Managing 'Threats': Uses of Social Media for Policing Domestic Extremism and Disorder in the UK

Project Report

Report Authors: Lina Dencik, Arne Hintz, Zoe Carey, Hina Pandya

Cardiff School of Journalism, Media and Cultural Studies

October, 2015

MEDIADEMOCRACYFUND

FORD FOUNDATION

OPEN SOCIETY FOUNDATIONS

CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

Executive Summary

This project examines the uses of social media for policing domestic extremism and disorder in the UK. The collection and analysis of social media data for the purposes of policing forms part of a broader shift from 'reactive' to 'proactive' forms of governance in which state bodies engage in big data analysis to predict, preempt and respond in real time to a range of social problems. However, there is a lack of research that accounts for the ways in which different state bodies are making use of big data, and how big data is changing the way states research, prioritize and act in relation to social and political issues. Although big data promises for more efficient, rational and objective decision-making, an important emerging body of work highlight that uses of big data for governance may also contribute to forms of suppression, inequality, and discrimination. What is more, whilst the collection of data may provide opportunities to identify problems and potential 'threats', the challenges of oversight, accountability and transparency involved in the collection and use of people's information have been identified as key concerns. This project engages with these debates by looking specifically at how social media data informs decision-making with regards to the policing of domestic extremism and disorder in the context of the United Kingdom.

The research examines two key areas of social media practices for policing: 1) the ways in which social media communication and data becomes identified as potential domestic 'threats' and 2) the ways in which the police engages with social media to manage and minimize those 'threats'. Our aim is to understand the nature of algorithmically-produced intelligence, what aspects of social media data are used to identify domestic extremism and disorder, and how the police actively communicate on social media platforms. In order to explore this we combined qualitative and quantitative research methods, carrying out semi-structured interviews with British police involved in the policing of domestic extremism and disorder together with big data analysis emulating practices of protest policing and analyzing police engagement on social media.

Our research found that the use of social media for policing domestic extremism and disorder is a relatively recent development that still constitutes an emerging practice within British police. However, there is increasing emphasis on the use of so-called Open Source Intelligence (OSINT) that stems predominantly from social media data available to view without bypassing privacy settings. The collection and analysis of this data is perceived to be a more proportionate and fair form of intelligence gathering than other tactics and provides a substantial resource for 'situation awareness', particularly in the lead-up to major events, such as protests and demonstrations. Although this data is considered 'public', police interpretation of the Regulation of Investigatory Powers Act (RIPA) includes a number of restrictions for the collection and retention of this data for policing purposes. This relates particularly to repeated viewings of profiles and the length of time data can be kept by police.

The programmes and tools employed by police for OSINT are predominantly developed by the commercial sector, and often marketing-driven. The police do not develop their own software and are not involved in the design and development of algorithms that produce predictive analytics. Our research found that police uses these

tools in similar ways to what we may be familiar with from marketing and social science research. In particular, for the policing of protests data is filtered and analysed in order to identify: keywords and 'threat words' (e.g. 'guns', 'flares', 'knife'); risk assessment and resourcing (particularly relating to who and how many people will attend); and influencers and organisers (not always clearly distinguished). To a lesser extent, police will also carry out sentiment analysis (mood of the crowd) and geo-location analysis (particular areas of gatherings). Importantly, our research found that big data analysis of this kind is not an isolated practice but is integrated with other forms of police intelligence, such as human intelligence and existing databases. Also, an emphasis on human assessment of any data analysis and 'professional judgement' in interpreting the significance of identified data was prevalent in our research, highlighting the role of discretion in predictive analytics.

Our research found that social media is used by police to inform strategies such as pre-emptive arrests, interception of activities, approaching particular individuals and groups, or change of tactics during events. Police uses of social media to actively engage and communicate with potential 'threats' are much less prevalent. Rather, police engagement on social media is very differentiated across different forces, done on an ad hoc basis, and largely concerns public information announcements or petty crime. The use of big data for identifying community needs and concern for engagement purposes as a way of managing domestic extremism and disorder is as of yet not part of police practice and raises concerns within police about the level of overlap between intelligence and engagement.

Although uses of social media may facilitate possibilities for pre-empting forms of criminality, the research for this project also highlights a number of challenges in the use of big social media data for the purposes of policing domestic extremism and disorder. Assumptions regarding the 'public' nature of social media communication do not consider important questions about the user's intent and the nature of consent in data collection and analysis. Furthermore, the lack of knowledge regarding the algorithms that produce predictive analytics for policing purposes raises concerns regarding the accountability of police tactics employed on the basis of such algorithms. Linked to this, the use of commercial and marketing-driven software for law enforcement needs introduces questions regarding the types of analyses that such software provide and the extent to which such 'knowledge' is suitable for policing purposes. Moreover, the role of human input both in terms of designing algorithms as well as any analysis and interpretation of such data remains central in data-driven policing. Whilst this helps correct the imperfections of the technology, it opens up possibilities for pre-existing human biases to enter predictive policing under the pretense of 'objective' and 'neutral' data analysis. In particular, questions around the interpretation of any unpredictability in predictive analytics as 'risk' can invite unnecessarily extensive forms of intervention by police. Finally, the use of social media by police (re)introduces debates around the nature of state-corporate relations and particularly the role of social media companies in carrying out police functions. This is further complicated by the proliferation of big data analysis being carried out by a host of different non-state actors who are not subject to the same restrictions as the police.

In carrying out this project, we therefore hope to illustrate the need for further discussion and research into this emerging area of police practice.

Disclaimer-Copyright

Table of Contents

List of Tables

List of Figures

Introduction: Big Data and Predictive Policing

The generation, analysis, use and consequences of 'big data' are key concerns in contemporary research as well as public administration, management and, increasingly, public discourse. The digitization of vast areas of everyday life has led to the availability of detailed data about a range of processes and activities, from industrial production to individual health monitoring, and from public finances to inter-personal communication. The analysis and use of this data promises more efficient delivery of public services, a better response to social problems, and better allocation of resources. Based on scientifically-generated and value-neutral information, data may reduce political influences and subjective judgements, and thus may offer a more rational, impartial, reliable and legitimate way of decision-making. Algorithms – automated instructions to process data and produce an output – have become increasingly prominent in the debates on big data as they have the potential to facilitate and enhance its promises. Through seemingly objective data processing they allow for not just the understanding of previous occurrences but for predictive analytics and the foreseeing of future behaviour, facilitating possibilities for pre-emptive action. However, the increasing use of big data comes with serious challenges for data protection and civic rights, such as the right to privacy. Further, selective data collection and use do not necessarily eliminate human partiality and algorithms, typically operated as a technical 'black box', lack accountability and transparency and may conceal both old and new forms of discrimination.

In the context of this emerging debate, the project 'Managing Threats' investigates how data analysis is used to inform the policing of potential public order 'threats'. Particularly, it explores how algorithms are used to both define and identify what has recently been termed in the UK as 'domestic extremism and disorder'. Originally used for radical animal rights activists, this concept has been expanded in practice and incorporated in police debate, even though it lacks a clear definition. The project thus addresses the question of how data informs policing in the context of a) the broader debate on state uses of big data analysis and b) uncertainty regarding contemporary understandings of dissent.

Big Data, Consequences and Concerns

As more and more aspects of our lives are mediated and organized by digital devices and networked systems, vast volumes of data are generated (Kitchin 2014a). The increased availability of this 'big data' has been celebrated as enabling more efficient public services (e.g., Mayer-Schönberger 2014). However, as we will outline in this section it has also raised a number of concerns.

To start with, the claims to **objectivity**, **impartiality**, reliability and legitimacy of big data and its algorithmic collection and analysis have been questioned (Gillespie, 2011/2014; Elmer et al 2015) and criticized as "carefully crafted fictions" (Kitchin 2014b: 9). As data collection is initiated, and algorithms are developed, by humans, a great deal of expertise, judgement, choice and constraints are reproduced in the data. Algorithms are created for a purpose, typically to identify, sort and classify people, or to collate and categorize processes, and they are therefore highly contextual and contingent. Rather than mere technical tools that represent facts, they have an active quality of shaping how we understand the world – "they are engines, not cameras" (Kitchin 2014b: 11; Mackenzie 2008). As algorithms adjudicate more and more

consequential decisions in our lives, scholars claim that they are "the new power brokers in society" (Mackenzie 2007: 93). Although there has been much discussion of the integration of big data into various forms of governance structures, there is little research which provides an actual account of how different bodies are making use of big data, and how big data is changing the way states research, prioritize, and act in relation to social and political issues (Cook 2014, Bertot et al. 2014, Bhushan 2014, Margetts and Sutcliffe 2013).

Key challenges regarding data collection and analysis relate to **privacy and surveillance**. In contemporary digital environments, 'all manner of everyday activities are recorded, checked, traced and monitored' (Lyon, 2007: 454). Data is often generated by the users of digital communication networks and their devices, processed by commercial intermediaries, and analysed by both commercial actors and state agencies (Trottier, 2015). The Snowden revelations that were initially published in June 2013 have pointed to diverse practices of state surveillance in the digital age (The Guardian, 2015; Fidler, 2015), but the 'big data' generated through social media platforms for commercial profit is at the heart of current surveillance trends (Lyon, 2014). Communication on social media has created an immediate and heightened visibility of social life, both in terms of the volume of communication and the public nature of many of these communications (Trottier, 2015). This is facilitated by corporate services like Facebook and Google operate on the basis of a business model of collecting and analysing user data. Detailed knowledge about user locations, activities, brand preferences and political orientations, as well as those of their friends and networks, is the foundation of their market value, and so they are designed to maximise (corporate) surveillance (Trottier and Lyon, 2012). Users are tracked as they move across the web, required to identify themselves (e.g., through Facebook's 'real name' policy), and subjected to automatic facial recognition.

The '**data mine**' (Andrejevic, 2012: 71) of social media and other commercial internet platforms has raised significant interest by state agencies. Programmes such as Prism allow security agencies to tap into the data collected by internet companies, complementing thousands of official requests for user data by government agencies every day (Google Transparency Report, 2014). This '**grey intelligence**' (Hoogenboom 2006; Walby and Monaghan 2011) based on information shared between public and private entities has created concern with regard to policing, not least as data may be collected before the relevance, use, or the user's role as a suspect have been determined. Open Source Intelligence (OSINT) collected from social media is supposed to be transparent, yet the means of collection and analysis remain obscure (Trottier, 2012/2015). In social media environments, much of the data is made visible by one's peers rather than the user himself or herself, thus blurring the lines between private communication and public data further. Identities can be reconstructed even if a user volunteers only partial data (Gross, 2015). The quasi-public sphere of open and ephemeral social media is further undermined by the indefinite archiving of data and its transformation into public record. Classic concepts that have informed policing in the past, such as probable cause or suspicion, may be less clear in a context of data collection through social media monitoring, and an increased emphasis on preventing rather than prosecuting crime provides challenges for judicial review and oversight of policing practices (Swain, 2013).

The algorithms that are at the centre of big data analysis and that categorize people in order to make predictions about their behavior (as well as recommendations of products, treatments, and courses of action) may replicate classic forms of **discrimination** and establish new categories of differential treatment. An important emerging body of work warns that big data processes may contribute to poverty, inequality, and social exclusion (Eubanks 2014, Boyd et al. 2014, Andrews 2013, Lerman 2013; Pena Gangadharan 2012). Often data is being reduced to the capabilities of old technologies and categories of analysis (Robertson and Travaglia. 2015). Developing algorithms on the basis of staid categories may create self-fulfilling prophecies whereby the targeting of certain groups in the initial analysis raises their visibility in all future calculations while obscuring other forces at play (Edwards 2015).

The 'social media assemblage' (Haggerty and Ericson, 2000) of social media-based data gathering brings together personal information, private platforms, and police organizational cultures. Information that is meant for friends and acquaintances in a user's personal network is combined with commercial information and analysed by police and security services according to their specific needs and frameworks. In addition to the surveillance concerns mentioned above, this raises questions regarding the **representativeness of data** and the **accuracy of predictions** derived from it. Social media websites emerged in the private sphere and users have developed specific cultures in relation to the platform, which are often very different from the cultures, interactions and types of communication found offline (or on other platforms), and investigations that are not rooted in these cultures will likely lead to misinterpretations. Conclusions about offline society drawn from social media thus may not be valid. Furthermore, users with more followers and retweets/shares are more likely to appear in social media samples, thus over-representing certain opinions. Social media platforms also contain limitations to expression that may alter the intended meaning of a user. Typically users do not always represent themselves accurately online, or follow through in real life with what they claim they will do online.

With the increasing digitization of everyday life, our notions of real and representation are becoming blurred between the digital and the analogue (Robertson and Travaglia 2015). While a strict dichotomy between the 'real' analogue and digital representation may not exist, a shift from policing the analogue to policing the digital has important implications for our concepts of selves and society. Social networks are different from personal networks, and the strength of social network ties does not always translate from social media to real life. The **social context** of data generation is thus crucial for its interpretation. Big data is not raw data, it is shaped by the way it is created, collected, stored, and interpreted (Halford 2015). These problems of data representativeness lead to challenges for data analysis, which Lyon (2014) argues are more serious than those related to data sourcing. This demonstrates the importance of understanding how patterns are identified and explicated by state agencies, including the police. Moreover, these challenges speak to a persistent concern regarding the lack of oversight, **accountability**, and transparency involved in the collection and use of (particularly, personal) data.

Predictive Policing

Police have been using information to try to predict risk long before the advent of big data. Defined by Her Majesty's Inspectorate Constabulary (HMIC) as "methods used by police forces to use and analyse data on past crimes to predict future patterns of crime and vulnerable areas," predictive policing differs from preemptive policing, which is primarily focused on preventing crimes and may not make use of predictive technologies (HMIC Report 2014 p.72.). However, big data analysis has restructured predictive policing, adding big data analysis to community-oriented policing, targeted surveillance, and other traditional approaches.

There are two dominant theories informing predictive policing: **near repeat theory**, which includes both flag theory and boost theory, and **risk terrain modeling**. Near repeat theory recognizes that similar crimes often occur in the same location and proposes that once a crime has occurred in a specific location it is statistically more likely that similar crimes will recur (Ferguson, 2012). This theory is especially popular for theft and burglary, as houses on the same street may have the same layout and weaknesses in security and escape routes may be learned and perfected. Flag theory and boost theory are two separate explanations for the phenomenon of near repeat theory. Flag theory states that some properties or locations are 'marked' and signal their vulnerability to observant criminals. Examples may be weak locks or poorly lit entrances. Boost theory assumes that criminals gain information about a location while committing a crime, which makes repeating the crime in the same location easier. Sometimes linked to the Offender as Forager theory, boost theory has also been shown to follow models that track the spread of contagious diseases (Johnson et al 2008). Risk terrain modeling, on the other hand, measures the strength of various risk factors across a larger geographic area and then compiles and weighs all of these to determine the probability of various types of crime occurring in more specific locations within the larger map. Risk terrain modeling can measure the probability of a variety of crimes by focusing on risk factors, whereas near repeat theory can only predict the likelihood of the same type of criminal activity being repeated. Furthermore, near repeat theory can only identify locations that may be susceptible to crime, whereas risk terrain modeling can identify locations and individuals who may be involved in criminal activity (Ferguson, 2012).

Police in the United States began using **big data in predictive policing** in the 1990s, with programmes like **Compstat** in New York City and **Palantir** in Los Angeles. Introduced by Police Commissioner William Bratton, Compstat used historical crime records to map hot spots for various types of crime in New York City. By mapping the specific time and location where various types of crime were likely to occur, police resources could more efficiently be mobilized to prevent, or more rapidly respond to, crime. Compstat was particularly helpful in the policing of 'quality of life crimes', which Commissioner Bratton targeted as precursors to more serious offenses under the Broken Windows model of policing (Kelling and Bratton, 1998). While Compstat and many other predictive policing programs focus on hotspots of criminal activity, the software Palantir predicts the likelihood that a known suspect or individual with a criminal record will commit a crime. Relying primarily on police car-mounted license plate readers, the license plate scans are linked with numerous disparate datasets - such as, known associates, cell phone numbers, and arrest records - at the LAPD Real Time Analysis and Critical Response Division (Kelly, 2014). This

shift from predicting hot spots to tracking individual activity is significant, not least regarding its privacy implications.

Other uses of big data analytics in predictive policing in the United States have included: enhancing the identification of slumlords and unsafe building conditions (NYC); honing in on businesses selling bootlegged cigarettes and pharmacies that over-distribute oxycontin (NYC); locating illegal dumping or waste disposal (NYC); uncovering instances of business license flipping (NYC); reducing property crimes by identifying the times that specific locations are at greater risk (Santa Cruz, CA); and outfitting officers on the street with real-time access to all data collected by the police force (Memphis) (Howard, 2012; Koehn, 2012; Badger, 2012). The use of big data predictive policing increased precipitously following the terrorist attacks on September 11, 2001 in New York City and subsequent terrorist attacks on a commuter train in Madrid in 2004 and the public transit system in London in 2005. In the United States and the United Kingdom, national security threats and the fear of terrorism created an environment generally more open to forms of surveillance that were previously considered unlawful. While the United States set a precedent for the analysis of big data for predictive policing, the United Kingdom has been at the forefront of surveillance technologies. The system of Closed Captioned Televisions (CCTV) served as a model for the Domain Awareness System developed by the NYPD and Microsoft, which established a network of private and police surveillance cameras in lower Manhattan. The UK Police have also been pioneers in social media surveillance. Expanding the scope to the entire European Union, Trottier found police monitored social media for a variety of practices, including identifying terrorists, assessing public opinion after environmental crises, identifying instances of child exploitation, tracking illegal protests, and cases of copyright infringement (Trottier, 2015).

Algorithms are designed to recognize patterns in large data sets and are thus essential to the analysis of big data. Existing predictive policing algorithms consider and weigh a variety of factors in determining, for example, the likelihood of crime in a certain area. Some algorithmic factors that are currently used include event-based concerns (frequency of arrests, emergency phone calls, incident reports, and complaints); place-based concerns (known addresses of criminal suspects, locations of gang activity, places where crime is common); the types of crime that are typically reported (violent, property); information about individuals (suspects, convicted criminals, individuals with links to criminal networks), gang activities, traffic patterns, and environmental factors (poor lighting, lack of police surveillance, easy escape routes, infrequent pedestrian traffic, etc.). Police forces often collaborate with **academia** to pool expertise and develop or improve predictive algorithms. While collaborating with the Santa Cruz police department, for example, mathematician George Mohler from UCLA analysed crime maps which resembled seismic maps of aftershocks following earthquakes (Moehler et al, 2011). Mohler's model predicted the location of property crimes, and burglaries dropped by 19% the year after it was introduced in Santa Cruz (Kelly, 2014). Mohler subsequently founded the company PredPol, marketing his predictive software under the same name and expanding the types of crimes that PredPol can predict. After trial runs in Greater Manchester, Kent, West Midlands, West Yorkshire, and the Met Police, HMIC recommended in 2014 that police forces throughout the United Kingdom adopt similar programs to improve efficiency (Jones, 2014). Today, Promap is the primary prospective crime mapping

tool in the UK (Edwards, 2015). Burnap and Williams (2015) have developed an algorithm to predict when hate crimes are likely to follow trigger events. Benefitting from the immediate response to trigger events permitted by social media, they collected and analysed tweets following the murder of drummer Lee Rigby by Islamic extremists. Based on their findings, they developed an algorithm to predict when hate crimes might occur. The software filters the grammatical and linguistic variation that often limits textual big data's usefulness to policymakers.

As used in this project, the Twitter analysis tool **COSMOS** – developed by researchers at Cardiff University – applies classic conversation analysis methods to investigate meanings and networks of Tweets. In particular, it focuses on three analytical concepts: membership categorization devices (MCDs), membership categories (MCs), and category bound activities/attributions (CBAs). MCs are often titles or personal categories, like 'mother' or 'police officer'. MCDs are larger frameworks in which a particular MC has meaning, so the MC 'mother' makes sense in the MCD 'family' and the MC 'burglar' has meaning in the MCD 'criminal justice system'. CBAs are actions, attitudes, or other attributions that are linked to specific MCs and make sense within the MCD. Thus the CBA of breast feeding can be logically linked to the mother if she is breastfeeding her child (another MC in the MCD family), and the CBA of arresting a suspect makes sense for a police officer operating within the criminal justice system. The COSMOS algorithm searches for tweets where an MC and CBA co-occur, thus turning the event that prompted the tweet into the MCD and allowing researchers to understand how natural language practice and sentiment may be communicated through tweets (Williams et al, 2013).

Policing Protests and Public Order

The logic of predictive policing was developed to deter criminal activity but becomes more controversial when applied to protests and legal demonstrations. Outside digital environments, police have adopted forms of predictive (or preemptive) action such as disrupting protests by implementing checkpoints and searches that discourage attendees from participating, or using preemptive arrest and 'kettling' (containing a crowd within a limited area) to upset the network of organizers by removing strategic influencers (Swain 2013). While some OSINT is helpful in targeting key figures and establishing barriers that disrupt the flow of demonstrators to a protest, these tactics are also a source of intelligence. Often, names and addresses of detainees are recorded and photos are collected during the demonstration. After the event, personal profiles are developed, including "habits, lifestyle, modus operandi, addresses, places frequented, family-tree chart, photographs, risks to public, ability to protect him/herself, and related information" (National Intelligence Model, quoted in Swain, 2013). These tactics have developed from a longer history of protest policing as the below table indicates (see Vitale, 2006):

Name	Time / Location	Description
Escalated Force	1960s – 1970s in USA	Militancy of protesters met and surpassed by militancy of police
Negotiated Management	1980s, became dominant in the USA and Europe	Effort to avoid violence through cooperation, protection of free speech, limited arrests, limited force, and a tolerance of community disruption
Command and Control	New York City; Million Youth March (1988), Matthew Shepherd Emergency Demonstration (1989), World Economic Forum (2002), Anti-War Rally (2003)	Linked to "broken windows" model of policing. Little discretion granted to police on the ground. Demonstration carefully orchestrated by police without cooperation with demonstrators. Punitive arrests.
Miami Model	Free Trade Area of the Americas protests (2003)	Linked to "paramilitary policing" – surveillance, denial of permits, deployment of defensive equipment and specialized police forces, non-lethal force used on demonstrators, and preemptive arrests.

Table 1: History of protest policing

Social media has proven an important resource in disseminating information in the early phases of disasters and crises. Many social movements make use of social media to coordinate demonstrations and protests, perhaps most notoriously in the Arab Spring uprisings (Khamis and Vaughn, 2011; Diamond and Plattner, 2012; Allan 2013). **Social movement use** of social media in times of crisis can also be tapped to identify and publicly shame rioters in digital platforms as further outlined below. Police can apply all of these civilian uses of social media in regulating protests – they monitor real-time tweets for comprehensive information as events unfold and to track the movements of demonstrators, they use social media shaming to identify rioters, and they broadcast information, instructions, and available resources to affected communities.

Electronic intelligence gathering about protests can start with the monitoring of phone calls and text messages. At an anti-fascist protest in Dresden, Germany, in 2011, police monitored the geolocation of calls and texts to digitally situate individuals suspected of public order offenses. To do this, they had to monitor the radio cells that relayed phone calls and sms, and to lay this digital grid over the actual grid of the city to trace protesters in real-time, as well as to collect evidence against them should further breaches of the law occur (Paasche, 2013). This practice required monitoring everyone's calls and texts within the grid (including inhabitants and passers-by). The near ubiquity of social media as a means of communication has significantly improved the ability of police to monitor real-time communications. With its open

platform, Twitter has become a particularly useful source for both researchers and police to monitor social phenomena, and the brevity of tweets and the hashtag, retweet, and mention functions facilitate the rapid mapping of information and events as they unfold.

In the UK, the 2011 riots served as a starting-point for a more systematic analysis of social media feeds in public order situations. Several research projects collected tweets related to the riots and mapped the flow of information as well as user types, tone of message, content, etc. (e.g., Procter et al, 2013a). A preliminary analysis showed that Twitter was integral, in particular, to the mobilization of cleanup campaigns after the riots, contradicting the largely negative portrayal of Twitter as an accelerant and enabler of the riots in the media and among politicians. Analysis of how police used social media during the riots revealed that the types of tweets differed for local Twitter accounts and larger, regional accounts, with local Twitter accounts providing more situation reports and larger accounts responding to civilian information or enquiries (Procter et al, 2013b). Police appeared to use Twitter primarily as a tool for broadcasting facts, information about available resources, and instructions to avoid risk.

Further research of Twitter messages has included, for example, research on the size and survival of information flows following the terrorist attack in Woolwich, London on May 23, 2013. Burnap et al. (2014) found that high tension tweets had a lower survival rate than low tension tweets, which matches findings that antagonistic tweets are less successful in creating an information flow. A similar study on verbal racial attacks during a football match in 2011 led to the development of an algorithm for the assessment of tension on social media. The development process involved police officers as human coders due to their knowledge of assessing tension in real life communities. Inter-coder agreement (Krippendorf's alpha coefficient) for these coders were particularly high, compared to other coders (Burnap et al 2015), which may point to the coders' expertise or, alternatively, to a common perception and understanding. The algorithm may thus be an example for how such understandings are incorporated in software development. The algorithm proved to be more accurate than SentiStrength, an algorithm that the authors noted could easily be adapted to the study of tension in policing.

While Twitter may be the easiest way to monitor real-time communication at protests, police have also used other social media to aid in policing, often turning to Facebook to gather information and evidence after a crime has occurred. During riots in Canada following a hockey game, Facebook groups emerged to share photos and videos of the riots and to thereby identify and publicly shame rioters. Facebook users saw public naming and shaming as an alternative method to holding rioters accountable without resorting to formalized police intervention and legal sanctions. While police were criticized for not making better use of social media to trace and predict the riots, they did use these groups to identify rioters after the event. Controversially, they sometimes presented names and judgments of guilt before identifying what criminal acts potential rioters had participated in (Trottier, 2012).

As such, police increasingly incorporate social media analysis in predictive policing, and they are expanding this practice from crime investigations to protest and public order situations. Yet a growing range of academic research is critically reviewing the

nature and assumptions of big data, and is pointing to a number of challenges. It is within this context that we are looking to examine the uses of social media by police for policing domestic extremism and disorder in the UK in order to consider the implications of these practices and possible recommendations for the way forward.

Case study

This project analyses social media data collection and use by the British police, particularly pertaining to the policing of domestic extremism and disorder. The focus of this case study is informed by the establishment of the National Domestic Extremism and Disorder Intelligence Unit (NDEDIU), previously the National Domestic Extremism Unit, which was created following a merger of the National Public Order Intelligence Unit (NPOIU), the National Domestic Extremism Team (NDET) and the National Extremism Tactical Coordination Unit (NETCU), and was placed under the lead of the Metropolitan Police Service's Counter Terrorism Command in 2011. The lead-up to establishing domestic extremism units within the police came from a period of militant animal rights campaigning in the late 1990s and 2000s in the UK, particularly aimed at targeting animal testing laboratories. Up until then, intelligence gathering and policing of animal rights activists had been done through regional forces Special Branch units. In 2001, a new unit was set up within the National Crime Squad to police 'animal rights extremism'. This was followed with increased focus on forms of militant activism that stretched beyond animal rights activists to take in other protest movements by placing the National Public Order Intelligence Unit under the remit of policing 'domestic extremism' in 2004. Further restructuring in the years that followed led to the eventual creation of the NDEDIU that had particular emphasis on gathering and understanding intelligence around domestic extremism in order to combine that with prevention and enforcement in the policing of domestic extremism and strategic public order issues in the UK.

The term 'domestic extremism' was controversial and ambiguous from the outset and continues to be so today; it is most frequently described as 'serious criminality'. It is intended to refer to forms of extremism that pertain to domestic policy as opposed to, for example, extremist views related to Islamist fundamentalism. However, as the policing of domestic extremism has increasingly been placed under the remit of counter-terrorism units which have proliferated at national and regional levels in the last few years, these distinctions are not always clear. Moreover, in recent years the UK government has repeatedly foregrounded concerns with forms of 'extremism' and, particularly since the election of a Conservative government in 2015, it has been actively expanding the meaning of extremism to include both violent and non-violent extremist 'ideology', framing this in terms of 'values'.[1] This has created further ambiguity around how forms of extremism are defined and distinguished, expressed also from within the police (cf. Dodd 2014). The counter-terrorism programme CONTEST that has been advanced over recent years include strategies such as PREVENT, for example, developed by the Home Office with the primary aim 'to reduce the threat of terror preventing people from being drawn into it, and given advice and support, whilst responding to those who promote the ideological face of terrorism.' Although PREVENT is not exclusively a police programme and incorporates sectors such as health services and education, it operationalizes a number of PREVENT officers situated within police run counter terrorism units. The programme is implemented primarily in areas of 'known domestic extremist activity' based on previous arrests. Importantly, some media reports have illustrated the

[1] See for example Prime Minister David Cameron's speech in July 2015 announcing of a new anti-extremism bill: http://www.independent.co.uk/news/uk/politics/david-cameron-extremism-speech-read-the-transcript-in-full-10401948.html

ambiguous meaning of 'domestic extremism' under the counter-terrorism framework in which, for example, police presentations on preparing for terror threats obtained through Freedom of Information requests have included pictures of Occupy protestors in the same context as al-Qaida and IRA under the heading of 'domestic extremism'.[2] This has drawn criticism from civil society organisations such as the police watchdog Netpol, who has argued that these types of occurrences are the 'result of including ill-defined labels, like 'domestic extremism', within the language and strategies of counter-terrorism…Programmes like the government's Prevent strategy overwhelmingly target and stigmatise Muslim communities, but…they also provide plenty of scope to include almost any group of political activists that the police dislike or consider an inconvenience.' (Kevin Blowe, a co-ordinator of Netpol, quoted in Quinn 2015)

Following the riots that took place in London and elsewhere during the summer of 2011, the police actively sought to incorporate the uses of social media for policing domestic extremism and disorder to a much greater degree than they had previously done. NDEDIU is reported to have a team of 17 people working in SOCMINT (Social Media Intelligence) as part of a strategy to investigate trends across social media (Wright 2013). The collection, engagement and uses of social media for policing purposes falls under the regulatory framework of the Regulation of Investigatory Powers Act (RIPA) from 2000. Within the RIPA policy framework, the Data Retention and Investigatory Powers Act 2014 (DRIPA), and the DRR Data Retention regulations from 2014 provide specific updates on data-based investigations. However, this regulatory framework has been widely criticized for being ill-suited for the current digital age, described by David Anderson QC, commissioned to review current terrorism legislation, as 'incomprehensible and undemocratic'. It therefore provides little governance for this growing use of social media in policing domestic extremism and disorder.

The objectives of this project are therefore to explore two key areas of social media practices for policing: 1) the ways in which social media communication and data becomes identified as potential domestic 'threats' and 2) the ways in which the police engages with social media to manage and minimize potential domestic 'threats'. The project explores the following three research questions:

1) What data is being collected from social media platforms by police, particularly with regards to policing domestic extremism and disorder?
2) How does social media data collection and analysis inform 'predictive policing', and with what consequences?
3) When and how does British police engage in social media activity to address 'threats'?

Our aim is to understand the nature of algorithmically defined 'threats', what aspects of social media data are used to identify domestic extremism and disorder, and how the police actively communicate on social media platforms. In doing this we are particularly concerned with implications for policies regarding the policing and managing of protest and dissent, both offline and online, and the potential

[2] See http://www.theguardian.com/uk-news/2015/jul/19/occupy-london-counter-terrorism-presentation-al-qaida

discriminatory and suppressive consequences of data analysis to silence and pre-empt protest. Moreover, the project explores implications for civil rights, issues around personal and public data, due process and democratic practice.

The research for this project builds on the 18-month project 'Digital Citizenship and Surveillance Society' funded by the Economic Social and Research Council (ESRC) which analyses the implications of the Snowden leaks for state-media-citizen relations in the UK. One of the four strands of this project concerns responses of civil society and activism to the reality of mass surveillance. Key debates emerging from the research include the use of social media data for surveillance and management of political activism. The case study pursued here complements this ongoing research by addressing both the monitoring of, and engagement with, citizens' social media activity by the police.

Methodology

To explore our research questions, we combined qualitative and quantitative research methods in order to gain an in-depth and practitioner-oriented understanding of police practices. In particular, we sought to combine semi-structured interviews with British police involved in the policing of domestic extremism and disorder together with big data analysis. The sample for the interviews consisted of 5 senior members of the British police force. We initially planned to conduct all five interviews with members of the National Domestic and Extremism Intelligence Unit (NDEIU) but our research process indicated that this approach would be limited by the structure and covert nature of the unit. We therefore decided to broaden our sample to include senior members of the police force who are involved with the use of social media for policing domestic extremism and disorder in other capacities. This allowed us to explore our research questions from both an intelligence and engagement angle, as well as at different levels and across different bodies and units within the British police, all related to social media practices and domestic extremism and disorder. Our final sample therefore consisted of:

- Head of Open Source and Social Media, National Counter Terrorism Police Functions Command [Interviewee A]
- Head of Digital Engagement at the College of Policing [Interviewee B]
- Previous Head of NDEDIU and now the Chief Officer Lead for the National Police Co-ordination Centre (NPoCC) [Interviewee C]
- Head of the Communications Data Investigators team [Interviewee D]
- Regional Prevent Officer leading a social media taskforce [Interviewee E]

All these interviews were organized and conducted within the short time-frame of the project and were carried out, in person, lasting on average around 90 minutes, during August and September, 2015. The interviews were structured around a number of key themes aimed at exploring the uses of social media for policing, both from an intelligence and engagement perspective. These included: 1) Purposes and treatment of social media data; 2) Types of analysis conducted with social media data; 3) Developments of algorithms and software; 4) Operationalising of analyses for police strategies and tactics; 5) Purposes and nature of engagement by police on social media. To focus the interviews on concrete police practices, the interviews were particularly concerned with the ways in which social media is used for policing protests. Doing these interviews provided us with a number of key insights into the

operations of the police and the ways in which social media is integrated into police practices. The level of access we were granted to very senior members of the police meant that we were able to explore the broader rationale, visions and challenges in using social media for policing purposes. By including an interview with a regional officer in our sample, we were also able to explore some of the issues that police 'on the ground' may face on a day-to-day basis when implementing uses of social media for policing domestic extremism and disorder.

Secondly, we supported this research with our own social media data analysis. We used big data analysis software to explore two different dimensions of police practices: a) based on prior research on social media intelligence and policing as well as information from our interviews, we emulated the practices of police based on collection of social media data in the lead up to protests and public order events in order to experience and examine potential challenges and issues with predictive analytics and algorithmic definitions of extremism and threats; and b) we examined if and how police engages online with potential 'threats', particularly with regards to the prevention of extremist and disorder activity by collecting social media data from police accounts and analyzing the nature of communication and the concerns and people they engage with.

In order to operationalize these methods, we used a combination of different tools. For part a) we used the big data analysis tool COSMOS which was developed by Cardiff University to analyse social media in the contexts of human safety and security. Focusing on Twitter activity, we collected and analysed data based on a number of relevant hash-tags and keywords in the lead-up to and during 3 major protests and demonstrations that took place in London during the time-frame of the project: 1) Anti-Austerity March on 20 June 2015; 2) the Anti-DSEI (Arms Fair) protests during September 2015; and 3) the Refugees Solidarity March on 12 September 2015. Using COSMOS, we then analysed the data, looking for potential threat-words, organisers, networks and influencers, as well as sentiment and geospatial clustering (full details of this are in the report chapters).

For part b) we collected data from active police accounts based on information from interviews as well as geopolitical information about arrest rates for domestic extremism in different regions (see details of full sample in the report chapters). Data from these accounts were collected and analysed using COSMOS, identifying keywords used, users communicated with, and prominent locations, and mapping retweet and mention networks. This analysis was supplemented by further analysis using the software Twitonomy for a smaller sample based on volume of tweets and followers aimed at further identifying hashtags and users engaged with, as well as platforms used for the purposes of engagement (full details of sample and analysis are in the report chapters).

Finally, we supplemented our interview and big data research with a dedicated one-day workshop called 'Social Media Intelligence and Policing Domestic Extremism and Disorder in the UK' with key stakeholders that included important civil society organisations involved in the area, such as Netpol, Privacy International, Open Rights Group, Greennet, and Article 19, senior members of the Metropolitan Police and the College of Policing, and leading scholars with expertise from a range of different perspectives. The purpose of this workshop was to discuss and debate relevant issues

with key stakeholders and practitioners as a form of engaged research practice that would help inform our findings and recommendations.

Limitations and Ethical Considerations

Overall, the combination of methods employed for this study and the level of access to police that we were able to obtain allows us to provide some much-needed evidence on the uses of social media by the police, particularly for the policing of protests, and renders transparent what are often hidden and obscure practices involved in policing forms of domestic extremism and disorder. However, it also needs to be recognized that for a project of this nature, our findings are informed primarily by what we are told in interviews and does not include other ways of examining police uses of social media that may either confirm or contradict the reflections from the police that we have included in this study. Although our sample includes a cross-section of key actors within the British police force, a larger sample may provide more representative reflections on police practices. Also, other methods of examining police uses of social media, such as through analysis of evidence from court cases, observation and ethnography that have been carried out in other contexts (cf. Swain 2015 and Schäfer 2014) may further inform our understanding of how social media is used for policing domestic extremism and disorder.

Furthermore, emulating predictive forms of policing through employing our own big data analysis of social media data in the lead up to major protests provides us with key insights into the opportunities and challenges of operationalizing big data analysis for the purposes of policing in the context of real events. However, without access to the exact tools and software that the police use, we are only provided with a limited sense of this practice and this understanding can only be enhanced by employing further tools and software in our analysis. We have also been limited in this instance by being dependent on getting data from external actors, meaning that time-lines, technical obstacles, and the operationalizing of tools has been outside of our control which has compromised our sample. Employing a multitude of different programmes to collect data, and having these programmes in-house longer-term, would provide a more comprehensive data-set that is more closely representative of the data resources available to the police. This would also require using tools that can incorporate data from a multitude of social media platforms, rather than being limited to Twitter as has been the case in our analysis.

Moreover, doing a study of this nature also includes a number of ethical considerations that have been considered in the research process of this project. Firstly, consent to carry out interviews with police was obtained by all the interviewees on the recordings of the interviews. In the instance of the first interviewee, we were not allowed to record the interview and agreement was made to keep the name of this individual anonymous due to the sensitivity of their role. During the interviews that were recorded, some statements were noted as being ‘off the record’ in that they may include information about police tactics or capabilities that could be considered a risk to reveal. We have sought to consider this by ensuring the confidentiality of any transcripts produced from the interviews and by not including such statements in any works resulting from the research.

Secondly, for our big data analysis, collecting social media data on individuals or

groups relating to a protest may put some subjects at risk, either directly or indirectly. We have considered this by relying on analysis of aggregated data where individuals cannot be identified. Where individuals may be identified, either explicitly or implicitly, we have considered the nature of the account to assess whether referring to these accounts in our analysis may place anyone at risk.

Thirdly, organizing a workshop with civil society organisations, activists and senior members of the police on the topic of social media intelligence potentially involves sharing sensitive information, forms of intimidation and even risk for participating members. We therefore ensured that everyone was properly informed beforehand of who would be present at the workshop and decided to conduct the workshop under Chatham House Rules in which information from the workshop can be used, but no one present at the workshop must be referenced or quoted, either directly or indirectly.

As is standard at Cardiff University, the research process has gone through the Ethics Committee at the School of Journalism, Media and Cultural Studies for approval.

Interviews: Police Practices

In this chapter, we will outline the ways in which police engage with social media for the purposes of policing domestic extremism and disorder. The analysis presented here is based on 5 semi-structured interviews with senior members of the British police force that work with or in relation to social media and the policing of domestic extremism and disorder. The chapter outlines a number of key themes that emerged from the interviews and will provide an overview of the nature of social media practices by the police as they relate to the policing of domestic extremism and disorder, particularly around direct action, protests and activism.

Social media use as emerging phenomena

The use of social media for policing, in all aspects, is an emerging development in the UK that is still relatively recent. Partly attributed to an institutional culture and a demographic make-up 'dominated by 40-plus white males, rightly or wrongly, that haven't grown up on social media' (Interviewee D), integrating social media into broader police practice is still a 'learning curve'. Within the operations of NDEDIU, the use of social media as a regular police practice only began to develop in 2012. In particular, the so-called 'London Riots' which kicked off during the summer of 2011 became the turning point for how the police thought about social media:

'I think we recognized, and the HMIC report subsequently called Rules of Engagement recognized, that we hadn't been good at all on social media. We hadn't even thought about it properly, and that's probably to do with the police service at that point and there's a bit of that today still.' (Interviewee C)

Although there is debate, on whether coordinated campaigns were orchestrated on social media to create crime, the London Riots revealed to the police a lack of knowledge with regards to the potential relevance of social media. In addition, during this time the police had been confronted with 'critical views' about the nature of their intelligence gathering tactics, in particular with regards to a number of incidents relating to undercover policing that had been revealed publicly. According to one of our interviewees, who was leading the domestic extremism unit at the time, this contributed to a reconsideration of tactics within the police:

'it made certainly me and others think is there another way we can gather information which is more proportional? For me it's always about…recognizing that if you want to have legitimacy amongst the public, you've got to be able to gather information which the public can go, that's not unreasonable.'

As such, uses of social media for policing domestic extremism and disorder emerged as a response to a combination of events that not only highlighted the potential role of social media in organizing and mobilizing forms of protests and uprisings, but also a perception of social media as a more legitimate resource for intelligence gathering than other tactics employed to gather this intelligence.[3]

[3] We were not informed whether any of these other tactics have been scaled down or stopped as a result. There is currently a public inquiry taking place into undercover policing in the British police that may discuss this further (see www.ucpi.org.uk).

‘Open source’ data collection

The emphasis on Open Source Intelligence (OSINT) as the dominant feature of Social Media Intelligence (SOCMINT) used by the police is integral to the perceived legitimacy of employing this tactic. The social media data gathered by the police in this sense is dependent on what is available ‘publicly’ and does not go beyond the privacy settings in place on social media platforms. As will be discussed further below, practices that involve bypassing privacy settings or getting communications data through other means also form part of intelligence gathering but these are considered separate from the more general practice of big social data collection from social media platforms and require particular processes and authorization. It is clear from our interviews with police that there is a strong sense that data available from social media platforms without having to get such authorization (so-called ‘open source’ data) provide a substantial intelligence resource. It allows for a number of different aspects to be included in any policing strategy, including an understanding of likely occurrences at larger happenings such as protests. It is a ‘way of understanding what people are thinking and saying in certain events.’ (Interviewee C)

However, the perceived ‘open’ or ‘public’ nature of this data has been the subject of negotiation within the police and there have been attempts to distinguish between the police as a purveyor of social media communication and other actors who might look at such communication (such as ordinary citizens). As such, there has been recognition within the police that collecting data for policing purposes is a particular practice that has implications for understandings of the public and private nature of such data:

‘We're very aware of privacy issues around social media and up until a couple of years ago, the joint thinking – not just in the police but across a lot of the organisations – was that if you saw it on social media, it’s open to anybody, then there’s no privacy issues. We fought for a long time, we fought for more governance and we said that’s not right, there are privacy issues here.’ (Interviewee D)

As such, over the past couple of years guidelines for how social media data can used for policing purposes have been produced by the police. In particular, there have been attempts to interpret legislation such as the Regulation of Investigatory Powers Act (RIPA) which has largely been discredited as unsuitable for contemporary developments in digital infrastructures and communication (see above). In such circumstances, the police expressed to us in interviews, as well as in our dedicated project workshop, the wish to interpret legislation in the way that they thought it was intended: ‘you’re in this uncontrolled space so you’ve got to try and interpret the law to try and work out what did the spirit of the law intend?’ (Interviewee C) This has meant a significant shift in police understandings of the nature of social media data. As an interviewee, who was part of developing the guidelines for interpretations of RIPA for OSINT, stated:

‘if we're surveilling you, then we need authority. Maybe you’re doing it in the open, you’re speaking in the open, making comments in the open – we still need authority because it’s surveillance. Whereas four or five years ago, that would have been considered if you made that comment out in the open, it’s open source and anybody can view it. So it’s quite a turnaround for the police.’ (Interviewee D)

RIPA therefore has implications for two main forms of social media practice: the monitoring of individual accounts and profiles, and the length of time data is retained. Repeated viewings of profiles and retention of data for longer periods of time both require forms of authorization, regardless of it being 'open source'.[4] This provides a level of accountability towards the public in what is perceived to be a grey area open to abuse: 'We need a cause, we need a reason, it needs to be justified and it needs to be necessary and it needs to be lawful.' (Interviewee D)

Commercial programmes and tools

The integration of OSINT and SOCMINT into police practice for policing domestic extremism and disorder has come through a process of bringing in external programmes and tools. The NDEDIU does not house software developers and engineers that develop own software for the police. Nor is there one specific provider of tools catered for police and law enforcement purposes. Rather, the police has bought a host of programmes and tools from different companies. These are a combination of 'off-the-shelf' tools that are already available and programmes that have been purchased through a procurement model. There is some scope for the police to make suggestions for changes and amendments of these programmes to better suit their needs, but there is no active involvement by the police with the design or development of the actual software. The collection and analysis of data, however, is all done in-house. Training for how to use the software and training for use of social media data more generally is provided by the software developers or private 'accredited training companies'. This means that the police do not design or necessarily have knowledge of the algorithms behind the software they are using for collection and analysis:

'we know what queries we want to create and mostly we'll try the query and see what comes back and then we'll tweak it. Behind that obviously is an algorithm that the company's got software to develop. We're not really seeing it at that level, we're just knowing that we're looking for A if it's associated with B and also has C in it, then we'll write that query and we'll see what comes back and then we'll tweak it and we'll add in exclusions or inclusions. So the actual algorithm sits behind it and it's beyond us.' (Interviewee D)

Adopting a tool or piece of software for policing results from a process of live testing with real events. From our interviews, it also seems that biases in algorithms are accounted for by using a multitude of different tools for any given event. However, as will be further outlined below, these processes and considerations are still at a nascent stage within everyday practices of the police and there is a continued reliance on human assessment and more traditional police tactics to navigate and 'correct' algorithmically produced intelligence.

Predominantly, therefore, the tools used by the police are commercial tools that have, more often than not, been developed out of marketing rather than law enforcement needs. The police adapt these tools for their own purposes:

[4] We were not provided with exact numbers in interviews beyond 'more than a couple of times' for viewing and 'more than days' for retention of data. However, some guidelines are now publicly available online with further details of how relevant legislation is interpreted by police: http://www.uk-osint.net/documents/ACPO-OSIW-&-Research.pdf

'A lot of stuff came out of marketing because marketing were using social media to understand what people were saying about their product...We wanted to understand what people were saying so it's almost using it in reverse.' (Interviewee C)

Due to sensitivity about revealing police capacity and tactics, we were not provided with names and details of the exact software that the police use. However, it emerged from interviews that a substantial part of their social media practices involve fairly mainstream tools that would be familiar to everyday users, such as Tweetdeck and Hootsuite. Moreover, as will be further outlined below, the nature of the programmes that police use for policing domestic extremism and disorder seem similar to tools that we are familiar with for basic research needs. As we were told in one interview, 'all our tools can do nothing more than Google.'

Purposes of social media monitoring

Most events are monitored by police on social media. This will involve collecting social media data in the week or so leading up to any event, such as a protest or demonstration, as well as monitoring social media activity during the event. As such, social media monitoring is used for both pre-emptive as well as real-time police tactics and responses. Most of the time, police will decide to monitor events based on prior information that an event is happening either through other forms of intelligence or from the media. This could also include knowledge of community tension somewhere, or if something has happened that might trigger reactions from certain groups. However, the members of the police we interviewed said that monitoring at the moment does not include general monitoring of community activities, although as will be discussed further below, social media data collection and analysis may be used for ascertaining community needs and concerns in the future as a way to enhance police 'engagement' with communities. In addition, we were told in an interview that the police are also currently looking into employing social media monitoring for potential tension surrounding the police, or hostile mentions of the police, what was described as 'looking for reputational risk for the force.'

Monitoring social media activity for the purposes of policing protests was predominantly described as aiding 'situation awareness' for any given event. Mostly, the focus of policing protests as expressed in the interviews concerns potential disruption or violence at protests: 'what we're looking for is somebody that's going to go there, either to cause disruption against the protest or use the protest as cover for further activity.' (Interviewee D) However, as we outline here, numerous different aspects form part of the data analysis used for devising strategies of predictive policing of protests. Below we highlight the main uses of social media data for policing domestic extremism and disorder as discussed in our interviews from an intelligence perspective.

- <u>Keywords and threat words</u>

  The most dominant practice in the uses of social media for policing is based around keyword searching. That is, using tools to filter large data sets relating to a particular event by identifying a list of keywords, including weighting certain keywords, and searching for potential threats. 'Threats' in this context would be particular words associated with violence or disruption ('threat words'), and to then make an assessment as to whether further action is

needed to identify individuals. List of keywords and threat words are context-specific and different lists of keywords and threat words are developed depending on the nature of the event, the location, and the people it is likely to attract (particularly to include sensibilities of language and dual meaning words, e.g. 'flared trousers as opposed to a flare being set off'). As such, algorithms are used to 'filter the noise' in terms of particular words that allows police to assess only highlighted data:

'The systems that we use produce reports, they'll produce PDF reports and so we'll look for keywords…we'll look for people talking about guns or whatever at protests and it'll produce a PDF document to say all these posts have got all the criteria you're looking for, and we'll look through them and then there's one in there that actually is of interest to us. We'll take that and we'll put that into an intelligence report.' (Interviewee D)

- Risk assessment and resourcing

  Another important use of social media data that informs police practice for policing domestic extremism and disorder is to gather a sense of who and how many people will be attending an event and how militant it might be. As such, 'threats' might be identified in this context by ascertaining whether certain groups and individuals are attending the event, and what their intent of going there might be:

  'Lots of events are organized on Facebook publicly and that gives you a good feeling for how many people are going, and I don't think that's unreasonable for the police to understand how many people are coming to an event, and whether or not the words they're saying, the symbols they're using suggest violence or otherwise.' (Interviewee C)

  The interest in whether 'risk' individuals or groups are planning to attend an event is frequently informed by prior knowledge about those people. In other words, groups will often be well known to the police based on previous intelligence kept on databases: 'you can work out there are some groups that come and protest and they don't protest peacefully and they never have.' Monitoring the social media activity of these groups in particular in the lead up to an event that could be of interest to them will be part of police practice in the planning for that event.

- Influencers

  Linked to that, social media data analysis is used to identify what was referred to in interviews as 'influencers.' One of the outcomes of the investigation into the London Riots was to identify what sort of individuals may be very influential in certain contexts (e.g. DJs proved to be influential individuals during the London Riots). In several instances, the notion of 'influencers' was intertwined with 'organisers' in interviews and it is not entirely clear how distinct these categories are. However, it did emerge from our interviews that influencers may not necessarily be organisers but may also be identified through online reach and following. Software examples such as Clout were

mentioned as the kind of tool that may help police identify influential individuals or groups, in which the amount tweets, re-tweets and followers will highlight particular accounts. This may be of interest to the police in terms of engaging with such individuals and groups before an event or for identifying potential criminal activity resulting from the nature of influencer communication:

'There are some really influential people on Twitter, have thousands and thousands of followers and they say something and it gets repeated a thousand times and that word or that feeling's been repeated 10 or 20,000 times. That's quite powerful and quite fast. So those people can influence what happens if they're people who people listen to…If you've got someone who is saying lots of things that suggest let's be violent and that's been retweeted by lots of other people, that person you could argue is starting to influence the people who are coming, you're starting to plant seeds in their minds. So the influencers are quite important I think because it helps us to understand and actually is this person a threat? Is this person conspiring to cause serious criminality? Do we need to look at this person in depth?' (Interviewee C)

However, there is also recognition that definitions of 'threats' in terms of influencers (or organisers) on these terms can be problematic:

'I think you have to be careful with that one because being an influencer, does that make you a bad person? Does that make you someone the police should be interested in? If you're influencing a crowd to do something that's unlawful, absolutely but if you're just an influencer, then I think you have to be careful.' (Interviewee D)

- <u>Sentiment analysis</u>

Even though marketing-driven software has placed much emphasis on sentiment analysis of big data, it remains a marginal aspect of police social media practices. It may serve as a source of information for more longer-term developments, but the level of sophistication of sentiment analysis is not high enough for it to serve much purpose for informing real-time police tactics (described as 'over-rated' in one interview). However, basic analysis of the mood of a crowd might help alert any potential tension with the police:

'I suppose things like if you're dealing with a large event and you've got crowds, is this crowd happy or are they cross or are they angry? Are they saying things that the language is really angry and really cross and they're not happy with the police, or is it really positive about the police because you'd argue if the sentiment was really negative about the police, we might change our tactics.' (Interviewee C)

However, doing such analysis may require contextual knowledge of language that can account for different demographics, places and cultures. Algorithmic intelligence on sentiment for policing purposes is therefore still not a major practice yet, although it may become so as the level of sophistication of algorithms increases.

- Geo-location

  Despite the uses of some software by the police that is particularly concerned with geo-location, the limited availability of geo-location on major social media platforms such as Twitter and Facebook makes it a marginal aspect of big data analysis. Less than 2% of tweets have geo-tagging on them, for example, making it problematic to rely on this to gather information about the location of crowds or individuals. Instead, potential locations for gatherings of crowds are identified through keyword searches as mentioned above.[5]

Integrated intelligence and human assessment

Importantly, a prevalent theme that emerged from our interviews is that SOCMINT is not treated as an isolated practice within policing. Rather, it is integrated with other forms of intelligence-gathering practices:

'Social media isn't the only tool you'd use to understand the dynamics of large scale protests which may become unlawful. There'll be other intelligence means, of course there will be. There'll be an understanding of what's happened before, what happened the last time this group protested. So social media I think is just one tool in the box of many.' (Interviewee C)

The integration of SOCMINT with other forms of intelligence is exemplified by the structure of the NDEDIU. As part of this unit, SOCMINT sits under the creation of an 'all source hub', which integrates social media data with other forms of intelligence (human intelligence, undercover work, etc.) and existing databases. In this way, the policing of domestic extremism and disorder is comprised of three elements: big data, intelligence, and databases. This means that SOCMINT is 'cross-checked' with other forms of intelligence.

Linked to this, our interviews highlighted the extent to which use of social media data for policing is intimately dependent on human assessment and discretions. Partly this is due to the learning process that continues to be involved in the development of algorithms: 'you still need a human at the back of it to go, yes that's good or it's got it wrong and we need to start again because algorithms work and they learn.' (Interviewee C) That is, algorithms are not yet at a level where police work can be an automated process. Certain aspects of police work may be automated, but the actual assessment of any data that may inform police tactics and strategies requires human intervention. As one of our interviewees noted:

'Algorithms aren't always right and when you're dealing with public safety, I think you've still got to have that human assessment and judgement of a professional person who goes, I don't agree with that. So they only shape our thoughts rather than make the decision.' (Interviewee C)

[5] Of course, this does not rule out the use of geo-location gathered through other means such as mobile data. However, for this project we are only focusing on the uses of 'open source' social media data.

As such, automated processes serve to filter data down and provide particular patterns of data that are then humanly assessed. This also means that much social media data is actually read by human eyes in order to assess its relevance for intelligence. The example we were provided with was the collection of tweets for the 2012 Olympics. For this event, 31 million tweets were collected pertaining to keywords relating to the Olympics. Through various automated processes, these were filtered down to 15,000 tweets. From these tweets, analysis and assessment of this data resulted in 2,500 intelligence reports. As such, a substantial amount of human resources still underpin the integration of social media uses for policing. Moreover, the relevance of collected data is a matter of 'professional judgement' (Interviewee C). As our next chapter on our big data analysis also highlights, the levels of human discretion still involved even in the types of analyses of data that we have outlined here is much greater than what might have been suggested from discussions on big data.

Pre-emptive tactics

A significant part of integrating uses of social media into police practices has been to shift the focus of policing from reactive to proactive policing. Using social media data to predict activity also leads to some forms of pre-emptive policing, such as pre-emptive arrests and/or interception of actions:

'If someone's discussing something openly online and we've come across it, and they've said I'm going to go to the protest tomorrow and I'm going to set off flares, then if there's something criminal in that – i.e. is it illegal to possess what he's saying he's going to set off if it's a flare or whatever. If it is, then that might be something that we take action against at his house and arrest him before he actually gets there…Or if actually what he's talking about is I'm going to go down there and I'm going to cause mayhem and all the rest of it, we actually might go round and say to him, we know you're planning to go and can we suggest another course of action for you, and if he turns up there, then you would look for some other disrupting tactics, deal with that if it becomes an issue because you know he's likely to cause problems.' (Interviewee D)

Moreover, it may lead police to seek out certain individuals or groups prior to an event:

'What we have in the past done is we've identified organisers of a protest and we've gone round and spoken to them, not because we think that they're going to do anything criminal but to say, we know you're going to have a protest, how can we help you make sure that that protest goes off safely?' (Interviewee D)

As such, predictive analytics will in some instances lead to pre-emptive tactics such as seeking out or confronting particular groups before any activity has occurred. A key aspect of predictive policing in this respect is to identify potential trigger points that might lead to disorder before it happens. Moreover, data will inform what strategies will be used for policing any given event depending on size, nature and militancy of the crowd. As mentioned above, this might mean softer or more forceful forms of policing as well as being able to navigate activities as they are about to happen so as to respond in real time with changed tactics.

Engagement

Linked to this is a broader theme of how SOCMINT relates to engagement. Although there are diverging views amongst the people we interviewed regarding the relationship between intelligence and engagement, it is clear that social media introduces a context in which these different aspects of policing may become increasingly integrated. Currently, police engagement on social media is relatively limited and is largely done on an ad hoc basis with substantial discrepancy between different forces in terms of how much and in what way police will engage. For some, this is the way engagement on social media should be conducted and any attempt to standardize uses nationally would be to the detriment of local knowledge and context:

'I think locality based decision making is really good. I'd hate for it to be a national policy because I think that's suicide. We got a lot of pressure to have a national strategy…You've got to own it locally. So when I look at 43 different forces I see 43 different flavours, I see some people being really good at it and some people being really cool on it. I'm okay because that's local ownership. Any more than you could standardise the way police officers talk to you in the street.' (Interviewee B)

As is further highlighted in our big data analysis, this emphasis on local ownership is partly because police communication on social media is dominated by information dissemination about police activities and warnings and mostly concerned with engaging with crime such as theft and missing persons. Engagement concerning domestic extremism and disorder via social media remains a marginal activity. Although the London Riots highlighted the need for police to respond in real-time to rumours that emerged on social media regarding forms of disorder (e.g. 'a lot of people were talking about it's kicking off somewhere and the police were able to say no.'), actively responding to activities happening online is really police engagement in 'exceptional circumstances' and is not a common practice.

As such, police engagement on social media involves few automated processes at this stage and, moreover, 'threats' as understood in terms of domestic extremism and disorder are not engaged with on social media in the form of police communication. However, a key theme that emerged in the interviews is the ways in which social media data may serve to inform strategies for engagement in future. Although it in principle falls outside the remit of engagement officers (such as those involved in the PREVENT programme) to scan social media for extremist content[6], social media is increasingly a feature of the Channel programme which outlines a number of indicators for assessing 'radicalisation'. That is, activity on social media is increasingly a key indicator of 'radicalisation' which would require a police response as part of the PREVENT programme. The danger of incorporating social media into engagement strategies in this way, according to some of people we interviewed, is that it would effectively serve to turn engagement into intelligence. Importantly, there is an explicit distinction between engagement stemming from 'neighbourhood policing'[7] and policing concerned with threats, harms and risks:

[6] This type of activity would involve the Counter Terrorism Internet Referral Unit (CTIRU) which we did not discuss in interviews. However, it was mentioned in interviews that engagement officers will at times be alerted to social media profiles that contain extremist content and they will then refer these profiles to CTIRU who will decide on a course of action.

[7] For more details of the idea behind 'neighbourhood policing' cf. Lowe and Innes (2012).

‘I think you’ve got kind of like two models…One is what are the threats, harm and risks in our community, where are our at risk people and where are our risky people? How do we use social media to find and locate all that kind of stuff? But there’s a more general form of engagement which is what’s your experience of policing like? What’s your locality like? What issues are you raising? Do you have contact with the police?’ (Interviewee B)

The first of these models speaks to intelligence needs or what this interviewee described in his interview as ‘engagement with an agenda’ whereas the second is ‘engagement for engagement’. However, on the front-line of policing it is clear that engagement and intelligence are closely intertwined:

‘They’re not separate camps. They can’t be, can they, because if I’m chatting to you and you’re a protestor or you’ve come to London…So there I am chatting to you and I’m engaging with you and I’m listening to you and you start to say, I’m doing things and I go, oh that’s not quite right…So I am starting to gain information now, thinking actually she was talking about some of the stuff that sounded bad. She’s talking about her friends.’ (Interviewee C)

Moreover, as social media use within the police increases further police engagement on social media will have implications for intelligence gathering. As police moves towards using big social media data for understanding community concerns and needs as part of their engagement strategies, this will also serve to provide intelligence for predicting forms of risks and threats, particularly around domestic extremism and disorder:

I think there’s a lot that can be done to say there’s a hotspot and if it gives concern to the community, then it should be giving concern to us because we’re there to provide a service to the community.’ (Interviewee D)

One of the key debates that emerged in our interviews was the extent to which the police could have predicted the London Riots from monitoring social media and identifying tensions:

‘Should the police have been doing better social media monitoring, looking at community tension following the Duggan shooting?[8] The Duggan shooting was Thursday, the riots kicked off on Saturday. The report I saw from the Met which is public was that they used the old ‘our community context tell us everything is fine’ type of stuff. So there’s kind of link an unwritten hypothesis that had they been doing better social media monitoring, would they have picked up on certainly the escalating tensions in the community? I suspect they would have.’ (Interviewee B)

Making use of social media data in this way borrows from predictive policing familiar in data-driven crime-prevention and incorporates it into community engagement (recognizing the ambiguity around what this means in practice) for the prevention of domestic extremism and disorder. As a regional PREVENT officer said:

[8] The Duggan shooting refers to the shooting of 29-year old Mark Duggan, a black resident of Tottenham North London, by a police officer. This event is said to have led to conflict with the police and the eventual escalation into riots across London and other English cities.

‘As far as PREVENT is concerned, partly in the sense that if you can gauge a trend, then you can perhaps intervene at an early stage…for example, the Draw Mohammed cartoons event that’s just been cancelled, it’s quite possible that through social media you can gain a fairly quick picture of each state of its organization but more importantly, the public response to it…and then that can also help you in terms of how you go about formulating your own counter narratives because you respond to what the most current threat is.’ (Interviewee E)

As such, analysis of social media data comes to inform police engagement strategies, illustrating how intelligence and engagement intersect. What is more, it may introduce further forms of engagement on social media directly by the police. As our interviewee continued:

‘you often find in the aftermath of, say, a significant event like Woolwich or the Charlie Hebdo massacre is that although that may happen elsewhere…that can affect the local situation…You often find Islamophobic incidents, for example, on a local level would rise – whether it was graffiti or hate crime or things like that. So the response to a big event, even if it’s simple community messaging on social media, appeals for calm or encourage reporting of Islamophobic incidents or anything like that. After Charlie Hebdo, for example, social media went absolutely crazy. If in amongst that, as a police officer you can get in a little message there in terms of let’s appeal for calm…I think there’s a benefit to responding to that.’ (Interviewee E)

In one interview the example of ‘Web Constables’ in Estonia and Finland was presented as a way of thinking about how police officers might actively engage on social media based on social media monitoring and analysis. Such ‘web constables’ are entirely focused on working online, incorporating intelligence-gathering with investigation and engagement, all within the digital environment. In these instances they also actively communicate and carry out policing online:

‘they’ll go into chatrooms and they’ll watch a chat and they’ll say that’s not very nice, why did you say that? Probably what we would do in the street…The online community is just another extension of the real world.’ (Interviewee D)

Such levels of monitoring, intelligence-gathering, and engagement is problematic for a country like the UK however, as linguistic boundaries are not demarcated along any sensible jurisdiction when the language is English (unlike Finland or Estonia where this might be more feasible). Putting the role of ‘web constables’ into practice in the UK might therefore be very challenging as long as it is not possible to have better data on geo-location than what currently exists from open source social media data. Nonetheless, the notion of policing being carried out increasingly within a social media context, across intelligence and engagement, for the purposes of policing domestic extremism and disorder (amongst other things) seems clear. As our next chapter will further illustrate, these types of police practices introduce particular ways of identifying and defining potential ‘threats’ that need more examination.

Partnerships with private actors

As the ability to collect and analyse large-scale social media data has become easier, a significant question for police that emerged in our interviews is what future relations

with other actors who engage in this activity might be. Companies will often monitor social media activity in order to identify potential risks to their brand, including forms of anti-corporate activism and dissent. Moreover, they engage in evermore extensive monitoring and profiling of their customers that include intelligence gathering that could have potential relevance for police. As one of our interviewees noted:

'The whole question of industry information in the future is one we need to confront because there's no doubt about it, industry has probably just as good a means of gathering some information than we do, particularly around their own customers.' (Interviewee C)

The key issue that emerged in our interviews is the extent to which such actors are able to collect and analyse social media data without the restrictions that apply to data collection for police purposes as outlined in the guidelines under RIPA:

'you shouldn't be able to but if you're a private enterprise, you're not having to follow the Regulation of Investigatory Powers Act, you're not having to follow all these laws that we do. They can do what they like.' (Interviewee C)

In terms of what this means for policing, our interviews highlighted the concerns that arise from how police might engage with intelligence that has been gathered by private actors beyond the control or oversight of the police themselves. However, as these types of practices extend to not just companies, but also research institutions, think tanks and civil society organisations, there is ambivalence around the extent to which the police might also wish to take advantage of this potential intelligence resource:

'There's no accountability or governance mechanism that can tell us it's been gathered proportionately or fairly, it's not been gathered as a grudge or a personal dispute; you don't know, do you?...Unless you can be sure and there's some form of agreement where we share for the right reasons, it's difficult. It's not impossible, but it's difficult.' (Interviewee C)

This ambivalence is also prevalent in questions around the relationship with social media companies who host the communication that police are seeking to monitor. Although recognizing that social media companies 'don't want to be associated with policing' and that police 'don't want to be seen to be taking mass data from social media companies without justification' (Interviewee D), there is an expressed wish for social media companies to have a more direct line of communication with police in which flagged material and content is directly reported to the police, particularly as these companies have more resources and capacity to monitor their platforms than the police:

'the relationship with the State and social media companies is always going to be good but tense because we want to see more. So if you look at Facebook, Facebook has probably got the best algorithmic detection systems than anything because it's got billions of people to try it on. It's got a huge sample size. So it takes down stuff using algorithms. So people talk about threats of violence and killing, it takes it offline but they don't tell the police about it. So those people may go out to kill someone and we might have been able to stop it. So Facebook are taking violence off Facebook

automatically, it's automated, but where's the, you might need to know about this, this guy's talking about killing and shooting.' (Interviewee C)

This was described as a 'moral duty' on behalf of these companies that steps outside questions of law and compliance, and suggests that the policing of extremism and disorder is not confined to the practices of police but necessarily incorporates other actors:

'the police are just a part player here, we're the end of the line, aren't we? It's gone through all the layers and we pick it up. The piece in the middle for economic crime it's banks. For content it's the internet providers. That bit in the middle is the bit where the activity should be focused and this is where I talk about their moral duty to do so, and I think they have a moral duty to do so.' (Interviewee C)

At the same time, our interviews also indicated that any such developments in relation to intelligence gathering need to be perceived as legitimate and justifiable amongst the public: 'we need to take the public with us.' (Interviewee D)

Big Data Analysis: Predictive Policing and Engagement

This chapter will outline the findings from the big data analyses that we conducted as part of our project. This included two different types of analyses: firstly, we 'emulated' police practices with regards to the use of social media data for policing protests by collecting and analyzing social media data in the lead-up and during three major protests in the UK during the research period: the Anti-Austerity protest in June 2015; the Solidary with Refugees demonstration in September 2015; and the DSEI Arms Fair protest also in September 2015. As we will outline in this chapter, we used software that carried out similar types of analyses as had been described in our interviews with police in order to illustrate how potential 'threats' may come to be identified in relation to protest policing. Secondly, we carried out analyses of police engagement on social media by collecting and analyzing communication from police accounts in order to examine the extent to which 'threats' pertaining to domestic extremism and disorder are actively managed or engaged with via social media. This chapter will outline the findings of each part of the study separately before providing a summary of the findings.

I. Analysis of Protest Tweets

Methodology

Tweets related to the Anti-Austerity Protest, which occured in London on Saturday, June 20, 2015, were collected over one week from 16-23 June, 2015. 322,005 tweets were gathered, the data was uploaded to the Cosmos analysis software and a systematic sample produced 53,688 tweets for analysis.

Analysis consisted of mapping the geolocation data from tweets to identify hotspots of activity and movement during the demonstration; an analysis of popular terms, including hashtags, user handles, and keywords; and a mapping of the retweet and mention networks. The list of popular terms was subsequently manually filtered for irrelevant or unclear words (it's, live, via, I'm) as well as duplicates (march/marching, London/#london) to produce a list of the top 100 terms. A sentiment analysis was conducted for the general tone of the dataset, as well as for the specific sentiment found in a selection of 12 key terms. Further analysis consisted of mapping the timestamp of the dataset to obtain the frequency distribution of tweets during the week the data was collected.

Tweets for the Solidarity with Refugees and DSEI Arms Fair protests, both planned for September 12, 2015, were collected for the three days leading up to the demonstrations, based on hashtags that were identified as trending. While this limits the analysis of how police could potentially use Twitter to monitor protests, it focuses on the analysis of how police might monitor Twitter for predictive policing, identifying 'threats' to prevent crimes or disruptive behavior during the protests.

	Anti-Austerity	Solidarity with Refugees	DSEI Arms Fair Protest
Search Terms	#junedemo #endausteritynow #anti-austerity	#solidaritywithrefugees #refugeeswelcomeUK	#stopdsei #stoparmingisrael #occupydsei #stopthearmsfair #stoparmstrade #wheelstopdsei
Dates of Collection	16-23 June 2015	9-11 September, 2015	9-11 September, 2015
Total tweets	322,005	245	6,800

Table 2: Data collection of protest tweets

Location

Some of the basic identifying features were missing from the data. From the anti-austerity protest sample, only 70 tweets included geolocation metadata, or about 0.1%, with 66 of those tweets coming from the United Kingdom. This is notable, as it indicates police could not use social media to pinpoint individuals and track their movements in conjunction with the demonstration. This finding is further supported by the absence of specific locations from popular keywords, with London and Glasgow the most popular geographic markers along with a few very general location words (square, streets.) The words parliament and church also appear, but it is unclear if these words are used in reference to the institution or location.

Because of the limited use of geolocation information in the previous analysis, as well as the focus on the preparation period rather than the actual protests, geolocation was not analysed for the refugees demonstration and the DSEI protest.

Keywords, threat words and popular accounts

The top 20 **keywords** of the anti-austerity demonstration referred to protest words (march/marching, protest, demo, #junedemo, rally) as well as characteristics of the protest (people, london, thousands, 250000, now) and the subject of the protest (#endausteritynow, antiausterity, austerity, cuts, tory). A sentiment analysis of these terms was inconclusive, as there was typically a wide distribution of sentiments, both positive and negative, for each keyword.

Established political parties feature prominently in conversations on Twitter. Tories topped the list of political parties mentioned, with both tory and tories ranking higher than any other party. The Labour Party also featured prominently, with Jeremy Corbyn singled out in popular accounts and keywords (@jeremycorbyn, @corbyn4leader, corbyn), and so did Unite, the largest trade union in the UK and Ireland. The Green Party was frequently referenced directly (@thegreenparty), along with Green Party Politician Caroline Lucas (@carolinelucas). Other popular accounts include media conglomerates (@bbcnews, @independent) as well as individual journalists (@georgeaylett, @owenjones84, @chunkymark). Popular celebrities sympathetic to the demonstration, like Charlotte Church and Russel Brand, were also referred to.

The most popular account to be referenced directly was for the Peoples Assembly (@pplsassembly), a UK wide organization that is dedicated to fighting austerity and is unaffiliated with an official political party. The Peoples Assembly were the organisers

of the demonstration, although it is not apparent from the Twitter data which individuals or groups in this horizontally-networked organization were involved in organizing the demonstration.

The most popular keywords in the refugees demonstration included basic logistical language (tomorrow, Saturday, march/marching, london, refugees) and variations/hashtags that used refugee and solidarity: worldsolidarity, glasgowsolidarity, refugees, refugeeswelcome, refugeecrisis, and syrianrefugees. Prominent references were made to two accounts spearheaded the organization (@the45storm and @actionaiduk).

For the DSEI protest, most tweets included related keywords such as stopdsei, arms, dealers, fair, and referred to two accounts working to stop the arms trade in the UK (@stopthearmsfair and @caatuk). Green Party MP Caroline Lucas' Twitter profile was a further popular account appearing in tweets, as she is likely an influencer and outspoken opponent of the arms fair. Interestingly, #pmqs appears in the top 10 keywords, indicating that people tweeting about the arms fair related to questions to the Prime Minister or may even raise their criticism of the arms fair to national political debate.

The anti-austerity protest sample included positive mentions of refugees under the hashtags #worldrefugeeday (9), #refugeesarewellcome [sic] (6), #refugeeswelcome (3), #refugeescontribute (2) and #refugeeweek (2). There was no mention of the DSEI arms far this early on, though army was mentioned 14 times, antiwar 13 and warfare 25 times. War appeared 466 times, but this could be used in conjunction with any number of causes. Many tweets at the DSEI protest included the hashtag #refugeeswelcome, which may refer to the Solidarity with Refugees protest, the refugee crisis more generally, or a link between the refugee crisis and the international arms trade. Tweets about the Solidarity with Refugees protest, on the other hand, lacked mentions of arms or DSEI.

The complete list of words that appeared in tweets was searched for potentially **threatening terms**. In the anti-austerity sample, 'kick' and 'kicking' were popular terms, which might be linked to kick off/kicking off, though it is impossible to know word combinations with our dataset. Threat/threaten/threatening/threatened had a total of 37 appearances in tweets, potentially piquing the interest of the police. Violence appeared a total of 69 times, with violent following in 19 tweets; police may pay attention to this to track incidents of violence or the potential eruption of violence during the protest. Escalate/escalating, flare/flaring, erupt/erupting and overthrow did not appear frequently in tweets.

The terms #occupy/occupy appeared 35 times, #occupylondon 6 and #occupycentral 3 times. Given the anti-corporate, anti-neoliberal discourse surrounding Occupy Wall Street and subsequent occupations, it is not surprising variations on the occupy brand would be popular at the anti-austerity protest. Police might pay attention to these keywords and accounts. As police have monitored Black Lives Matter protesters in the United States[9], it is possible UK police would also pay attention to the 7 mentions of #blacklivesmatter.

[9] Cf. Joseph (2015)

As for more general protest language, black block and mob were popular terms in various forms. Perhaps more alarming for the police would be mentions of #blackbloc (44), with potentially more coming from separate mentions of black (67) and bloc (514). Other sections of the protest used the 'bloc' terminology as well, although for different tactics and concerns: #greenbloc (202), #justicebloc (67), #privacybloc (61), #housingbloc (43), #barnetbloc (39), #greenpartybloc (7), #redbloc (5). Further, police might track keywords such as rentamob (23), mob (23) or mobbing (9), and there were numerous uses of variations on mobilise/mobilization (59).

Because of the history of the concept 'domestic extremism', references to animal rights and environmental activism might be used to flag up potential 'threats'. Fracking/#fracking/frack appeared in 86 tweets, with mention of anti-fracking in Lancaster an additional 22 times (#dontfracklancs/@frackfreelancs/#frackfreelancs/@nofracklancs). There were only 3 mentions of oil, 2 mentions of #animalrights and 3 mentions of the user @network4animals.

Many of the potential threat terms identified for the anti-austerity protest did not appear in the refugee protest sample. Instead, march/marching were the most popular (102 tweets) followed by protest (7), demo (13), and rally (2). Depending on previous intelligence or experience, this varying terminology might be useful for police in assessing the tone of the event.

Both black and bloc(k) appeared in the sample for anti-DSEI protests, but never together as blackbloc. Blockade/blockading/blockades were mentioned several times, which may be in reference to specific plans for the main day of action, or may refer to smaller actions in the week leading up to the larger demo. Occupy appeared in connection with a number of different hashtags (#occupydsei, #occupydemocracy, occupy) and accounts (occupylondon, occupynn, occupydemocracy, occupymurdoch, etc.). Other protest specific language did not appear as frequently (kick, threat, flashmobs) or at all (escalate, riot). Given the nature of the protest, keywords such as weapons, arms, bomb or guns may not register with police as suspicious language or a potential threat.

While a Twitter analysis through the use of a tool such as COSMOS allows the researcher to search for common terms that may appear problematic or even threatening during a protest, police may identify different threat words based on experience and information gathered by other means. The choice of keywords and the interpretation of potential threat words thus require significant discretion by the person analyzing the data. This points to the potential for bias to enter predictive policing through social media. Police who regularly monitor social media would likely become familiar with certain terms that might be cause for concern, but the search for keywords will inevitably generate a large number of tweets that do not relate to any actual threats. The lack of context when searching for tweets using keywords raises further concerns that police may be looking at a much larger population than those sections from which violent or criminal behaviour might emerge.

Networks

Mapping networks of retweets and mentions provides information about popular accounts. In addition to Peoples Assembly and Artist Taxi Driver (@chunkymark) mentioned above, Mellisa B. (@mellberr) was mentioned frequently in connection to the anti-austerity protest. As the graphic analysis of the mention network shows, however, she is somewhat isolated from other active accounts, as people that retweet or mention her tweets typically do not engage with any other accounts that mention the protests. As someone who tweets frequently about a wide variety of news, she likely has influence among her large audience (35.8K followers) but would not be considered an organiser or actively involved in the protest.

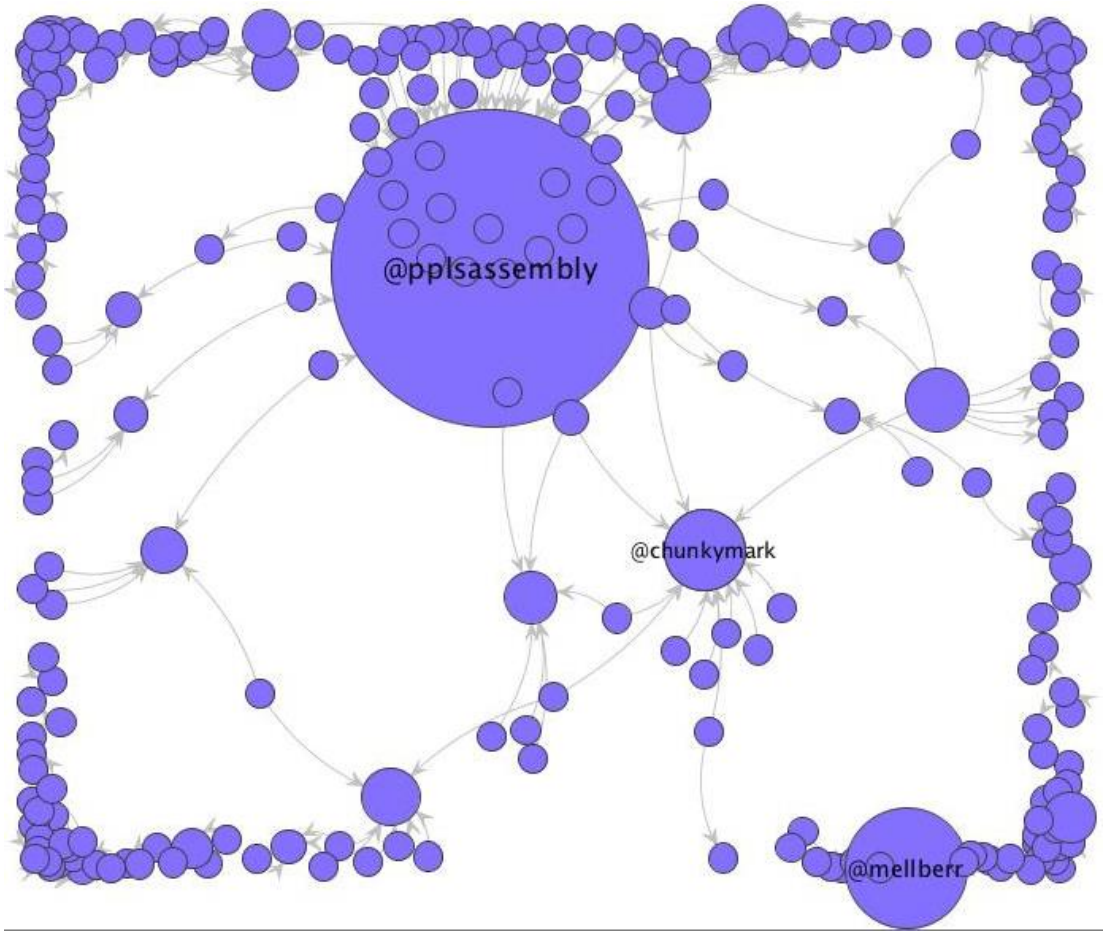

Fig. 1 Anti-Austerity Protest Mention Network Map

The Retweet Network Map reveals a more dynamic network with further individual accounts, several of which are linked to internet personalities who have a large following (@rednorthuk, @harryslaststand, @angrysalmond, @heardinlondon,@imajsaclaimant, @bravemany). While all of these users are influencers and trendsetters, it is unlikely that any of them served as organizers.

Other organizations are, like Peoples Assembly, dedicated to related social justice issues and came out strongly in support of the Anti-Austerity Protest. Wow Petition (@wowpetition) fights cuts that specifically impact the sick and disabled, and Anti-Racism Day (@antiracismday) is a similarly loosely organized group that planned the Anti-Racism Protest in London in March 2014. These two organizations share with

Peoples Assembly a relative (online) anonymity that may shield organizers and supporters from direct surveillance.

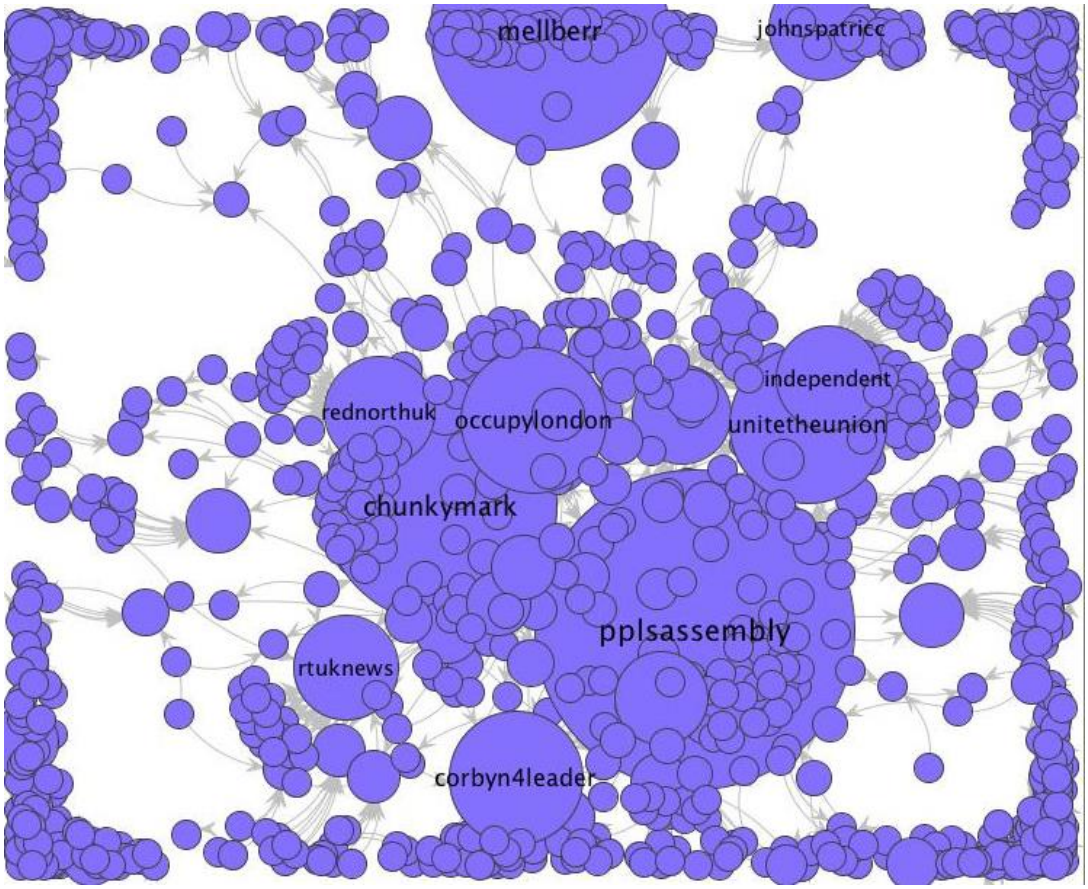

Fig. 2 Anti-Austerity Protest Retweet Network Map

Understanding the range of interest groups engaging with the anti-austerity protest on Twitter requires both contextual knowledge and a closer look at the actual tweets. Disability activists came out in strong support of the protest, as proposed cuts may disproportionately impact those that rely on NHS. Proposed cuts to the NHS were also linked to migrants and rising xenophobia in the UK, thus indicating renewed nationalist sentiment undercutting the austerity cuts. Groups like these may be disproportionately targeted because of their strong online support for the anti-austerity protest.

In contrast to the anti-austerity protest, the network map of the refugee demonstration reveals a lack of central organizers or influencers. The accounts concentrate on the border of the frame and few profiles appear in the middle of the map. Some accounts are frequently retweeted and mentioned and thus are depicted with a large circle, but they have separate communities of followers and do not seem to be in dialogue with each other. This may be characteristic of the short timeline for organizing and assembling the demonstration which was responding to a crisis moment.

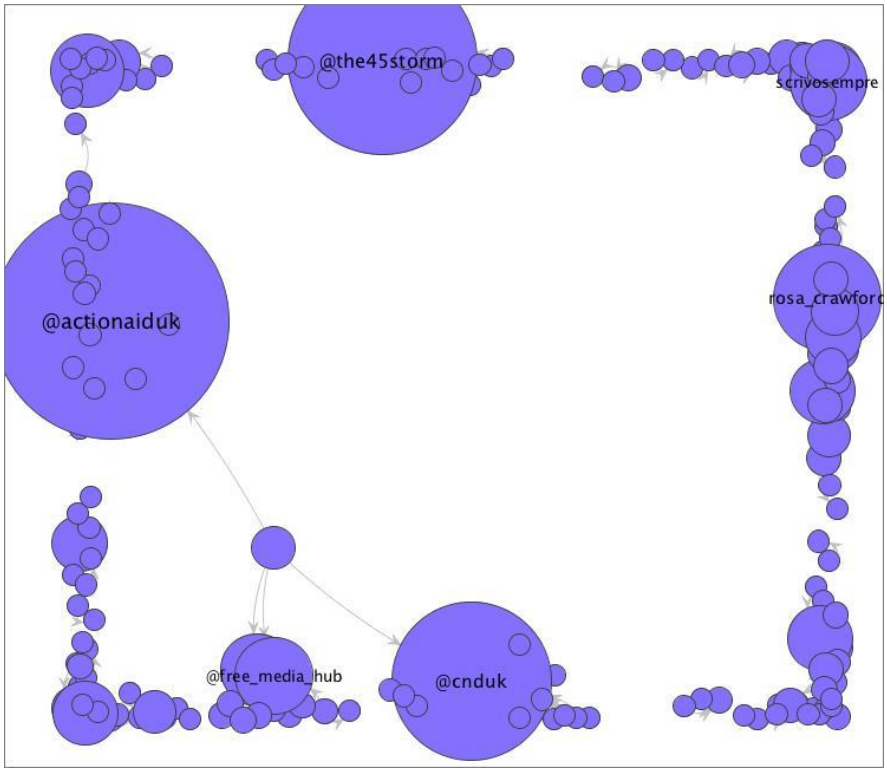

Fig. 3 Solidarity with Refugees Mention Network Map

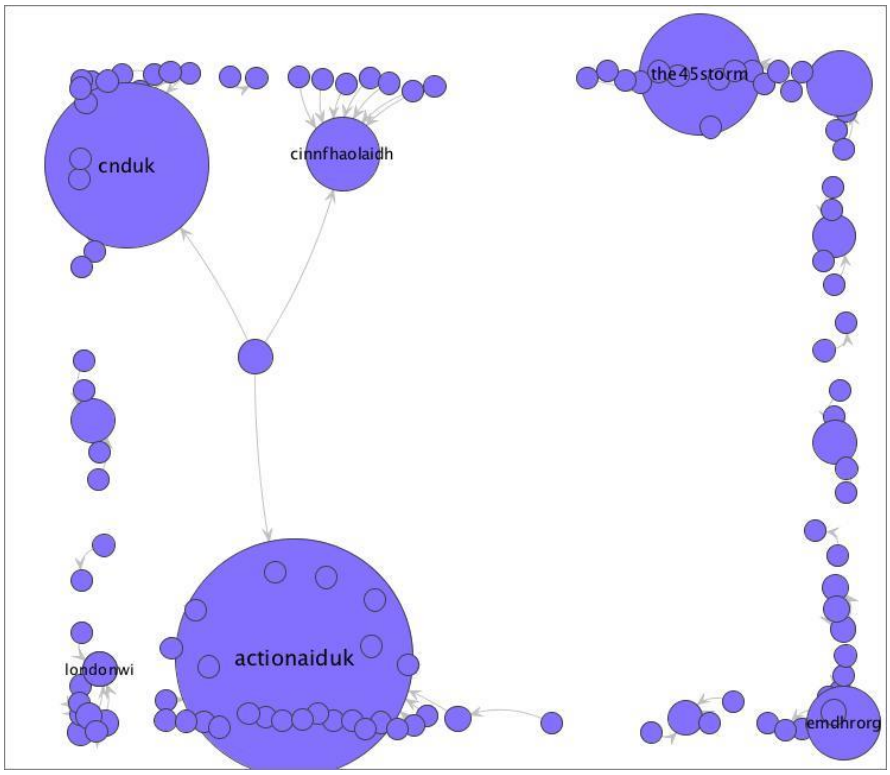

Fig. 4 Solidarity with Refugees Retweet Network Map

The DSEI protest sample, on the other hand, demonstrates a highly centralized network of social media communication. The two most retweeted accounts are stopthearmsfair and caatuk, which were already identified in the keywords analysis as likely organizers. Other popular accounts that were retweeted inclue dcipalestine, thegreenparty, caroline lucas, and occupylondon, each with separate communities of

followers that retweet their tweets. These accounts would most likely be identified as influencers within their respective networks.

Between the two poles of stopthearmsfair and caatuk are accounts that have retweeted both major organizers but do not seem to engage with other accounts. Others (e.g., Mburnettstuart, jamesclayton, sarahreader0, and jcrawl_) have retweeted a variety of accounts, including the two major organizers, as well as other individual accounts. All of these are dedicated to various social justice causes, some explicitly referencing their opposition to arms trade in their twitter description (Mburnettstuart, sarahreader0).

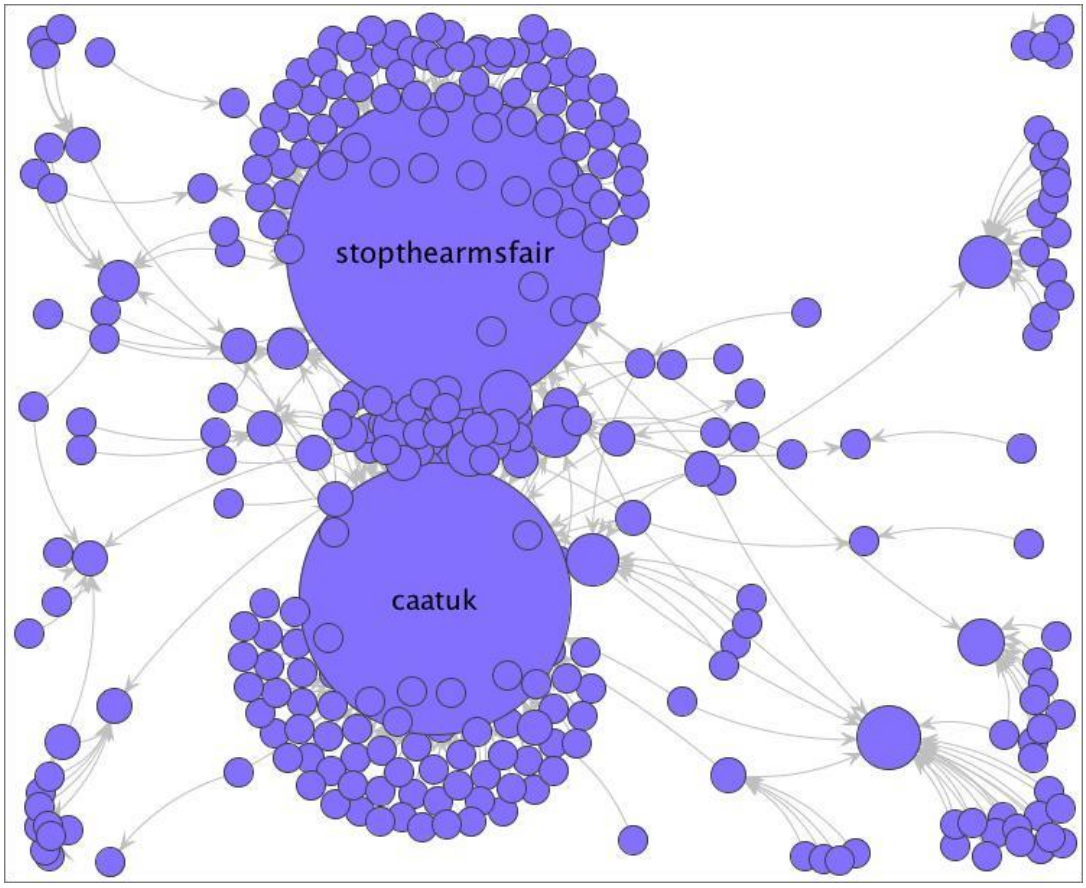

Fig. 5 DSEI Arms Fair Protest Retweet Network Map

Frequency distribution

For the anti-austerity protest, an analysis of the frequency distribution of tweets reveals that the vast majority (75%) of tweets occurred on Saturday, June 20th, the day of the protest. Only 5% of the total dataset was generated before the day of the protest, with 20% generated in the following 3 days. This seems to indicate that tweets about the protest were more oriented toward sharing news about the protest than coordinating or organizing for the protest. In this case, Twitter would not have been a rich source of intelligence for predictive policing.

II. Analysis of Police Accounts

Methodology

A selection of 143 police accounts across the UK was compiled, incorporating both official police accounts and further accounts shared by interviewees of this project. It focused on police forces that were particularly active on Twitter, including the Met Police, Surrey, and Greater Manchester Police.

All of the tweets for these 143 accounts were collected over one week, 1-8 September, 2015. In total, 57,870 tweets were collected and analysed with Cosmos. The analysis identified keywords used by UK police forces (over 600 mentions) which were then coded for reference to Twitter users, locations, types of crimes, and other words related to policing. Engagement with communities was further assessed by mapping the retweet and mention networks.

The sample period may not be reflective of tweeting behavior at other times of the year. For example, police might use Twitter differently during large events like the Olympics, pride festivals, or protests. Similarly, police use of Twitter would likely vary in response to unintended events (riots, attacks). Tweets collected during the sample period demonstrate regular Twitter activity by police in the absence of major actions, events, or attacks.

In order to better understand how individual police accounts engage their communities, a supplementary analysis of 19 accounts from the broader sample were analyzed using the online tool Twitonomy. This sample included the regional accounts for London, Surrey, Essex, Greater Manchester, and West Midlands, as well as a selection of local accounts in these regions. Local accounts were selected based on the volume of tweets and followers within regions indicated as areas of interest for domestic extremism.

Analysis with Twitonomy consisted of identifying the top 10 hashtags for each account, as well as the Twitter users most frequently engaged via retweets, replies, and mentions. Finally, the platforms most frequently used to tweet from each account were reviewed for a general understanding of the ways police engage with Twitter. The sample period for this part of the analysis was not limited to one week but stretched across the full lifetime of the Twitter accounts.

Keywords and Hashtags

Police accounts sometimes reference each other and often refer to news articles. Manchester and London are the only two locations mentioned by name, and only 63 of the total tweets included geolocation metadata. This is in accordance with the protest tweets (see above) but it may also indicate police may be communicating with a more local audience about specific locales, and that those who follow their account are familiar with the geographic region in which they operate.

The types of crimes most frequently mentioned include assaults and theft (punching, stolen, assault, burglary, robbery), as well as traffic accidents, with road, car, and collision among the top 100 terms. Both ‘please’ and ‘thankyou’ ranked in the top 30 terms police used, revealing a concern with polite communication with the public.

The analysis of specific police accounts with Twitonomy confirms the engagement of local communities through neighborhood/borough hashtags. Regional police accounts tend to use these hashtags more frequently to communicate local information throughout the police force. For example, 9 of the top 10 hashtags used by @metpoliceuk refer to boroughs, 8 of 10 for both @EssexPoliceUK and @SurreyPolice, and 7 of 10 for @WMPolice.

Local police forces, on the other hand, use hashtags to highlight specific local issues, and they also develop hashtags for crimes that occur in specific locales, including drunk driving (#nonefortheroad, #drinkordrive, #drinkaware, #alcoholharm), speeding (#speedwatch, #watchyourspeed, #slowdown), and theft (#loveyourphone, #burglary, #shoplifting, #giveathiefgrief, #60secondsecurity). Specific programs or campaigns pursued by local police forces are also communicated via hashtags on Twitter (#weedemoutweek, #mcrpride, #dagenhamopenday, #poga2015 – the Pride of Gorton Awards).

Twitter is used to engage communities and make police operations accessible and transparent. Hashtags like #trackacop and #tweetalong take the place of earlier 'ride alongs', allowing people to follow along with the routine activities of police. In Barking and Dagenham, police used the hashtag #dagenhamopenday to communicate with the public about their open house, which featured presentations by different police divisions about their work and allowed people to explore the inside of a holding cell or to try on police paraphanelia. Feedback from the public was largely positive, and the @MPSBarkDag engaged many twitter users who used this hashtag. Some local police forces have even developed hashtags for officers (#opcairo and #opolympus in Cheetham, #opmandera and #opramsey in Manchester City Centre).

Police use twitter to communicate with the public about major events such as sporting matches (#worldcup, #dagenhamopenday) and community celebrations (#prideweekend, #mcrpride, #communityactionday, #stpatricksday, #vfestival). In 2013 the Cheetham police force connected the hashtag #treacle2013 to the Treacle campaign, which "highlights the dangers and consequences of antisocial behaviour, criminal damage and the misuse of fireworks, and involves police working closely with Greater Manchester Fire and Rescue Service and local authorities."[10] While the police used this hashtag to warn of the dangers of unlicensed fireworks and report incidents that required police response during the festival, various local fire departments (@salfordfireteam, @Wiganfireteam, @manchesterfire, @manchesterfireteam) were far more active in using the hashtag.

While the Twitonomy analysis was more comprehensive temporally, its limitation concerned uncertainties of Twitter behaviour over time. It was unclear if individual accounts used Twitter consistently, or primarily for engagement around certain events. The popularity of event-specific hashtags indicates that some accounts may have engaged via Twitter more during these events. For other accounts, the popularity of hashtags that address types of petty crime indicate that police may be tweeting

[10] Greater Manchester Police, 2014. "Treacle." Accessed here: http://www.gmp.police.uk/content/section.html?readform&s=DA5C989DAD645FF380257C07004BB471

more consistently about recurrent problems, rather than focusing their efforts on specific days or occasions.

There do not seem to be any police hashtags referencing domestic exremism, disorder, or threats.

Networks

The Met Police are the most retweeted and mentioned police account in the UK. Other popular accounts include the Essex Police, West Midlands Police, and Greater Manchester Police. While most police accounts rarely retweet other accounts, they do retweet specific police- or crime-related tweets. Sussex_police retweeted sussex999events; essexpoliceuk retweeted ccessexpolice; nwpolice retweeted nwpcybercrime and nwprpu; leicspolice retweeted, among others, hatecrimeleics, emas_lgbt, and policediversity. The account @policinguk retweets police accounts around the UK.

The Manchester Evening News account @mennewsdesk features prominently in the Retweet Network Map as it retweets, particularly, civilian accounts on police matters. Several civilian accounts (e.g., human rights activist and freelance journalist @wnicholasgomes) retweet a large number of police tweets. Other active retweeters include retired police officers (e.g., @mptrg) and people with family members in the police force (e.g., @dl95ches).

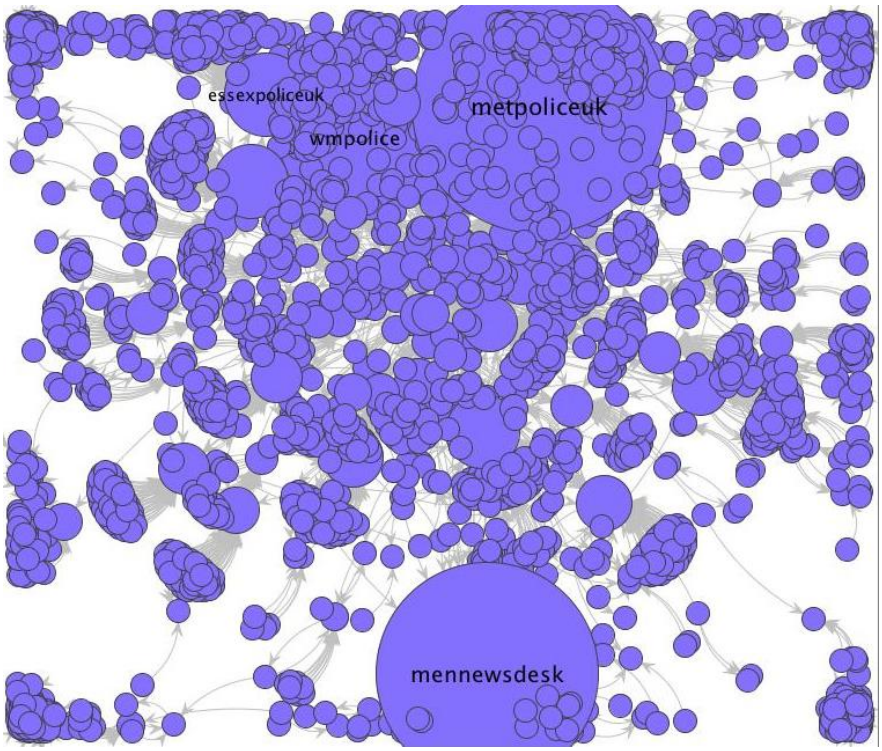

Fig. 6 Police Retweet Network Map

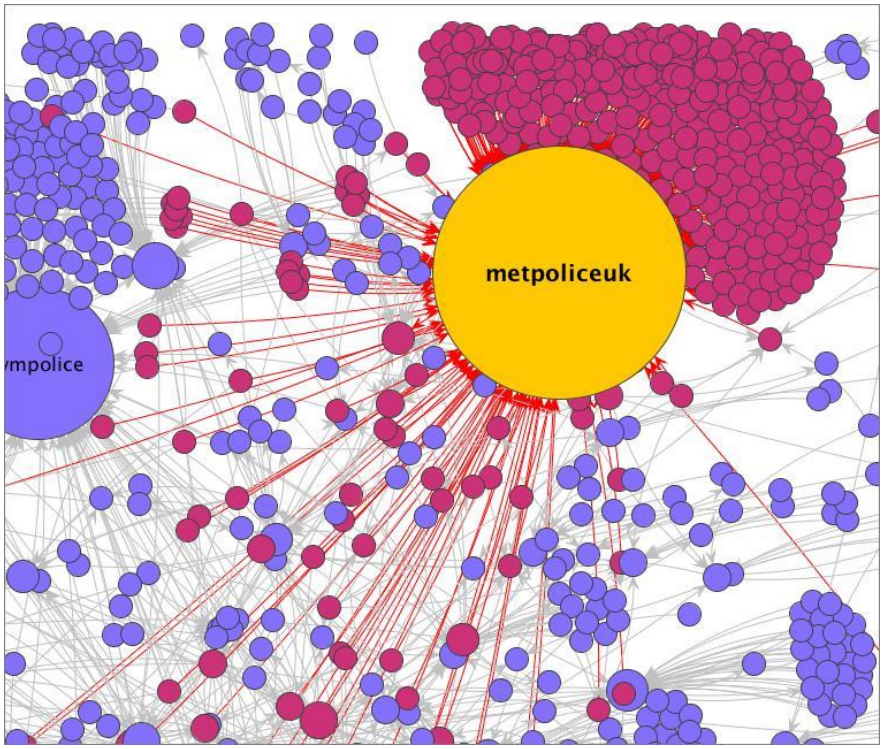

Fig. 7 @metpoliceuk Retweet Network Map
@metpoliceuk is the most retweeted account in the UK policing network. Pink circles represent twitter accounts that retweeted @metpoliceuk posts, with a red arrow pointing from the retweeting account to the retweeted @metpoliceuk.

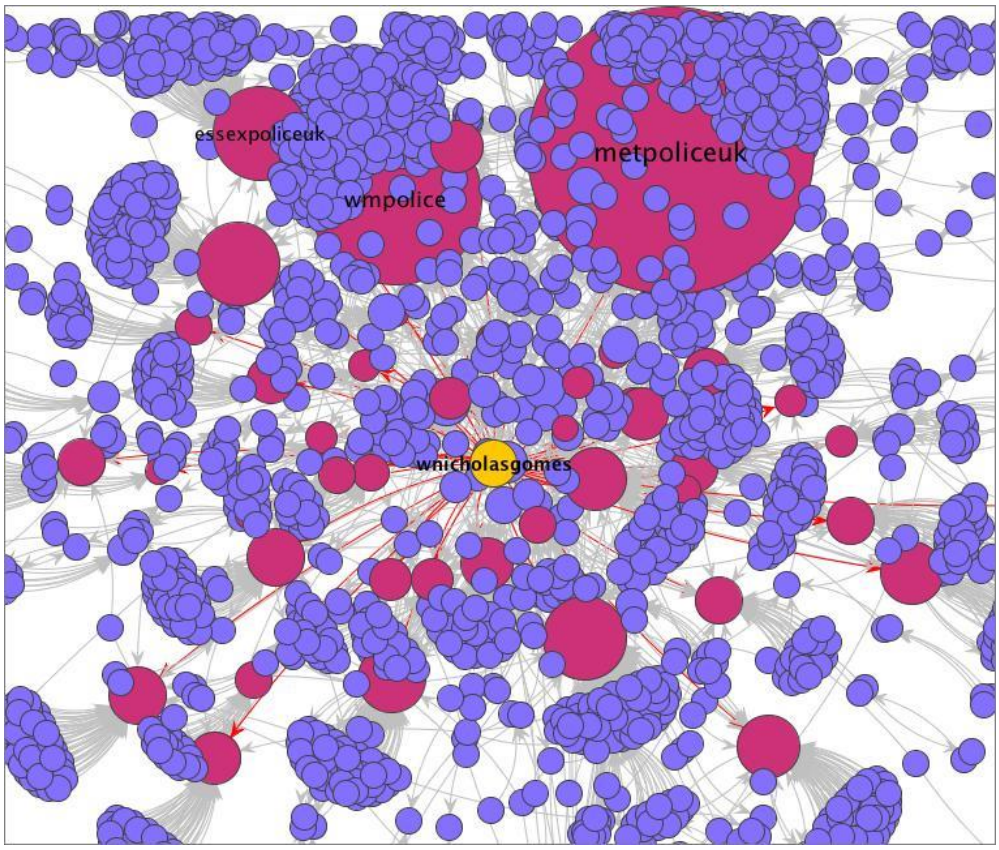

Fig. 8 Civilian Retweet Network Map for engagement with police accounts: @wnicholasgomes retweets police accounts across the UK. His active retweeting of

numerous police accounts situates him at the center of the network map, overlapping with other accounts that retweet numerous different police accounts.

The practice of mentioning other accounts follows closely the findings on retweeting. The same accounts that were frequently retweeted were also frequently mentioned in posts (@metpoliceuk, @wmpolice, @gmpolice, @sussex_police, @leicspolice, @policescotland). @Mennewsdesk appears to be the only popular non-police account that is mentioned in regard to police matters.

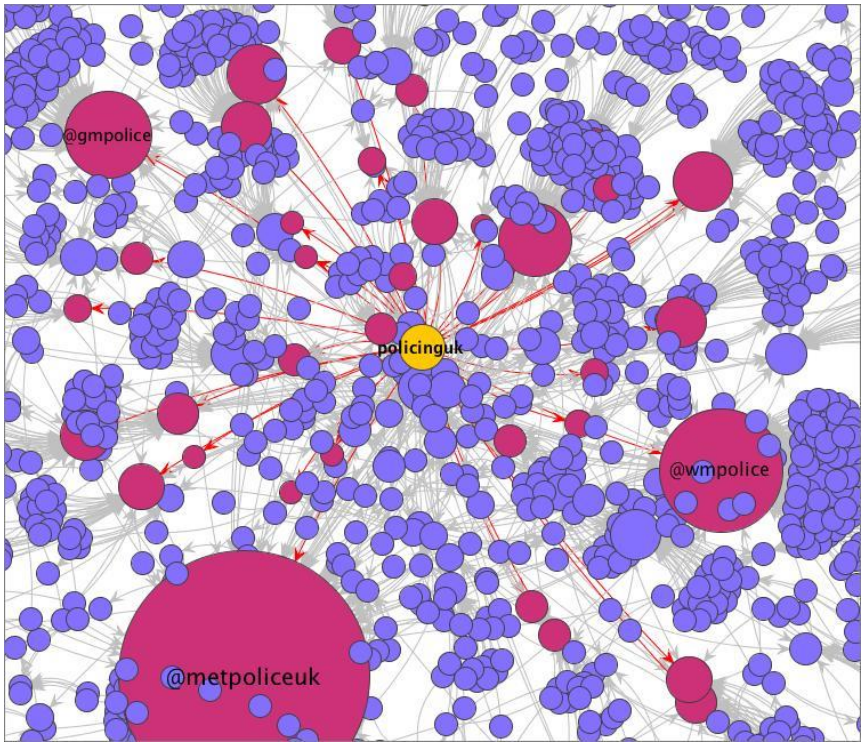

Fig. 9 Police Mention Network Map: @policinguk is at the center of both the retweet and mention network maps as it aggregates police accounts from the entire UK.

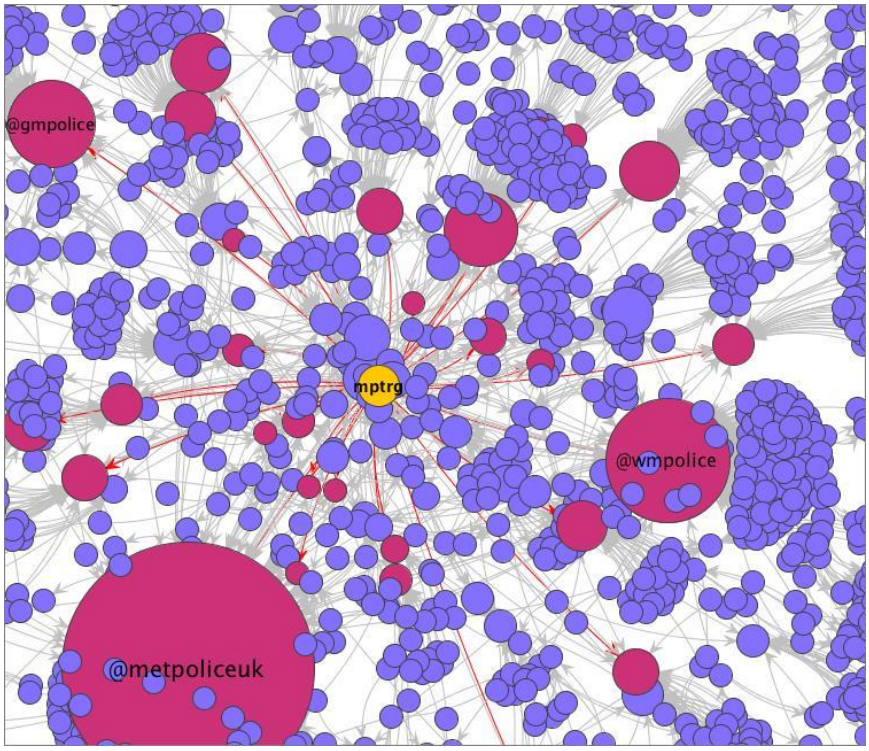

Fig. 10 Civilian Mention Network Map for engagement with police accounts: A close up of @mptrg shows that @mptrg is retweeting and mentioning other accounts, but none of these engage with @mptrg in exchange.

From the Twitonomy analysis, it appears that police accounts primarily engage civilians through the reply function, followed by mentions, and only rarely do police retweet posts shared by civilian users. In contrast, police accounts frequently engage local media through the retweet function.

Local police forces often retweet neighboring police accounts, as well as private accounts by prominent police officials. Regional police forces also retweet local police accounts within their force, as well as news sources and local interest and entertainment accounts. Significant differences in Twitter practices are apparent across the different police forces.

Platforms

Most police forces tweeted from phones – a mixture of iphones and androids - while some used ipads and computers. Third party programs like Hootsuite, Echofon, Tweet Deck, Crowd Control HQ, My News Desk, and Muster Point, are used widely. Often tweets from the same account originated from multiple platforms, which may mean that several officers have access to the account.

Discussion

In this section we will outline some of the key implications of our findings, both from the interviews as well as the big data analysis, and outline how these relate to on-going debates in the literature on big data and predictive policing as well as the current regulatory framework as outlined in earlier parts of the report. Although uses of social media for policing have been said to increase situation awareness and possibilities for pre-empting forms of violence and serious criminality, our findings also highlight a number of challenges that are associated with uses of social media data for policing. In this section we will focus on how current social media police practices introduce some key concerns regarding forms of algorithmic decision-making and automated processes in the policing of domestic extremism and disorder in the UK, particularly around questions of privacy, freedom of expression and accountability. Moreover, it will question some of the promises of big data for governance that have been prevalent in much debate, particularly around notions of objectivity and efficiency. As our findings indicate, such debate tends to underestimate the continued presence of human discretion and judgement in any algorithmically-informed decision-making (for good or bad). Finally, we will conclude by outlining how the advent of big social media data collection for policing domestic extremism and disorder might potentially incorporate new actors and forms of knowledge in the future that will have significant implications for how policing is organized and carried out.

Public and Private Data

As our findings indicate, a key area of contention is the issue of what constitutes public and private data on social media platforms. This is a familiar debate that has become increasingly pertinent in a context in which multiple actors collect and use social media data for multiple purposes that may not align with the user's initial intention. What has emerged from our interviews is the extent to which this continues to be a negotiation and an area of contention within the police. Although there is a perception within the police that 'open source intelligence' is a more fair and proportionate, and ultimately more legitimate, way of gathering intelligence than many previous tactics employed by police, it relies on a perception of 'public' information in an ambiguous way. As the stated police interpretation of RIPA indicates, the extent of this 'publicness' is limited as there is a recognition that monitoring data for police purposes is markedly different than monitoring data in a 'social' meaning, more closely aligned to the user's intent. This is also why repeated viewings of 'public' communication for police purposes is seen to be an invasion of privacy and requires further authorisation. As such, the integration of social media practices into policing has (re)introduced significant ambiguity around an individual's right to privacy and a right not to be under surveillance without reasonable suspicion of criminal activity. Moreover, collecting and using data that has been produced voluntarily in a way that fundamentally subverts the intended purposes of the production of that data speaks to what Andrejevic (2012) has described as the alienating dimension of the 'digital enclosure' in which online activity always has a dual character: the conscious action and the captured information. A social media platform might be an arena that facilitates for intentional actions of alerting others to certain information or posting a photo for a campaign or sharing stories of experiences with friends, but all these actions also generate additional unintentional

information: data about user-behaviour captured by the platform – and in this case, a third party actor in the form of the police. This dynamic introduces questions around how to conceptualise the public and the private that is able to also consider intent, particularly pertinent, as we have seen, for data-driven policing.

Accountability

Although police are increasingly trained to use social media for policing domestic extremism and disorder, a particularly interesting part of our findings speaks to the lack of knowledge and understanding of the algorithms in place for both the collection and analysis of social media data within the police. Partly this is due to the fact that police do not have their own software developers but rely on others to develop software for them, and also acquire tools and programmes through procurement processes, which means they cannot actively develop these tools and programmes with companies. The so-called 'black-box' of algorithms is a central feature of contemporary society, a system whose workings are mysterious, one in which we can observe its inputs and outputs, but we cannot tell how one becomes the other. Commonly, this is discussed in relation to the disempowering impact of the secrecy that surrounds algorithms for citizens who are tracked but have no clear idea of just how far much of this information can travel, how it is used, or its consequences (Pasquale 2015). However, with the advent of big data analysis for governance purposes by various state actors, the 'black-box' of algorithms takes on further meaning beyond the relationship between user and platform. The fact that actors making use of big data analysis for governance purposes, and perhaps particularly for policing purposes, without knowledge of the algorithms in place that produce the patterns, trends and networks that come to inform police strategies and pre-emptive tactics, introduces a significant issue regarding the lack of accountability. To some extent, our interviews indicate that it is recognized within police that some knowledge of how algorithms work is important for understanding the data properly. However, predominantly this knowledge is gathered from live-testing different programmes and tools over time. Knowledge of algorithms in this way remains superficial (speaking to the 'knowledge problem' that Pasquale identifies in a 'black-box society'), partly based on what has become a familiar conundrum in a digital age; the extent to which sectors making use of software need also to be able to understand and develop software. Certainly police officers cannot, at the same time, be software engineers. However, if predictive analytics is becoming a growing part of how policing is to be carried out, as our research indicates it is, then understanding the inner workings of predictive analytics as well as the results that they produce may be necessary in order to ensure police remain accountable regarding the tactics that they employ. The fact that these algorithms are designed by (often private and commercial) actors that lack public accountability and are informed by a set of interests that do not necessarily align with the broader context of law enforcement (and with that, protecting freedom of expression and freedom of assembly) only further highlights this concern.

Marketing-driven knowledge

As alluded to above, our findings highlight the significant part that is played by private and commercial actors in automated processes in the policing of protests. Most of the tools used by police are commercial tools either obtained 'off-the-shelf' or commissioned through a procurement model. As we have mentioned, this

introduces issues around accountability and the involvement of non-state actors in police practices. Moreover, the dominance of marketing-driven software development, which informs much of the commercial tools and programmes available for predictive analytics, also produces a particular type of data, and ultimately, knowledge. Debates in the emerging field of 'data science' have indicated the extent to which big data introduces a new epistemology and a new way of categorizing social phenomena. Our findings for this project illustrate the extent to which the algorithms that are developed and the categories that are used to order data are catered towards marketing needs and language. Our interviews, data analysis, and workshop highlighted this further by all integrating elements of terminology and salient categories of subjects and communication derived from the field of marketing. Notions such as 'sentiment' and 'influencers' are predominantly defined and identified on terms that speak to information that is important for marketing purposes. These same categories, and the basis upon which they are defined and identified (whether through reach or through negative or positive language), are being transferred onto analyses of data for entirely different purposes, such as law enforcement. Although some of these categories may be informative for police, they shift understandings of 'threats' towards particular communicative practices that have original meaning in a very different context. This is at its most obvious, perhaps, in the wish to incorporate big data analysis for monitoring reputational risks for the police. These practices may lead to a reinterpretation of 'threats' on quite alien terms that will have significant implications for freedom of expression and freedom of assembly.

Discretion

As our findings highlight, police continue to emphasize the role of 'human assessment' in any outcome produced by automated processes such as the analysis of big social media data. The 'biases' of any algorithmically produced pattern or identification of networks, groups and individuals, are 'corrected' within the police by integrating big data analysis with human intelligence and existing data bases. Thus, any action or tactic employed continues to rely on what was described as 'professional judgement'. Our own big data analysis based on the type of police practices involved in predictive protest policing, confirmed the extent to which analysis and interpretation of data depends on human discretion. As such, big data analysis is not an automated process in the way that is frequently assumed in debates on big data. The role of human input both in terms of designing the algorithms as well as any analysis and interpretation of such data remains central in data-driven governance. The notion that big data may absolve human errors and allow for 'objective' or 'efficient' forms of governance, therefore, is largely mythical in the context of this study at least. Rather, big data is predominantly used to identify patterns that are subjectively (humanly) interpreted and assessed, not least in the identification of any anomalies within these patterns. Thus, discretion (and as outlined in our background section, assumptions and ideology) is a key feature in data-driven policing. In particular, pre-existing knowledge, intelligence and broader societal understandings of events continue to shape and determine big data analyses. Whilst this helps correct the imperfections of the technology, it opens up possibilities for pre-existing human biases to enter predictive policing, resulting in the discriminatory implications that several researchers have highlighted (cf. Pena Gangadharan 2012). The use of supposedly 'objective' and 'neutral' data analysis may conceal that bias

and limits the potential for understanding the politics of data-driven forms of policing.

Interpreting unpredictability

Social media data analysis captures a wide range of communicative exchanges that are deemed 'public' (see above) and include vast volumes of personal and, most likely, harmless communication. This communication (as we have outlined previously) is contingent on particular contexts in which it emerges and is often difficult to understand in its decontextualized form as data. From this follows that much of this data will remain inconclusive and will not lead to predictable results. In this sense, ascertaining the probability of something happening (which relies on knowing all information) is not the same as ascertaining the predictability of something happening. A key question for practices of predictive policing is therefore how to deal with the uncertainty and unpredictability that remains with much (if not most) of social media data. If the goals and promises of predictive policing lead officers to interpret unpredictability as 'risk', this can become conducive to an environment of 'over-intervention' by the police. In other words, will what is (and inevitably will remain) 'possible' be interpreted as 'probable' and therefore lead to pre-emptive tactics? If the assumed possibility of predictive policing to pre-empt and therefore eliminate an increasing range of criminality means that a risk becomes interpreted as a possible threat, monitoring of, and intervention into, activity based on social media data is likely to expand, with implications for freedom of expression and assembly.

Intelligence and engagement

Our findings highlight an element of tension within the police with regards to how intelligence and engagement are related, with some expressing a concern with how these two areas overlap and the implications this might have for trust and community relations with the police. However, it is also clear that intelligence and engagement intersect significantly with the growth of social media practices in policing. In a context in which ambivalent understandings and definitions of extremism have come to the fore, the ways in which police interact with communities is evermore significant and complex. As social media serves as both a source of intelligence as well as a source of engagement, the policing of 'threats' and 'extremist' activity and communication blurs these boundaries and there is a sense in which further police engagement via social media may increasingly become an extension of intelligence gathering (noting that this has been argued by organisations such as Netpol to have always been the case, particularly around the policing of protests and political activism). As our big data analysis of police tweets indicates, currently police engagement on social media is fairly limited in terms of engaging with potential threats of domestic extremism. However, with programmes such as PREVENT becoming more encompassing, extending existing police practices in managing 'threats' to social media is already happening. This has been noted with high-profile cases of YouTube videos, for example, produced to provide a counter-narrative to Islamist fundamentalism. Such counter-narratives are also being provided on an ad hoc basis from different forces that are more active in terms of social media engagement, particularly around criticisms of police practices. The mention of 'Web Constables' in our interviews, currently in place in Estonia and Finland, that would engage with forums further highlights the potential direction of police engagement on

social media, as does the use of social media data for identifying reputational risks to the police as well as the possibility of identifying community concerns and needs. Although our findings indicate an expressed concern with the dangers of such developments, these types of social media practices would necessarily combine forms of surveillance and intelligence gathering with forms of engagement, unless there are clear guidelines and institutional structures preventing overlaps. As social media blurs these distinctions within policing, it (re)introduces significant questions regarding the role and nature of policing, particularly around protest, dissent and activism.

Out-sourcing of policing

Finally, our findings indicate a significant development in policing that further advances existing debates around state-corporate relations in governance structures, particularly around digital communications. A key area of contention in the debate that followed the Snowden leaks, for example, is the role that private intermediaries, such as in the form of social media companies, have come to play in carrying out state functions and particularly in the monitoring of citizens (cf. Hintz 2015). As we have found in our research with the police, this debate continues to be very prevalent in the use of social media data for policing. Although conflicting views emerged in our interviews on the question of the relationship between police and social media companies, it is clear that social media companies have come to be a significant actor in how policing around domestic extremism is carried out. Currently, much of this relationship is enacted through police requests for data from social media companies on a case-by-case basis. As such, potential 'threats' are identified by police monitoring social media platforms and subsequently making requests for data through the legal processes in place to which social media companies may or may not choose to comply. However, there is a significant sense in which this relationship may be changing with the British government calling for closer relations between police and major internet service providers, with more automatic processes in place for identifying potential 'threats' on social media platforms. Although Facebook's UK public policy director has recently stated that there is no algorithm in place for identifying content on its platform that violates its 'community standards' (i.e. material needs to be reported by someone in order for it to be removed), a key theme emerging in our interviews is the wish amongst parts of the police force for there to be a much more direct line of communication with the police regarding any such content, rather than the police having to first identify the threat. As was also recognized in our interviews, this would significantly alter the role of social media companies, effectively making them an extension of the police force.

Linked to this issue of the out-sourcing of policing, particularly with regards to domestic extremism and disorder, is the growth of a host of actors who are engaging in similar practices as the police when it comes to collecting and analyzing social media data. As we have demonstrated in this study with our own big data analysis of protests, academics and research institutions are increasingly engaging in social media research. This has a particular significance for policing as police practices of monitoring social media communication are subject to restrictions (e.g. repeated viewings of profiles, retention of data, as well as the use of false identities online) that are not in place for private bodies or research institutions. Such actors may be able to engage in more intrusive monitoring of social media activity. As both universities and corporations engage in intelligence gathering around domestic extremism and

disorder their activity introduces significant questions as to how such data collection and analysis may be, if at all, used by the police. This would further outsource police practices to non-state actors and would enhance the lack of both transparency and accountability around algorithmically-produced intelligence for policing purposes.

Recommendations/ Way forward

In this final section of the report we will outline some recommendations and way forward in considering the implications of social media uses for policing domestic extremism and disorder in the UK. These suggestions have emerged from the research results discussed in the previous sections and they draw on discussions that developed during our workshop with key civil society organisations, scholars in the field, and senior members of the British police force in September 2015. Building on our Discussion section, we want to highlight some of the areas that need consideration and debate as uses of big data become further integrated into policing.

Regulation and Governance

A key theme in the uses of social media by police concerns the inadequacies of existing regulation. As new counter-terrorism legislation is being drafted pertaining also to the use of social media for policing domestic extremism and disorder in the UK, a number of issues need clarification both within legislation and the organizational structure of policing:

- Definitions of 'domestic extremism' in relation to counter-terrorism legislation: As discussed earlier, current definitions are vague, have evolved over time, and boundaries between civil society-based protest and activism, on the one side, and terrorism, on the other, are unclear. In particular, definitions need to be clarified as the concept of 'extremism' is expanded to include non-violent activity.
- Distinctions between public, private and personal social media data: A more contextual understanding of the notion of 'public' interaction and communication is needed that takes account of contemporary social media practices. Rather than equating public with public record, an understanding of 'public' data needs to incorporate informed consent to data sharing across platforms and transfer of data in contexts beyond the user's original intentions.
- The oversight bodies in place to oversee police practices in a coherent way, particularly in light of the fragmented and ad hoc nature in which different forces engage in different social media practices: Despite existing guidelines by the Office of Surveillance Commissioners a strong and clear framework is currently lacking.
- The terms upon which authorization to carry out surveillance of particular individuals or groups on social media is granted: Clarification should consider requirements raised by different stakeholders, including police and civil society organisations, take into account the need for crime prevention as well as privacy protection, and limit data gathering to what is necessary and proportionate. Data collection and analysis needs to be conducted on the basis of internationally agreed human rights.
- Distinctions between social media uses for intelligence gathering and for police engagement: In order to avoid 'corrupted' transactions between communities and police, user trust in online environments would benefit from a clearer consensus amongst the police around the extent to which these practices are separate, if indeed they are.
- The role that non-state actors have in monitoring social media activity for the purposes of policing domestic extremism and disorder, and the rules that apply to them, particularly as they interact with state authorities: This includes private corporations, research institutions as well as social media companies.

- The legal and judicial avenues in place for challenging police uses of social media data, particularly for predictive policing of protests: Current frameworks for understanding rights to freedom of assembly and freedom of expression are considered by many civil society organisations as too vague and ambivalent to be adequate in this context.
- The extent to which decision-making informed by big data analyses within the police are considered in the context of other types of evidence, recognizing the structural biases and flaws of predictive analytics, as well as its inherent limitations regarding representativeness, accuracy and identification of 'risk'.

Transparency

The issue of accountability that we identified in our discussion speaks to a broader concern with a lack of transparency in how police uses social media and how social media activity comes to inform police strategies and tactics. Here we suggest a number of areas that need greater transparency:

- The algorithms in place in the software used by police to identify patterns and risks. These algorithms need to be known and understood by police in order for the actions they might take based on the intelligence produced by these algorithms to be accountable.
- The basis for any data collection, amount of data and retention of data collected from social media platforms, including information about how data is kept (for example, a requirement to disclose surveillance is already in place in Germany).
- The criteria in place for which activities, groups and individuals come to be identified as potential 'threats' and are entered into databases and intelligence reports relating to domestic extremism (and related to this, the difference between data, information and intelligence).

Ethics

As a host of non-state actors, both private and public, as well as police come to make use of big data analysis for different purposes and interests, a number of ethical issues have become pertinent that need to be integrated into the ethical guidelines for collection and analysis of data:

- Consideration that private information and metadata is typically connected to public messages on social media, and that private/public distinctions are rarely straightforward. This includes consideration for the user's intent and level of understanding that content is likely to be shared and used.
- Proper reasoning for engaging in covert surveillance, such as creating fake profiles or repeated viewings of profiles and accounts. This includes consideration for public interest and broader social implications of how research is conducted and the results produced.
- Sufficient oversight of practices that have relevance for police, including clear guidelines for how to store and handle sensitive data.
- Clear framework for relationship with law enforcement.
- Recognition that social media data is not representative.
- Transparency around algorithms and software used to collect and analyse data.
- Consideration for the contingency and uncertainty of much social media data, and recognition that unpredictability cannot be equated with 'risk'.

Future research

The role of algorithms and automated processes in structures of governance, particularly around predictive policing and pre-emptive tactics needs much more research. Here we suggest a number of possible questions for exploration:

- How do we evaluate the efficacy of social media monitoring for policing and to what extent is this practice more effective than other means?
- How should policing differ online and offline?
- What are the further implications of the Internet of Things and facial recognition software for policing?
- What are existing 'best practices' in using third party analytics for policing?
- What is the potential role of civic intermediaries in data management and analytics?

References

Allan, Stuart (2013) *Citizen Witnessing: Revisioning Journalism in Times of Crisis.* Cambridge: Polity Press.

Andrejevic, M. (2012) "Exploitation in the Data Mine." In C. Fuchs, K. Boersma, A. Albrechtslund and M. Sandoval, eds., *Internet and Surveillance: The Challenges of Web 2.0 and Social Media*, 71–88. Abingdon: Routledge.

Andrews, Lori (2013) *I Know Who You Are and I Saw What You Did: Social Networks and the Death of Privacy*. New York: Free Press.

Badger, Emily (2012) "How to Catch a Criminal With Data." CityLab from *The Atlantic*. March 14, 2012. http://www.citylab.com/tech/2012/03/how-catch-criminal-data/1477/

Bertot, John Carlo, Gorham, Ursula, Jaeger, Paul T., Sarin, Lindsay C. an Choi, Heeyoon (2014) "Big Data: Open Government and E-Government: Issues, Policies and Recommendations." *Information Polity* 19: 5-16.

Bhushan, Aniket (2014) "Fast Data, Slow Policy: Making the Most of Disruptive Innovation." *SAIS Review of International Affairs* 34(1): 93-107.

Boyd, Danah, Karen, Levy and Marwick, Alice (2014) "The Networked Nature of Algorithmic Discrimination." In: Gangadharan, Seeta Peña, Eubanks, Virginia and Barocas, Solon (Eds.) *Data and Discrimination: Collected Essays*. Open Technology Institute, New America. Accessed May 19, 2015. http://newamerica.org/downloads/OTI-Data-an-Discrimination-FINAL-small.pdf

Burnap, Pete, Matthew L. Williams, Luke Sloan, Omer Rana, William Housley, Adam Edwards, Vincent Knight, Rob Procter, and Alex Voss (2014) "Tweeting the Terror: Modelling the Social Media Reaction to the Woolwich Terrorist Attack." *Social Network Analysis and Mining* 4 (1): 206.

Burnap, Pete and Matthew L. Williams (2015) "Cyber Hate Speech on Twitter: An Application of Machine Classification and Statistical Modeling for Policy and Decision Making." *Policy and Internet* 7(2): 223-242.

Cook, Thomas D. (2014) "'Big Data' in Research on Social Policy." *Journal of Policy Analysis and Management* 33(2): 544-547.

Diamond, Larry, and Marc F. Plattner (2012) *Liberation Technology: Social Media and the Struggle for Democracy*. The Johns Hopkins University Press.

Dodd, V. (2014) "Chief constable warns against 'drift towards police state', *The Guardian*, 5 December, http://www.theguardian.com/uk-news/2014/dec/05/peter-fahy-police-state-warning

Edwards, Adam. (2015) "Big Data, Predictive Machines and Security: Enthusiasts, Critics and Sceptics." *Discover Society*. July 28, 2015.

http://discoversociety.org/2015/07/28/big-data-predictive-machines-and-security-enthusiasts-critics-and-sceptics/

Elmer, Greg, Ganaele Langlois and Joanna Redden, eds. (2015) *Compromised Data: From Social Media to Big Data,* New York: Bloomsbury.

Eubanks, Virginia (2014) "Big Data and Human Rights." In: Gangadharan, Seeta Peña, Eubanks, Virginia and Barocas, Solon (Eds.) *Data and Discrimination: Collected Essays.* Open Technology Institute, New America. Accessed May 19, 2015 http://newamerica.org/downloads/OTI-Data-an- Discrimination-FINAL-small.pdf

Ferguson, Andrew Guthrie (2012) "Predictive Policing and Reasonable Suspicion." *Emory Law Journal* 62.

Fidler, D.P., ed. (2015) *The Snowden Reader.* Bloomington: Indiana University Press.

Ganghadaran, Seeta Pena (2012) "Digital Inclusion and Data Profiling". *First Monday* 17(5), http://firstmonday.org/ojs/index.php/fm/article/view/3821/3199

Gillespie, Tarleton (2011) "Can an algorithm be wrong? Twitter Trends, the specter of censorship, and our faith in the algorithms around us." *Culture Digitally,* October 19, 2011.

Gillespie, Tarleton (2014) "The relevance of algorithms". In: T. Gillespie, P.J. Boczkowski and K.A. Foot (eds.) *Media Technologies: Essays on Communication, Materiality, and Society.* Cambridge: MIT Press, pp. 167-193.

Google (2014) Transparency Report 2014. Retrieved from http://www.google.co.uk/transparencyreport/.

Gross, Ana (2015) "The Domesticated Aboutness of Big Data Types." *Discover Society.* July 30, 2015. http://discoversociety.org/2015/07/30/the-domesticated-aboutness-of-big-data-types/

Haggerty, K.D., and R.V. Ericson (2000) "The surveillant assemblage". *British Journal of Sociology*, *51*(4), 605-622.

Halford, Susan (2015) "Big Data and the Politics of Discipline." *Discover Society.* July 30, 2015. http://discoversociety.org/2015/07/30/big-data-and-the-politics-of-discipline/

Her Majesty's Inspectorate of Constabulary (2014) "Policing in Austerity: Rising to the Challenge. Compendium." https://www.justiceinspectorates.gov.uk/hmic/wp-content/uploads/2014/03/valuing-the-police-compendium.pdf

Hintz, Arne (2015) "Social Media Censorship, Privatised Regulation and New Restrictions to Protest and Dissent." In: Lina Dencik and Oliver Leistert (eds.) *Critical Perspectives on Social Media and Protest.* Rowman & Littlefield.

Hoogenboom, B. (2006) "Grey Intelligence." *Crime Law Soc Change*, 45: 373-381

Howard, Alex (2012) “Predictive Data Analytics is Saving Lives and Taxpayer Dollars in New York City.” *O’Reilly Radar*. June 26, 2012. http://radar.oreilly.com/2012/06/predictive-data-analytics-big-data-nyc.html

Johnson, Shane D., Lucia Summers, and Ken Pease (2008) “Offender as Forager? A Direct Test of the Boost Account of Victimization.” *Journal of Quantiative Criminology* 25: 181-200.

Jones, Chris (2014) “Predictive Policing: Mapping the Future of Policing?” *Open Democracy*. https://www.opendemocracy.net/opensecurity/chris-jones/predictive-policing-mapping-future-of-policing

Joseph, George. (2015) “Undercover Police Have Regularly Spied on Black Lives Matter Activists in New York.” *The Intercept*. Accessed here: https://theintercept.com/2015/08/18/undercover-police-spied-on-ny-black-lives-matter/

Kelly, Heather (2014) “Police Embracing Tech that Predicts Crimes.” *CNN*. May 26, 2014. http://edition.cnn.com/2012/07/09/tech/innovation/police-tech/

Kelling, George L. and William Bratton (1998) “Declining Crime Rates: Insiders’ Views of the New York City Story.” *Journal of Law and Criminology* 88(4): 1217-1232.

Khamis, Sahar, and Katherine Vaughn (2011) “Cyberactivism in the Egyptian Revolution: How Civic Engagement and Citizen Journalism Tilted the Balance.” *Arab Media & Society,* http://www.arabmediasociety.com/?article=769

Kitchin, Rob (2014a) *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences.* London: Sage.

Kitchin, Rob (2014b) “Thrinking critically about and researching algorithms”. The Programmable City, Working Paper 5, 28 October 2014.

Koehn, Josh (2012) “Algorithmic Crimefighting.” *San Jose.com*. February 22, 2012. http://www.sanjose.com/2012/02/22/sheriffs_office_fights_property_crimes_with_predictive_policing/

Lerman, Jonas (2013) “Big Data and Its Exclusions.” Stanford Law Review Online 66(55): 55-63.

Lowe, T. and Innes, M. (2012) “Can we speak in confidence? Community intelligence and neighbourhood policing v2.0.” *Policing and Society: An International Journal of Research and Policy,* 22(3): 295-316.

Lyon, D. (2007) “Surveillance, Power, and Everyday Life”. In R. Mansell, C. Anthi Avgerou, D. Quah and R. Silverstone, eds., *The Oxford Handbook of Information and Communication Technologies.* Oxford/New York: Oxford University Press, 449-472.

Lyon, D. (2014) “Surveillance, Snowden, and Big Data: Capacities, Consequences, Critique.” *Big Data & Society* (July–December 2014), 1–13.

Mackenzie, A. (2007) “Protocols and the irreducible traces of embodiment: The Viterbi Algorithm and the mosaic of machine time.” In: R. Hassan and R.E. Purser (eds.) *24/7: Time and Temporality in the Network Society.* Stanford: Stanford University Press, pp. 89-106.

Mackenzie, D. (2008) *An Engine, Not a Camera. How Financial Models Shape Markets.* Cambridge: MIT Press.

Margetts, Helen and David Sutcliffe (2013) “Special Issue: Potentials and Challenges of Big Data.” Policy & Internet 5(2): 139-146.

Mayer-Schoenberger, Viktor, and Kenneth Cukier (2013) *Big Data: A Revolution That Will Transform How We Live, Work and Think,* New York: John Murray.

Mohler, G. O., M.B. Short, P.J. Brantingham, F.P. Schoenberg, and G.E. Tita (2011) “Self-Exciting Point Process Modeling of Crime.” *Journal of the American Statistical Association* 106 (493) 100-108.

Paasche, Till F. (2013) “Coded Police Territories: ‘Detective Software’ Investigates.” *Area* 45 (3): 314–20.

Pasquale, F. (2015) *The Black Box Society*. Cambridge, MA and London: Harvard University Press

Procter, Rob, Jeremy Crump, Susanne Karstedt, Alex Voss, and Marta Cantijoch (2013a) “Reading the Riots: What Were the Police Doing on Twitter?” *Policing and Society* 23 (4): 413–36.

Procter, Rob, Farida Vis, and Alex Voss (2013b) “Reading the Riots on Twitter: Methodological Innovation for the Analysis of Big Data.” *International Journal of Social Research Methodology* 16 (3): 197–214.

Quinn, B. (2015) “City of London police put Occupy London on counter-terrorism presentation with al-Qaida”, *The Guardian*, 19 July. http://www.theguardian.com/uk-news/2015/jul/19/occupy-london-counter-terrorism-presentation-al-qaida

Robertson, Hamish and Joanne Travaglia (2015) “A Politics of Counting – Putting People Back into Big Data.” *Discover Society.* July 30, 2015. http://discoversociety.org/2015/07/30/a-politics-of-counting-putting-people-back-into-big-data/

Schäfer, M. (2014) “Policing the Social Media. Control and Communication in a networked Public Sphere”. Paper presented at Social Media and the Transformation of Public Space conference, University of Amsterdam, The Netherlands

Swain, Val (2013) “Disruption Policing: Surveillance and the Right to Protest.” Open Democracy. https://www.opendemocracy.net/opensecurity/val-swain/disruption-policing-surveillance-and-right-to-protest

Swain, Val (2015), PhD thesis, University of East Anglia, UK

The Guardian (2015) “The NSA Files.” Retrieved from http://www.theguardian.com/us-news/the-nsa-files.

Trottier, D., and Lyon, D. (2012) “Key Features of Social Media Surveillance.” In C. Fuchs, K. Boersma, A. Albrechtslund and M. Sandoval, eds., *Internet and Surveillance: The Challenges of Web 2.0 and Social Media*, 89–105. Abingdon: Routledge.

Trottier, Daniel (2015) “Open Source Intelligence, Social Media and Law Enforcement: Visions, Constraints and Critiques.” *European Journal of Cultural Studies* 18 (4-5): 530-547.

----- (2012) “Policing Social Media.” *Canadian Review of Sociology* 49 (4): 411–25.

Vitale, Alex (2006) “From Negotiated Management to Command and Control: How the New York Police Department Polices Protests.” *Policing and Society* 15 (3): 283-304.

Walby, Kevin, and Jeffrey Monaghan (2011) “Private Eyes and Public Order: Policing and Surveillance in the Suppression of Animal Rights Activists in Canada.” *Social Movement Studies* 10 (1): 21–37.

Williams, Matthew L., Adam Edwards, William Housley, Peter Burnap, Omer Rana, Nick Avis, Jeffrey Morgan, and Luke Sloan (2013) “Policing Cyber-Neighbourhoods: Tension Monitoring and Social Media Networks.” *Policing and Society* 23 (4): 461–81.

Wright, P. (2013) “Meet Prism’s little brother: Socmint.” *Wired*, 26 June. http://www.wired.co.uk/news/archive/2013-06/26/socmint

Abbreviations

COSMOS – Collaborative Online Social Media Observatory

CTIRU – Counter Terrorism Internet Referral Unit

DRIPA – Data Retention and Investigatory Powers Act

DSEI – Defence and Security Equipment International

ESRC – Economic and Social Research Council

NDEDIU – National Domestic Extremism and Disorder Intelligence Unit

NDET – National Domestic Extremism Team

NETCU – National Extremism Tactical Coordination Unit

NHS – National Health Service

NPoCC – National Police Co-ordination Centre

NPOIU – National Public Order Intelligence Unit

OSINT – Open Source Intelligence

RIPA – Regulation of Investigatory Powers Act

SOCMINT – Social Media Intelligence

Acknowledgments

We would like to thank the Media Democracy Fund, Open Society Foundations and Ford Foundation for funding the project. We would also like to thank all the interviewees in our sample for participating in our research and the COSMOS team, particularly Dr Jeffrey Morgan, for assisting with our big data analysis. Further thanks to Professor Martin Innes, Dr Joanna Redden, and all the participants at our workshop for their guidance and advice.