instagram photos embedded within the tweet. This search returned 11,964 Instagram images that we manually coded for study.

Coding methods

Due to the large volume of images we needed to interpret, our coding methods focused on motif categories rather than trying to interpret the nuances and idiosyncrasies of the images or their ‘performative power’ (Chouliaraki/Blaagaard 2013: 254). Following established methods of categorizing tweets into disaster-related bins (Vieweg et al. 2010), these images were then coded into separate disaster-related motif categories (see Figure 2 for a list of these categories). A team of six coded images following the same coding framework. Images that fit multiple codes were double coded, yielding a total of 15,924 coded instances.

The categories we used for coding were designed to relate to all three phases of Sandy: pre-storm, when Sandy made US landfall, and Sandy’s aftermath. The categories in alphabetical order that we employed are detailed in Table 1. Figure 1 provides examples of posted images by coding category.

Code	Description
ad	images depicting commercial advertisements (not inclusive of relief-related campaigns which were categorized as ‘relief’)
animals	images depicting animals of any kind
damage	images depicting storm-related damage to the built environment or otherwise.
drink	images depicting beverages of any kind, alcoholic and non-alcoholic
food	images depicting perishable food, generally prepared food (not inclusive of canned goods or other non-perishable food; inclusive of restaurants)
gear	images depicting equipment and supplies (inclusive of water and non-perishable food used as emergency supplies)
macro	images depicting humorous macros, images that have a picture superimposed with text with a specific purpose of being funny.
other	images depicting anything not fitting into other specified coding categories
outside	images depicting the built environment, nature, or spaces/places not indoors
people	images depicting people (not inclusive of cartoon depictions of people; inclusive of selfies, individuals, and groups of people)
relief	images depicting relief efforts and relief campaigns (inclusive of screenshots of relief campaigns)

Table 1: Image coding categories.

Results

Frequency Results

Our data clearly illustrates that storm-related Instagram images are concentrated over a three-day period (see Figure 2). On the day of the storm, an increase in activity on Twitter occurred as users weighed in on the storm, including those far outside its path. This fits with previous work on Twitter and disasters which indicates Twitter frequency is generally highest the day a disaster hits (Hughes/Palen 2009). Unsurprisingly, the modal frequency of tweets was immediately after the US landfall of Hurricane Sandy (see Figure 2). Users at the peak of the storm, which occurred when Sandy made US landfall, were likely to post photos at a higher rate. On October 28, 2012, Instagram users posted photos regarding supplies for when the hurricane hit, ‘Sandy parties’ in restaurants and homes, and of friends and family. Sandy parties were common subjects of Instagram images. They featured groups of people eating and drinking before they had to face the predicted impact of Sandy. In terms of supplies to prepare for the storm (what we categorized as ‘gear’), it was clear that users felt anxious about how long the storm would be impacting the Eastern Seaboard. Many images were posted of storm rations, food, and forms of entertainment to keep people busy as they stayed indoors.



Figure 1: Examples of Instagram images by category.

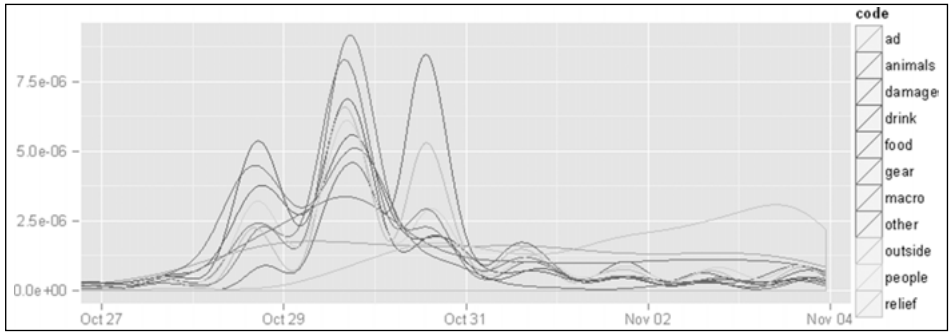


Figure 2: Frequency of selected coded categories by date.

On October 29, 2012, the day Sandy made US landfall, humorous images (what we term ‘macro’ as the majority of images in this category were image macros), were widely circulated. As Figure 2 indicates, the modal category on October 29 was macro. Some humorous images started ‘trending’, becoming the most circulated images. This is an important finding as macro images have been previously found to be important to defining events and movements on social media. For example, Milner (2013: 2358) cites the example of a macro-like poster image of a ballerina dancing on top of Wall Street’s Charging Bull statue, which became emblematic for the Occupy Wall Street movement.

Instagram images perhaps also enabled users to collectively cope by making light the severity of what was to come, similar to the use of humour in other disasters including 9/11 (Kuipers 2005) and the Challenger explosion (Smyth 1986). Other work on images circulating on social media during Sandy indicate the prevalence of humorous images (e.g. with Godzilla and Jaws in front of the Statue of Liberty), but also highlight the circulation of dramatic, non-fiction images (Burgess/Vis/Bruns 2012).

On October 30, 2012, the aftermath of Sandy, users began to focus on the built environment, with specific attention paid to the categories of ‘damage’ and ‘outside’, findings, consistent with other work on images during Sandy (Burgess/Vis/Bruns 2012). We found a discernible trend of circulating iconic images, where particular images became signatures of the storm on Instagram. This parallels work on visual cultures where, during 9/11, the ‘visual impact’ of the image of planes crashing into the twin towers was tremendous, both in the short and long-term (Young 2007). In the case of Sandy, it was of the crane from outside the new World Trade Center tower and the façade of a Chelsea apartment that had been ripped off. These became iconic images of the storm seen from the vantage point of Instagram images posted on Twitter.

In images depicting external damage (what we term ‘outside’), the focus was of trees coming down, flooding, and damage to the built environment. What is interesting to note is that these images are prevalent in the immediate aftermath

of this disaster, providing a glimpse into the vantage point of those living within disaster-affected areas. In terms of images related to ‘relief’ efforts, we saw a jump in frequency on October 30 that did not fully peak until November 3. Some users took pictures of themselves at the Sandy telethon. Others took screenshots of their Red Cross donations and posted them on Instagram. The latter was particularly popular and resembles peer-motivated behavior seen in other events such as the Obama badges used by users on their social media profiles during the 2008 American presidential election (Harfoush 2009: 27).

The longevity of categories such as damage, outside, and people (as seen in Figure 2) illustrates that the categories had different meanings both in terms of when they were posted (specifically the stage of Sandy) as well as the fact that these categories remained important throughout the disaster. Users posting Instagram images on Twitter therefore had an interest and commitment to documenting their own disaster experiences (whether these were lighthearted prior to Sandy’s landfall, about getting back to daily routines (such as getting pets out of the house), or commenting on damage in Sandy’s aftermath). In other words, these users kept these types of image categories persistent through Sandy, representing an important collective effort.

	Frequency	Percent
ad	247	1.6%
animals	397	2.5%
damage	2035	12.8%
drink	1217	7.6%
food	1026	6.4%
gear	208	1.3%
macro	381	2.4%
other	1155	7.3%
outside	5623	35.3%
people	3366	21.1%
relief	269	1.7%
Total	15924	100.0%

Table 2: Frequencies of coded images.

Table 2 illustrates the frequency of codes applied to images. These frequencies are important as aggregated frequencies in some ways help tell alternative stories of Sandy than mainstream accounts. Specifically, there is a very high frequency of people, food, and drink for example and low frequencies for relief and gear. Additionally, what is noteworthy is that categories such as relief and gear do have frequencies that are lower than we might expect. However, damage does have a

relatively high frequency and this indicates that individuals are keen to document their disaster experience, something also supported by the higher levels of frequencies of outside and people. The macro coded category has a relatively high frequency, when one considers this category is based on macro images, images with superimposed text, rather than humorous images as a whole. Indeed, macro images have a higher frequency than both relief and gear. The same, unexpectedly, is true for animals. These image frequencies help sketch an outline of the types of narratives unfolding during Sandy and the types of things people are focusing on representing in their images, potentially suggesting unique views of the disaster, that do include the destruction wreaked by Sandy, but also the human, everyday side such as selfies with friends, pet images, and sharing funny macros.

Category Correlations

As part of our analysis, we correlated categories in our data to investigate whether particular categories were co-occurring. This analysis not only produced some useful insights into exploring Instagram use during Hurricane Sandy, but also provided evidence of how images often perform multiple tasks. As Table 3 and Figure 3 indicate, we found images correlated across coded categories.

People	Outside	$r = 0.86^{**}$
Food	Drink	$r = 0.85^{**}$
Outside	Damage	$r = 0.84^{**}$
People	Food	$r = 0.81^{**}$
People	Drink	$r = 0.78^{**}$
Animals	Outside	$r = 0.73^{**}$
Food	Outside	$r = 0.73^{**}$
Relief	Macro	$r = 0.0$
Food	Relief	$r = -0.06$
Relief	Drink	$r = -0.09$

(* $p < .05$, ** $p < .001$)

Table 3. Select coded category correlations.

Table 3 highlights selected correlations and Figure 3 illustrates all correlations. Our results indicate discernible patterns of co-occurrence within Instagram images we coded. Of particular note is that more than a quarter of coded images tend to have co-occurrence rather than involving single categories. Indeed, the vast majority of categories show significant co-occurrence (at $p < .05$ and $p < .001$). This is important as many images do not tend to be singular in their focus, but are representations of multiple aspects of the disaster experience.



Figure 3: Correlations of all coded categories.

There are also several important instances of an absence of correlation. In particular, the lack of correlation between relief and macro is noteworthy as it signals a sort of taboo or no-go area regarding the Sandy relief efforts (especially the telethon). What emerges is an underlying understanding that relief efforts were a serious endeavour and key to larger, long-term post-Sandy recovery. The lack of correlation of relief with food and drink was expected as Sandy parties generally took place before US landfall. Every other category except for ‘ad’ has a significant level of correlation with macro, indicating that humour is widespread across Sandy-related motif categories, but not in the case of relief-oriented images. This is an important finding.

Temporality

Ultimately, we found that the temporality of tweets was a fundamental, constant variable across our research questions for understanding the context of tweets and serves as a key means to understand how Twitter users responded to Hurricane Sandy at each phase. We found that the highest levels of Twitter activity occurred when Sandy made US landfall. After October 29, there is a significant drop off in storm-related tweets, suggesting that Twitter users prioritize the here and now. Twitter's prompt of 'What's happening' encourages a focus on one's immediate environment and to tweet about one's experience in the present. Our collected data suggests that particular types of subjects are captured in posted images and, collectively, certain narratives emerge at particular points of time (e.g. the importance of meeting with friends prior to Sandy's US landfall and the need to get on with daily routines such as walking one's dog after Sandy made landfall). In other words, the highly temporal nature of social media platforms means that there are many narratives unfolding, but that there are also group-level narratives and perspectives that can be identified. As is the case with other events on Twitter (Murthy 2013), most users were quick to move on to a new story once the storm had passed. This was partially a function of the 2012 presidential election on the horizon, but is more attributable to the fact that Twitter is event/trend driven (Naaman, Becker/Gravano 2011).

However, once Sandy made US landfall, the focus of tweets moved from humorous macros to serious reflections of the real damage caused by Sandy. Many tweets during this time included information from weather sites that tracked the path of the storm as it made its way up the US eastern seaboard. As we used geolocated tweets, we were also able to see that Twitter users outside of the Tri-State area (New York, New Jersey, and Connecticut) generally tweeted prior and during US landfall, but their tweeting is never, relatively speaking, frequent. We compared several unaffected US cities to New York and saw that their tweeting pattern was largely unchanged during Sandy (Murthy/Gross 2016). However, to some extent, this confirms a certain necessary co-presence of tweets during disasters or, put another way, out of sight, out of mind. In addition, we wanted to contextualize the images we collected both in terms of time and space. Specifically, we are claiming that these images help tell a story of how Sandy was experienced by its victims in the Tri-State area. The fact that the modal frequency in our data is when Sandy hit the Tri-State area and our data confirms the majority of location-enabled tweets as originating from there as well. Ultimately, our empirical data indicates that users posting Instagram images on Twitter followed a collective temporal pattern, marking particular stages of Sandy: pre-US landfall, US landfall, and aftermath. These stages also correspond to traditional temporal demarcations of hurricanes (Brommer/Senkbeil 2010).

Discussion

For the US, Hurricane Sandy was one of the most devastating storms since Hurricane Katrina. In the days leading to Sandy's US landfall, professional and citizen meteorologists were actively tweeting within Sandy-related hashtags offering forecasts of when Sandy would make landfall and what types of inclement weather the storm would bring. All major news outlets were covering the storm well before US landfall and there was a general climate of anxiety in this coverage. Residents of the Tri-State area responded by stockpiling storm supplies such as clean water, dried food, and changes of clothes. Tweets from official weather services such as the National Oceanic and Atmospheric Administration's Twitter account mixed with unofficial weather reports on Twitter. Often, official and unofficial reports were blurred, inaccurate, or confusing and this caused palpable anxiety amongst Twitter users. It is also unsurprising that humour, a genre generally popular on Twitter and Instagram was deployed as part of a mode to engage with this anxiety.

The role of self and time

Previous studies of Instagram highlight that selfies, images of friends, and images of activities being undertaken are the top three types of images found in Instagram data (Hu, Manikonda, and Kambhampati 2014). These three categories tend to place primacy on representing what an Instagram user is experiencing from their vantage point. The role of 'selfies', the taking of photographs, which include oneself (Tiidenberg 2015), is important to the placement of self within these contexts in Sandy-related images. Images of the outside, including damage to the built environment tended to include people, as our correlations found. This indicates that the experience of placing users within the physical context of the disaster was considered important. Of note, in our findings is that the marked high frequency of humorous tweets occurs before Sandy's US landfall during this heightened state of anxiety, indicating that significant numbers of sampled users were responding to Sandy's forecasted damage through humour. Interestingly, this was simultaneously semi-private and public. The former mostly involved selfies, mentions of friends, and images of friends, family, and local places. The latter primarily involved image macros.

An important finding of this study is that embedded Instagram images in tweets during Sandy often had a co-occurrence of coded categories. For example, images including pets outside were also frequent during Sandy's aftermath as users began to navigate affected areas with their pets. The posting of pet-related also came from people confined at home with their pets providing a diversion as the disaster unfolded. This is an important conclusion in that the images be posted tend to engage with multiple aspects of the lived disaster experience. Another finding is that Instagram images posted during Sandy exhibit identifiable trends before, during, and after Sandy made US landfall. This illustrates that

social media including Instagram images posted on Twitter are sensitive to the distinct phases of disasters. Users are likely synchronously posting their content to social media from their mobile devices. This provides new ways of meaningfully narrating one's disaster experience visually in real-time.

The role of humour and inaccurate/deceptive/misleading images

Though many first-hand experiences are clearly being portrayed, Instagram images posted on Twitter also included significant numbers of inaccurate/deceptive/misleading images, especially on the day Sandy made US landfall (Gupta et al. 2013, Madrigal 2012). Though beyond the remit of this study, we also anecdotally saw a reasonable frequency of these types of images in our data. However, we did not code for this category, but believe that future work in this area would be beneficial for not only understanding when these types of images were circulating (i.e. stage of the disaster), but also their uses and impacts.

The importance of humour within our collected data was an interesting finding. As our case study was based on the convergence between Instagram and Twitter, we were able to see humour play into this cross-platform space and become networked into Twitter via hashtags. For example, some image macros featured celebrities (e.g. Sean Bean – Eddard Stark from *Game of Thrones*) and had hashtags on Twitter such as #GameOfThrones; #ImminentNed (the name given to this macro); and #BraceYourself. SpongeBob SquarePants image macros eventually joined this pre-landfall stream via the fictional character Sandy Cheeks and SpongeBob-related hashtags such as #sandycheeks parodying Sandy were appended. Other images played to the humorous hashtag '#frankenstorm' as the storm took place around Halloween in the US and image macros included clips from the 1931 original *Frankenstein* film. Many images also parodied the character Sandy from the classic hit movie/musical *Grease* and also incorporated humorous hashtags. Even on the day of US landfall, we expected to find more images chronicling the destruction and impact on the built environment or of trees falling down, humorous images were more frequent, highlighting a different, lighthearted, representation of their disaster experiences. (That being said, images of damage caused by the disaster were the most frequent images across the entire data set).

Implications for visual social media research

The landscape for conducting visual social media research is always a moving target. This is partially why research has often avoided interpreting visual content. In addition, image-oriented social media platforms such as Instagram and Snapchat have restricted or inaccessible APIs. This case study was conducted at a unique point in time when Instagram content was being integrated with Twitter feeds and researchers could easily collect these data by accessing the Twitter API. After its acquisition by Facebook in April 2012, Instagram began a slow divorce

with Twitter shutting off friend finding functions in July 2012 and ultimately completely shutting off image sharing into Twitter feeds in December, 2012. Therefore, just a month after Sandy, Instagram and Twitter fully parted ways. Perceived commercial competition in the social media market and its acquisition by Facebook likely influenced Instagram's decision to close its API except for paid access in June, 2016.

Through this empirical case study, we have argued that part of considering the politics of big data in regards to visual social media involves us taking stock of the real external challenges to these types of research, both in terms of stricter API access as well as having to innovate research methods to overcome these difficulties. But, a payoff of this is that alternative narratives that emerge may help tell different stories to those portrayed by professional journalists, relief organizations, or institutional bodies. It is important for these voices and stories to be represented alongside mainstream accounts in order to provide a more holistic narrative of the disaster experience, that includes humour, pets, and even drunken parties before Sandy made landfall. However, social researchers may be placed in the situation of having to purchase visual social media data to accomplish these ends and this has real implications for what types of researchers can actually undertake these studies.

Conclusions

This article uses a case study of geolocated Instagram images shared on Twitter immediately before, during, and after Hurricane Sandy to explore the behavioural dynamic of disaster-related image sharing behavior during the storm. We studied the content of 11,964 images and explored how and what types of images users shared during the storm. Hurricane Sandy-related images were posted at a rate of 10 per second on Instagram (Laird 2012). Our case study is unique in that Sandy is perhaps the first major US disaster where Instagram played a major role in victims' social experience, and Instagram images were integrated with Twitter. This case study enabled a large set of images to be collected that were networked images, crossing both Twitter and Instagram. They became part of larger Twitter and Instagram discourses and these converged data were collected prior to Instagram tightening and eventually closing down its application programming interface (API). Most social media platforms are walled gardens for commercial reasons and this case study reveals how users are behaving in a somewhat convergent social media space. In addition, posted image volume may have been higher as a product of this confluence.

Images perform multiple tasks during a disaster. By coding images, we were able to better understand some of the multiple functions they performed. These images emphasize the human side of Sandy and how people experienced the disaster first hand. The high frequency of images with both food and drink as

well as using animals in humorous images illustrates individuals, in some ways, carried on with our daily lives even in the face of large-scale destruction. However, empirical studies of what types of images are being produced and consumed remain limited. This is attributable to both the challenges of obtaining visual social media data (such as Instagram's closing of its API) to the difficulty of interpreting social media (e.g. the lack of easy content classification). Because the use of images during disasters has implications ranging from emergency services provisioning to understanding people's varied disaster experiences to understanding user content and practices, it is important to develop ways of tackling these obstacles.

Unlike other work on disasters, our study finds that in the case of Hurricane Sandy, many Instagram images posted on Twitter reflect the vantage point of disaster victims rather than official responders as Mills et al. (2009) found in previous disasters. This is likely due to users posting images right as Sandy made US landfall. In this sense, we found that time is a crucial variable. Specifically, pre-social media reporting usually defined the narrative of disasters, but it was often difficult for journalists to get to the scene first. With contemporary social media platforms, everyone is potentially a citizen journalist and time becomes critical to what images end up representing the disaster. Images can circulate virally on social media before even a single reporter arrives at the scene. This is not only true in highly visible types of damage like the façade of a Chelsea apartment that had been ripped off, but also is true in more banal disaster experiences such as electricity shortages. In the case of the latter, images of New Yorkers stringing together extension cords to charge their phones and laptops helps us see Sandy through the eyes of those experiencing the disaster first hand. This potential shift of representation to social media users away from journalists is significant.

It is also clear that these images provide a sort of counter-narrative to mainstream media accounts. By studying these large sets of images, we are able to potentially help tell a different part of the Sandy narrative that may also provide more humanizing and holistic accounts. For example, these images tell a story of how those in the New York area were anxious prior to Sandy's landfall and threw Sandy parties, met with friends to have dinner or grab drinks. Then when Sandy made US landfall, there was major damage, but disaster victims also took their pets outside, took selfies in front of smashed cars, and struggled with power, water, and food shortages. Social researchers should pursue work to give voice to these alternative depictions of disasters and other crises.

Our findings also highlight that Twitter's use – in all phases of Sandy – is generally not geared towards interacting with major relief organizations such as the Red Cross. Rather, the medium affords different uses. For example, the ease by which users can share images and check into locations via Foursquare made it seamless for users to not only continue these routine habits, but actually perhaps increase these behaviours – especially photo sharing. Disaster sociology literature emphasizes how our use of the word 'disaster' refers to 'a distinct event

that disrupts the accustomed flow of everyday life' (Erikson cited in Bruhn 2011, 112). However, our findings suggest that social media-based routines are not particularly interrupted, though they do have different subjects for photos and locations for check-ins. These findings help highlight the importance of critical empirical-based big data work to better understanding what latent narratives may exist within visual social media data. Our work also reveals the potential richness of what can be derived from visual social media data, while simultaneously highlighting the implications of these types of work. Historically, disasters were understood through top-down perspectives, but large volumes of social media data can change the politics of representing disasters. Specifically, visual social media data during disasters can reflect a 'vernacular' (Murray 2013) understanding of the disaster, which rather than glossing over humour, pets, and other everyday experiences, draws our attention to them.

Acknowledgements

The authors are grateful for the generous feedback received by reviewers. We also benefited from questions and comments on a work in progress version of this article presented at the 6th International *Social Media & Society* conference held in Toronto, Canada, July 27–29, 2015.

References

- boyd, danah/Crawford, Kate (2012): "Critical Questions for Big Data." In: *Information, Communication & Society* 15/5, pp. 662–679.
- Brommer, David M./Senkbeil, Jason C. (2010): "Pre-landfall evacuee perception of the meteorological hazards associated with Hurricane Gustav." In: *Natural hazards* 55/2, pp. 353–369.
- Bruhn, John G. (2011): *The Sociology of Community Connections*, Berlin-Heidelberg: Springer.
- Burgess, Jean/Vis, Farida/Bruns, Axel (2012): "Hurricane Sandy: the most tweeted pictures." In: *The Guardian*, November 6 (<https://www.theguardian.com/news/datablog/gallery/2012/nov/06/hurricane-sandy-tweeted-pictures>)
- Chouliaraki, Lillie/Blaagaard, Bolette (2013): "Special Issue: The ethics of images." In: *Visual Communication* 12/3, pp. 253–259.
- Cook, James (2016): "Instagram made a change that stopped lots of third-party apps from working." In: *Business Insider*, 4 June (<http://www.businessinsider.de/instagram-made-a-change-that-stopped-lots-of-third-party-apps-from-working-2016-6>)
- Feng, Ling et al. (2015): "Competing for attention in social media under information overload conditions." In: *PloS one* 10/7.

- Gupta, Aditi/Lamba, Hemank/Kumaraguru, Ponnurangam/Joshi, Anupam (2013): "Faking Sandy: characterizing and identifying fake images on twitter during hurricane Sandy." In: Proceedings of the 22nd international conference on World Wide Web, pp. 729–736.
- Harfoush, Rahaf (2009): *Yes We Did! An inside look at how social media built the Obama brand*, San Francisco: New Riders.
- Hochman, Nadav/Schwartz, Raz (2012): "Visualizing instagram: Tracing cultural visual rhythms." In: Proceedings of the Workshop on Social Media Visualization (SocMedVis) in conjunction with the Sixth International AAAI Conference on Weblogs and Social Media (ICWSM–12).
- Hogan, Bernie/Quan-Haase, Anabel (2010): "Persistence and change in social media." In: *Bulletin of Science, Technology & Society* 30/5, pp. 309–315.
- Hu, Yuheng/Manikonda, Lydia/Kambhampati, Subbarao (2014): "What We Instagram: A First Analysis of Instagram Photo Content and User Types." In: Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media.
- Hughes, Amanda Lee/Palen, Leysia (2009): "Twitter adoption and use in mass convergence and emergency events." In: *International Journal of Emergency Management* 6/3, pp. 248–260.
- Instagram (2016): "Press News." September 02 (<https://www.instagram.com/press/>).
- Kaisler, Stephen/Armour, Frank/Espinosa, Juan Antonio/Money, William (2013): "Big data: Issues and challenges moving forward." In: 46th Hawaii International Conference on System Sciences (HICSS).
- Kitchin, Rob (2014): "Big Data, new epistemologies and paradigm shifts." In: *Big Data & Society* 1/1.
- Kuipers, Giseline (2005): "'Where Was King Kong When We Needed Him?' Public Discourse, Digital Disaster Jokes, and the Functions of Laughter after 9/11." In: *The Journal of American Culture* 28/1, pp. 70–84.
- Laird, S. (2012): "Instagram Users Share 10 Hurricane Sandy Photos Per Second." In: Mashable (<http://mashable.com/2012/10/29/instagram-hurricane-sandy/#AOTCvPj4Zkqu>).
- Lister, Martin (2013): *The photographic image in digital culture*, London: Routledge.
- Lupton, Deborah (2015): *Digital sociology*, London: Routledge.
- Madrigal, Alexis C. (2012): "Sorting the Real Sandy Photos From the Fakes." In: *The Atlantic*, Oct. 29.
- Manovich, Lev (2012): "Trending: the promises and challenges of big social data." In: Matthew K. Gold (ed.): *Debates in the digital humanities*, Minneapolis: University of Minnesota Press, pp. 460–475.
- Mendoza, Marcelo/Poblete, Barbara/Castillo, Carlos (2010): "Twitter Under Crisis: Can we trust what we RT?" In: Proceedings of the first workshop on social media analytics.

- Mills, Alexander/Rui, Chen/Lee, Jin-Kyu/Rao, H. Raghav (2009): "Web 2.0 emergency applications: How useful can Twitter be for emergency response?" In: *Journal of Information Privacy and Security* 5/3, pp. 3–26.
- Milner, Ryan M. (2013): "Pop polyvocality: Internet memes, public participation, and the Occupy Wall Street movement." In: *International Journal of Communication* 7/34.
- Montgomery, Kathryn C. (2015): "Youth and surveillance in the Facebook era: Policy interventions and social implications." In: *Telecommunications Policy* 39/9, pp. 771–786.
- Morgan, Elizabeth M/Snelson, Chareen/Elison-Bowers, Patt (2010): "Image and video disclosure of substance use on social media websites." In: *Computers in Human Behavior* 26/6, pp. 1405–1411.
- Murray, Susan (2013): "New media and vernacular photography: Revisiting Flickr." In: Martin Lister (ed.): *The photographic image in digital culture*, London: Routledge, pp. 165–182.
- Murthy, Dhiraj (2013): *Twitter: Social communication in the Twitter age*, Chichester: John Wiley & Sons.
- Murthy, Dhiraj (2016): "The ontology of tweets: Mixed methods approaches to the study of Twitter" In: Luke Sloan/Anabel Quan-Hasse (eds.): *The SAGE Handbook of Social Media Research*, London: Sage.
- Murthy, Dhiraj/Bowman, Sawyer A. (2014): "Big Data solutions on a small scale: Evaluating accessible high-performance computing for social research." In: *Big Data & Society* 1/2.
- Murthy, Dhiraj/Gross, Alexander J. (2016): "Social media processes in disasters: Implications of emergent technology use." In: *Social Science Research*, in press, doi: <http://dx.doi.org/10.1016/j.ssresearch.2016.09.015>.
- Naaman, Mor/Becker, Hila/Gravano, Luis (2011): "Hip and trendy: Characterizing emerging trends on Twitter." In: *Journal of the American Society for Information Science and Technology* 62/5, pp. 902–918.
- Rashtchian, Cyrus/Young, Peter/Hodosh, Micah/Hockenmaier, Julia (2010): "Collecting image annotations using Amazon's Mechanical Turk." In: *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk*, Los Angeles, California.
- Ringrose, Jessica/Harvey, Laura (2015): "Boobs, back-off, six packs and bits: Mediated body parts, gendered reward, and sexual shame in teens' sexting images." In: *Continuum* 29/2, pp. 205–217.
- Ruppert, Evelyn (2015): "Big data economies and ecologies." In: Linda McKie/Louise Ryan (eds.): *An End to the Crisis of Empirical Sociology? Trends and Challenges in Social Research*, Abingdon-Oxford: Routledge, pp. 13–28.
- Seltzer, Emily K./Jean, N. S./Kramer-Golinkoff, Emily/Asch, David A./Merchant, R. M. (2015): "The content of social media's shared images about Ebola: a retrospective study." In: *Public Health* 129/9, pp. 1273–1277.

- Shifman, Limor (2012): "An anatomy of a YouTube meme." In: *New media & Society* 14/2, pp. 187–203.
- Shifman, Limor (2014): "The cultural logic of photo-based meme genres." In: *Journal of Visual Culture* 13/3, pp. 340–358.
- Smyth, Willie (1986): "Challenger Jokes and the Humor of Disaster." In: *Western Folklore* 45/4, pp. 243–260.
- Tiidenberg, Katrin (2015): "Boundaries and conflict in a NSFW community on tumblr: The meanings and uses of selfies." In: *New Media & Society*, pp. 1–16.
- Veinott, Beth/Cox, Donald/Mueller, Shane (2009): "Social media supporting disaster response: Evaluation of a lightweight collaborative tool." In: *Proceedings of NDM9, the 9th International Conference on Naturalistic Decision Making*.
- Verhage, Julie (2016): "Credit Suisse says Instagram is going to have a huge year." In: *Bloomberg Markets*, April 19 (<http://www.bloomberg.com/news/articles/2016-04-19/credit-suisse-says-instagram-will-generate-over-three-times-as-much-revenue-as-facebook-paid-to-acquire-it>).
- Vieweg, Sarah/Hughes, Amanda L./Starbird, Kate/Palen, Leysia (2010): "Microblogging during two natural hazards events: What twitter may contribute to situational awareness." In: *Proceedings of the SIGCHI conference on human factors in computing systems*.
- Vis, Farida (2013): "A critical reflection on Big Data: Considering APIs, researchers and tools as data makers." In: *First Monday* 18/10.
- Wilson, Steven G./United States Bureau of the Census (2012): *Patterns of Metropolitan and Micropolitan Population Change: 2000 to 2010*.
- Xinpeng, L. Liao/Pradip, Chitrakar/Chengcui, Zhang/Gary, Warner (2015): "Object-of-Interest Retrieval in Social Media Image Databases for e-Crime Forum Detection." In: *International Journal of Multimedia Data Engineering and Management (IJMDEM)* 6/3, pp. 32–50.
- Yates, Dave/Paquette, Scott (2011): "Emergency knowledge management and social media technologies: A case study of the 2010 Haitian earthquake." In: *International journal of information management* 31/1, pp. 6–13.
- Young, Alison (2007): "Images in the aftermath of trauma: Responding to September 11th." In: *Crime, Media, Culture* 3/1, pp. 30–48.
- Zappavigna, Michele (2016): "Social media photography: construing subjectivity in Instagram images." In: *Visual Communication* 15/3, pp. 271–292.
- Zhenguo, Yang et al. (2015): "Semi-Supervised Multimodal Fusion Model for Social Event Detection on Web Image Collections." In: *International Journal of Multimedia Data Engineering and Management (IJMDEM)* 6/4, pp. 1–22.

