the state of the art 2015 a literature review of social media intelligence capabilities for counter-terrorism

Jamie Bartlett
Louis Reynolds

September 2015

DEMOS

Published by Demos September 2015

Unit 1, Lloyds Wharf
2-3 Mill Street
London
SE1 2BD

hello@demos.co.uk
www.demos.co.uk

This paper was funded by Public Safety Canada's Kanishka Project research initiative. From Demos we would like to thank Theo Bass, Alex Krasodomski-Jones, Carl Miller, and above all Ali Fisher, for their support on this project. The views expressed are those of the authors and do not necessarily reflect those of Public Safety Canada. All errors or omissions are ours.

CONTENTS

INTRODUCTION

This paper is a review of how information and insight can be drawn from open social media sources. It focuses on the specific research techniques that have emerged, the capabilities they provide, the possible insights they offer, and the ethical and legal questions they raise. These techniques are considered relevant and valuable in so far as they can help to maintain public safety by preventing terrorism, preparing for it, protecting the public from it and pursuing its perpetrators. The report also considers how far this can be achieved against the backdrop of radically changing technology and public attitudes towards surveillance.

This is an updated version of a 2013 report paper on the same subject, *State of the Art*. Since 2013, there have been significant changes in social media, how it is used by terrorist groups, and the methods being developed to make sense of it. In particular, the context in which this research took place has changed in a number of significant ways. First, there have been continuing concerns about internet surveillance and privacy following the revelations of NSA contractor Edward Snowden. Second, and partly as a result, there have been changes in the way that people use social media, and indeed in social media companies and platforms themselves. Third, so-called Islamic State (IS) made social media a central component of their *modus operandi* – particularly for propagandistic purposes. Finally, the methods and software used to make sense of social media have continued to improve.

Social media research has emerged as a practice, but is still not yet a coherent academic discipline or distinctive intelligence tradecraft. It is neither a distinct area of study, nor driven by a united research community. It is conducted across the public, private and academic sectors, spanning disciplines from the computer sciences and ethnography to advertising and brand management. Its aims range from understanding the topography of social networks comprising millions of individuals to the deep, textured knowledge of the social worlds of individuals and small groups.

As such, techniques and approaches often reflect specific disciplinary traditions and rarely refer to those found elsewhere. Social media research is also fragmented by platform. There is already a distinct nascent discipline surrounding Twitter, driven by free access to millions of tweets, an easily available Application Programming Interface (API) and fewer concerns about privacy and intrusion. Since 2008, the corpus of work on 'Twitterology' has grown from a handful to hundreds of research papers, covering everything from topic identification to event detection and political forecasting.

Research on Facebook – either about it or using it – has struggled in the face of technological difficulties in acquiring the data and Facebook's corporate orientation towards advertising rather than research. As of 2011, there were 42 peer reviewed journal articles about Facebook research, although this number is rising quickly. However, since 2013, this has changed, and there are a growing number of academic papers related to Facebook too.[1]

The overall aim of this review is to describe the emerging contours of social media research to codify the capabilities that have emerged and the opportunities they have created, and the risks and hurdles that they must commonly face and surmount –methodological, legal and ethical – in order to usefully contribute towards countering terrorism in a way that is publicly supported and effective.

A semi-systematic literature review methodology was employed. The purpose of the original review was defined with an explicit statement of focus and further refined following a series of short meetings with a small group of likely consumers of that paper in March 2013. On the basis of these meetings, studies were consistently included and excluded on the basis of agreed criteria.'[2] In total, 112 papers were analysed, and their key contribution to the question of counter-terrorism capability was identified and recorded. Further notes were incidentally made on location, date, method, and overall thesis. The results were synthesised into categories of capability, as set out below.

In May–June 2015 Demos researchers conducted a second literature review using the same criteria, and in the same subject areas, with all relevant areas updated. Additional research was included to:

- Update all statistics relating to social media usage
- Update public opinion polling and other research relating to attitudes about monitoring and surveillance;
- Review trends in changing social media use by terrorist groups; specifically IS
- Examine the growth in new encrypted social media platforms and other technological development – such as 'dark net hidden services' – and their potential uses.

In total another 133 papers were analysed. Where it was felt that areas of interest had not significantly changed since the original review, the text of the report has not changed.

Caveats

It is notable that very little social media research found was directly related to counter-terrorism work, but much had, when extrapolated, implications for counter-terrorism. Therefore, we have provided reflections where necessary based on our research and judgment. We have made this clear throughout. Secondly, we noted in 2013 that there was a large difference between current capabilities and capabilities that had been published at the time of writing. This is still the case, and is likely to remain so in the future. We do not have access to a great deal of use-cases – including novel techniques, novel applications of techniques or substantive findings – that are either in development or extant but unpublished. Academic peer-reviewed publishing can take anywhere from six months to two years, while many commercial capabilities are proprietary.

Furthermore, much social media research is conducted either by or on behalf of the social media platforms themselves and never made public. The growing distance between development and publishing, the increasing role of proprietary methodologies and private sector ownership, and exploitation of focal data sets are important characteristics of the social media research environment.

Finally, this paper does not consider techniques to acquire or use closed or private information, or methods by which detailed profiles of individuals can be built. Particularly since the Edward Snowden revelations, more is known about some of these methods now than in 2013. However, these techniques are more readily situated within the gamut of secret intelligence work rather than research, and beyond the scope of the authors' expertise.

Structure

The paper is structured as follows:

Part 1 is an overview of social media use, focused on how it is used by groups of interest to those involved in counter-terrorism. This includes new sections on trends of social media platforms; and a new section on Islamic State (IS).

Part 2 provides an introduction to the key approaches of social media intelligence (henceforth 'SOCMINT') for counter-terrorism.

Part 3 sets out a series of SOCMINT techniques. For each technique a series of capabilities and insights are considered, the validity and reliability of the method is considered, and how they might be applied to counter-terrorism work explored. The techniques examined in this manner are:

- Machine learning & Natural Language Processing
- Event detection
- Predictive analytics (notably non-machine learning based)

- Network Analysis
- Manual analysis / ‘netnography’
- Solicited / ‘crowd sourced’ insight

Part 4 outlines a number of important legal, ethical and practical considerations when undertaking SOCMINT work.

PART 1: OVERVIEW OF SOCIAL MEDIA USE

Summary

- Social media use continues to grow in both volume and diversity of available platforms.

- Since the Edward Snowden revelations (and even before then), the number of platforms and messaging systems which include some form of default encryption has increased.

- These systems are increasingly popular among several (perfectly legitimate) groups and individuals, and are also widely discussed and (we suspect) used by terrorist groups.

- IS, and other groups such as Jabhat al Nusra, have demonstrated both the difficulty in censoring and removing social media content, and the potential uses of social media to reach very large audiences at low cost. They use a variety of platforms and strategies to remain always online, and consider social media an important part of their 'jihad'.

- We anticipate a growth in the availability of default encrypted social media services, anonymous social networks, and decentralised distributed social networks, which are run without centralised servers or administrators (sometimes called a 'distributed trust' network). The implication is that censorship will become more difficult.

- We anticipate the increasing integration of such services within the multi-platform networks of jihadist groups and their supporters. This would represent an increased network resilience.

- Given this increased difficulty of censorship, more attention has been placed on 'counter-speech', which comprises efforts by individuals and groups to counter extremist or terrorist messaging online. However, the likely effectiveness of these measures is not yet clear.

Trends in use

Loosely grouped, 'social' media provide the means for the way in which the internet is increasingly being used: to participate, to create and to share information about ourselves and our friends, our likes and dislikes, movements, thoughts and transactions.

Although social media can be ‘closed’ (ie not publically viewable), the underlying infrastructure, philosophy and logic of social media make it, to varying extents, ‘open’: viewable by certain publics as defined by the user, the user’s network of relationships, or anyone.

The most well-known platforms are Facebook (the largest, with around 1.4 billion users), YouTube and Twitter. Estimates for Q1 2015 suggest there are two billion active social media accounts worldwide, the equivalent of an account for two in every three internet users. On an average day, Facebook users spend 9.7 billion minutes on the site, share 4 billion pieces of content and upload 250 million photos. Facebook is further integrated with 7 million websites and apps.

However, a much more diverse (linguistically, culturally and functionally) family of platforms spans social bookmarking, micromedia, niche networks, video aggregation and social curation. The specialist business network LinkedIn has 200 million users, the Russian-language VK network 190 million, and the Chinese QQ network 700 million. Platforms such as Reddit (which reported 400 million unique visitors in 2012) and Tumblr, which has just reached 100 million blogs, can support extremely niche communities based on mutual interest.

Eighty-seven per cent of Canadian households are connected to the internet and spend on average 17.2 hours online every week, which includes watching an average of one hour of online videos every day (80 per cent of it on YouTube).[3] Furthermore, 57 per cent of all Canadians owned a Smartphone in 2013, and more than a quarter (26 per cent) used a mobile device to access social media services.[4] The most recent study on the subject found that general internet usage is higher among Anglophone than Francophone Canadians. This gap, slowly diminishing, is more pronounced in older age groups, and non-existent in the 18–34 age group.

Canadians are among the earliest and most enthusiastic adopters of social media. In 2013, it was reported that almost 24 million Canadians (69 percent of the population) visited a social media site at least once a year.[5] Canadians spend on average over two hours a day on social media platforms, slightly below the global average of 2.4 hours a day.[6] This is double the time spent in 2013. The social media agency WeAreSocial estimates that almost half of Canadians have a social media account, of which the vast majority are mobile-enabled.[7]

What was just a few years ago thought a ‘new’ form of media – social, open to certain publics, personalised – is now the dominant form of digital media. Three of the top ten most visited sites worldwide are social media platforms, and seven of Canada’s top 20.[8]

The past two years have seen a 'levelling off' in demographic change on the largest social media channels as younger users, traditionally early adopters, move to newer platforms and laggards catch up. The fastest-growing demographics on both Facebook and Twitter are the over-65s. The number of over-65s on Pinterest and Twitter doubled between 2013 and 2014.[9]

Age, unsurprisingly, strongly characterises social media use in Canada: 18 to 25-year-olds spend almost twice as much time on social media network sites as those over 55. (Nonetheless, every age group in Canada is above the worldwide average). In the younger age groups, male and female users are roughly similarly represented, but in older age cohorts women tend to use social media in significantly higher numbers than men.

In terms of use, 61 per cent of Canadians use social media to stay connected with friends and family, 39 per cent to stay connected with professional contacts, and 55 per cent to stay up to date on news and general items of interest. In any typical month, 44 per cent update their status on one platform or another, 38 per cent post photos, 17 per cent post videos, and 14 per cent share their GPS location on a social media network.

As in many other countries, Facebook is the most popular social media platform, although the precise numbers, especially when concerned with actual use rather than formal membership, are controversial. A recent AskCanadians survey found that 73 per cent of Canadian social media users were on Facebook, 35 per cent use YouTube, 21 per cent use LinkedIn and Twitter, 19 per cent use Google+, 5.3 per cent use Pinterest and Flickr, 3.3 per cent use Tumblr, 3 per cent use Instagram, 2.4 per cent use MySpace, and 1.7 per cent use Foursquare. (However, a recent survey by Forum Research found 25 percent of Canadians surveyed now use Twitter, just below the 30 per cent that used LinkedIn. According to a 2015 poll by eMarketer, 23 percent of users reported going on Twitter at least once a day.)[10]

Changing types of social media platforms

Social media is not a static set of technologies – and there are signs that the infrastructure of social media is likely to change. At the heart of the evolution of terrorist use of social media for propaganda purposes are changes in social media itself.

Partly as a result of the Edward Snowden revelations, and partly due to growing awareness of the value of personal information, there is increased public concern about personal data and privacy. (We have reviewed these statistics below). The result could have significant ramifications for how social media works, and how people use it. There are four components to the way in which citizens and companies have responded to these concerns.

First, social media users have started to behave in slightly more private ways. A 2013 survey by Ask Your Target Market reported that 46 per cent of social media users said that all of their social media profiles are set to private so people cannot search for them. Only nine per cent of respondents said that they would continue to use a site that did not allow you to enable a private profile, compared to the 46 per cent who said they probably would not use the platform anymore.[11] A study by Consumer Reports in the US also suggested that 37 percent of the Facebook app users have used privacy tools to customise how much information the app is allowed to see.[12]

Second, social media companies themselves have started to introduce new default encryption services. For example, Facebook recently allowed users to add a PGP ('Pretty Good Privacy') key to their messenger services, which enables users to send encrypted text-based messages to each other. Others, including Apple and Google, have introduced default 'end-to-end' SSL (Secure Sockets Layer) encryption. We believe this to be partly an attempt to assure their users that they are responding to concerns about data privacy.

Third, there has been an increase in the use of various types of encryption by internet users (which will inevitably include social media users). Anonymous browsers like 'Tor' are used to browse the net without giving away the user's location. Such browsers can then be used to access the 'Hidden Services', an encrypted network of sites that uses a non-standard protocol, making it close to impossible for websites or people who use them to be tracked. These tools are becoming ever-more popular: there are now around between 2 and 3 million daily users.[13] Facebook recently allowed users to access the site via a Tor Hidden Service.[14]

The fourth component, and perhaps the most far-reaching of all, is the range of new types of social media platforms being created, such as the 'anti-Facebook', ad-free social network site Ello. 'Collecting and selling your personal data, reading your posts to your friends, and mapping your social connections for profit is both creepy and unethical,' states the site. More significant is the growth in new privacy-enhanced software and social media platforms. Soon there will be a new generation of easy-to-use, auto-encryption internet services, such as MailPile, and Dark Mail, both of which are email services where everything is automatically encrypted.

There are even more revolutionary plans in the pipeline. One important development is how the protocol behind the crypto-currency bitcoin is being applied to social networking sites. Bitcoin creates an immutable, unchangeable public copy of every transaction ever made by its users, which is hosted and verified by every computer that downloads the software. This public copy is called the 'blockchain'. However, this public copy can also be used for other applications, not just currency transactions. The Ethereum project is dedicated to creating a new, blockchain-operated internet. Ethereum's developers hope the system will herald a revolution in the way we use the net – creating a network of computers that are distributed and encrypted, making it very difficult to censor and with no single point of failure (sometimes called a 'distributed trust model'). There is already a domain name system that cannot be removed called Namecoin,and an untraceable email system call Bitmessage, which we expect to grow rapidly over the next 2–3 years.

The use of social media by extremist and terrorist groups

Extremist and terrorist groups from across the ideological spectrum have long used the internet for a wide range of purposes including community and operational communication and propaganda, technical information sharing and intelligence gathering, recruitment, training, financing and equipment procurement. However, terrorists' and extremists' use of the internet has evolved rapidly as a result of new technological opportunities, the proliferation of social media platforms, developments in online policing and a number of other factors.

By 1999 nearly all known terrorist groups had established a presence on the internet. By the mid-2000s, most terrorist and extremist groups had completed the transition from text-heavy websites to interactive forums. In turn, by the late 2000s the use of interactive forums by many groups had begun to stagnate or decline, and mainstream social media platforms like Twitter and Facebook started to become more important avenues for propaganda, recruitment and information-sharing within the various Islamist extremist communities.[15] For example, jihadi forums like *al-Falluja*, *al-Fidaa*, *al-Shmukh* and *Ansar al-Mujahideen Arabic Forum* have experienced declines in or the stagnation of activity.[16] This relative decline in the importance of internet forums within some extremist movements – as well as merely mirroring the decline in use of internet forums in general – has been catalysed by declining levels of general trust as the policing of established forums has increased.

However, for many groups, online forums remain important communication tools. There is often a strong relationship between newer profiles and pages on social media platforms and dedicated forums, with extremist groups' presence on social media platforms often acting as a public-facing gateway to forums, complementing the more insular environment of dedicated websites with a more public facing presence.[17] For example, many right-wing groups use social media as a way to redirect a broader audience towards their dedicated forums, and as a result of continual deletion of White Nationalist pages and accounts on Facebook and Twitter, certain older forums like *Stormfront* remain popular.[18] Many Islamist forums also remain active, but in many cases the type of content has changed, and much of the communication which takes place on Islamist extremist forums is low-stakes activity, such as propaganda sharing (eg isdarat.tv) and ideological discussions.[19] Instead of the increased use of social media platforms by extremist groups supplanting the use of online forums, the trend has been towards a diversification of the online resources (see the section on IS, below).

While a select few, high profile uses of social media by extremist groups – like Al-Shabaab's and IS's use of Twitter[20] – have captured significant media attention, the increasing use of social media platforms by terrorist and extremist groups over the last few years has reached across ideological dividing lines and around the globe.

The Ogaden National Liberation Front in Somalia, Patani separatists in Thailand,[21] Chechen terror groups in the Caucasus, Uyghur militants in China,[22] White Nationalists in Sweden and Islamist extremists in the UK all use Youtube to praise martyrs, share speeches, distribute music videos, react to political events and generally spread propaganda.[23] Al-Muhajiroun exploits its international network of YouTube channels in a particularly sophisticated manner, employing cohesive, emotive Islamist messaging featuring local emirs and Islamist activists in nationally-targeted videos.[24] In Africa alone, The Mujahideen Youth Movement, Al-Qaeda in the Lands of the Islamic Maghreb, the Oromo Liberation Front, the Movement for the Emancipation of the Niger Delta and Boko Haram all maintain Twitter profiles (at least intermittently), often providing content in English, focused on the media and Western audiences.[25] Even more heavily censored sites like Facebook are regularly used by extremists as diverse as Hizb ut Tahrir, Mouvement Pour L'Unicite et le Jihad en Afrique de L'Ouest, Boko Haram, Bulgarian anti-Roma groups,[26] Austrian Neo-Nazis,[27] and British Islamists.[28] Al Qaeda and its affiliates, as well as IS and Chechen extremist communities, have used Instagram and Flikr to circulate propaganda, [29] while IS has used Tumblr to recruit Western women to jihad in Syria and Iraq, Ask.FM to host ideological discussions, and JustPaste.It to host content taken down from other platforms.[30]

Not only are numerous extremist groups using a wide range of social media platforms, but they are also, in some cases and by some measures, producing large volumes of social media content. Indeed, whereas production of propaganda for terrorist movements was historically controlled by the movement itself, social media has allowed individuals from around the world to be part of the production and distribution of content. This, it is broadly agreed, has resulted in a significant increase in the volume of terrorist related content available online.

The UK's Counter Terrorism Internet Referral Unit, formed in early 2010, removed 49,000 pieces of extremist online content between its formation and October 2014; yet the Centre reported that it had removed 30,000 pieces of content between December 2013 and October 2014 alone, the vast majority of which referred to Syria and Iraq.[31] The Simon Wiesenthal Center's 2014 Report on Digital Terrorism and Hate recorded more than 30,000 websites, forums, and social media accounts promoting terrorism in the US and abroad, with 'a shocking rise in the use of social networking by extremists for recruitment and to denigrate "the enemy".'[32] A recent study by the Brookings Center for Middle East Policy found that between September and December 2014, 46,000–70,000 IS supporter accounts were active on Twitter alone, with the average account having 1,000 followers.[33]

These figures are hard to verify, partly because of the speed with which accounts are removed and then re-created. While studies such as those conducted by the Simon Wiesenthal Center and the Brookings Institute can provide snapshots of discrete portions of extremist activity online, it remains exceptionally difficult to undertake comprehensive assessments of the volume of extremist activity, hate speech or terrorist group representation on social media.

Most extremist and terrorist organisations use public-facing social media platforms for a common core of purposes: as a means of promoting group cohesion as well as facilitating informal communication and socialisation between members; and as a way of spreading propaganda outside of core group membership and into other online communities, particularly semi-radicalised individuals, extremist sympathisers, people vulnerable to radicalisation and the media.

However, patterns of social media often differ significantly between groups. For example, White Nationalist groups often use social media as a means of marketing extremist material, music and merchandise (a key propaganda and revenue function) and of distributing music,[34] while amongst Islamist terrorist groups, there is often a sharp delineation between social media activity targeted at Arabic speakers and activity targeted at European language speakers, often in English.[35] Islamist extremist organisations often distribute videos praising martyrs as a means of targeting vulnerable individuals. [36] It is, of course, entirely possible that this will change in the future, and these distinctions should not be considered as set in stone.

The extent to which propagandistic material has an effect on individuals who view it remains unclear. It certainly increases the number of potential recruits that terrorist groups can reach.[37] It is generally believed that prolonged exposure to violent or graphic imagery catalyses desensitisation towards violence, while social media networks can act as an 'echo chamber', allowing users to surround themselves with material which reinforces their views and pushes them towards more extreme positions.[38] Furthermore, social media also presents a more user-friendly way of distributing information and resources amongst extremists, from advice on how best to reach a foreign conflict zone to operational security information or training manuals and videos.[39]

One key area of concern is the extent to which the increasing online reach of extremist organisations has reinforced the 'lone wolf' terrorist phenomenon. A large portion of terrorists who have launched attacks in the West over the last two years have undergone some degree of online radicalisation or used social media to gather information.[40] However, the precise influence of online content relative to other factors is not clear, and very difficult to ascertain.[41]

The advantages offered have resulted in a seismic shift towards the medium, breaking the monopoly of discussion forums and establishing a new battlefield for counter-terrorism efforts.[42] According to cyber-terrorism expert Evan Kohlmann, 'ninety per cent of terrorist activity on the internet takes place using social networking tools'.[43]

Evolving platforms and uses

In the face of the increased use of social media by extremist and terrorist groups, social media companies themselves– sometimes under pressure from the governments – have made more proactive efforts to remove or reduce the impact of hate-speech on their platforms, police content more actively, and remove offending accounts or material more effectively.[44] Increased vigilance in both policing and more active social media platform administration has led to higher rates of page, profile and account deletion, stimulating significant changes in the online habits of extremist and terrorist groups. [45] For example, IS accounts have been observed collectively operating as a 'swarmcast' – disseminating content across multiple platforms and dispersed networks that actively react to account deletions by creating new accounts and reconnecting rapidly with other members of the community.[46] We discuss this further below.

As well as changing their behaviours, extremists on social media are increasingly using new technologies to adapt to more difficult online environments, though as their activities become further removed from more public social media platforms, so the scale of this activity becomes more difficult to estimate. Extremist groups increasingly use more anonymous or privacy-enhanced social media platforms (including those based in regions considered to be less cooperative with Western law enforcement agencies) to communicate, like Vkontakte, Kik and Snapchat.[47] IS sympathisers, for example, are known to have used Frendica, VK, and Quitter, though again, the scale of this use is unclear. [48] According to Shiraz Maher from the International Centre for the Study of Radicalisation, it is typical for IS operatives to find possible recruits on public social media platforms, before then inviting them to join more closed and private forums for further instructions.[49]

This will probably continue to become more widespread. The development of increasingly sophisticated and user-friendly decentralised or privacy-enhanced social media platforms like Twister, BitChirp and Bitmessage has also opened up a new communication avenue for potential exploitation by extremists. For instance, Diaspora, a decentralised social media network running off private servers, is known to have been exploited by jihadists.[50]

Islamist extremist groups have continued the trend of producing their own encryption tools, for example through the September 2013 release by GIMF of Tashfeer al-Jawwal, a mobile SMS encryption tool for Android and Symbian. Tor, anonymous browsers and anonymity best practice are frequently discussed by terrorists on forums and websites, and not only by Islamic extremists.[51] Anders Breivik, the Norwegian terrorist who murdered 77 people in 2011, wrote a manual for others to follow his action: it contained best practice recommendations regarding the use of Tor and the Virtual Private Network service IPredator.[52] A Tor Hidden Service is also believed to have been used by Al-Qaeda leaders Ayman al-Zawahiri and Nasir al-Wuhayshi to discuss high-level strategy, a communication that, when detected, led the US to temporarily close down 21 embassies.[53] There is some evidence that IS sympathisers are familiar with the various opportunities of encryption software.[54] However, in practice, these deceptively complex tools might not have an entirely positive effect on the security of extremists' communications.[55]

Islamic State and social media

The group IS probably makes the most active use of social media, and offers a good example of how a modern terrorist organisation views the opportunities presented by social media. Research conducted by the authors – and a review of literature on the subject – suggest there are several reasons the group uses social media.

First, jihadist groups including IS spread propaganda to recruit newcomers and funders to the cause. They do this through the production and distribution of high-quality content. The physical 'frontier' of holy war is shifting to encompass the 'armchair jihadists' on the virtual front – with professional media teams embedded with fighting units as well as the global network of media supporters.[56] Prior to 2011, Al-Qaeda (AQ) had established a 'jihadist cloud' which, Nico Prucha argued, allowed AQ to remain resilient within "its virtual spaces and niches on the Internet", despite setbacks on physical fronts.[57] Since 2011, members of jihadist forums have issued media strategies that encourage the development of this 'media mujahideen'.

Over the last 18 months, the most frequently used public-facing social media platform for IS has been Twitter. A large quantity of propaganda is posted daily on Twitter each day from the US, the UK, Saudi Arabia, India and Russia.[58] Individuals sympathetic to IS – whether based in Syria/Iraq or not – have organised hashtag campaigns on Twitter to generate internet traffic, and have been able to get those hashtags 'trending', which increases their possible reach. (This is a marketing ploy. On Twitter, people often include a relevant hashtag – simply a word such as #worldcup – which allows others to find their content more easily). IS sympathisers also released, on the Google Store, a now banned Android app called 'The Dawn of Glad Tidings'. Once registered, users automatically posted a stream of tweets carefully selected by social media operatives, released at irregular intervals to outwit the Twitter anti-spam filter.

According to JM Berger and Jonathan Morgan, between September and December 2014, at least 46,000 Twitter accounts were used by IS supporters (although not all of them were active at the same time). The researchers also note that sympathisers are active in creating accounts around the same time account suspensions take place. The report is too lengthy to cover in detail, and includes very useful data on tweeting patterns, bot and app use, and other insights. We consider this to be a very useful study, although there are some remaining difficulties and questions over elements of the analysis. For example, the selection method appears to have limited the potential of collecting emergent and new accounts around hashtags and video releases. The authors selected accounts to be in the sample manually, and then added accounts followed by at least one account already identified as sympathetic to IS (this approach has both strengths and weaknesses).[59] The report is especially useful in its identification of bots and apps designed to extend or inflate the group's perceived size and reach – at which they are sometimes effective. Nevertheless, it is worth bearing in mind that two key purposes of jihadist groups on Twitter are firstly to cultivate and strengthen group cohesion within the mujahid vanguard, and secondly to propagate awareness among the general public in the hope of mobilising elements among it. Bots would only be useful for the latter of these aims – otherwise they are spamming themselves – and even then using bots is about awareness, not encouraging sympathisers to engage further.[60]

Video content has been a staple of Islamic terrorism over the past decade, but IS have moved away from lengthy theologically output towards live-action media content produced to high standards. In September 2014, IS's al-Hayat media wing released *Flames of War*, an hour-long film that spread like wildfire, amassing thousands of views on YouTube before being taken down. (It was reposted across the internet: the copy on LiveLeak has over 100,000 views alone). In it, historical explanations of the group's origins and opposition to Western states are cut with footage captured in raids on Syrian and Iraqi positions and their bloody outcomes. IS sympathisers use blogs to write narratives on the caliphate; the process of its establishments, its theological and political underpinnings, its enemies and so on. Based on an analysis of 1700 pieces of propaganda produced by the group, researchers at the Quilliam Foundation have distilled the messaging into six narratives: brutality, mercy, victimhood, war, belonging and utopianism.[61] One important distinction from Al-Qaeda propaganda is the shift from theological debates to live action, and the existence of a 'functioning' state.

Links to hundreds of different blogs scattered across the web are shared on Twitter and other social media platforms. Common blog hosts like Blogspot (http://gareeb-alikhwan.blogspot.co.uk/) and Wordpress (https://akhbardawlatalislam.wordpress.com/) are both frequently seen in lists of IS content-sharing. Others are nicher still: http://www.shabakataljahad.com/, for example, sits behind a password, pointing towards the smaller and more secretive communities of jihadis online.

Perhaps the most valuable insight into IS's online behaviour is how they have been able to maintain a presence online. Internet censorship has always presented practical difficulties, but IS's use of social media has compounded such challenges. Typically, when a major platform – such as YouTube – deletes their content, IS sympathisers immediately post it on text-based sharing boards like justpaste.it, dump.to or elsewhere, and alert everyone to its new location – from where it is very quickly downloaded and re-posted across multiple sites. (In a study conducted by the authors, two per cent of links on IS sympathetic Twitter accounts linked to justpaste.it, which is generally viewed to have a lax policy in relation to extremist content).

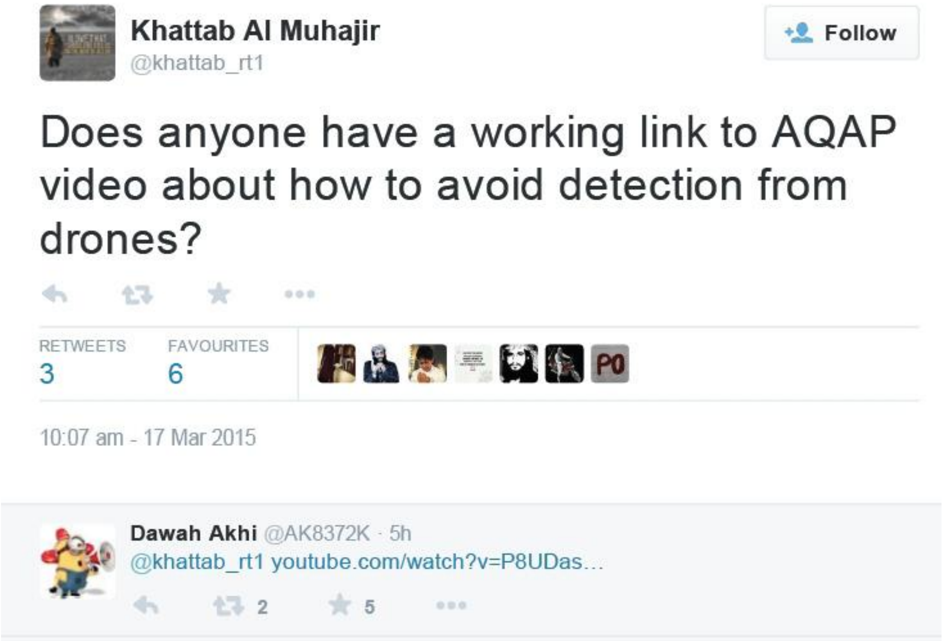
A similar pattern emerges when user accounts are shut down. IS sympathisers immediately start another one – or, more often, have multiple accounts ready to hand. The report authors have examined how IS are using Twitter. We found one user name, @nhnhna, who had 21 versions of his name, all lined up and ready to use (@nhnha7, @nhnha8, @nhnha9, and so on). Another posted tweets under the handle @Abu_Umar8246 for eight days before the account was shut down, at which point a new account, @Abu_Umar_8246, started posting. When that was closed, @AbuUmar__8246 began. Most impressively, as soon as the new AbuUmar account was set up, he or she picked up all his or her followers almost immediately. These users disseminate content through a network that is constantly reconfiguring, which thrives in the chaos of account suspensions and page deletions. We believe they have established complex networks of influential accounts across multiple platforms, which creates an inherent resilience, and renders the effect of account suspension little more than a temporary inconvenience – followers are quickly able to locate the account's designated replacement.[62,63] This ensures a rapid re-establishment of followers and creates resilience in the network, limiting the impact of individual account suspensions.[64]

According to JM Berger and Jonathan Morgan's report on IS's Twitter activity, 'account suspensions do have concrete effects in limiting the reach and scope of ISIS activities on social media'. Their analysis suggests that, following a wave of suspensions on the platform, the size of the ISIS supporting network was reduced significantly, and perhaps more internally focused. This suggests that account suspension may both tie the group up with trying to re-assemble the network; and that, although the group remains online and active, it is less able to reach as wide an audience. This suggests that, although account suspensions are not entirely successful, it does not mean they are not worth the effort. They can limit the ease with which content can be accessed. As noted by the authors, however, tampering with networks 'is a form of social engineering' and can create several unintended consequences which are not well understood.[65]

Of course, this production of content can also offer valuable insights into the operational tactics of the group, in particular where that content includes footage of previous successful attacks.[66] Some recent videos from IS in Syria and Iraq have also contained extensive footage of the group's newly acquired arsenal. While enabling IS powerfully to demonstrate their growing strength, such videos also give counter-terrorism operatives clear information about the weapons they might face.

Research has also found that IS propaganda online varies messaging campaigns according to the intended audience (international or domestic, supporters or adversaries).[67] The #twitterkurds hashtag offers an excellent example. Originally used three years ago by the account user @hevallo in discussing the question of Kurdish rights in Turkey, it has been used during the conflict to publicise the struggle of the Kurdistan Workers' Party (PKK) against IS. In response, IS sympathisers – including some believed to be in the territory itself – also started to post using the hashtag, to insult and mock the Kurds they were fighting. Much of the content is very graphic; hangings, executions and kidnappings are common themes. The intention, we believe, is to demoralise further an opposition already low on morale. This form of personalised propaganda has also been targeted at the West, playing on the dangers of 'boots on the ground' and tweeting graphic imagery of killed or injured personnel. Some tweets are instead aimed at the vanguard of IS itself; these take the form of messages reinforcing IS ideologies and celebrating those who are fighting for the organisation. The most common example is the celebration of those who have been killed or 'martyred', either as fighters or as civilians. News of recruits is often circulated, particularly in relation to foreign fighters whose decision to join IS is perceived as a victory over the rejected West. The strict application of Shariah law is also a common topic; the administration of harsh punishments – amputations, lashes and executions by hanging and stoning – is recorded and shared.[68]

A recent article on the nature of the Swarmcast, has demonstrated that in addition to the use of social media platforms, the media mujahedeen also utilise their own devices as part of a distributed storage system. This emergent behaviour allows specific content to be requested on Twitter. For example, one tweet asks: "does anyone have a link to AQAP video about how to avoid detection from drones?" The video was posted to YouTube within five hours.



Khattab Al Muhajir
@khattab_rt1

Follow

Does anyone have a working link to AQAP video about how to avoid detection from drones?

RETWEETS 3 FAVOURITES 6

10:07 am - 17 Mar 2015

Dawah Akhi @AK8372K · 5h
@khattab_rt1 youtube.com/watch?v=P8UDas...

2 5

This means that even if it were theoretically possible to wipe all jihadist content from social media and the internet more broadly, the media mujahedeen would be able to repopulate those platforms with the content stored locally. The ability to repopulate content provides an additional level of resilience to the Swarmcast. This capability will become increasingly potent if greater use is made of encrypted social media.

As noted above, we anticipate a growth in the availability of default encrypted social media services, anonymous social networks, and decentralised distributed social networks, which are run without centralised servers or administrators (sometimes called a 'distributed trust' network). The implication is that when extremist material is posted there is no authority who can be contacted by the police to request its removal. IS's propaganda wing has used Diaspora in much the same way as Twitter, posting images, text and video, including footage of the recent beheading of American journalist James Foley.[69] Jihadists are also known to have used Frendica, VK and Quitter, although it is not clear how many accounts have been since taken down.[70] However, rather than a complete migration to distributed social media, we should perhaps anticipate the increasing integration of such services within the multi-platform networks of jihadist groups and their supporters. This would represent an extension of the 'swarmcast' tactic identified and would further increase network resilience. The inclusion of social network sites which are inherently difficult to censor – either on 'dark nets' or distributed open platforms – could further ensure that certain core content remains widely available even where mainstream sites have taken steps to remove it.

Social media and law enforcement

More generally, social media use is affecting other types of law enforcement activity: criminal organisations and gangs exploit the internet and social media. While this is not directly related to the terrorists' use of social media, it provides some useful background. Well-organised and longstanding groups have stable social media presences, usually used for advertising their organisation, or in some cases 'cyber banging' – levying threats against rival groups, or individuals. Indeed, in November 2012, the British Justice Secretary announced a crackdown on the use of social media by criminals to intimidate witnesses. Additionally, the amount of personal information posted on social media has been shown to influence the risks of individuals to burglary.

Social media is also of growing relevance to public disorder policing. A common tendency was identified in the August 2011 riots in the UK, and in Vancouver following the Stanley Cup the same year. During the early stages of disorder, participants and uninvolved observers recorded and shared information about the event. As the disorder increased, information describing the apparent impunity of the rioters, visibly shared on social media, may have escalated the disorder further. In the aftermath, similar themes of a united community coming together to condemn the riots and organise a clear-up were seen both in London and Vancouver. A 2014 analysis of the use of Twitter and Facebook during political protests in Ukraine in 2013 and 2014 showed that during dramatic or violent incidents, the use of Facebook and Twitter increased significantly. It also found that a significant number of people tweeting about the protests had joined Twitter during the course of them, and suggested that protestors were specifically joining social media platforms in order to share and access information about the protests.[71] Moreover, the confusion of modified digital content, rumour and hearsay were noted as having slowed down the policing procedures following both riots. An analysis of the 2014 Ferguson Riots showed that a high proportion of popular tweets were 'rumourous'. Researchers identified 42 different rumours circulating with regard to Ferguson on social media between 9 August and 25 August, including that Israel had been involved in the training of two of the four police departments in Ferguson, and that the Pentagon had sent St. Louis County police military-grade weapons. Around 25 per cent of all tweets about Ferguson with over 100 retweets and around 27 per cent of all tweets with over 250 retweets were rumours. They found that in the case of Ferguson, rumours tended to attract slightly more replies than non-rumours.[72]

More generally, there is a growing evidence base to suggest that an increasing proportion of crime is taking place online. In 2015, 38 out of 45 UK police forces saw a rise in the number of crime reports that involved Facebook.[73] There are various initiatives in place to help respond. In April 2015, chief constables from the UK agreed a new approach run by a 'Capabilities Management Group' (CMG). The CMG is a multi-agency effort to bring digital intelligence and investigation into the mainstream of police activity.[74]

A new response: 'counter-speech'

In recent months there has been increased concern about various extremist groups using social media. As a result, there have been some calls for greater monitoring and censorship of content, as well as calls to move beyond the manual, flag-based system currently used by most social networks. For example, the Online Hate Prevention Institute has urged Facebook to review its reporting system, suggesting solutions aimed at lightening the load on administrators such as automation and IP address-based bans.

However, there is little evidence that censoring or removing content has an effect (or indeed, on what that effect might be). Indeed, based on our analysis above, a mobile-enabled, reactive 'swarm' will rapidly reconfigure itself to counter censorship and account suspension. (It is the authors' view that forcing extremists to keep re-posting content and creating fresh accounts adds energy to the network, provoking new ways of avoiding moderation and spreading their message. A preferable response is a small number of strategic mass take-down efforts, which would make the network harder to reconstruct and allow analysts to study the effect it has on the network.)

There has been a slowly emerging consensus that confronting hate speech with 'counter-speech' is a potentially more fruitful approach. A recent White Paper by the Quilliam Foundation describes "censorship and filtering initiatives" as "ineffective", and emphasises the critical role of counter-speech in "challenging the sources of extremism and terrorist-material online".[75]

Practical, state-funded social media counter-speech efforts have been undertaken in both the US and Canada. The US Department of State's high profile Think Again Turn Away campaign was launched in December 2013 in order to contest jihadi theology and highlight the realities of jihadi terrorism and extremism on social media. The Department of State manages this counter-speech presence across multiple platforms, including Twitter, Facebook and Tumblr, and publishes in Urdu, Arabic and English.[76] As well as publishing counter-extremist material, the account also engages in arguments with high profile jihadist accounts. Some have disputed the effectiveness of this technique, arguing that an antagonistic engagement presents jihadists with a platform to air their views.[77]

In Canada, Extreme Dialogue, a project funded by the Kanishka Project and created by the Institute for Strategic Dialogue (ISD), Duckrabbit and the Tim Parry Jonathan Ball Foundation for Peace, aims to present counter-speech through targeted social media adverts. Launched in February 2015, the project centres on a series of short documentary films concerning Canadians who have been effected by extremism. These include, for instance, the mother of an IS fighter who was killed in Syria and a far-right extremist turned anti-far right activist. The project also promotes educational resources intended to build resilience to extremism through active discussion and critical thinking.

The ISD approach (which is also taking place outside Canada) distinguishes between 'upstream' and 'downstream' content, producing different interventions for each. 'Upstream' content – for a general cohort – is educational content, disseminated both online and offline, that tries to create critical consumption skills and question propaganda. With more downstream individuals – for example radical forums – the ISD uses profile data, based on browser cookie data, to create more targeted adverts at individuals (and measures the amount of engagement with the material). Further downstream still, ISD works with former extremists and victims of extremism who contact high-risk individuals directly through (for example) Facebook messenger, and offer the opportunity to talk. These projects are ongoing, and results for how effective they are at reaching and affecting certain individuals are expected in 2015/2016. Early results are promising.[78] However, a recent internal White House memo suggested that, despite 'counter-messaging' the Islamic State being a vital part of the US government's strategy, current efforts are not having the desired effect.[79] According to a briefing produced by the Danish Institute for International Studies, there is little evidence that counter-narratives have much impact (and may, suggests the author, even be counter-productive). However, this report does not appear to contain any empirical research.[80]

Counter-speech need not be organised; often, it is an organic reaction to exposure to extremist material. Indeed, according to a recent statement from Facebook, 'we've also found that posting insensitive or cruel content often results in many more people denouncing it than supporting it on Facebook.'[81]

Combating extremism in this way has some advantages: it is faster, more flexible and responsive, capable of dealing with extremism from anywhere and in any language, and allows government to retain a commitment to free and open public spaces for debate. However, the forms taken by 'counter-speech' are as varied as the extremism it confronts. It is also likely that counter-speech is not always as effective as it could be –some types of counter-speech could even be counter-productive. There is growing interest in a more rigorous and evidence-led approach to counter-speech, particularly since the February 2015 White House Summit on countering violent extremism. There is an emerging body of evidence regarding what works in hate-speech policy and what does not; for example, some academics argue that changes to site architecture can have a profound effect on the volume of hate speech and the success of counter-speech.[82]

PART 2: AN INTRODUCTION TO SOCIAL MEDIA INTELLIGENCE

SOCMINT covers a wide range of applications, techniques and capabilities available through the collection and use of social media data. The term was first coined by the authors in a 2012 report, *#Intelligence*. Some analysts have suggested SOCMINT to be a branch of open source intelligence (OSINT), which has been defined as 'information that is publically available and can be lawfully obtained by request, purchase or observation'.

SOCMINT does not easily fit into the category of open or secret intelligence. SOCMINT is defined not by the openness of the information on which it is based but by its existence on a social media platform. As either open or closed intelligence, SOCMINT requires very specific considerations of validity and interpretation.

This paper does not discuss closed or secret SOCMINT, which by definition would require access to communications which are not publicly available. Instead, this paper focuses only on open SOCMINT as define above. We believe this type of SOCMINT is potentially a useful and important part of counter-terrorism and public safety efforts, a view that is echoed in a 2013 white paper report looking at social media as a security monitoring tool in the UK and beyond.[83]

In the United States, OSINT is considered to be of considerable and increasing value, covering commercial, procurement and trade data, expert opinion data and a variety of types of 'gray' literature produced by the private sector, government agencies and academics. The US Committee on Homeland Security considers OSINT to be a tool that federal state and local law enforcement agencies should use to develop timely, relevant and actionable intelligence, especially as a supplement to classified data.[84] More recently there have been calls for the US government and military to expand and deepen its use of OSINT as part of wider reforms to intelligence, national security and governance.[85]

There are many different types of open SOCMINT, ranging from very technical methods to quite general approaches. We believe the most significant – capable of reducing ignorance and improving decision making for the purposes of preventing, pursuing, protecting and preparing against terrorism – are the following:

- Natural language processing – a branch of artificial intelligence involving the computational analysis (often using machine learning methods) of 'natural' language as it is found on social media.

- Event detection – the statistical detection analysis of social media streams to identify offline 'events', whether natural, political, cultural, commercial or emergency to provide situational awareness, especially in dynamic and rapidly developing contexts. In counter-terrorism work, this is likely to be particularly valuable in the aftermath of a major terrorist incident (see case study, below).

- Data mining and predictive analytics – the statistical analysis or 'mining' of unprecedentedly large ('big data') data sets, including social media and other 'big' or open data sets (such as Census data, crime, health, environmental and transport data), to find the dynamics, interactions, feedback loops and causal connections between them.

- Social network analysis– the application of a suite of mathematical techniques to find the structure and topography of the social networks found on social media. These networks are then subjected to analysis, which can identify a range of implications and conclusions (including predictive ones) on the basis of the characteristics of the network structure and type.

- Manual analysis/'netnography' – drawn from qualitative sociology and ethnography, this is a broad collection of manual approaches to collecting and analysing data concerning social media data. It often aims for depth over breadth in order to reveal and untangle the hidden, obscured, overlooked or contingent social significances, meanings and subjectivities experienced by individuals on social media.

- Solicited/'crowd sourced' insight – insight garnered from the emerging technique, practised by a number of public and private agencies, to use social media to ask citizens or social media users for information directly.

Each of the methods can then be applied in different ways and contexts. In part 3, we set out the technical description of the methods themselves, and then where and in what way they might be employed.

PART 3: METHODS AND APPLICATIONS

Summary

We critically discuss the state of the art in each category of open SOCMINT. Each section considers capabilities generally and, where possible, specific applications for the purpose of countering terrorism.

- There are several useful ways to access and process very large data sets directly from social media platforms via Application Programming Interfaces (APIs). However, the terms of use, data format, and usefulness of data vary greatly from platform to platform. Nonetheless, given the increasing value and importance of big data analytics, this type of data collection is becoming more available across more platforms.

- Natural language processing is the discipline of collecting and processing large 'natural language' data sets, such as from social media. This is becoming increasingly accurate and sophisticated (no longer just based on simple distinctions of positive or negative sentiment) although significant methodological difficulties remain.

- Similarly, the field of network analysis and meta-data analysis are improving quickly, although, just with natural language processing, there remain significant challenges.

We consider the most promising uses of SOCMINT for the purposes of counter-terrorism are:

- SOCMINT in general – the collection, analysis and use of social media – is an increasingly important aspect of counter-terrorism work, in both intelligence collection and communications.

- In our view, in a counter-terrorism content, natural language processing is likely to be most useful at present in the immediate aftermath (or during) a terrorist incident. During major events there is typically now a vast amount of accompanying social media activity. Natural language processing can allow analysts quickly and quite accurately to identify and process important information as it is posted – which is useful for immediate response and intelligence collection purposes.

- Various forms of network analysis – such as constructing a network of sympathisers – are useful for understanding the broad structure and behaviour of an online network, although there is a lot of variety in how a relationship between

users is measured and understood. Network analysis is extremely valuable in gaining a better understanding of the sorts of information and ideas being shared within a group or movement.

- Alongside these 'big data' methods, more detailed 'netnography' – ethnography on the internet – is an emerging discipline which allows for rigorous study of smaller online communities. This has proven to be of value in finding and collecting valuable information about criminal and terrorist activity.

- More and more studies are using these methods in order to 'predict' offline activity on the basis of online data, although these remain of variable quality in practice.

- 'Citizen journalists' – those at the scene of an incident or collecting and collating other sources of social media information – are likely to be an increasingly important source of valuable insight. Methods to curate and use this data are likely to form a very important part of SOCMINT capabilities.

- Social media is also increasingly important as a means to communicate with large numbers of people, which is likely to be extremely valuable during and after any major terrorist incident.

- All these SOCMINT approaches should be treated as part of a significant new discipline of intelligence, with its own standards of evidence, technical methods, and dedicated expertise. The field of SOCMINT is most likely to evolve and improve with greater multi-disciplinary work between the computer and social sciences.

- However, all of these methods and approaches need to be undertaken with due consideration for privacy and public expectation, which are discussed further in the next section.

Social media data collection and retrieval

It is possible to collect social media data manually in a number of ways – copying, screen-grabbing, note-taking, and saving web-pages. However, where large volumes of data are involved, the most appropriate method is to collect the data automatically. This is done through connection to a platform's 'Application Programming Interface' (API).

The API is a portal that acts as a technical gatekeeper of the data held by the social media platform. They allow an external computer system to communicate with and acquire information from the social media platform. Each API differs in the rules it sets for this access: the type of data it allows researchers to access, the format in which it produces these data in, and the quantities in which it produces them.

Some APIs can deliver historical data stretching back months or years, while others only deliver very recent content. Some deliver a random selection of social media data taken from the platform, while others deliver data that match the queries stipulated by the researcher – usually keywords selected by the analyst. In general, all APIs produce data in a consistent, 'structured' format, and in large quantities. Facebook's and Twitter's APIs also produce 'meta-data' – information about the data itself, including information about the user, their followers, and profile. This meta-data can be a rich and valuable source of information for social media researchers, often containing information on everything from the sender's device type, to their account creation date, location and social media following.[86] Along with Facebook and Twitter, most major social media platforms allow API access for researchers in some form.

There are several types of API access to Facebook data, most of which have been designed for app makers, such as Public Feed API, a Keyword Insights API, a Marketing API and Atlas API.[87] The Facebook API relevant to social media research is the 'Graph API', which can be directly accessed online with Facebook's Graph API Explorer, or via Facebook-approved third party commercial re-sellers of data, like DiscoverText or DataSift. The difference between Graph API Explorer and a third party front end is that the third party software is designed to gather large amounts of data via the Explorer and present them in a way that is conducive to detailed analysis. There is no additional functionality, and Facebook retains all control over the kind and quantity of data that can be collected.

Graph API allows posted text, events, or URLs, plus any comments on posts to be accessed, along with metadata on user information, including gender and location. It operates like database interrogation software: a user asks it for information using the relevant coding language; Explorer finds where on Facebook that information is stored (ie the web address) and returns the information. Facebook API is sometimes considered opaque by researchers that use it. There is no detailed written record of how it works, which potentially introduces bias to any data gathered through the API. Access to all Facebook data is predicated on the user's settings and who has agreed to share information with them. Facebook's privacy structures are complex – potentially, any single user can have a distinct privacy setting for every piece of data they share. They can, for example, decide that only their 'close' friends (a user-defined group of 20 people) can see a single post, all posts, or posts on a particular page. API searches only return data that is public, and fail to quantify the information that has remained uncollected due to privacy restrictions. This is a significant methodological weakness.

A critical awareness of what data is produced by an API, and how this data is gathered, has been increasingly recognised as a requirement of professional social media research; in this sense, a lack of openness in social media platform APIs can be a barrier to more methodologically sound research.[88]

The most prolific and heavily researched provider of social media data for research is Twitter. Twitter has been operating since 2006 and has over 300 million active users; in late 2014, Twitter announced that it had indexed roughly half a trillion publically sent tweets.[89] As a platform experiencing extremely rapid growth, the demography – geography, language, age and wealth – of its users is constantly changing. Major studies, whilst struggling to keep pace with this rapid change, have found that over 100 languages are regularly used on Twitter. English accounts for around half of all tweets, with other popular languages being Mandarin Chinese, Japanese, Portuguese, Indonesian, and Spanish (accounting together for around 40 per cent of tweets). The number of languages used on Twitter has continued to grow. One recent area of NLP research has been the development of classifiers capable of accurately distinguishing between closely related, smaller language groups, such as the South-Slavic languages Bosnian, Croatian, Montenegrin and Serbian.[90]

These languages are geographically spread, with concentrations in Europe, the United States, Latin America and South East Asia. As of October 2013, 24 per cent of monthly active Twitter users were from the United States, while 9 per cent were from Japan, 7 per cent from Indonesia, 6 per cent were from the UK and 4 per cent were from Brazil.[91] While the US is set to continue to have the largest number of individuals with Twitter accounts over the next few years, the proportion of Twitter users from the Asia-Pacific region is increasing.[92]

To set up a stream or search to collect the data, it is typical to create a user interface which is built around the underlying API provided by Twitter. The API is a series of http 'end points' that return data according to the parameters that are provided with the request. Twitter has three different APIs that are available to researchers based on tweet content (although there are some other API access points based on searching for users). Twitter's 'search' API returns a collection of relevant tweets matching a specified query (word match) from an index that extends up to roughly a week in the past. Its 'filter' API streams tweets that contain one of a number of keywords in real time. Its 'sample' API returns a small number (approximately 1 per cent) of all public tweets in real time.

Each of these APIs (consistent with the vast majority of all social media platform APIs) is constrained by the quantity of data they will return. Twitter provides three volume limits. A public, free 'spritzer' account is able to collect one per cent of the total daily number of tweets. White-listed research accounts may collect 10 per cent of the total daily number of tweets (known informally as 'the garden hose') while the commercially available 'firehose' collects 100 per cent of daily tweets.

With daily tweet volumes averaging roughly 500 million, many papers do not find any of these restrictions to be limiting to the number of tweets they collect (or need) on any particular topic.

Each of Twitter's APIs produces varying amounts of meta-data with each tweet (far exceeding in length the content of the tweet), including (if it exists) the geo-location of the author (expressed as longitude-latitude coordinates), their profile's free-form text location, their time-zone, the number of followers they have, the number of tweets they've sent, the tweet's creation date, the author's URL, the creation date of the account, even the author's wallpaper on their Twitter homepage. There are typically around 30 pieces of meta-data, depending on the account.

One of the key advantages of acquiring data via a social media platform's API is the consistent, 'structured' nature of the data that is provided. This advantage becomes important when gathering high volumes of data from dynamic platforms such as Twitter. Alongside direct API access, a number of licensed providers make available raw data to multiple APIs, Including DataSift, Gnip and DiscoverText. However, over the last few years, third party provider access to Twitter data has been reduced as Twitter has consolidated access to its data. For example, Twitter has now bought Gnip, terminated DataSift's access agreement, and ended NTT Data's access to the Firehose. [93] API access changes are announced with appropriate warning through the Twitter developer blog. There is a schedule for the phasing out of features so companies have some time to adapt.

Youtube also uses multiple APIs. The 'Data API' allows you to gather publicly available data on a video or channel, displaying detailed information on, for example, views, descriptions, comments and ratings – it will also enable you, when authenticated, to post videos and moderate comments on videos which you own.

This API offers a valuable chance to quickly uncover viewing trends, assess popularity and examine discourse on given topics. YouTube also offers an 'Analytics API,' again requiring authentication as a user, which is designed to analyse the impact of videos for their owners – especially those who wish to better market themselves or their videos on the site. It will enable you to track geographical data and obtain information on the demographic of users watching a video, but only if you can get the permission of the video's owner.

Although less popular at the moment in terms of user volumes, many other social media platforms include API access, and the sort of data available will depend on the sort of data that is created by users of the site. Examples include Foursquare, which includes location-based data, and Pinterest which includes image-based data. Helpfully, licensed third party data providers such as DataSift provide single API access to at least 24 different social network sites.[94] Although it is beyond the scope of this paper, we suggest a review of all publicly available API-based data as a useful starting point for research into this area.

In addition to these options, companies known as 'data brokers' collect data from multiple online and offline sources. One of the largest, Axciom Corporation, is believed to hold information about 500 million consumers around the world, with annual sales worth over $1 billion.

Web scrapers and crawlers

For the purpose of this overview, 'scrapers', 'crawlers', 'spiders' and 'bots' are all automated programs which are used to find and catalogue information stored on websites. This is typically achieved through transforming website data (usually expressed in a language called 'HyperText Markup Language', or html) into structured data sets.

A basic crawler of this type is usually a relatively simple piece of code that employs a technique to collect and process data on different websites. Programmers can use regular expressions (or 'regex') to define a pattern (for example, any acronym involving at least 3 letters) to enable the crawler to execute a particular predefined action when a match is identified on a webpage. This allows the crawler to copy or index specific information from any page on the World Wide Web to which it is directed. These and many other associated techniques are subject to constant development.

Someone with little experience can, in a short space of time, build their own bespoke crawler using only freely available programs and tutorials. A very basic crawler can be created with only a few lines of code on platforms such as Scraperwiki or using programming languages such as Python, Java, or PHP. These crawlers can be built very quickly and cheaply, and increasingly the code is open source.

Despite their relative simplicity, basic crawlers and the vastly more complex crawlers employed by commercial and public organisations have the potential to unlock data on the way communities interact, the information they seek, and the sources of information they use.

Information retrieval

In general, information retrieval refers to a body of techniques employed to identify those documents in a large collection that are relevant to a specific information need. Patterns and 'objects' within documents are often found by rules-based algorithms, which allow documents to be ranked according to their relevance.[95]

This is a rapidly developing area of work.[96] Retrieval techniques are designed to allow for more powerful meaning based searches. For example, running a search for conversations related to jihad and filtering the subsequent results based on clustered groups of identical or near-identical material highlights those retrieved items that include new material.[97]

Search engines still do not always effectively search social media content, even though it might be highly relevant. For example, photos with a relevant title or geo-location often contain little textual narrative making them difficult to search for. Improving the accuracy of social media searching is also an emergent field of considerable interest. Current developments focus on 'similarity clustering', which facilitates the identification of relevant clusters of social media data considered to be importantly similar, either in their content, or when or where they the content was posted. Another important area of research seeks to improve the retrieval of images on social media. Images are often retrieved based on the geo-location, text or tags accompanying the image, which can often be sparse or misleading, reducing the accuracy of image retrieval. The construction of innovative solutions to this problem – for example using the personal data stored in search logs and social media profiles to improve the relevance of returned images – is the focus of a significant body of research.[98] Other efforts focus on filtering out irrelevant results in information retrieval, for example by developing automated ways of detecting spam URLs in social media, [99] or automatically discriminating between personal and organisational accounts on Twitter.[100]

According to Tim Berners-Lee, automated search techniques require further development. Information embedded in a document is still not easy to find. Berners-Lee believes the Web will evolve from a 'web of documents' to a 'web of data' – underpinned by Universal Resource Identifiers (URIs) to allow for a consistent reference. Simple Protocol and Resource Description Framework Query Language will allow this semantic web to be searched.[101]

Technique: Natural Language Processing

The growth of NLP

Natural Language Processing (henceforth, NLP) is a long-established sub-field of artificial intelligence research. It combines approaches developed in the fields of computer science, applied mathematics and linguistics.

NLP is based upon machine learning. A number of different machine learning methods – most widely recognised within the field – are increasingly being applied to real-world problems. These include classifiers, such as Naïve Bayes, Logistic Regression, and Support Vector Machines. They also include sequence labellers, such as Hidden Markov Models and Conditional Random Fields, and also topic modellers such as Latent Dirichlet Allocation.

Most of these approaches seek to teach computers to 'classify' documents: to place a document in one of a fixed set of different options (or classes), based on the features of that document. Of particular relevance to SOCMINT is text classification, wherein the machine seeks to place a document (such as a Tweet, or Facebook post) within a particular category on the basis of the text that the document contains.

NLP is increasingly being applied to many areas beyond its academic origins. NLP algorithmic models look for statistical correlations between the language used and the meaning expressed. The technique therefore offers the practical capacity to teach algorithms which automatically detect the meaning of the 'natural' language used when people speak to each other. From online feedback forms to Tweets, NLP is being used as the technological response to a problem that is growingly faced across a number of different domains: too much language is produced for analysts to read themselves. Computational approaches are therefore sought in order to allow automated analysis capable of dealing with very large bodies of data. Social media research is one such domain, and NLP is increasingly applied to the natural language contained within social media data sets to analyze them.

Machine learning techniques are often grouped according to the nature of the interaction between a human being and the computer. Sometimes the items in the data set are 'labelled', where their features have already determined by a human. This is supervised machine learning. Sometimes none of the data is labelled – this is unsupervised machine learning. A popular approach to understanding social media for SOCMINT is semi-supervised machine learning: when a small number of documents (such as Tweets) are labelled into one of a number of different categories in order to teach the algorithm how to make the same kind of distinction on a much larger body of Tweets.

The semi-supervised training of NLP algorithms happens through a process called 'mark up'. Typically messages are presented to an analyst via an interface. The analyst reads each message, and decides which of a number of pre-assigned categories of meaning it best fits. After the analyst has made a decision, they 'annotate', the message, associating it with the category into which it best fits. This produces labelled data. The NLP algorithm then appraises the linguistic attributes of the corpus of labelled data, to determine significant correlations between each category, and the language contained within the examples in each category.

Numerous different NLP algorithms can be trained. NLP programmes vary in the way they make their decisions: some place more weight on specific words, others on structural or grammatical features. Each parses language differently, whether by words (or unigrams), collection of words (such as bigrams and trigrams), grammar, word order or emoticons.

These measured correlations provide the criteria for which the algorithm then proceeds to make additional automatic judgments about which category additional (and un-annotated) pieces of social media data best fit into. The statistical nature of this approach renders it notionally applicable to any language where there is a statistical correlation between language use and meaning.

The NLP classifier

The operational opportunity of NLP for countering terrorism is to use these algorithmic models as 'classifiers'. Classifiers are applied NLP algorithms that are trained to categorise each piece of social media data – each post or tweet – into one of a small number of pre-defined categories. The earliest and most widely applied example of this technology is 'sentiment analysis', wherein classifiers make decisions on whether a piece of social media data is broadly positive or negative in tone. However, the kinds of distinctions that a classifier can make are arbitrary, and can be determined by the analyst and the context.

The performance of NLP classifiers is often quantified by comparing a body of automatically classified data against a separate set of human classifications drawn from the same data set, but not included in the building of a model. On this measure, their accuracy – the ability of an NLP algorithm to classify any given message the same way a human would – varies considerably. There are many scenarios where 90 per cent accuracy would be expected. However, an accuracy of around 70–80 per cent in a three-way classification task would often be considered excellent.

Based on the experience of the authors, classifiers are sensitive to the specific vocabulary seen in the data used to train them. The best classifiers are therefore also highly bespoke and trained on a specific conversation at a specific time to understand context-specific significance and meaning. As language use and meaning constantly change, the classifier must be re-trained to maintain these levels of accuracy. The more generic and expansive the use of any NLP classifier, the more likely that it will misunderstand language use, misclassify text and return inaccurate results.

In many situations, the performance of these classifiers is sufficient to produce robust aggregate findings, even when the accuracy of any given singular classification is quite low. This arises because the data sets are sufficiently large that even relatively inaccurate individual estimates lead to an accurate assessment of the overall trend (although, subsequently, may be less good at identifying every relevant piece of data). Assessing when this technology tends to work well and when it does not is an area of active research.

A key area of active research is in the reduction of the time, effort and cost required to train and maintain an NLP classifier. It is typically very expensive to produce the labelled training data that these supervised machine learning algorithms require. In complex tasks, it would not be unusual for this to take multiple person-months of effort. The novel introduction of an information-theoretic technique called 'active learning' is, in some cases, beginning to allow classifiers to be built much more rapidly and cheaply – often in a matter of hours, and sufficiently quickly to meet changing operational requirements prompted by rapidly shifting situations and contexts.

One key development is improvements classifying texts into categories other than positive, negative and neutral: such as urgent, calm, violent or pacific. Some of the most significant advances in sentiment analysis have concerned the creation of a sentiment analysis system capable of distinguishing between emotions much more complex than the traditional 'positive', 'negative' and 'neutral' labels, or to make other distinctions much more complex than those regarding simple differences in topic or subject matter. Much research has focused on improving the capability of classifiers to distinguish between more nuanced emotional categories, such as 'shame', 'confusion' or 'sadness', or categorisations based on nuanced differences between the purpose and intent of tweets, for example distinguishing between casual racism and racially-aggravated threats on Twitter.[102] One of the enduring weaknesses of NLP is the difficulty of understanding the context in which a sentiment is expressed through automated means. This is a particularly problematic shortcoming on Twitter, where the text analysed is often brief and the context is often not established. Another focus of recent research in sentiment analysis is the development of systems capable of making more detailed distinctions, by, for example, distinguishing between multiple sentiments regarding different topics in the same piece of text, or making judgements about sentiment based on the association between descriptive words and their contexts.[103]

The attempt to produce classifiers capable of making complex judgements about whether a tweet contains hate speech – and what type of hate speech it is – has been a continuing focus of NLP research. Examples of such research include efforts to classify context-specific anti-Muslim hate speech after the murder of Drummer Lee Rigby,[104] to categorise different types of racist hate-speech on

Twitter,[105] and to create a general, lexicon-based classifier capable of detecting hate speech have all met with success.[106] Beyond hate-speech detection, more complex classifiers capable of undertaking task such as automatically detecting paedophiles grooming victims in chat rooms or extremists recruiting online are an ongoing area of research.[107]

Attitudinal research

Perhaps the largest body of attitudinal research on social media has focused on the use of NLP to understand citizen attitudes on Twitter. This research has been driven by the view – implicit or explicit in most of the research papers – that attitudinal data sets on Twitter are different to those gathered and understood by conventional attitudinal research. This is because the available data sets are huge, naturalistic (meaning that they are not exposed to observation bias) and constantly refreshing in real time. Furthermore, because of the increasing ease of data access and dramatic reductions in computing costs, these data sets are notably more analysable, and also cheaper to collect and analyse.

Harnessing social media data sets of this kind stands to have a transformative impact on our ability to understand sentiments and attitudes. However, no published output has yet been able to understand attitudes on social media using methods that satisfy the conventional methodological standards of attitudinal research in the social sciences, or the evidentiary standards of public policy decision makers. There remain a number of methodological problems and uncertainties associated with social media research, which have been the subject of a number of projects and programmes run by research councils and organisations like the UK's NatCen and the Economic and Social Research Centre.[108] The recent Demos report *Vox Digitas* presented an in-depth examination of the challenges of social media research and its exploitation.[109]

Perhaps the most important methodological challenge is sampling. Twitter's API delivers tweets that match a series of search terms. If searches are subjected to Boolean operators similar to search engines, searching for 'Canada' returns tweets that contain 'Canada' in either the username of the tweeter, or the text of the tweet. A good sample on Twitter must have both high recall and high precision. 'Recall' is the proportion of possibly relevant tweets on the whole of Twitter that any sampling strategy can find and collect. 'Precision' is the proportion of relevant tweets that any sampling strategy selects.

A high recall, high precision sampling strategy is therefore comprehensive, but does not contain many tweets that are irrelevant. Arriving at a series of search terms that return a good sample is far from easy. Language-use on Twitter is constantly changing, and subject to viral, short-term transformations in the way language is mobilised to describe any particular topic. Trending topics, '#' tags and memes change the landscape of language in ways that cannot be anticipated, but can crucially undermine the ability of any body of search terms to return a reasonably comprehensive and precise sample.

Current conventional sampling strategies on Twitter construct 'incidental' samples using search terms that are arbitrarily derived. They do not necessarily return precise, comprehensive samples, and certainly do not do so in a way that is transparent, systematic or reliable. Furthermore, it is becoming clear that the way Twitter is used poses first-order challenges to discerning people's genuine attitudes. This may be one explanatory factor behind Pew Research Centre's findings that reactions to major political events on Twitter measured through sentiment analysis often differ a great deal from public opinion as measured through traditional surveys.[110] Indeed, a lot of Twitter data does not actually include any attitude at all – it is often just general broadcasting or link shares.

It should be noted that work on sentiment analysis has begun to drawn upon other methodologies beyond NLP. Some studies have drawn upon network analytics (see below) and specifically theories of emotional contagion to inform sentiment analysis algorithms. In 2014, Nesta also launched a 'Political Futures Tracker' which aims to combine NLP with new tools that will enable topics and sentiments to be mapped over time.[111] Others have attempted to refine existing techniques, for example, by attempting to incorporate more user information in to their analyses.[112]

Latent insight versus explicit insight

A lot of social media data contains what is known as 'meta-data' which refers to data *about* the data. As noted above, each tweet includes up to 33 additional pieces of meta-data associated with it, including the time it was posted and any relevant biographical data the user has chosen to share.

There is a great deal of interest in meta-data analysis, because meta-data can provide a very rich data source. Recent social media research efforts have sought to use meta-data as a way of classifying accounts on Twitter by increasingly complex and detailed categories. A recent research paper from the Collaborative Social Media Observatory (COSMOS) sought to classify Twitter accounts against the UK's NS-SEC socio-economic status framework, with mixed success. If meta-data analysis can allow accounts to be precisely categorised against measures used in wider population studies, such as censuses and national surveys, then the value of social media research to governments – in relation to attitudes, sentiments, opinions and online discourse – could be augmented significantly.[113]

NLP works on the premise that certain features of a text can be statistically analysed to provide probabilistic measures of meaning. One rapidly emerging area of study in NLP is to run classifiers on large training data sets in order to generically reveal 'latent' meaning, especially features about the author – age, gender and other demographics – which are not revealed explicitly in the text or captured by the social media platform and provided as meta-data, but which can be probabilistically determined by the underlying structures of language use. The development of latent NLP classifiers is an area of intensive investigation by university research institutes and social media platforms themselves.

One university, for example, has developed a fairly accurate gender estimator, based on around 50 characteristics that tend to be associated with male or female language use (there is a free test interface available; http://stealthserver01.ece.stevens-tech.edu/index trained against a large data set of emails. On Twitter, the main way to spot gender is by user name, which is possible using an automated system and is correct around 85 per cent of time. The inclusion of meta-data in analysis, or the use of nationally-specific data sets or lexicons, can increase this accuracy to 92 or even 95 per cent.[114] On other social networks, such as LinkedIn, this accuracy can reach over 98 per cent.[115] The researchers on this project are currently testing algorithms to determine geography, gender, precise of children, and socio-economic background. Full results are expected in late 2015, but there are important differences in how easy it is to determine different kinds of demographic information.

Ascertaining information about a user's location is another important area of work. Around 2–3 per cent of tweets include latitudinal and longitudinal meta-data, allowing tweets to be located very precisely. A larger body of tweets is possibly resoluble to a location through the use of additional meta-data. One academic study found that approximately 15 per cent of tweets can be geo-located to a specific city, based on the cross-referencing of other fields of meta-data: location (of the account, recorded as free-form text) and time zone (structured). Another study demonstrated that resolving place names to longitude/latitude coordinates have been shown to increase the ability to geo-locate social media documents by a factor of 2.5. Recent research has explored the exploitation of an individual's friends' or followers' geo-tagged locations to their location if they themselves are not otherwise geo-located.[116] Other methods of inferring geo-location relate to the content of tweets, such as 'location indicative words' (LIWs), or a user's location description combined with LIWs or longitude and latitude.[117] The accuracy of geo-location models varies with the detail of an inference. For example, one geo-location tool has reported 91 per cent accuracy at country level and 55 per cent accuracy at city level against an incomplete sample of tweets, while another reported that it could geo-locate an individual based on their friend's geolocation with 46 per cent accuracy.[118]

Other techniques have been applied to determine latent details from online data. A 2013 report by Berkeley and Cambridge universities found that it was possible to deduce personal information about people through an analysis of their 'likes' on Facebook, including sexual orientations, ethnicity, religious and political views, and some personality traits. The model correctly discriminated between homosexual and heterosexual men in 88 per cent of cases, African Americans and Caucasian Americans in 95 per cent of cases, and between Democrat and Republican in 85 per cent of cases. Drugs use was successfully predicted from likes in 65 per cent of the time.[119] However, personality traits – such as conscientiousness or happiness – were less easily deduced. It appears that simple demographic data – especially dichotomous variables – are more amenable to this type of analysis, but behaviour less so. Some studies have found that personality can be predicted with a reasonable degree of accuracy on the basis of web browsing, music collection, or friend numbers and networks.

NLP represents only one way of employing algorithms to determine the latent characteristics of Twitter users. Other research efforts have focused not on user characteristics, but on commonalities across different pieces of text to detect authorship across posts and accounts, with the design of systems capable of identifying similarities in the authorship of text samples.[120]

The use of automated language recognition to spot certain types of 'risky' behaviour or criminal intent is also a developing application of the NLP. Some linguists argue that certain structural, underlying features of a sentence and related syntax can be broadly correlated to general behaviour types, such as anger or frustration, subconscious states of mind. A RAND report on the science of violent act prediction suggested that certain linguistic markers could be associated with distinct psychological states, and that this might, in combination with natural language processing, present a potential tool in the effort to predict violent acts in advance.[121] The Cardiff Online Social Media Observatory (COSMOS) has also run a recent study looking at how opinion mining applications and sentiment analysis can be used to judge levels of online 'tension' over time.[122]

Based on our experience training classifiers, the extent to which this might be amenable to practical application will depend on the existence of training data – the information fed into the classifier to allow it to spot pattern. Therefore there is no reason that a classifier with enough training data would not be able to spot language use known to be correlated with certain behaviours (for example criminal activity); and assess the confidence with which it had made these decisions on the basis of quantifiable values. This would allow an analyst to effectively target further investigation. Indeed, the automatic detection of online criminal activity continues to be an important area of research. For example, recent research supported by the Spanish government has sought to combine NLP and psycho-linguistic characteristics to detect text in chat rooms that carries features characteristic of paedophilic sexual predation online, a proposition that could raise significant ethical issues. [123] The potential implementation of NLP to identify lone wolf threats online, through the identification of authorship, detection of linguist markers signalling potential terrorist violence and calculation of 'fixation' – the extent to which an author online writes about only a single subjects – has long been a subject of research.[124]

Other promising research in this area concerns the use of NLP to detect the veracity of online text samples, distinguishing between likely truths and likely deceptions, for example detecting when an individual is being misleading about their gender online based on culture-specific common gendered characteristics of language. NLP is already used to comb through large quantities of data in fraud cases, and models have been designed to detect deceptive language in financial statements.[125]

Technique: event detection and situational awareness

Social media can be viewed as an information platform containing 'events', defined as discrete incidents of, for example, a political, cultural, commercial or emergency nature. These events may be intrinsic to social media, such as a particular type of conversation or trend; conversely, they might be indicators or proxies of events that have occurred offline.[126] During the 2011 Egyptian revolution, for instance, 32,000 new groups were formed and 14,000 new pages created on Facebook in Egypt.[127] In Ukraine, the EuroMaydan Facebook page (named after the loose coalition of anti-government protestors) was set up on in late November 2013 and grew rapidly, at one point by 70,000 followers in a week, becoming the most popular Facebook page in Ukraine and surpassing 250,000 likes as the President was ousted from power in February 2014.[128]

Event detection technology attempts to identify and characterize events by observing the profiles of word or phrase usage over time - usually anomalous spikes of certain words and phrases together – that indicate that an event may be occurring. Broadly there are two styles of positively identifying an event: query drive and data driven. Query driven event detection is akin to waiting for a fairly specific 'thing' to happen and report that it has when enough evidence that matches the event 'query' has been recorded over a short enough time period. A purely data-driven event detection system has no preconceived notion what type of event it is meant to report. Rather it has a preconceived notion of what an event 'looks like' in terms of the statistical characteristics that are exhibited by the text stream.

Events can be detected through the examination of a range of measures. Rapid or anomalous changes in sentiment, in the frequency with which certain key-terms are used, and the volume of posts or tweets within a population are all examples of measures that can indicate an event. One of the key considerations of event detection is the population group that a researcher examines or measures. Community groups on social media are rarely clearly defined, and a significant event within a sub-group of the population might well not be as significant amongst the wider population group. Recent research has looked at ways of identifying groups at the same time as identifying anomalous events within that group.[129]

Situational awareness / event characterisation via Twitter

Of all the uses of event detection technology, building situational awareness of rapidly developing and chaotic events – especially emergencies – has, we think, some clear applications for counter-terrorism.[130] Emerging events are often reported on Twitter (and often spike shortly thereafter as eye witnesses start to share information) as they occur.[131] Increasingly, NLP and machine learning systems designed to provide situational awareness through social media data are being built and implemented. For example, SUPER, funded under the EU's Seventh Framework Programme for Research, is an emergency and security incident management system for social media, which combines event detection and summarisation, sentiment analysis (including by pre-defined community groups), rumour identification and credibility analysis and a social media data set search engine, brought together under a single interface.[132]

Social media users (especially Twitter users) can play a number of different roles in detecting events through the exchange of information. They can generate information about events first-hand. They can request information about events. They can 'broker' information by responding to information requests, checking information and adding additional information from other sources, and they can propagate information that already exists within the social media stream.

Multimedia content embedded on social media platforms can add useful information – audio, pictures and video – which can help to characterise events. One crucial area of development has been to combine different types of social media information across different platforms. One study used YouTube, Flickr and Facebook, including pictures, user-provided annotations and automatically generated information to detect events and identify their type and scale.[133]

Generally speaking, untrue stories tend to be short lived due to some Twitter users acting as information brokers, who actively check and debunk information that they have found to be false or unreliable. One study, for instance, found that false rumours are questioned more on Twitter by other users than true reportage.[134] Using topically agnostic features from the tweet stream itself has shown an accuracy of about 85 per cent on the detection of newsworthy events.[135]

One 2010 paper, 'Twitter under crisis', asked whether it was possible to determine 'confirmed truth' tweets from 'false rumour' tweets in the immediate aftermath of the Chilean earthquake. The research found that Twitter did tend toward weeding out falsehoods: 95 per cent of 'confirmed truth' tweets, were 'affirmed' by users, while only 0.3 per cent were 'denied'. By contrast, around 50 per cent of false rumour tweets were 'denied' by users. Nevertheless, the research may have suffered a number of flaws. It is known, for example, that the mainstream media still drives traffic – and that tweets including URL links tend to be most re-tweeted, suggesting that many users may have simply been following mainstream media sources. Moreover, in emergency response, there tends to be more URL shares (approximately 40 per cent compared to an average of 25 per cent) and fewer 'conversation threads'.[136] The 2011 London riots were widely discussed – and perhaps partly organised – via social media networks. It does not appear that Twitter was able to 'dispel' misinformation quickly. Indeed, rumours spread rapidly, and although some disagreement was found, they were within different, sealed networks.[137]

Indeed, more recent studies, for example research in 2014 into rumours and misinformation on Twitter in the aftermath of the 2013 Boston Bombing, confirmed this challenge to the idea of social media as the 'self-correcting crowd.' That research examined three rumours – that an eight-year-old girl had died, that the bombing was a 'false-flag' government plot and that Sunil Tripathi was a bomber – and found that while the ratio of misinformation to correction varied, tweets spreading misinformation outweighed those issuing corrections by between 5:1 and 44:1, and the number of corrections did not always increase in line with an increased circulation of the falsehood they addressed.[138] The potential harm that can be caused by falsehoods circulated on social media has led to a number of efforts to create 'social media lie detectors' – algorithms designed to automatically detect false event-related statements on social media.[139] Other research has found that inaccurate information spreads further and faster than subsequent corrections.[140]

One important factor, especially important for situational awareness, is the ability to identify the geo-spatial characteristics of an event. Many of the techniques described above to infer the location data of social media content are also used in the field of event detection. These techniques could be combined with other NLP functions, for example systems designed to automatically detect demonstration planning on social media – a particularly useful tool in democratic states where protests are often publically planned online – in order to gather information both before events take place and as they take place.[141]

There are a number of new research projects currently underway to try to predict events before they occur (or very quickly after they do) funded through EU FP7 grants. Some of these projects are ongoing, and so conclusive results cannot be presented here. Of particular note is; technology to allow researchers to quickly identify all (or more than just keyword-based searches) Facebook and Twitter data on a particular subject;[142] an evaluation of technologies used during and after crises and testing their effectiveness ('the Contribution of Social Media In Crisis management);[143] 'Athena', which is testing how far social media can be a way of information sharing between the public who are present at a critical event and the authorities;[144] 'Emergent', which is examining how far it is possible to effectively identify and integrate valuable and reliable information from social media into emergency management processes;[145] and Project Slándáil which is researching how social media can be better used to spread messages about the worst affected areas during these natural crises.[146] It is the view of the authors that this represents one of the most important areas of research, including for counter-terrorism. As we note below, the immediate aftermath of a terrorist incident is likely to result in a very significant volume of real-time, on the ground reporting from witnesses. This information could be extremely valuable in the immediate emergency response, and any longer term investigation. Finding, prioritising and acting on the relevant information is one area that requires an automated solution, and one that is well suited to machine learning applications.

Use of social media during a terrorist incident

Since the last version of this paper was published, there have been three major terrorist incidents in Western democracies where social media has been an important way of understanding how the events unfold: the 2014 Sydney hostage crises, the 2014 shootings at parliament hill, Ottawa, and the 2015 Charlie Hebdo shootings.

The Sydney hostage crisis, in which Iranian-born Man Haron Monis held 18 people hostage in a café for 16 hours, was notable for the way in which the gunman used the hostages as social media go-betweens. In the event, the hostages used Facebook, Twitter and YouTube to communicate demands including Monis' desire to speak to Prime Minister Tony Abbott and for an Islamic State flag to be delivered to the café. Police involved in the operation in turn were able to glean valuable information from these channels that may have helped shape their rescue strategy. Following the incident, many came out in support of Muslim communities in Australia on social media, fearing a backlash as a result of the hostage crisis. According to Twitter Australia, 40,000 tweets used the hashtag #illridewithyou in just two hours, growing to 150,000 in four hours during the siege.[147]

Trending topics were also an important feature in the aftermath of the attack on the Charlie Hebdo Paris offices in 2015, when gunmen linked to Al Qaeda's branch in Yemen killed 12 people including a policer officer. Within 24 hours of the attack, the hashtag #JeSuisCharlie received over 3.4 million mentions in a display of solidarity that was matched by national demonstrations in the days after the attack.[148] Opinions echoed on social media, however, were interpreted by some as embodying anti-Muslim sentiments. Concurrently there was a spike in the use of the hashtag #RespectforMuslims in relation to the incident, which was used 160,000 times in 24 hours.

The 2014 attack on Parliament Hill in Ottawa, in which a Canadian soldier was shot while on sentry duty and the Parliament held up by a single gunman, demonstrated some of the potential uses of social media and citizen reporting in times of crisis. In the event, Twitter was flooded with real-time updates for people to avoid certain downtown areas, for example.

However, the Centre for International Policy Studies also cautioned the use of social media in emergencies. In the case of the Ottawa shooting, they draw attention to the fact that thousands of people in the surrounding remained unnecessarily 'locked down' as a result of the circulation of false reports on Twitter of a second gunman.[149] Concerns were also raised about the safety of officers as a result of citizen reporting. Ottawa Police Sgt. Iain Pidcock, for example, tweeted urging people not to disclose locations or photos of police on social media while responses were still underway.

Twitter following the 2014 attack on Parliament Hill, Ottawa
The research team writing this paper began collected data on Twitter during the 2014 attack on Parliament Hill in Ottawa. The first recorded tweet about the incident was at 13:53 UK time, which read: "Gunfire at the War Memorial in Downtown #Ottowa". There were 1,653 separate tweets within the first ten minutes. Between 21st – 30th October we collected around 1.3 million tweets on the subject (although this is unlikely to be the total figure, since we collected it based on popular hashtags and key words – which technique invariably misses a lot of data). The most popular content shows the importance of both news and official pronouncement. The two most popular tweets were both from the RCMP: "@rcmpgrcpolice: #RCMP advises if you are in downtown Ottawa to stay away from windows and off roofs due to ongoing police incident" (3,092 Retweets); and "@RCMPONT: Please do not post videos or photos of the on-going incident to ensure safety of first responders and the public. #Ottawa" (3,058 Retweets). The most popular link shared was the CNN's live Coverage (4,110 times).

Alongside commentary on social media usage in the recent terrorist incidents detailed above, there is now a greater interest in bringing in techniques of sentiment analysis as a means of predicting backlashes to terrorist events. A 2014 study by COSMOS looked at mapping sentiments on social media specifically following terrorist events as a way of predicting information size and survival. The authors concluded "using the case of the 2013 Woolwich attack in London, that sentiments expressed in tweets (either positive or negative) were significant predictors of both the amount of coverage on social media (size) and the amount of time the information flow lasted for (survival)".[150]

Visualisation and dashboards

Alongside the development of analytical techniques and technologies for understanding social media data has been the development of ways of presenting and understanding complex, large data sets in clear, intuitive ways. 'Data visualisation' describes a very large group of techniques and tools used to do this, ranging from freely available graphing and plotting software to bespoke coding platforms.

Data visualisation importantly allows large, complex data not only to be understood, but also interrogated – comparisons and overlays between different kinds of data, for patterns to emerge and trends spotted. A collection of visualisation dashboards, such as Qlik and Tableau, have emerged that allow data sets to be filtered and re-presented by the different variables and quantities that they contain. This has proven to be important in social media research, as these dashboards allow data to be filtered, combined, compared, and broken down.

Dashboards have also become important for decision-making based on large data sets, and so the operational application of social media analysis. Dashboards are used to convey the key findings and policy-relevant insights from the analysis in digestible and understandable forms, especially to non-specialist decision makers, and often in real time. Dashboards are increasingly understood to form important bridges – between the data and the analyst, and also between the technical and specialised challenge of analysing social media data, and the non-specialist decision maker who often needs the findings quickly, and in a format that can be understood at a glance.

Technique: data mining and prediction

Broadly, there is a growing sense that the 'big data' revolution – the ability of humans to make measurements about the world, record, store and analyse them in unprecedented quantities – is making new kinds of predictions possible. This, 'predictive analytics', brings together a wide range of intellectual and technical infrastructure, from modelling and machine learning to statistics and psychology.

The explosion of social media is part of the big data revolution. More and more of our intellectual, cultural and social activity is being captured in digital form on social media platforms. It represents the 'datafication' of social life. It renders social life measurable and recordable.

Interest in harnessing these social-digital traces by predictive analytics was sparked by a paper published in 2009 by Hal Varian, Google's chief economist, who argued that Google search terms can sometimes predict real world behaviour (such as searches relating to jobs preceding and predicting unemployment figures). Since then, there has been an interest in applying predictive analytics to social media data sets to predict a range of social behaviours and phenomena, from election results, to box office numbers, to product sales and stock market trends.[151] While some research has challenged the value of predictive analytics in certain areas – for example in relation to predictions of box office success – in other areas, like stock market trading, the practice of social media-based predictive analytics has become big business.[152]

Politics

Correlations of social media sentiment are also subject to predictive analytics. Eric Siegel, in *Predictive Analytics – The Power to Predict Who Will Click, Buy, Lie, or Die* explains how Obama's predictive analytics team predicted those 'swing voters' who had the greatest likelihood of being influenced to vote for Obama. They used data from Twitter and Facebook to predict which people were strong influencers of the swing voters, and targeted them, not the swing voters themselves (an example of the 'Persuasion Effect'). That approach is at the very cutting edge of predictive analytics today, largely because of its development and successful deployment within American electoral campaigns.[153]

Increasingly, academics and researchers are trying to examine the relationship between Twitter activity and voting intention. In 2010, the organisation Tweetminster estimated the likely overall vote share of each of the parties by adding up the total volume of mentions per candidate. Predictions for national results were more accurate (90.5 per cent) than predictions for regional results (87.5 per cent), which in turn were more accurate than results for individual candidates (69 per cent). Based on the overall vote share (although not seat share) this was comparable to predictions made by traditional polling techniques.

It's not always quite so accurate, however. A similar study was undertaken in the German Federal election of 2009, although these results were critically analysed by other researchers, who found that the relative frequency of mentions of political parties had no predictive power, and argued the results were contingent on the arbitrary choices of the researchers. On replication, the researchers included the online group the Pirate Party, which the original research team failed to do, and found that it secured the greatest share of Twitter mentions and yet failed to secure a single seat. Similarly, although with a higher degree of sophistication, researchers at the University of Indiana used mentions of candidates to accurately predict the outcome of US Congressional Elections – even when controlled for incumbency, district partisanship, media coverage, time, and demographic variables.[154]

Likewise, Nick Beauchamp from Northeastern University in 'Predicting and Interpolating State-level Polling using Twitter Textual Data' has found that it is possible to model the correlations between state-level polls and the text content of tweets posted from that state. He found that the text content of tweets can predict changes in fully representative opinion polls 'with a precision currently unfeasible with existing polling data'. (In short, this suggests that words found in tweets do correlate with changes in voting intention).[155]

Despite these significant improvements, the state of the art remains poor. Some of the successes of predicting voting using Twitter have been the result of what is called 'over-fitting', meaning predictions that are made after the event itself, and where several models were tested to find the one that worked.[156] There are significant challenges involved with using Twitter to predict election results (all of which are noted by the research groups mentioned above).

Health

One area that has received a lot of attention is the use of big data to understand the spread of infectious disease, known as 'public health monitoring'. Some analysts believe this will become a vital part of spotting and tracking health trends. It has famously been suggested that Google search terms for flu symptoms, can identify outbreaks faster than doctor's records, though that claim has been aggressively disputed.[157] More recent research has suggested that Twitter is a more accurate tool than Google for predicting flu outbreaks.[158] Twitter has predictive value across a range of health-related areas. For example, a person's Twitter content can be used to predict the likelihood of heart disease with greater predictive value than factors such as income,[159] and other research has shown that Twitter can be a valuable means with which to predict the number of asthma-related admissions to hospital that will occur in an area.[160]

One 2012 paper found that, based on an analysis of 2.5 million geo-tagged tweets, those with online ties to an infected person were more likely to be infected, particularly where geographically proximate (due, of course, to the increased incidence of physical transmission). The analysis was based on 6,237 'geo-active users', who were tweeting with geo-location enabled Twitter accounts more than 100 times per month. While the results are fairly obvious, the researchers suggest that these findings demonstrate that Twitter analysis can help model global epidemics.

This study was undertaken through the analysis of only open, geo-located Twitter accounts, and using machine learning as outlined above to identify tweets which appear indicative of flu. Some papers have suggested ways to geo-spatially characterise social media posts for health-related analysis, combining text features (eg tags as a prominent example of short, unstructured text labels) with spatial knowledge (eg geo-tags, coordinates of images and videos).[161]

Crime detection

Most of the work that has been done on criminal incident prediction relies primarily on historical crime records, geospatial information and demographic information, and does not take in to account the rich and rapidly expanding social media context that surrounds many incidents of interest. One paper presents a preliminary investigation of Twitter-based criminal incident prediction. The model analysed the tweets of a single feed (Charlottesville, Virginia news agency), but believed an adapted version could potentially be used for a larger-scale analysis of tweets. Rather than keyword volume analysis and sentiment analysis, which are unhelpful to predict discrete criminal incidents that are not mentioned ahead of time, the authors used NLP techniques to extract the semantic event content of the tweets. They then identified event-based topics and used these to predict future occurrences of criminal incidents. The performance of the predictive model that was built was evaluated using ground-truth criminal incident data, and compared favourably to the baseline model that used traditional time series methods to study hit-and-run incidents per day.[162] More recent research has sought to supplement criminal incident prediction models with social media data. Examining a range of 25 crimes, including prostitution, criminal damage and burglary, a recent paper found that the addition of Twitter data to a common crime prediction model (kernel density estimation) improved the accuracy of prediction for 19 of them.[163]

This technology can extend beyond the board prediction of crime trends. Motorola's Intelligent Data Portal produces intelligence for first responders, producing individual address-based threat scores and profiles for police officers derived from commercial and public data sources including social media – particularly 'offensive comments.'[164] A number of police forces in the United States, for example the Chicago area police, have begun to integrate social media into their situational awareness systems and first responder intelligence tools.[165]

The problem of prediction

Nate Silver has described how big data-driven predictions can succeed but also fail in his recent book *The Signal and the Noise*. He argues that 'prediction in the era of big data is not going very well'. Silver attributes this to our propensity for finding random patterns in noise, and suggests the amount of noise is increasing relative to the amount of signal, resulting in enormous data sets producing lots of correlative patterns which are ultimately neither causal, nor accurate, nor valuable.[166]

Correlations, without either sound theoretic underpinning or explanation, are common in many branches of social media research. Incidental correlations of this kind – such as an apparently strong relationship identified in one Facebook study between high levels of intelligence and the liking of 'Curly Fries' – add little insight or value.[167] Silver's suggestion is that we use more Bayesian mathematics: probabilistic predictions of real-world events based on clear expressions of prior beliefs, rather than statistical significance tests or dichotomous predictions. Interestingly, as Silver points out, big companies spend less time modelling than running hundreds of data experiments to test their hypotheses.[168]

Indeed, predictive analytics have rarely been used experimentally and then tested in reality. All studies cited in this paper have been based on a 'retrospective fit' – where researchers, acting with the benefit of hindsight, construct post-event analyses of pre-event data. At the same time, the most high performance classifiers are not those that are pre-prepared, and generically applied, but those that are created to fit a specific data set. This limitation remains a barrier to effective 'real-time' use of social media data for predictive analytics, sentiment analysis and situation awareness, and is currently the subject of extensive research efforts.

This is obviously ill suited to many of the operational needs of counter-terrorism agencies, who have to make time-dependent forecasts in chaotic, unpredictable and fundamentally uncertain circumstances.

Technique: network analysis

Introduction

Social network analysis (henceforth, SNA) is at its root a sociological and mathematical discipline that pre-dates the internet and social media. It aims to discern the nature, intensity, and frequency of social ties, often as complex networks. Its premise is that social ties influence individuals, their beliefs, and behaviours and experiences. By measuring, mapping, describing and modelling these ties, social network analysts attempt to explain and indeed predict the behaviour of the individuals who comprise the network.

In order to derive SOCMINT, SNA can be conducted on different types of data sets of online activities, initially focusing on blogs, news stories, discussion boards, and now most frequently social media sites. It attempts to measure and understand those 'network links' both explicitly and implicitly created by the features of the platform, and how the platform is used. These include: formal members of particular movements, those who follow or interact with specific Twitter feeds and Facebook pages, members of forums, communities of interests, and interactions between users. Sometimes these are referred to as 'explicit' or 'implicit' communities depending on the degree of involvement in or commitment to the group in question. A recent article examining the flow of information around global media events concluded that 'users form networks of influence via their interactions affecting the ways that information is shared about specific global events'.[169]

Explicit communities tend to refer to groups where members have made an explicit decision to join a blog-ring, Twitter cluster, group, or network, while implicit communities refer to the existence of broader interactions such as linking, or commenting. In addition, there is an increasing body of research showing some characteristics of online networks mirror the structure of social networks in the offline world.[170] One key aspect is an observed limit on the size of natural face-to-face social networks theorised to result from a combination of cognitive and time constraints.[171] In practice cognitive and time constraints drive humans toward gathering information from a small number of sources they consider reliable. For example, a USIP study of the conflict in Syria, the most socially mediated civil conflict in history, concluded that the "pattern in social media toward clustering into insular like-minded communities is unmistakable and has profound implications".[172] The ability to identify these clusters of closely interconnected individuals is a key element of SOCMINT, as different clusters may interpret information in different ways, and most importantly adopt different behaviours based on that interpretation of the information. However, the network characteristics of digital information are often measured using a technique, pre-dating the internet and the advent of social media; to analyse the relationships between members of a specific network.

The influence which the 'social' aspect of social media has on the flow of news information amongst a community[173] has been the focus of a range of studies, including the role of social media in political change,[174] in sharing information about elections,[175] and even attempts to forecast large-scale human behaviour.[176] These studies follow on from one of the classic questions in social science, articulated by Larsen and Hill: 'How does the news get around?'[177] This question has been asked in many forms: how people locate job opportunities,[178] or seek trusted information,[179] and how small communities connect to form large social structures which allow two individuals who have never met to send and receive information through long chains of acquaintances.[180]

Studies analysing how news and information flow through a community often return to the questions posed by Larsen and Hill: 'Who learned the news, when and by what means, and how much did such knowledge affect subsequent communications behaviour?'[181] While these questions have been addressed from many angles over the years, including polling in the aftermath of specific events such as the September 11th terrorist attacks,[182] the increasing ability to access news via websites, and more recently social media, provides greater opportunities for the use of SOCMINT to understand diffusion of information between the users of these platforms.

SNA in practice

The following section examines different ways SNA can be used as part of SOCMINT, to understand the relationship between websites, relationships on social media, the flow of information, and the diffusion of content.

For the analysis of websites, which has influenced the subsequent analysis of social media platforms, crawlers follow hypertext links from one site to the next, recording whether and how each page links to others. In general terms, a crawler tends to start from a small number of carefully selected seed sites and then continuously find the links from there to other sites. There is a range of methodologies for effective crawl 'depth' in research (meaning how many steps should be crawled from the seed sites). The design of the data capture and selection of seed sites for a web crawl stems from the perspective created by the research question.[183]

Borgatti, in his famous analysis of 200 Conservative bloggers, used a crawl depth of two in order to balance the risk of a sample being too shallow - a significant risk when the crawl depth is one - with the risk of a sample being too deep, introducing a high degree of noise, or mapping neighbouring issue networks. Indeed, a crawl depth of two was also used in a number of recent studies concerning a variety of political networks, including pro-gun control networks, the mapping of the Norwegian Blogosphere and conservative Bloggers active during the 2013 Presidential Election in Iran.[184]

Linkages can be split into three classes: content, structure and usage. The identification of these kinds of linkages allows the user to build a data set of online activities, whether they take place on blogs, news sites, discussion boards, or social media sites.[185] Once the data are gathered they can be used for a number of purposes, ranging from the analysis of how many individuals are engaged in a specific activity online, to the assessment of information flows and influence in complex systems.[186] Indeed, it is possible to map even covert networks using data from Open Source intelligence sources on the World Wide Web, as shown by researchers including Valdis Krebs.[187]

Typical activities include:

- Tracking increases in content produced about a specific issue or location
- Tracking the spread of a specific piece of information
- Tracking the sharing of information between individuals
- Understanding the complex structures created by the behaviour of individuals which influences the information other users receive, and subsequently the behaviours those communities adopt.

Numerous mathematical techniques can be used to understand and describe social networks expressed in social media data. Centrality analysis is a well-established technique that describes the position of any given node in a network to other nodes through three measures.

First, the 'degree' – or how many links a node has to other nodes. High-degree nodes are sometimes described as 'Achilles' heels' within a network, and often represent 'leaders' or 'influencers' of various types.

Second, 'betweenness' measures how far a node lies between other nodes in a network. Nodes with high betweenness are sometimes considered the gate-keepers between different, tighter clusters within a looser network, and act as important channels of influence between them.

Third, 'closeness' is measured as the sum of the length between a node and the other nodes (low scores means it may be hard to communicate).

Another commonly used type of analysis is known as 'community analysis', which is designed to identify social groups in a network. A 'community' is identified where members of a group have a higher density of links than with those outside a group; the specific limits of a group can be accurately divined by the establishment of a 'threshold' which determines at what point a node is part of a group.

Friends, Followers and affiliates – understanding the loose network
Several groups likely to be of interest host open social media accounts. The network of open account followers of Al-Shabaab – easily downloadable – is highly diverse, with many likely to be curious spectators, journalists, researchers or

analysts as well as supporters and ideologically aligned fellow-travellers. There is not, as far as we know, any technique for making these distinctions, beyond careful and manual reconstruction of each individual 'ego' and the analysis of each follower's network.

It is for this reason that the free, automated analytics tools of Twitter followers, such as followerwonk or foller.me can be highly misleading. When making policy decisions, it is often good practice to use systems that are transparent about the way influence or 'influencers' has been calculated. Some more detailed academic studies have been able to rank users' influence on a specific subject area, rather than more simplistic measures such as engagement and follower numbers. By analysing their followers, and whom they follow, on a thematic basis, it is possible to observe clustered relationships based on particular themes.[188]

A recent paper published an analysis of the 3,542 followers of 12 White Nationalist Twitter accounts, and a random sample of each of their 200 most recent tweets. It was found that around 50 per cent did not overtly subscribe to White Nationalist ideology (although these were not removed in the final analysis). The researchers created their own compound measure of network influence. Rather than using the existing centrality measures detailed above, they measured 'influence' through the combination of two metrics: 'engagement' (the number of times a user's tweets resulted in a response of any kind, for example in the form of a reply, retweet or favourite); and 'exposure' (the number of times a user responded to other people's tweets in the same way).

As noted above, it is possible to create new measures of understanding networks in this way through Twitter. This research found the most 'engaged' also tended to be the most overt supporters of White Nationalism: 93 of the 100 most engaged accounts were also those who appeared the most overt supporters of White Nationalism.[189] When the same method was applied to anarchist accounts, results were less clear-cut. The data set was less coherent, and there was less covert self-identification as anarchist; as a result, top engagement was not as closely correlated with active involvement.[190]

This research also found a large number of link shares. The authors argued that by identifying the key content among radical and extreme groups, through the links that they share, it would be possible to understand in greater detail their ideology. Furthermore, the paper recommended that targeting shared links for disruption through terms of service violation reporting would be an important potential counter-extremism tactic.

A similar study of White Nationalist Twitter accounts started with a core or seed set of accounts. In this case, social ties were measured through the phenomenon of one user mentioning another through the use of a Twitter handle (@<username>) in a tweet; in this context, reciprocal mentions can be considered a dialogue. A network was then created based on these collected reciprocations. A 'highly stable' network based on significant dialogue was thereby mapped out, and an analysis undertaken on common keywords employed, in order to determine the common themes of communication within the community. The research team then conducted analysis on the location of members, with some success.

The research found that the dialogue network tended to be among people from the same country, in contrast to a simple network of followers (although this allowed the researchers to identify a user acting as an English language translator for a Swedish nationalist group). However, the work has a caveat, recognising the likely incompleteness of data sets it used, presumably based on the imperfect choices made when selecting the initial core seed accounts.[191]

Relational networks

An important application of SNA for SOCMINT has been to estimate the strength of relationships on the basis of different forms of social media activity, as a precursor to more detailed understanding, often through predictive modelling, of how strong and loose networks influence individual behaviour.[192]

In very general terms, Facebook evidences high rates of connection: 92 per cent of users are connected by four degrees of separation. It also appears to support the 'weak ties' argument – many users relate intensively and constantly with a small group of friends (10–20) but follow more loosely a larger group (150–200). Other research has found that there is a tendency to join a community based on both the number of friends the user has in the community, but also, importantly, by how those friends are connected to one another.[193] Jure Leskovec's work has been especially prominent in this regard, applying social psychological principles to machine learning to interpret and predict the positive and negative feedback on the Epinions, Slashdot and Wikipedia platforms.[194]

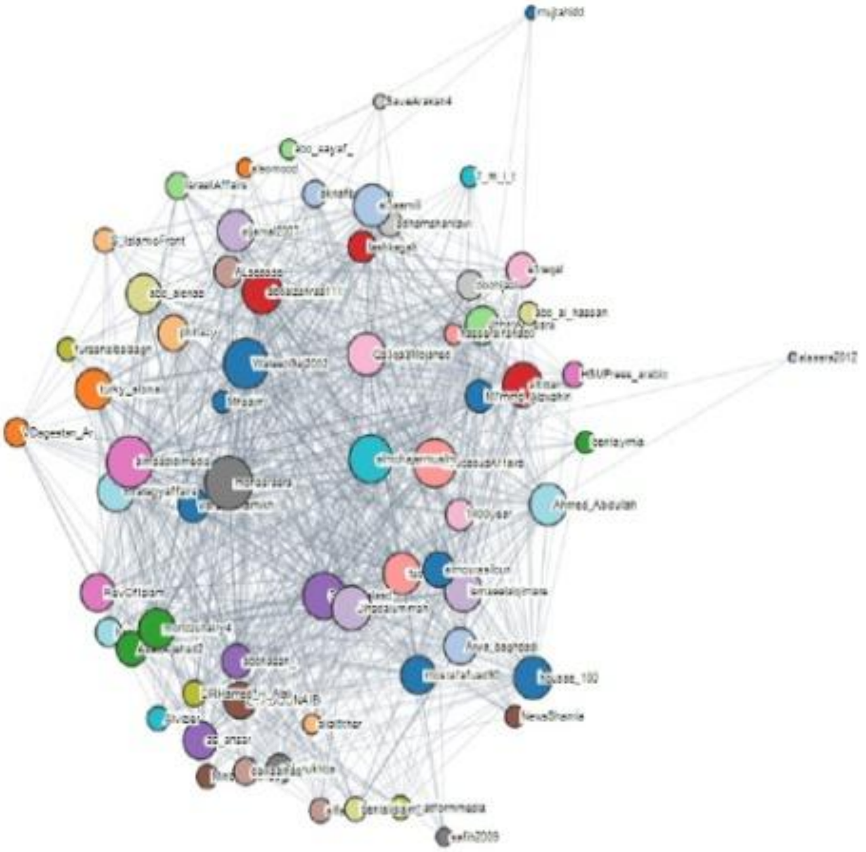
In one study of White Nationalist blog sites, researchers manually identified a series of seed blogs and blog-rings that used White Nationalist terminology in their title or description.[195] The subsequent crawl of linked sites identified 28 groups, comprising 820 individual bloggers, and found more blog groups than were listed on hate speech directories, suggesting that the use of blog spiders was as useful in identifying groups as understanding them.

The researchers noted how many users were active on each blog, and further extracted all profile information about each blogger, including their user ID, date of birth, city, and real name. This profile information is self-reported, and thus of dubious accuracy, although the 'blog creation date' is automatically recorded by the host site and is therefore a reliable source. In this case the 'link' analysed to understand the network was whether the blogger had subscribed to another blog-ring.

The researchers found the community was well connected internally – the average number of links that would link any member of the network to any other was only 2.89. The clustering coefficient – a measure of how tightly grouped together nodes in a network are – was 0.37, characteristic of a small, nascent community. This data set, similar to many others found on social media, was subject to a 'power-law distribution': the top bloggers had many more direct links than other members of the community.[196] Those with high in-degree scores might be usefully subject to further detailed analysis of their blogs to understand better how ideologies, motivations and messages are formed and spread throughout the group.

In this approach to SNA provides insight on a number of levels; providing a sense of the structure of the network as a whole showing the existence of specific clusters, as well as the ability to locate influential individuals. The ability to gain a tangible overview of the relationships between candidates in the UK general election was shown in a recent Demos analysis, featured on the BBC.[197] This demonstrated the ability to identify the clusters of users and their particular political affiliation as well as the extent to which popular tweets tended to be shared only by a specific cluster.

An example of SNA use in SOCMINT, an analysis published in 2014 showed the relationship between 66 important jihadist accounts on Twitter.[198] It found the users which these 66 accounts tended to follow.



Interactive version can be found here: http://bit.ly/1cFbjDg

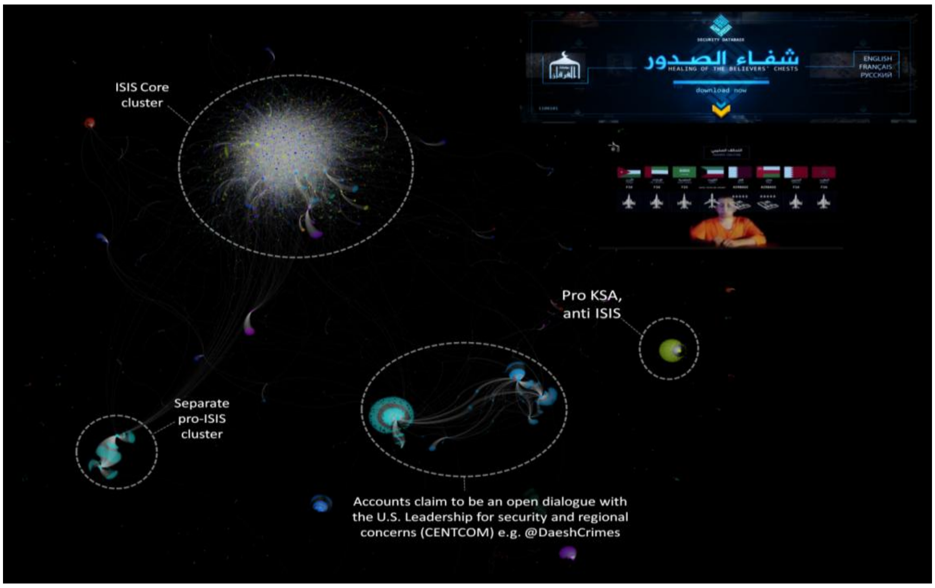
At the time of the analysis, the account of Jabhat al-Nusra was the account which was most popular amongst these important jihadist accounts.[199] This type of analysis has an advantage over services such as Klout which provide influence metrics across an entire platform, as it allows an analyst to identify accounts which are particularly influential amongst a specific target community. In SOCMINT this level of detail is often more important than the platform-wide influence as many uses of SOCMINT focus on probing specific niche themes and activity.

A strong sub-discipline within social media research has sought to identify why and how people are influenced (either ideologically or behaviorally) on social media. This has often been driven by a desire to market and reach 'key influencers' within a particular field. A 2010 study built a series of 'models of influence' that strongly predicted on a probabilistic basis whether a user would perform an action on social media on the basis of their position on a social graph.

Information flow and content dissemination

A significant component of an individual's information environment is the relationships that affect how they acquire information and knowledge.[200] Research in this field has focused on how information (and therefore influence) flows in multiple directions, and how it coordinates around hubs or focal points. Crawlers, API calls and network analysis are jointly being used to develop insight, locate influential individuals or communities of influence and understand the hubs around which these social media users coordinate.

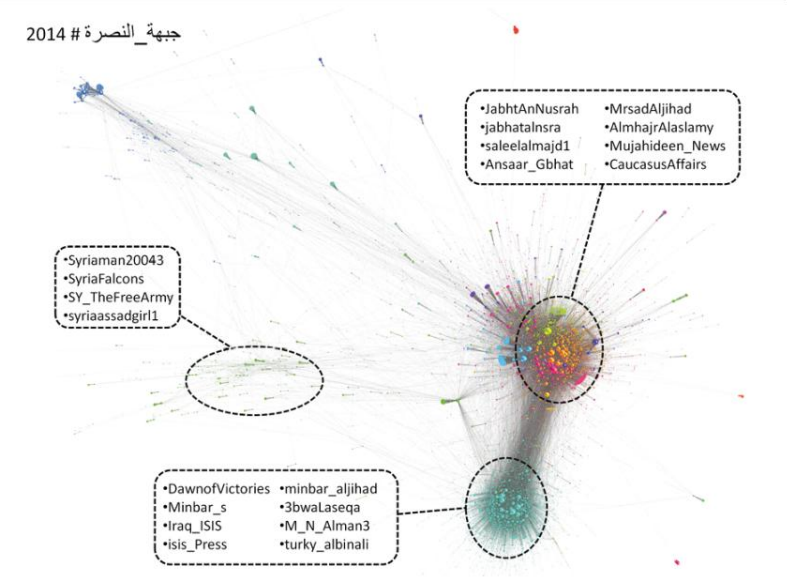
For example, information flow, if represented by a directional network, allows influential users and particularly interconnected clusters to be identified. It allows SOCMINT to address a question posed by Larsen and Hill: 'Who learned the news, when and by what means, and how much did such knowledge affect subsequent communications behaviour?'[201] For example, the release of a jihadist video, such as the execution of the Jordanian pilot, provides an opportunity to analyse the means through which content diffuses through Twitter. As the image below shows, this method allows an analyst to identify a cluster of users potentially sympathetic to jihadist groups and those who seek to counter ISIS messaging.[202] In effect, SNA can be used to evaluate whether counter-messaging is engaging the same group of users who are sharing content produced by extremist groups.



SECURITY DATABASE
شفاء الصدور
HEALING OF THE BELIEVERS' CHESTS
download now
ENGLISH
FRANÇAIS
РУССКИЙ
ISIS Core cluster
Pro KSA, anti ISIS
Separate pro-ISIS cluster
Accounts claim to be an open dialogue with the U.S. Leadership for security and regional concerns (CENTCOM) e.g. @DaeshCrimes

On a social media platform such as Twitter, users who are influential across the network as a whole can be discerned through a number of measures of 'centrality', as discussed above. All such measures are useful in different ways; 'degree centrality', for example, shows those with the greatest number of connections.[203]

An analyst using a combination of these measures and community detection algorithms can gain extensive insight into the interconnected communities which exist within the network, along with the level of influence of different members. It may, as demonstrated in a 2013 study of Jabhat al-Nusra on Twitter, provide insight into the extent to which an organisation can benefit from being retweeted by specific regional broadcasters.[204]



A repetition of the 2013 study during the same period in 2014 demonstrated how ISIS sympathisers used the same hashtag as Nusra sympathisers but still interacted with each other in identifiable and largely separate clusters.[205] This demonstrates that SNA can be used to investigate the potential nuance and particular allegiances between the sympathisers of specific extremist groups. This allows an analyst to identify accounts which fulfil important roles in specific groups within the wider flow of information.

In addition to using individual network metrics, SOCMINT can utilise the relationship between different network metrics to identify 'Key Actors' in a network. This builds on the work by Valente et al. which examined the correlation between network metrics, and used to examine audience interaction with BBC accounts during the London 2012 Olympic Games.[206]

To identify key actors two network metrics – Betweenness and PageRank – can be plotted against each other. Betweenness represents how important a user is in facilitating the flow of information to specific parts of the network. Individuals fulfilling this role are often known as bridges or gatekeepers and are valuable as they select and tailor information to users in a specific part of the network. Those with a high PageRank score are key members of the network because other important network members interact with them. These users are heavily invested in the activity of that network, and are usually recognised as important actors by other members.[207] When the position of each user is plotted against both metrics, the different roles which key actors fulfil can be identified.

The combination of network analysis and sentiment analysis has been used in order to identify the focal points for specific communities within a wider trending topic or complex issues. For example, research into the use of Twitter during the 2009 Iranian election revealed quite a distinct set of different communities that were using the tag #Iranelection, and very different topics that 'trended' within them.[208]

What are they sharing?

One of the most remarkable emerging areas of interest has centered around the analysis of what information is 'shared' – content that has been posted and then re-posted, re-tweeted or otherwise further disseminated by individual users. According to one analyst, more detailed understanding of what is being shared can provide insight into a group's changing beliefs and views, and is the most interesting element of social media analysis, with greater promise than social media sentiment analysis.[209]

Numerous free or cheap tools cam provide simple data on the trends of link sharing. Seesmic, for instance, provides a number of useful tools for understanding of media consumption behaviours within communities. Cascade, a piece of software developed by the *New York Times,* shows who shares each story and when, in order to understand the structure of sharing. This is done by analysing the trajectory of bit.ly URL shortened links. This helps identify influencers (who shares the most and who drives subsequent traffic); and which variables appear to affect this.

Unedited, user-generated content adds to the challenge as people often copy and paste entire articles or parts of articles into blog posts without providing a hyperlink to the source.[210] Inter-media have explored the nature of media consumption during the Arab Spring by examining patterns of sharing. They found that some journalists were important in driving traffic to particular news stories, blogs and tweets, and became information brokers, aggregating, filtering and disseminating relevant content.[211]

Discussion

There has not yet, however, been a full and detailed study into the sociology of the phenomenon of sharing. The social significance of sharing content – news stories or otherwise – is little understood. Given this, any extrapolation from content sharing into the purposes or motivations of the sharers must be treated with care.[212]

This has important implications for understanding the research design of SOCMINT SNA. As with machine learning, determining the nature of a network depends partly on the initial decisions of an analyst in deciding what sort of link is important and selecting seed accounts/sites. One key, recurring theme in network analysis study is the extent to which study design captures everyone in a network. One of the key challenges in analysing a network is how and where to define the boundary of the network. There is rarely a simple boundary to a network, and the larger the networks, the less likely there is to be a clear boundary, as recognised by Malcolm Sparrow.[213] The case for embracing the concept of more fluid network of relations rather than conceiving of groups in a formal, rigid structure has been further reiterated by the 'Fijnaut Group' and others studying organised crime in the Netherlands.[214] Indeed, as Borgatti has argued: 'the choice of nodes should be dictated by the research question and one's explanatory theory', rather than arbitrary, inflexible conditions.[215]

Automated network analysis can produce both strategic and tactical insight, but only in the appropriate context. Real-time monitoring, which tracks shifts in the volume of terms and content produced around a specific issue, can have significant tactical value. It can, for example provide real-time intelligence regarding changes in locations being mentioned by groups seeking to create or exploit public disorder.

One of the most useful aspects of automated network analysis is the identification of information, groups and individuals. The rapid identification of the most engaged individuals in certain ideas is an extremely simple and cheap type of analysis, which can be done via access to APIs without any machine learning or NLP. However, this would still require a great deal of analytical review because of the lack of clarity about who these people are. On the strategic level, on the other hand, analysis can also be usefully done through big data sets which comprise content aggregated from a range of tactical sources and over a longer time period than (near) real time. At this level, it is possible to analyse the wider information system through the fluctuations in volume flow, and in doing so identify users who have different influential roles within that system.

Technique: netnography

'Netnography' broadly refers to the application of ethnographic and qualitative sociological methodologies to the study of social media data. Consistent with the theoretical commitments of these disciplines, netnography usually avoids the quantification or numerical measurement of social media data and instead sees it as part of an individual's social and cultural life that is textured, complex and often only understandable when studied in depth. Netnography often practically takes the form of 'participant observation' – sustained contacts between the researcher and members of a digital community.

The careful study of behaviours within forums of 'communities of interest' is a potentially powerful way of gaining insight into attitude formation and behaviour. One recent detailed study of jihadi forums examined the role that a lack of trust between extremists in online forums played in their interactions, what made them suspicious of each other, and how they built inter-personal trust in an environment of surveillance, as well as examining how the nature of trust in their interpersonal relationships changed with forum architecture. Such ethnographic studies can have significant counter-terrorism policy implications, because they can identify opportunities for law enforcement. This research, for example, demonstrated how high-profile agent provocateur activities within online communities could have a significant and long-lasting disruptive effect on terror networks.[216]

The kinds of social space that forums represent are changeable. Many chat room forums include several sub-forums, some of which are public and others private. For counter-terrorism purposes, and more generally the study of discussions based on socially problematic or stigmatised views, closed forums are often more valuable than open ones. One report concerning al-Qaeda forums found that 'it is not possible to have a rounded sense of what is taking place through only the public sites' – although such a level of access does offer useful insight into the 'zeitgeist' of the movements, including broad ideological shifts.

Reconstructions of offline groups from social media

In some instances, very careful reconstruction of the social interactions contained within a forum can help researchers understand the specific membership and hierarchy of a group. In 2009, Strathclyde Police launched Operation Access, which used social networking sites such as Facebook to uncover criminal activity by identifying weapons carriers, especially in the context of urban gang memberships and inter-gang feuding. As part of the programme, police officers searched through images to find users who had posted pictures of themselves with weapons. The Superintendent in charge of the operation stated that as a result, 400 people were questioned. More current examples include the work of the Canadian think tank the SecDev Foundation. The foundation monitors the social media activity of drug cartels in Latin America on Twitter, Facebook, Instagram, Tumblr and a range of other platforms on a large scale and for a range of purposes. They map the relationship between cartels and gangs to help measure their changing levels of influence, to help build and maintain an active glossary of slang terms, and to better understand the internal structure and operation of the cartels, and to map gang activity – an undertaking facilitated by an incautious approach amongst many gang members to geo-tagging. Recently, the SecDev Foundation has expanded its projects to work with the police in California.[217]

Growth of digital sociology / computational social media sciences

Over the last five years, we have seen a significant growth in the academic study of 'big' social data. This has focused on the methodological combination of social science and computer science disciplines together to collect, combine and understand very large bodies of social data. Predicated on the proposition that leveraging new, and very large amounts of data about social life will uncover new knowledge of social processes and dynamics, a number of large institutes and departments within universities have formed to develop and apply new methods to do this. They have a variety of different focuses, academic preoccupations and technological expertise. UK institutions include: the Urban Laboratory at University College London, the Oxford Internet Institute, the Centre for Interdisciplinary Methodologies at the University of Warwick, the Digital Sociology programme at the University of Goldsmiths, the Social Data Lab at the University of Cardiff, the Visual Social Data Lab based at the University of Sheffield, and the Computational Sociology Centre at the University of Surrey.

This body of new academic effort is changing academic practice in a number of ways. It is causing current disciplinary boundaries within universities to be challenged, especially as funding bodies have begun to emphasise the importance of not only multi-disciplinary, but inter-disciplinary work across quantitative and qualitative disciplines. It is also leading to new research partnerships, as researchers work with data scientists from the private sector, and the social media platforms themselves, in accessing and analysing social data. It has also led to a range of new advisory roles for academics within government as a range of departments and public-sector bodies attempt to incorporate and harness emerging analytical techniques emerging from academia and the private sector.

Several recent projects have also been conducted into how to use social media interactions as a basis for anti-extremist education and training. The Kanishka Project, set up in 2011, is a five-year initiative by the Canadian government which offers funding to research in how counter terrorism policies can be made more effective. Recent research proposals under the initiative include a project that intends to monitor the social media impact of films created to counter narratives of violent extremism. This will build on SecDev's previous study in to the feasibility and appropriateness of using open source information as an early warning system for youth radicalisation. Other studies in the fifth round of the Kanishka Project proposals include plans to develop and evaluate anti-extremism teaching materials, based on longer-term observations of how hate speech can lead to violence among young people.[218]

Data verification problems

There are lots of examples of inaccurate information or misinformation being widely believed, even by subject specialists. Tweets by imposter accounts have been picked up, believed and reported on by major news outlets. A number of Facebook studies have asked whether users tend to portray an accurate picture of themselves on the site, and a literature review of these publications suggests that Facebook profiles convey fairly accurate personality impressions of users. This may be because, unlike other online groups, people tend to become Facebook friends only after being offline friends.

In many respects, determining the veracity of any single source is much the same as usual – and would require the same standards and methods applied in any human intelligence source: track record, known capabilities, motivations, and so on. As such, most important techniques appear to be fairly obvious: images can be cross-referenced against known landmarks, and through the checking of unique URLs. Social media adds some technological components that might be more at home in intelligence fields such as imagery intelligence (or, IMINT). For example, a basic knowledge of current capabilities of widespread imagery manipulation software such as WARP – a perspective modification tool – is necessary. Other techniques might be specifically related to certain social media platforms. Producing visually convincing photography of inauthentic tweets and Facebook content is straightforward and has been used in the past.

Technique: 'Crowd sourced' information

It has been recognised in recent years that public safety requires the involvement of a large number of different actors. For example, the British counter-terrorism strategy relies on the active engagement of citizens. In the US, the gang prevention initiatives that work most effectively are those that have 'all-community' involvement from the police, social support services, charities, youth groups, local churches, parents' organisations, rehabilitation centres and schools. There is significant potential for the police to create and curate networks of citizens cooperating to keep their community safe. Indeed, this 'co-production' of safety and security has already developed in many areas, often at the instigation and insistence of civilian participants, not the police.

Social media can be, and often is, used to inform the public rapidly and directly, and as a way of directly asking the public for help and assistance in keeping the public safe. Possible applications are as diverse as policing itself, from reporting successes and providing reassurance to promoting community activities and engagement or delivering statements, particularly following a major incident (for example, the 2013 terrorist attack in Boston).

A more controversial method of official engagement with the public is to dispel rumours and conspiracy theories, including by proactively intervening in discussions and conversations. Police forces in the UK regularly dismiss rumours regarding subjects as diverse as terror threats, riots, demonstrations and abductions through Twitter and Facebook, as well as responding to queries from the public and investigating complaints made over social media.[219] For example, West Midlands Police Force used social media, particularly Twitter, to counter rumours of an attack on the police station by posting 'Twitpics' of officers standing outside the station. This use of images to respond to rumours and misinformation is recommended in Defence Science and Technological Laboratories guidelines regarding social media use in emergency management.[220]

Direct solicitation for information

Perhaps the greatest way crowd sourcing is currently being used to collect information on individuals is the simplest: asking the public. Recently, some US police forces have also used content sharing sites, such as Pinterest, to ask for the public's assistance in identifying criminals. Following the 2011 UK riots, police uploaded photos to a Flickr stream and a dedicated website that compiled images of people thought to be involved in looting. As a result, 770 people were arrested and 167 charged. Furthermore, up to 2,800 images were uploaded to the smartphone app Facewatch ID, created in partnership with Crimestoppers, which allowed users to sort the images via postcode and then inform on those they recognise by sending a name and address to the police. The app also included 2,000 or more images of people wanted for offenses not connected to the riots. Similarly, in the aftermath, citizens organised themselves using #riotcleanup, and staged public demonstrations in condemnation of the criminality and the violence. However, there are potential problems with this kind of crowd-sourced policing, as the FBI's appeal for public information after the 2013 Boston bombing revealed, when the misidentification of the bombers by users of Reddit and 4Chan led to their harassment and vilification.[221] Interestingly, the platforms within which these crowd sourced investigations took place may have influenced the quality of their output: on Reddit it is possible to 'upvote' and 'downvote' content based on its merit, whereas on 4Chan it is not, removing a useful faculty with which to hide poor information and promote useful information.[222]

The development of mobile applications for the purpose of supporting incident reporting has been an area of significant activity. In the US, further development has focused on making the use of these apps more simple and ensuring more accurate reporting of incidents, events, and tip-offs (known as e-tips). Anonymous reporting functions have also been developed in an effort to extend popular participation. The continued development of useful reporting apps could help create a citizens' network of reliable human intelligence sources for event detection. Anonymisation might encourage more accurate and more problematic reporting – but does not help in verifying the information itself. Possible areas for further development in the area of direct information solicitation include crowd sourcing rumour verification or dismissal during emergencies.[223]

This approach has also become more popular among private companies. This is a wide area and examples are diverse, and often exist outside of social media. One example of how social media is used to achieve this sort of 'co-created' research is provided by Burberry, which has developed an online discussion forum where users and product developers have conversations about what kind of products they expect next season.[224]

Challenges of two-way communications

The challenge is to set the right balance of central control. Counter-terrorism police operations increasingly work to secure positive citizen engagement, but police forces have understandably sought to limit the risks of this new environment by issuing guidance and establishing internal control procedures. Police forces usually issue strict guidance which requires police officers to protect the reputation of the force and to pay proper attention to operational considerations such as the protection of the identities of victims and witnesses, the protection of the integrity of current operations, and the avoidance of comment which might be prejudicial to legal proceedings.

However well controlled, the opening of direct channels of communication between the public and the police poses inherent risks. Responsibility rests with the police to respond to emergency calls, and with a lesser degree of urgency, other non-emergency forms of contact by the public. Most police forces reviewed by the authors have taken the view that tweets directed at an official account should not be treated with the same degree of urgency as other forms of communication – indeed, most police sites on Twitter contain a warning not to use the channel to report crime. Twitter feeds are not routinely staffed 24 hours a day or integrated into force control centres. Nevertheless, numerous forces are reporting a significant increase in the number of information requests coming to them through social media – and there are not, as far as we know, systems in place to manage and filter these requests. While this is not a significant problem yet, we anticipate it might become one in the near future.

Social media can also disrupt other forms of communication and engagement. Social media is challenging both for the press and for force press officers – journalists and reporters increasingly find breaking stories online, and seek police verification before the force is ready to confirm or deny a particular instance.

PART 4: LEGAL ETHICAL AND PRACTICAL CONSIDERATIONS

Summary

- Overall, open SOCMINT does pose new challenges to existing legal frameworks that govern intelligence work, and ethical challenges for research work. If SOCMINT is to be used as a valuable and legitimate form of insight, we believe it must be based on a clear legal, publicly argued footing.

- The main difficulty facing most law enforcement agencies when collecting information of any kind is the extent to which certain types of data collection might require a legal authorisation. In making this decision, the very broad principle to which most legal frameworks adhere is that of 'reasonable expectation' of privacy. This is a useful guide, although needs to be based on a detailed understanding of the platforms and data being analysed. In general terms – and in our judgement – simply because something is publicly available or accessible does not mean that no legal authorisation is required.

- One useful way to determine an individual's expectation of privacy on social media is by reference to whether that individual has made any explicit effort or decision in order to ensure that third parties cannot access this information. This can be done in a variety of ways.

- Social media analysis software and tools allow for far greater surveillance than ever before, with concomitant risks and opportunities. The increased use of automated software to collect and analyse information (inevitable in the age of terabytes of unstructured data) poses additional risks of misuse. Therefore outsourcing data analytics to third parties could potentially result in breaches of the law if the data collection efforts are not clear.

- While the guidance which governs lawful access and academic research ethics are based on similar underlying principles, they are fairly different in practice. For academics, we consider reasonably assumed consent and minimisation of harm to research subjects to be paramount.

- Overall, given increased public concerns about privacy, public acceptability and proportionality should inform any decisions taken in respect of even open SOCMINT – even where a decision has been made that no warrant is required. Agencies undertaking this type of research work should try to conduct open SOCMINT work according to good ethical and professional research standards:
    - Being explicit and public about the research aims and methods used where possible.

- Considering whether the measures taken might reasonably be seen as proportionate by those potentially monitored, and could be defended as such. Even where data is anonymised, there is increasing public concern about 'pseudo' anonymous data, where individuals can be identified by cross-referencing data sets.
- Assessing if any measures might undermine the existence of a free and open internet, which would cause damage to the economic and social well-being of the nation. It is our view the benefits of collecting and storing large quantities of open data in a general, non-targeted way, should be carefully weighed against this possible risk; and whether such measures are an effective use of public money.

Lawful access and social media

Like all intelligence work, SOCMINT work must be carried out within a legal framework. Most OECD countries have legislation that covers the collection and use of private information, which is intended to ensure that state agencies can only access citizens' private information in a legal, proportionate way, with various mechanisms of oversight and scrutiny designed to minimise potential abuses of power. Different countries have different legal frameworks underpinned by slightly different principles. In the UK for example, the collection of information that might reasonably be considered 'private' requires the utilisation of a strict authorisation process and oversight by legitimate bodies, as well as that the intelligence is used for appropriate purposes and gathered using appropriate methods.

The main difficulty facing most law enforcement agencies when collecting information of any kind is the extent to which certain types of data collection might require a legal authorisation. In making this decision, the very broad principle to which most legal frameworks adhere is that of 'reasonable expectation' of privacy. Of course, there is sometimes a difference in how this principle is applied. In the UK, RIPA authorisation is required where there is a likelihood that 'private information' will be obtained, even if it comes from a public source. In Canada under the 'Intercept' parts of the Criminal Code, it appears the key consideration is whether the communication itself – rather than the content – might be reasonably considered private.[225] In Belgium, whether 'technical means' are used has a bearing on whether a judicial or department warrant is necessary.

Despite these different considerations 'reasonable expectation' is a useful starting point in respect of social media. Based on our previous work on the subject, we believe that the specifics of any judgment about reasonable expectation will rely upon a number of distinctions and assessments – from what is proportionate to what is a private space – that are contextual, mutable and a matter of degree. In respect of social media, these assessments are extremely difficult to make because:

- SOCMINT covers both open-source data and closed networks. Sometimes, however, the distinction is not clear. For example, Facebook accounts and groups often have varying degrees of openness, and different platforms often have quite different terms and conditions and norms of use that might determine the degree of intrusion. There is therefore no clear definition of what might be considered private information.
- Social media analysis software and tools allow for far greater surveillance than ever before, with concomitant risks and opportunities. The increased use of automated software to collect and analyse information (inevitable in the age of terabytes of unstructured data) poses additional risks of misuse.
- Public attitudes toward data sovereignty and privacy (even on open platforms) change quickly, and there are reputational risks for law enforcement agencies seen as 'snooping' online.

Determining 'reasonable expectation'

The academic Susan Brenner has highlighted two questions which are specific to US law enforcement, but which can provide a useful format through which to frame the consideration of reasonable expectation of privacy. Writing about 'search' under the 4th Amendment, Brenner draws on Katz v. US, 389 US 347 (1967),[226] and suggests that someone has a reasonable expectation of privacy in a place/thing if two conditions are met: (i) he thinks it is private; and (ii) society accepts as objectively reasonable his belief that the place/thing is private.

This was re-affirmed as an important principle in a 2012 US Supreme Court decision relating to the FBI's use of warrantless GPS tracking devices placed on the underside of cars parked in public places: although the Court did not rule on the reasonable expectation consideration (being limited to whether the use of GPS constituted a 'search'), the principle was discussed in detail in the ruling and affirmed as the key principle at stake by Justices Alito and Sotomayer.[227] Increasingly, questions are being raised about reasonable expectation in relation to communications content held in storage by third parties, where the laws such as the Stored Communications Act (1986) have not been updated to reflect the technology landscape of today.[228]

Since this is remains a question of degrees and judgement, below we discuss some considerations that may help determine both an individual's and society's reasonable expectation of privacy.

Individual expectation

One way to determine an individual's expectation of privacy on social media is by reference to whether that individual has made any explicit effort or decision in order to ensure that third parties cannot access this information. This could be manifested in a series of ways:

- Any data coming from closed accounts, or any account or group where a restriction has been placed limiting the access (for example 'friends only' settings). This suggests that an explicit decision has been made to limit the access of outside parties and the material can thus be considered a 'private communication' even if the group involved is extremely large; and could thus be considered to be reasonably private.
- Where a password is required in order to enter a site.
- Any robot.txt restriction that has been placed by the site administrator in order to prevent permission for a search bot or scraper to access data.

An individual's reasonable expectation can also be determined by reference to the terms and conditions of use of a forum or site and the typical behaviour of users, as these will often help to shape the expectations that an individual has about the nature of the interaction. Not all 'open' platforms are the same in terms of the reasonable expectations of the user. This too can be manifested in several ways:

- Some chat room forums and threads have fairly explicit instructions that request that users sign in to take part, and that data and conversations are not shared outside of the group, while others (such as Twitter) make it clear that they will encourage people's personal information to be widely shared –and that users should be prepared for that.
- Often, targeting certain individuals results in the obtainment of information on other, non-targeted individuals who interact with the targeted individual, information which these non-targeted individuals may reasonably consider to be private.
- Most terms and conditions of sites – including Facebook – clearly state that users are expected to be honest about their profile information. Therefore, the creation of fake/pseudo social media accounts (on Facebook these are sometimes called 'ferret' accounts) in order to join a closed group or chat room, including when an individual joins using a blank or anonymous account, might be considered unreasonable – and in some senses a privacy breach.
- Similarly, any direct interaction in any forum – open or closed – in which an officer seeks to elicit information and are not explicit about their real identity can be problematic.

Social expectation

Brenner also argued that a reasonable expectation is driven by society's view of what is acceptable. Recent debates in relation to internet surveillance – in particular those sparked by the Edward Snowden revelations – have demonstrated that public acceptability is fundamental to any measures being undertaken. Polling data reveals that online privacy have become an issue of growing public concern.[229] The CIG-Ipsos Global Survey on Internet Security and Trust suggested that two thirds of users (64 per cent) were more concerned about internet security compared to one year ago, and only 36 per cent believed that private information on the internet is very secure.[230]

Polling undertaken in 2012 suggests that the erosion of privacy is the second most important worry Canadians have, just behind the global financial crisis (but ahead of climate change and terrorism). Seventy-two per cent of Canadians express concern about this compared to 73 per cent who worry about the financial crisis.[231] They worry specifically about online privacy too: a 2013 survey prepared for the Office of the Privacy Commissioner of Canada found that one quarter of respondents said they were extremely concerned about the protection of their privacy. In terms of the risks posed to personal privacy, most Canadians cited financial information/bank fraud (23 per cent), followed closely by computer and internet privacy (21 per cent) and identity theft (21 per cent).[232]

These sentiments have been pronounced for a number of years. A 2011 poll commissioned by Canada's federal privacy watchdog found that 82 per cent of Canadians were against giving police and spy agencies the power to access emails and online data without court authorisation.[233] Increasing concerns were also reported in a 2014 poll, in which 52 per cent of Canadians reported that they were concerned about government and law enforcement agencies monitoring their internet activity.[234]

At the same time however, citizens expect security services and law enforcement agencies to have the necessary powers to fulfil their obligations in regards to public safety and security. The challenge posed to authorities, therefore, is to balance these potentially conflicting societal expectations of promoting safety and security while at the same time protecting citizens' online privacy. There are indications that there is general support for the idea that law enforcement agencies should be able to access social media data for public safety and security purposes.

For example, a 2008 Eurobarometer poll found that 80 per cent of European citizens trust the use of citizens' personal information in a proper way by police, while the same survey found that a majority supported the monitoring of internet activity to protect society against terrorism. Similarly, a poll undertaken by consultancy firm Accenture in five countries, including Canada, found that 72 per cent of respondents believe social media can aid in criminal investigations and prosecutions.[235] However, a 2013 survey clarified that while only a quarter (27 per cent) were very concerned with Canadian law enforcement or security agencies using personal information *with* a warrant, this rose to 55 per cent who were very concerned at such use of personal information *without* a warrant.[236]

One key challenge therefore is in identifying what type of SOCMINT data can reasonably considered to be open, and what might be considered in some senses private and therefore require lawful warranting to access. We consider that the following types of SOCMINT collection might be reasonably considered as genuinely open-source and non-intrusive, where there is little or no expectation of privacy:

- Volunteered crowd-source intelligence through direct and explicit solicitation. This should be employed wherever possible instead of 'listening in' technologies and techniques. In these instances, the expectation of privacy can reasonably be considered to be very low.
- Where social media users have no reasonable expectation of a right to privacy, because they understand this content is likely to be shared and used. That condition is met if any terms of agreement establish that content uploaded is public and will be made available through an Application Programme Interface access. A good example is Twitter, which makes it clear that it will actively encourage sharing, which means that data collected from open Twitter accounts would not require authorisation. In addition, no privacy blocks or walls (often effected through the use of robot.txt restriction) or password requirements should exist.
- Network analysis through the use of 'crawlers' or 'spiders' (automated programmes to map a network of individual accounts), providing no individuals are named, no private information about an individual is collected, and providing API access is granted through robot.txt, and it is made clear in the terms and conditions that data are shared. The use of automated systems can make these decisions more difficult. As a robots.txt file is not enforceable, they can be ignored when crawling. However, many would consider that 'bad manners'. On the other hand, this is not a privacy control, as the page is viewable publicly and the purpose for which open source intelligence is being used may outweigh the desire to 'be polite'.
- Furthermore, the legal and ethical dimensions shift when the content or behaviours on a specific site are considered criminal. Additionally, every port requires a txt file, or else the bot assumes no restrictions are in place. Since 2011, for example, Facebook has banned the collection of data using automated means (without their explicit approval). Crawlers will tend to issue multiple requests to the site for information. Sometime a site administrator will notice multiple requests and then decide to prevent access. However, it is possible – and inexpensive – to use Virtual Private Networks which reroute requests, to make it appear to the site that the request is coming from multiple different sources. Again, this may result in serious privacy concerns.

In addition to robot.txt, the server and API etiquette is important. Many systems pause, or sleep, briefly between calls. This is to prevent the crawl, or API calls, from putting unnecessary strain on the servers where the data is stored. The reason for this is partly etiquette, and partly practical. Some sites have the capacity to ban users that are making too great a demand on a server from accessing further information on that server. This is done through an IP ban, denying anyone using a specified IP address from accessing data on a server, or rate limiting when data is accessed through a platform API. There are numerous ways around these limitations, but the decision to get round an IP ban should be taken after consideration of the ethics of overriding the attempt by a system administrator to block further data collection – along with any legal considerations which relate within the relevant jurisdiction.

Given the reputational considerations set out above, public acceptability and proportionality should inform any decisions taken in respect of even open SOCMINT – even where a decision has been made that no warrant is required. Agencies undertaking this type of research work should try to conduct open SOCMINT work according to good ethical and professional research standards:

- Being explicit and public about the research aims and methods used where possible.
- Considering whether the measures taken might reasonably be seen as proportionate by those potentially monitored, and could be defended as such. Even where data is anonymised, there is increasing public concern about 'pseudo' anonymous data, where individuals can be identified by cross-referencing data sets.
- Assessing if any measures might undermine the existence of a free and open internet, which would cause damage to the economic and social well-being of the nation. It is our view the benefits of collecting and storing large amounts of open data in a general, non-targeted way, should be carefully weighed against this possible risk; and whether such measures are an effective use of public money.

Research ethics / regulatory bodies guidance

Research ethics are not legally binding. Rather, they are a set of commonly agreed principles according to which academic research institutions undertake research. Similar to legislation covering intelligence work, they aim to measure and minimise harm, and in this instance balance the need to undertake socially useful research against possible risks to those involved. They cover more varied harm considerations than the law, and are usually given effect by university ethics committees or institutional review boards.

The extent to which legal or ethical considerations come into play is likely to be driven by which organisation is conducting the work, and for what purposes. For example, professional regulatory bodies – such as market research regulators – also adopt similar principles, albeit with slightly different considerations.

Anderson Review into surveillance powers

In June 2015 David Anderson QC, the independent reviewer of terrorism legislation in the UK, published a review into the effectiveness of existing legislation relating to investigatory powers, and to examine the case for a new or amending law, entitled *A Question of Trust*. It contained 124 recommendations for reform.

Based on his review – in which he had full unrestricted access to the top secret material – Anderson recommended that a new piece of legislation be drafted, replacing 'the multitude of current powers and providing for clear limits and safeguards on any intrusive power' He also recommended that the definitions of 'content' and 'communications' data be reviewed, clarified, and brought up to date. This does not mean substantially weakening the ability of the state to conduct surveillance where necessary. Anderson argued, for example, that the UK government should still require service providers to retain communications data for a period of time; to conduct 'bulk' collection of intercepted material subject to warrants. (Both have been widely criticised by campaigners).

Anderson characterised the key issue as one of trust, and based his recommendations on five broad principles: minimise no-go areas, limited powers, rights compliance, clarity and unified approach. The overarching aim is to minimise the amount of unnecessary surveillance undertaken, and ensure the law which governs these powers is clear and widely understood. Anderson stresses that any effective surveillance regime needs to be based on the broad understanding, consent and trust of those subject to it.

One recommendation of how to improve trust is greater oversight. Anderson called for Judicial rather than Ministerial warrants for interception; and a single Independent Surveillance and Intelligence Commissioners (at present, there are three separate Commissioners, each covering slightly different aspects of surveillance work). The government has not implemented these proposals – but as of writing is committed to proposing new legislation in the Autumn of 2015.

Principles of research ethics and ethical treatment of persons are codified in a number of policies and accepted documents such as the UN Declaration of Human Rights and the Declaration of Helsinki, which aim to uphold the principles of human dignity, safety and respect for individuals, and maximise benefits while minimising harms. In the UK, the standard best practice is the Economic and Social Research Council (ESRC) ethical framework, composed of six principles that express this broader moral-ethical doctrine. The ESRC will not offer funding to research projects that do not demonstrably adhere to these principles. Social media research is a new field, and the extent to which (and how) these ethical guidelines apply practically to research taking place on social media is as of yet unclear. Because the nature of social media research is highly varied – ranging from large quantitative data analysis down to very detailed anthropology – there is no single approach that can be applied.

Principles and approaches for research work have been set out by the Canadian Panel on Research Ethics, in their 2014 'Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans'.[237] The principles themselves are different from those espoused by the ESRC – the Tri-Council has three core principles, 'Respect for Persons'; 'Concern for Welfare'; and 'Justice'. However, the underlying values and implications are very similar. For example, an 'important mechanism for respecting participants' autonomy in research is the requirement to seek their free, informed and ongoing consent'. Similarly, 'Concern for Welfare' means that researchers and Research Ethics Boards should aim to protect the welfare of participants, and, in some circumstances, to promote that welfare in view of any foreseeable risks associated with the research. And, as above, the purpose of the Policy Statement is to help researchers and Research Ethics Board make difficult judgements about balancing these principles against the legitimate needs and social benefits of conducting research.

The Association of Internet Researchers (AoIR) also outlines seven different inquiries that constitute internet research, some of which are applicable to social media research. These include utilising the internet to collect data or information through data scraping; study of how people use and access the internet including participating on social networks; employing visual and textual analytics to study images, writing and media forms.[238] The definition of internet research is too broad to deal specifically with social media research but the AoIR guideline does provide an initial distinction within social media research. This is to do with how researchers collect their data. Some types of social media research rely on automation due to dealing with vast quantities of data, whilst others are mostly observing a small number of users. A rough analogy might be drawn here with quantitative and qualitative research –one gleans information from large data sets, whilst the other inquires into individual experiences.

Market research regulators

ESOMAR – an international market research regulator – published guidelines for social media research in 2011.[239] In this they outline several different kinds of social media research to which their guidelines apply:

- Social media monitoring (which encompasses everything from basic desk research to automated sentiment analysis)
- Ethnographic analysis (this includes observation of behaviour online and may include direct contact such as 'friending' subjects)
- Online communities (these may be created by the researcher, or pre-existing)Co-creational research techniques (feeding the ideas of users directly into new products).

In another similar publication, CASRO – the largest American market research membership body – listed similar methodologies within social media research.[240]

Very generally, most guidance for market researchers makes a distinction between private and public social media. This is because some areas online (sometimes known as 'walled gardens') may give the user an expectation of some kind of privacy. Private social media may include areas where privacy settings are set up to prevent individuals seeing the user's profile or posts. A good example of this is Facebook, which is public social media – visible to anyone with a Facebook account – until a user tightens their privacy settings to prevent people who are not their 'friends' from seeing their details. At this point their details become visible only to those who are friends with them and are classed by guidelines as 'private' social media.[241] Different types of social media platform and forums often offer very different types of privacy settings, often blurring the distinction between public and private spaces.

The AoIR released its latest ethical framework in March 2013.[242] This guidance – first issued in 2002 and frequently updated – is commonly used by institutional review boards when making decisions. The AoIR notes that ethical principles cannot be applied universally, but must rather be understood inductively and through the use of applied practical judgments. Crucially, they note that because all digital information at some point involves individuals, consideration of principles related to research on human subjects may be necessary even if it is not immediately apparent how and where persons are involved in the research data.[243]

We consider, drawing on the ESRC model, that the most commonly applied principles for human subject research are as follows: a) any possible harms to participants must be measured, managed, and minimised; b) informed consent should be sought when and where possible. These guidelines are considered below. The issue of whether 'informed consent' is required on open public social media data sets – and how that can be reasonably achieved – remains perhaps the biggest and as yet unresolved debate in social media research.

Harm to participants

There is a broadly agreed obligation for researchers to avoid research that is harmful to its subjects, irrespective of how research is collected. Harm is difficult to measure in respect of social media research. There may be new harms related to mass data extraction, such as a loss of confidence in the platform. Extraction tools need to be designed to avoid accidental extraction from non-public accounts, and new forms of collection – such as extracting profile information – might in some instances require explicit consent. The 2015 ESRC framework for research ethics changes 'harm to research participants and researchers must be avoided in all instances' (in the 2012 framework) to the more nuanced 'research should be worthwhile and provide value that outweighs any risk or harm. Researchers should aim to maximise the benefit of the research and minimise potential risk of harm to participants and researchers. All potential risk and harm should be mitigated by robust precautions.'

This type of ethical consideration is important when using crawlers and other automated bots, particularly when one considers how the digital medium can potentially dehumanise research subjects. Crawlers can collect information about user profiles, types, videos uploaded, and place individuals within a network. Individuals within that network – especially on the fringes – might very reasonably consider this to be a significant harm. If the crawler is looking at linked sites, it will quickly fall upon rather more moderate forums and individuals who may find themselves on the fringes of a network of extremist sites.

The presentation of the data is a critical consideration. According to the British Psychological Society, researchers should avoid using quotes that are traceable to an individual posting via a search engine unless the participant has fully understood and consented to this. As such, quotes should be paraphrased, and not linked to the forum they were gathered from.[244]

According to the 'Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans', privacy is an important element of the principle of 'Respect for Persons'. It advises that researchers are 'expected to determine whether the information or data proposed in research may reasonably be expected to identify an individual'. If it might identify an individual, the ethical concerns are heightened, and consent becomes more important. The Policy Statement notes that it is not always possible to use anonymous or anonymised data (although this is usually preferable) and that it is often possible to re-identify individuals even where it might appear that measures have been taken to avoid that risk. In those cases 'the ethical duty of confidentiality and the use of appropriate measures to safeguard information become paramount. This Policy generally requires more stringent protections in research involving identifiable information.'

Informed consent

In traditional research methods, the principle of informed consent refers to the need for researchers to be open about who they are, about the purpose of their work and about how it will be disseminated. Informed consent is considered of vital importance as a way to minimise harm, and is meant to ensure that there is no explicit or implicit coercion so that research subjects can make an informed and free decision on their involvement in the research. (This is mentioned in both the Tri-Council Statement and the ESRC framework.) They should therefore be informed about the fact that information they share is being used for research purposes. Informed consent is not always necessary, however, and in certain cases such consent is widely acknowledged to be impracticable or meaningless, for example in research on crowd behaviour.[245]

While such research should not be undertaken without particular caution and consideration, research without informed consent can be justified when no details about an individual are likely to be divulged, and where the risk of harm to research subjects is fully minimised. (However, in this case, it is perhaps unwise to refer to 'informed' consent, since this tends to have a very specific meaning, which cannot be realised with proxy assumptions of consent. One interesting instance was the recent Facebook study which 'manipulated' users' emotions.)

The application of informed consent is likely to vary depending on the type of research being undertaken. The developing field of internet research poses various new challenges to this basic research principle because of the ambiguity of the concepts of privacy and informed consent in online settings, and the difficulties of establishing the real identity of research subjects and of obtaining their consent.

One outstanding question is whether or not an individual signing a service's terms and conditions is a sufficient proxy for informed consent. Upon signing up to a social media service, users subscribe to a privacy policy terms of use that govern the users access to that service. In the past, these documents have been accused of being overly long and complex. Research has shown that the majority of users do not read these documents. Both Facebook and Twitter's terms of use are very far-reaching in what they permit the companies to do with public data. The data can be shared and processed in ways that are close to being unlimited. Given that users sign up to these terms, it may be assumed that the social media researchers receive consent through the terms of use.

Facebook's 'emotional contagion' experiment

In January 2012, Facebook conducted a week-long study with academics from Cornell and the University of California in which it manipulated the newsfeeds of 689,003 users, removing either all of the positive posts or all of the negative posts to see how it affected their moods. Given that users have already agreed to the site's data use policy, it was deemed unnecessary to get participants to sign consent forms for the experiment. The policy itself says that information will be used for "internal operations, including troubleshooting, data analysis, testing, research, and service improvement".

Researchers concluded that emotions – or at least language expressing emotion – were contagious. It was also noted that when 'emotional' posts were removed from a person's news feed, they became "less expressive", for example, writing fewer status updates.

Following the announcement of its study, Facebook came under widespread criticism, with critics suggesting that it may have breached ethical and legal guidelines by not informing its users that they were being manipulated for experimental purposes. Facebook responded to the controversy over 'informed consent' by saying that it would change the way it undertakes research in future, but stopped short of offering an apology for the study itself.

The approach of the market research regulators appears to be fairly permissive in this respect. In 2014 the Market Research Society produced a new code of conduct, with two new clauses about informed consent:[246]

16) Members must ensure that participants give their informed consent where personal data are collected directly from them.

17) Members must ensure that they have a fair and lawful basis for the collection and processing of personal data from sources other than the data subject themselves.

The inclusion of consent in cases where personal data is not collected 'directly' from the research subject allows passive data collection to take place on social media without the 'informed consent' of the individuals involved. Clause 17 backs this up by saying expressly that processing personal data from third parties may be allowed so long as it is lawful and in line with terms of use.

The extent to which all of these principles are applicable will partly depend on the extent to which the platform is open or closed. As with the legal framework, one useful heuristic is the social networks' own privacy policies, for example, Facebook's restriction of the use of web-crawlers in March 2011. This is useful because it is an important indication of the expectation of privacy. As it stands, it is generally agreed that Twitter data are in the public domain and can therefore be treated as carrying implicit informed consent.

Even with open source data, however, certain conditions still ought to be met. Because this area of research is so changeable and often difficult to interpret, principles rather than hard and fast rules are most suitable (for example, the use of 'situational ethics'). These principles need to take into account frequent technological changes, the medium involved and the expectations of the research subjects.

NOTES

[1] RE Wilson, SD Gosling & LT Graham (2012), ‚A review of Facebook research in the social sciences', *Perspectives on Psychological Science*, 7(3), 203-220. doi: 10.1177/1745691612442904. [CHECK]

[2] The paper had to have used 'social media' as the primary or sole focus. Following a relatively consensual definition, social media was defined as both the technology and the use of a varied category of internet services inspired by 'the participatory web' or 'web 2.0' which enable users to create and share digital content,

whether textual, audio or video. Where possible, the paper should have been published within the last three years, especially if it was related to newly emerging techniques or areas of rapid development. The paper had to suggest a method, capability, technique or usage considered by the researchers to be broadly relevant to the purposes of countering terrorism (with particular stress on the prevention of violent extremism and the broader task of understanding the phenomenon). Notably, given the diversity of the literature, no judgment of

relevance was made a priori regarding research design, methodology or technique. All studies found to meet the criteria above were incorporated and considered. These were found through a variety of techniques: a) Scholarly searches using keywords relevant and terms relevant to the purpose as defined above. Experts in the field were approached for the purpose of bibliographical recommendation.

[3] Cira Factbook: http://cira.ca/factbook/2014/the-canadian-internet.html

[4] TNS Mobile Life: http://cira.ca/factbook/2014/the-canadian-internet.html

[5] http://cira.ca/factbook/2014/the-canadian-internet.html

[6] http://www.statista.com/statistics/270229/usage-duration-of-social-networks-by-country/

[7]http://www.slideshare.net/wearesocialsg/digital-social-mobile-in-2015?ref=http://wearesocial.net/blog/2015/01/digital-social-mobile-worldwide-2015/

[8] http://www.alexa.com/topsites/countries/CA

[9] http://www.pewinternet.org/2015/01/09/demographics-of-key-social-networking-platforms-2/

[10] http://www.adweek.com/socialtimes/canada-social-media-study/614360

[11] https://aytm.com/blog/daily-survey-results/private-profiles-survey/

[12] http://www.zdnet.com/article/13-million-us-facebook-users-dont-change-privacy-settings/

[13] Source: Tor user metrics, which are available here: https://metrics.torproject.org/userstats-relay-country.html?graph=userstats-relay-country&start=2013-01-25&end=2015-04-25&country=all&events=off

[14] See for example http://thenextweb.com/facebook/2014/10/31/facebook-now-available-tor-network/

[15] T Hegghammer, 'Interpersonal Trust on Jihadi Internet Forums', *Norwegian Defence Research Establishment,* 19 February 2014, p.4.

[16] T Hegghammer, 'Interpersonal Trust on Jihadi Internet Forums', *Norwegian Defence Research Establishment,* 19 February 2014, p.5.

[17] D O'Callaghan et al., *Uncovering the Wider Structure of Extreme Right Communities Spanning Popular Online Networks*, 2013 .

[18] C King, D Leonard, *Beyond Hate: White Power and Popular Culture* (Ashgate Publishing: 2014), p.14.

[19] T Hegghammer, 'Interpersonal Trust on Jihadi Internet Forums', *Norwegian Defence Research Establishment,* 19 February 2014, p.1.

[20] E Goldberg Knox, 'The Slippery Slope of Material Support Prosecutions: Social Media Support to Terrorists', *Hastings Law Journal*, vol.66, no.1, 2014 ; J Farwell, 'The Media Strategy of ISIS', *Survival: Global Politics and Strategy*, vol.56, no.6, pp.49-55.

[21] V Andre, 'The Janus Face of New Media Propaganda: The Case of Patani Neojihadist Youtube Warfare and Its Islamophobic Effect on Cyber-Actors', *Islam and Christian-Muslim Relations*, vol.25, no.3, pp.335-356.

[22] M Vergani, D Zuev, 'Neojihadist Visual Politics: Comparing YouTube Videos of North Caucasus and Uyghur Militants', *Asian Studies Review*, vol.39, no.1, 2015, pp.1-22.

[23] M Ekman, 'The dark side of online activism: Swedish right-wing extremist video activism on Youtube', *Journal of media and communication research*, vol.56, pp.79-99.

[24] Jyette Klausen et al., 'The YouTube Jihadists: A Social Network Analysis of Al-Muhajiroun's Propaganda Campaign', *Perspectives on* Terrorism (Vol. 6, No. 1, 2012).

[25] S Bertram, K Ellison, 'Sub Saharan African Terrorist Group's use of the Internet', *The Centre for the Study of Terrorism and Political Violence: Journal of Terrorism Research*, vol.5, no.1, 2014.

[26] L Fekete, *Pedlars of hate: the violent impact of the European far right* (Institute of Race Relations: 2012)

[27] 'State Intelligence Agencies and the Far Right: A Review of developments in Germany, Hungary and Austria,' *Institute of Race Relations European Research Programme*, Briefing Paper No.6, 2013, p.1.

[28] E Saltman, G Hussain, *Jihad Trending: A Comprehensive Analysis of Online Extremism and How to Counter it* (Quilliam: 2014), p.44.

[29] G Weimann, *Terrorism in Cyberspace: The Next Generation* (), p.136 ; G Weimann, *New Terrorism and New Media* (Wilson Center: 2014), p.13.

[30] T Neer, M O'Toole, 'The Violence of the Islamic State of Syria (ISIS): A Behavioural Perspective', *Violence and Gender*, vol.4, no.1, pp.145-156.

[31] http://www.ft.com/cms/s/0/a0266d5e-489e-11e4-9d04-00144feab7de.html

[32]http://www.wiesenthal.com/site/apps/nlnet/content.aspx?c=lsKWLbPJLnF&b=8776547&ct=13928897

[33] J Berger, J Morgan*, The ISIS Twitter Census: Defining and Describing the population of ISIS supporters on Twitter*, (Brookings Institute: 2015).

[34] D O'Callaghan, Derek Greene, Maura Conway, Joe Carthy & Padraig Cunningham, 'Down the (White) Rabbit Hole: The Extreme Right and Online Recommender Systems', *Social Science Computer Review*, 2014, pp.1-20; *OSCE Background Note – Jihadist Use of the Internet: Lessons for the far right?* (Institute for Strategic Dialogue: 2014), p.2.

[35] J Berger, J Morgan*, The ISIS Twitter Census: Defining and Describing the population of ISIS supporters on Twitter*, (Brookings Institute: 2015).

[36] E Saltman, G Hussain, *Jihad Trending: A Comprehensive Analysis of Online Extremism and How to Counter it* (Quilliam: 2014), p.44.

[37] Ines von Behr et al., 'Radicalisation in the Digital Era', op. cit.

[38] Ibid.; Ines von Behr et al., 'Radicalisation in the Digital Era', op. cit.

[39] *The Daily Telegraph*, 'Britain's Involvement in Islamic Extremism', 20 August 2014, available at http://www.telegraph.co.uk/news/worldnews/middleeast/iraq/11045347/Britains-involvement-in-Islamic-extremism.html, accessed 07 September 2014; *The Daily Telegraph*, 'What every jihadi in Syria needs: hair gel, an iPad and Kit-Kats', 26 November 2013, available at http://www.telegraph.co.uk/news/worldnews/middleeast/syria/10476333/What-every-jihadi-in-Syria-needs-hair-gel-an-iPad-and-Kit-Kats.html, accessed 07 September 2014.

[40] G Weimann, *New Terrorism and New Media* (Wilson Center: 2014), p.14.

[41] Raffaello Pantucci (2015) *We Love Death as You Love Life* (Hurst & Co).

[42] Rita Katz and Margaret Foster, 'Al-Nusra Front: Breaking al-Qaeda's Forum Monopoly', op. cit.

[43] Evan Kohlmann, cited in Yuki Noguchi, 'Tracking Terrorists Online', *Washington Post* video report, April 19 2006, available at http://www.washingtonpost.com/wp-dyn/content/custom/2005/08/05/CU2005080501141.html?whichDay=1, available at 05 September 2014.

[44] J White Montoya, F Hofstetter, ISIL's *Utilization of Multimedia to Fulful Their Quest of Creating a New Islamic State*, 2014. Also see for example http://www.newsweek.com/twitter-shuts-down-2000-isis-linked-accounts-310969

[45] E Saltman, J Russsell, 'White Paper – The role of prevent in countering online extremism', (*Quilliam*: 2014), p.6

[46] Ali Fisher, 'How Jihadist Networks Maintain a Persistent Online Presence.' *Perspectives on Terrorism* 9.3 (2015).

[47] E Saltman, J Russsell, 'White Paper – The role of prevent in countering online extremism', (*Quilliam*: 2014), p.6.

[48] Rita Katz, Insite Blog on Extremism and Terrorism posting entitled 'IS: From Twitter.com to Friendica.eu', SITE Intelligence Group, 17 July 2014, available at http://news.siteintelgroup.com/blog/index.php/entry/210-is-from-twitter-com-to-friendica-eu, accessed 08 September 2014.

[49] Interview by the author with Shiraz Maher.

[50] FAQ section of BitChirp website, https://bitchirp.org/faq/

[51] Manuel R. Torres Soriano, 'The Vulnerabilities of Online Terrorism', *Studies in Conflict and Terrorism* (Vol. 35, No. 4, March 2012).

[52] Extract from Anders Breivik's manifesto, available at https://publicintelligence.net/wp-content/uploads/2011/07/AndersBehringBreivikManifesto_Page_0012.jpg, accessed 5 September 2014.

[53] Institute for National Security Studies, 'Backdoor Plots: The Darknet as a Field for Terrorism', Insight no. 464, September 10, 2014.

[54] We recently discovered a strong piece of evidence that demonstrates that Islamic State sympathisers are up to speed with the latest in counter-surveillance software. We found this on the text sharing board justpaste.it (a website that allows people to upload text and image documents anonymously) and appears to be the strongest evidence yet that demonstrates both a) how significant Islamic State sympathisers considered the internet to be as part of their struggle; and b) the extent to which they are aware of tools that allow them to mask their identity. See http://www.demos.co.uk/files/Islamic_State_and_Encryption.pdf?1426713922 for a full description.

[55] Michael Mimoso, 'Terror Group's Choice of Homegrown Crypto Likely Aids US Intelligence', *Threat Post*, 15 May 2014, available at http://threatpost.com/terror-groups-choice-of-homegrown-crypto-likely-aids-us-intelligence, accessed 04 September 2014.

[56]Ali Fisher and Nico Prucha, ‘The Call-Up: The Root of a Resilient and Persistent Jihadist Presence on Twitter’, *The Combatting Terrorism Exchange* (Vol. 4, No. 3, August 2014), available at https://globalecco.org/en_GB/the-call-up-the-roots-of-a-resilient-and-persistent-jihadist-presence-on-twitter, accessed 5 September 2014.

[57] Nico Prucha, ‘Online Territories of Terror – Utilizing the Internet for Jihadist Endeavors’, ORIENT IV (2011)

58 Unpublished research conducted by the authors.

[59] JM Berger and Jonathan Morgan ‘ISIS Twitter Census’, 2014. Available here: http://www.brookings.edu/~/media/research/files/papers/2015/03/isis-twitter-census-berger-morgan/isis_twitter_census_berger_morgan.pdf

[60] Ali Fisher and Nico Prucha (2014), ‘The Call-up: The Roots of a Resilient and Persistent Jihadist Presence on Twitter’. CTX, 4(3), 73-88.

[61] Charlie Wintour, ‘Islamic State Propaganda’, *Terrorism Monitor* Volume: 13 Issue: 12

[62] Ali Fisher and Nico Prucha, ‘Is this the most successful release of a jihadist video ever?’, Jihadica: Documenting the Global Jihad (website), available at http://www.jihadica.com/is-this-the-most-successful-release-of-a-jihadist-video-ever/, accessed 5 September 2014.

[63] Peter Neumann, ‘Options and Strategies for Countering Online Radicalization in the United States’, op. cit.; Klausen et. al, ‘“Open Source Jihad”’, op. cit.

[64] Ali Fisher, ‘Eye of the Swarm, : The Rise of ISIS and the Media Mujahedeen’ USC Centre on Public Diplomacy web blog, 8 08 July 2014, available at http://uscpublicdiplomacy.org/blog/eye-swarm-rise-isis-and-media-mujahedeen, accessed 05 September 2014

65 JM Berger and Jonathan Morgan ‘ISIS Twitter Census’, 2014. Available here: http://www.brookings.edu/~/media/research/files/papers/2015/03/isis-twitter-census-berger-morgan/isis_twitter_census_berger_morgan.pdf

66John Curtis Amble, ‘Combatting Terrorism in the New Media Environment’, *Studies in Conflict and Terrorism* (Vol.35, No. 5, April 2012)

[67] Kori Schake, ‘@ISIS Is #Winning’, *Foreign Policy*, 9 July 2014, available at http://www.foreignpolicy.com/articles/2014/07/09/isis_is_winning_social_media_hashtag_diplomacy, accessed 05 September 2014

[68] Unpublished research by the authors.

[69] SITE Intelligence Group, ‘Islamic State Releases Video on Diaspora Showing Beheading of U.S. Journalist James Foley’, 19 August 2014, available at http://news.siteintelgroup.com/blog/index.php/entry/236-islamic-state-releases-video-showing-beheading-of-u-s-journalist-james-foley,-threatens-to-kill-another-prisoner, accessed 07 September 2014

[70] Rita Katz, Insite Blog on Extremism and Terrorism posting entitled ‘IS: From Twitter.com to Friendica.eu’, SITE Intelligence Group, 17 July 2014, available at http://news.siteintelgroup.com/blog/index.php/entry/210-is-from-twitter-com-to-friendica-eu, accessed 08 September 2014.

[71] SMapp Lab Data Report, Ukraine Protests 2013-2014: Preliminary Results,’ *Social Media and Political Participation Lab, New York University*, 2014.

[72] A Zubiaga, M Liakata, R Procter, K Bontcheva, P Tolmie, ‘Towards Detecting Rumours in Social Media’, 2015, (http://arxiv.org/abs/1504.04712).

[73] See: http://news.sky.com/story/1496511/big-increase-in-facebook-and-twitter-crimes

[74] See: http://news.npcc.police.uk/releases/chief-constable-stephen-kavanagh-we-have-to-think-digital

[75] http://www.quilliamfoundation.org/wp/wp-content/uploads/publications/free/white-paper-the-role-of-prevent-in-countering-online-extremism.pdf

[76] http://www.latimes.com/nation/nationnow/la-na-state-department-islamic-social-media-20140906-story.html ; http://time.com/3387065/isis-twitter-war-state-department/

[77] http://time.com/3387065/isis-twitter-war-state-department/

[78] Interview with Dr. Erin Saltman, ISD.

[79] M Mazzetti, 'ISIS is winning social media war, US concludes', New York Times, June 12 2015 Available here: http://www.nytimes.com/2015/06/13/world/middleeast/isis-is-winning-message-war-us-concludes.html?_r=1

[80] T Gemmerli, 'Campaigns'. Available: http://www.diis.dk/en/research/campaigns-targeting-extremism-on-the-internet-no-documented-effect

[81] https://www.facebook.com/notes/facebook-safety/controversial-harmful-and-hateful-speech-on-facebook/574430655911054

[82] https://cyber.law.harvard.edu/events/luncheon/2014/03/benesch

[83] https://www.eurosint.eu/system/files/employing_social_media_monitoring_tools_as_an_osint_platform_for_intelligence_defence_security.pdf

[84] Stephen C. Mercado, CIA Center for the Study of Intelligence, Studies in Intelligence, Vol. 48

No.3, https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/csistudies/studies/vol48no3/article05.html Also Lynda Peters 2012 'utilising osical media to further the national suspicious activity reporting initiative'.

[85] Eg Robert David Steele's 'Open Source Everything Manifesto', 2014. http://www.phibetaiota.net/2014/05/robert-steele-at-libtechnyc-the-open-source-everything-manifesto/

[86] C Miller, 'Social Action on Social Media', *Nesta Working Paper, Working Paper Series* (Nesta: 2015)

[87] https://developers.facebook.com/docs/chat

[88] F Vis, 'A critical reflection on Big Data: Considering APIs, researchers and tools', *First Monday*, vo.18, no.10, October 2013

[89] https://investor.twitterinc.com/releasedetail.cfm?releaseid=909177 ;

[90] N Ljubesic, 'Discriminating Between Closely Related Languages on Twitter', *Informatica*, No.39, 2015, pp.1-8

[91] http://mashable.com/2013/11/20/twitter-users-countries/

[92] http://www.computerworld.com/article/2489798/social-media/twitter-user-base-to-grow-by--24-4--in-2014--researcher-says.html

[93] http://fortune.com/2014/04/16/social-data-as-a-business-is-dead-long-live-big-data-services/ ; http://blog.datasift.com/2015/04/13/twitter-datasift-gap/

[94] http://datasift.com/platform/datasources/

[95] M Chau & J Xu, 'Mining communities and their relationships in blogs: A study of online hate groups' *International Journal of Human-Computer Studies*, p62.

[96] Mining Social Media: Tracking Content and Predicting Behavior, Manos Tsagkias. Ph.D. thesis, University of Amsterdam 20.

[97] Stuart Shulman 'Keeping Humans in the Machine Learning Loop' (March, 28, 2013). *Paper presented to the social text workshop (closed)*, University of Birmingham, 28 March 2013.

[98] S Liu, P Cui, H Luan, W Zhu, S Yang, Q Tian, 'Social-oriented visual image search', *Computer Vision and Image Understanding*, vol.118, 2014, pp.30-39 ; M Tsai, C Aggarwal, T Huang, 'Ranking in heterogeneous social media', *Proceedings of WSDM 2014, 7th Annual International Conference on Web Search and Data Mining, 2014,* pp.613-622

[99] C Cao, J Caverlee, 'Detecting Spam URLs in Social Media via Behavioural Analysis', *Department of Computer Science and Engineering, Texas A&M University* (Springer: 2015), pp.703-714

[100] R Oentaryo, J Low, E Lim, 'Chalk and Cheese in Twitter: Discriminating Personal and Organization Accounts', *European Conference on Information Retrieval (ECIR)*, 2015.

[101] K O'Hara., T Berner-Lee, W Hall& N Shadbolt (2011) "The use of the semantic web in eResearch", in Dutton, W & Jeffreys, P (ed.) *World Wide Research: Reshaping the Sciences and Humanities*

Mining Social Media: Tracking Content and Predicting Behavior, Manos Tsagkias. Ph.D.

thesis, University of Amsterdam.

[102] M Sykora, T Jackson, A O'Brien, S Elayan, 'Emotive: Extracting the Meaning of Terse Information in a Visualisation of Emotion', *National Security and Social Media Monitoring, EISIC* 2013 ; J Bartlett, J Reffin, N Rumball, S Williamson, *Anti-social Media*, 2014.

[103] A Coden, D Gruhl, N Lewis, P Mendes, M Nagarajan, C Ramakrishnan, S Welch, 'Semantic Lexicon Expansion for Concept-based Aspect-aware sentiment analysis' *IBM Research USA,* 2014

[104] P Burnap, M Williams, 'Hate Speech, Machine Classification and Staticistlca Modlling of Information Flows on Twitter: Interpretation and Communication for Policy Decision Making', *Cardiff School of Computer Science and Informatics, and Cardiff School of Social Sciences*, 2014.

[105] J Barlett, J Reffin, N Rumball, S Williamson, *Anti-social Media* (Demos: 2014).

[106] H Gitari, Z Zuping, H Damien, J Long, 'A Lexicon-based Approach for Hate Speech Detection', *International Journal of Multimedia and Ubiquitous Engineering*, vol.10, no.4 (2015), pp.215-230.

[107] J Scanlon, M Gerber, 'Automatic detection of cyber-recruitment by violent extremists', *Security Informatics*, no.3, vol.5, 2014 ; J Parapar, D Losada, A Barreiro, 'Combining Psycho-linguistics, Content-based and Chat-based Features to Detect Predation in Chatrooms', Journal of Universal Computer Science, vol. 20, no.2 (2014), pp.213-239.

[108] http://www.ncrm.ac.uk/RMF2014/programme/session.php?id=E8 ; http://nsmnss.blogspot.co.uk/2014/02/new-social-media-new-social-science-and.html ; http://www.natcen.ac.uk/news-media/press-releases/2014/october/natcen-publishes-innovative-book-of-blogs-on-social-media-research/

[109] J Bartlett, C Miller, D Weir, J Reffin, S Wibberly, *Vox Digitas* (Demos: 2014)

[110] http://www.pewresearch.org/2013/03/04/twitter-reaction-to-events-often-at-odds-with-overall-public-opinion/

[111] http://www.nesta.org.uk/blog/introducing-political-futures-tracker

[112] Eg Li et al. (2014) on supervised user-item based topic models for sentiment analysis.(http://www.aaai.org/ocs/index.php/AAAI/AAAI14/paper/view/8663/8620); S Finn, E Mustafaraj (2012) on distinguishing 'opinion-makers' from 'opinion holders' on social media for political opinion analysis (http://www.aaai.org/ocs/index.php/AAAI/AAAI14/paper/view/8663/8620).

[113] L Sloan, J Morgan, P Burnap, M Williams, ‘Who Tweets? Deriving the Demographic Characteristics of Age, Occupation and Social Class from Twitter User Meta-Data’, *PLoS ONE*, vol.10, no.3, 2015.

T Ugheoke, ‘Detecting the Gender of a Tweet Sender’, *Master of Science thesis, Department of Computer Science University of Regina* (2014).

[115] A Kokkos, T Tzou,’A robust gender inference model for online social networks and its application to Linkedin’, *First Monday*, vol.19, no.9, 2014.

[116] R Compton, C Lee, J Xu, L Artieda-Moncada, T Lu, L De Silva, M Macy, ‘Using publically visible social media to build detailed forecasts of civil unrest’, *Security Informatics*, vol.4 no.3, 2014.

[117] B Han, P Cook, T Baldwin, ‘Geolocation Prediction in Social Media Data by Finding Location Indicative Words’, Proceedings of COLING 2012 ; B Han, P Cook, T Baldwin, ‘Text-based Twitter User Geolocation Prediction’, *Journal of Artificial Intelligence Research*, no.49, 2014.

[118] M Dredze, M Paul, H Tran and S Bergsma, ‘Carmen: A Twitter Geolocation System with Applications to Public Health’, *Expanding the Boundaries of Health Informatics Using Artificial Intelligence AAAI,* 2013.

[119] M Kosinski, D Stillwell, T Graepel, ‘Private traits and Attributes are predictable from digital records of human behaviour’, *Proceedings of the National Academy of Sciences of the United States of America*, vol.110, no.15, 2013 pp.5802-5805.

[120] V Benjamin, W Chung, A Abbasi, J Chuang, C Larson, H Chen, ‘Evaluating text visualization for authorship analysis’, *Security Informatics*, vol.3 no.10 (2014).

[121] P Davis, W Perry, R Brown, D Yeung, P Roshan, P Voorhies, *Using Behavioural Indicator to Help Detect Potential Violent Acts: A review of the science base*, (RAND: 2013).

[122] Burnap et al., ‘Detecting tension in online Twitter communities with computational Twitter analysis’, Technological Forecasting and Social Change, Vol. 95, June 2015, pp. 96-108 (http://www.sciencedirect.com/science/article/pii/S0040162513000899).

[123] J Parapar, D Losada, A Barreiro, ‘Combining Psycho-linguistic, Content-based and Chat-based

Features to Detect Predation in Chatrooms,’ *Journal of Universal Computer Science*, vol. 20, no. 2 (2014), 213-239.

[124] K Cohen, F Johansson, L Kaati, J Mork, ‘Detecting Linguistic Markers for Radical Violence in Social Media’, Terrorism and Political Violence, vol.26, no1, pp.245-256.

[125] M Banerveld, N Le-Khac, M Kechadi, ‘Performance Evaluation of a Natural Language Processing Approach Applied in White Collar Crime Investigation’ *Future Data and Security Engineering, Lecture Notes in Computer Science*, Vol. 8860, 2014 ; W Dong, S Liao, B Fang, X Cheng, Z Chen, ‘The Detection of Fraudulent Financial Statements: An integrated language model’, *Pacific Asia Conference on Information Systems*, 2014.

[126] Xuning Tang, Christopher C Yang , TUT: a statistical model for detecting trends, topics and user interests in social media, International Conference on Information and Knowledge Management 2012

[127] A review of Facebook research in the social sciences, Perspectives on Psychological Sceince 7(3) 203-220

[128] http://motherboard.vice.com/blog/revolution-is-ukraines-most-liked-facebook-page ; http://techpresident.com/news/wegov/24790/how-EuroMaidan-play-out-online

[129] R Yu, X He, Y Liu, ‘GLAD: Group Anomaly Detection in Social Media Analysis,’ *Department of Computer Science, University of Southern California,* 2014.

[130] T Sakaki, M Okazaki & Y Matsuo 'Earthquake Shakes Twitter Users: Real-Time Event Detection By Social Sensors', International Conference on the World Wide Web, 2010

[131] Fabian Abel, Claudia Hauff, Geert-Jan Houben, Richard Stronkman and Ke Tao, Fighting Fire with Information from Social Web Streams, International Conference on the World Wide Web, 2012.

[132] R McCreadie, K Kappler, M Kardara, A Kaltenbrunner, C Macdonald, J Soldatos, I Ounis, 'SUPER: Towards the use of Social Sensors for Security Assessments and Proactive Management of Emergencies', *International World Wide Web Conference Committee (IW3C2)*, 2015.

[133] Learning similarity metrics for event identification in social media, WSDM Conference, Hila Becker, Mor Naaman and Luis Gravano, 2010.

[134] Mendoza, M., Poblete, B., Castillo, C , Twitter Under Crisis: Can we Trust What we RT?, KDD Workshop on Social Media Analytics, 2010.

[135] C Castillo, M Mendoza, & P Poblete (2011). Information credibility on twitter. Proceedings of the 20th international conference on World wide web - WWW '11, 675. doi:10.1145/1963405.1963500

[136] A Hughes "Twitter adoption and use in mass convergence and emergency events" IGCRAM, 2009

[137] M Burke (2009) Feed me: motivating newcomer contribution in social network sites in Proceedings of the 27th Conference on human factors in computing systems.

[138] K Starbird, J Maddock, M Orand, P Acheterman, R Mason, *Rumors, False Flags, and Digital Vigilantes: Misinformation on Twitter after the 2013 Boston Marathon Bombing*, iConference 2014.

[139] A Zubiaga, M Liakata, R Procter, K Bontcheva, P Tolmie, 'Towards Detecting Rumours in Social Media', Association for the Advancement of Artificial Intelligence, 2015 ; http://www.sheffield.ac.uk/news/nr/lie-detector-social-media-sheffield-twitter-facebook-1.354715

[140] See for example: http://www.poynter.org/news/mediawire/165654/visualized-incorrect-information-travels-farther-faster-on-twitter-than-corrections/

[141] S Muthiah, B Huang, J Arredondo, D Mares, L Getoor, G Katz, N Ramakrishnan, 'Planned Protest Modelling in News and Social Media', *Association for the Advancement of Artificial Intelligence*, 2015.

[142] http://fp7-steer.eu/technologies/adaptive-event-profiler/

[143] http://cordis.europa.eu/result/rcn/153278_en.html

[144] http://www.westyorkshire.police.uk/athena

[145] http://www.fp7-emergent.eu/

[146] http://slandail.eu/project/

[147] http://www.bbc.co.uk/news/blogs-trending-30479306

[148] http://www.theguardian.com/media/2015/jan/11/charlie-hebdo-social-media-news-readers

[149] http://cips.uottawa.ca/twitter-and-the-ottawa-attacks/

[150] Burnap et al., 'Tweeting the terror: modelling the social media reaction to the Woolwich terrorist attack', Social Network Analysis and Mining, 2014. (http://download.springer.com/static/pdf/6/art%253A10.1007%252Fs13278-014-0206-4.pdf?originUrl=http%3A%2F%2Flink.springer.com%2Farticle%2F10.1007%2Fs13278-014-0206-4&token2=exp=1434030198~acl=%2Fstatic%2Fpdf%2F6%2Fart%25253A10.1007%25252Fs13278-014-0206-

4.pdf%3ForiginUrl%3Dhttp%253A%252F%252Flink.springer.com%252Farticle%252F10.1007%252Fs13278-014-0206-4*~hmac=9b7cad34cc65cfe8a4b35f02845f7d7b80a8ad4c086e551abae378891e281433)

[151] H Schoen, D Gayo-Avello, P Metaxas, E Mustafaraj, M Strohmaier, 'The Power of Prediction with Social Media', Internet Research, Vol. 23 Iss: 5 (2013), pp.528- 543.

[152] S Nann, J Krauss, D Schoder, 'Predictive Analytics on Public Data – The Case of Stock Markets', *Proceedings of the 21st European Conference on Information Systems*, 2013 ; F Wong, S Sen, M Chiang, 'Why Watching Movie Tweets Won't Tell the Whole Story?', *Princeton* 2013. (http://arxiv.org/abs/1203.4642)

[153] J Bartlett, J Birdwell, L Reynolds, *Like, Share, Vote* (Demos: 2014).

[154] J DiGrazia, 'More Tweets, More Votes: Social Media as a Quantitative Indicator of Political Behaviour'

[155] Nick Beauchamp (September 2013), 'Predicting and Interpolating State-level Polling using Twitter Textual Data'

[156] Ibid.

[157] http://www.nature.com/news/when-google-got-flu-wrong-1.12413; and http://www.forbes.com/sites/stevensalzberg/2014/03/23/why-google-flu-is-a-failure/

[158] A Aslam, M Tsou, B Spitzberg, L An, M Gawron, D Gupta, K Peddecord, A Nagel, C Allen, J Yang, S Lindsay, 'The Reliability of Tweets as a Supplementary Method of Seasonal Influenza Surveillance', *Journal of Medical Internet Research*, vol.16, no.11, 2014.

[159] http://newsroom.melbourne.edu/news/twitter-can-predict-hot-spots-coronary-heart-disease

[160] http://www.newsweek.com/twitter-predict-emergency-room-rush-hours-323360

[161] M Dredze, M Paul, H Tran and S Bergsma, 'Carmen: A Twitter Geolocation System with Applications to Public Health', *Expanding the Boundaries of Health Informatics Using Artificial Intelligence AAAI,* 2013.

[162] Wang et al (2012) *Automatic Crime Prediction Using Events Extracted from Twitter* Posts

[163] M Gerber, 'Predicting Crime Using Twitter and Kernel Density Estimation', *Decision Support Systems*, 2014.

[164] http://urgentcomm.com/intrado/intrado-steve-reed-talks-about-integrating-beware-tool-motorola-solutions-idp-platform ; http://abc7chicago.com/news/new-dragnet-cops-scouring-social-media-for-red-flags/525690/

[165] http://abc7chicago.com/news/new-dragnet-cops-scouring-social-media-for-red-flags/525690/

[166] N Silver (2012) *The Signal and The Noise*, Penguin: London, p.452.

[167] M Kosinskia, D Stillwella & T Graepelb (2013) 'Private traits and attributes are predictable from digital records of human behavior', *Proceedings of the National Academy of Science*, (submission)

[168] Silver, *The Signal and The Noise*, p.452.

[169] Willis, Alistair, Ali Fisher, and Ilia Lvov, 'Mapping networks of influence: Tracking Twitter conversations through time and space' (2015).

[170] Dunbar, R. I. M., et al. 'The structure of online social networks mirrors those in the offline world.' *Social Networks* 43 (2015): 39-47.

[171] A. Sutcliffe, RIM Dunbar, J Binder, H Arrow, 'Relationships and the social brain: integrating psychological and evolutionary perspectives', *Br. J. Psychol.*, 103 (2) (2012), pp. 149–168.

R.I.M. Dunbar, 'Evolutionary basis of the social brain', in J Decety, J Cacioppo (Eds.), *Oxford Handbook of Social Neuroscience*, Oxford University Press, Oxford (2011), pp. 28–38.

[172] Lynch, Marc, Deen Freelon, and Sean Aday. ‘Syria’s socially mediated civil war’ *United States Institute Of Peace* 91.1 (2014): 1-35.

[173] Boyd, Danah, and Nicole B. Ellison. ‘Social network sites: Definition, history, and scholarship in Journal of Computer-Mediated Communication, 13 (1), article 11’ (2007).

Newman, Nic, William H. Dutton and Grant Blank, ‘Social media in the changing ecology of news: the fourth and fifth estates in Britain’, *International Journal of Internet Science*, 7(1), 2012, pp. 6- 22

Xu, Weiai Wayne, Yoonmo Sang, Stacy Blasiola and Han Woo Park, ‘Predicting opinion leaders in Twitter activism networks: the case of the Wisconsin recall election’, *American Behavioral Scientist*, 2014.

[174] Shirky, Clay, ‘The political power of social media: technology, the public sphere, and political change’, *Foreign Affairs*, 90(1), 2011, pp. 28-41.

[175] M Mäkinen and MW Kuira, ‘Social media and postelection crisis in Kenya’, *The International Journal of Press/Politics*, 13(3), 2008, pp. 328-335.

Metzgar, E. and A. Maruggi, ‘Social media and the 2008 US presidential election’, *Journal of New Communications Research*, 4(1), 2009, pp. 141-165

[176] K Leetaru (2011) ‘Culturomics 2.0: Forecasting large-scale human behavior using global news media tone in time and space’, *First Monday*, 16(9), [WWW document] URL http://firstmonday.org/article/view/3663/3040 [accessed 5 May 2015].

[177] Otto N Larsen and Richard J. Hill, ‘Mass media and interpersonal communication in the diffusion of a news event’, *American Sociological Review*, 19, 1954, pp. 426-433.

[178] MS Granovette, ‘The strength of weak ties’, *American Journal of Sociology*, 78(6), 1973, pp.1360– 1380.

[179] DZ Levin and R Cross, ‘The strength of weak ties you can trust: the mediating role of trust in effective knowledge transfer’, *Management Science*, 50(11), 2004, pp.1477–1490.

[180] S Milgram, ‘The small world problem’, *Psychology Today*, 2(1), 1967, pp.60–67

[181] Otto N Larsen and Richard J Hill, ‘Mass media and interpersonal communication in the diffusion of a news event’, *American Sociological Review*, 19, 1954, pp. 426-433.

[182] EM Rogers, ‘Diffusion of news of the September 11 terrorist attacks’, in AM Noll (ed.), *Crisis communications: Lessons from September 11*, Lanham, MD: Rowman & Littlefield, 2003, pp. 17–30.

[183] Borgatti, Stephen P., and Daniel S. Halgin. "On network theory." *Organization Science* 22.5 (2011): 1168-1181

[184]Axel Bruns (2007) ‘Methodologies for Mapping the Political Blogosphere: An Exploration Using the IssueCrawler Research Tool’. *First Monday* 12(5). Zachary Devereaux, Wendy Cukier, Peter Ryan, and Neil Thomlinson, ‘"Using the Issue Crawler to Map Gun Control Issue Networks’ (2009); H Moe, ‘Mapping the Norwegian blogosphere: Methodological challenges in internationalizing internet research' *Social Science Computer Review* 29.3 (2011): 313- 326. Unmasking the Arzeshi, “Iran’s Conservative Cyber-Activists and the 2013 Presidential Election“, Small Media (2014) http://www.unmaskthearzeshi.com/

[185] As discussed later, on many social media platforms, calls to the API (application programming interface) are frequently more effective for gathering structured data than crawlers.

[186]Malcolm K Sparrow, ‘Network vulnerabilities and strategic intelligence in law enforcement’ *International Journal of Intelligence and Counter Intelligence* 5.3 (1991): 255- 274.

[187] Valdis Krebs, ‘Uncloaking Terrorist Networks’ *First Monday* [Online], Volume 7 Number 4 (1 April 2002). Valdis Krebs, ‘Mapping Networks of Terrorist Cells’ *CONNECTIONS* 24(3): 43-52 2002. Richard Medina and

George Hepner 'Advancing the Understanding of Sociospatial Dependencies in Terrorist Networks', *Transactions in GIS*, vol. 15, no 5, 2011.

[188] J Weng, E Lim & J Jiang (2010) 'TwitterRank: Finding Topic Sensitive Influential Twitterers' *WSDM* 10, February 2010.

[189] Berger, J & Strathearn, B (2013) *Who Matters Online*, ICSR: London.

[190] Ibid.

[191] D O'Callaghan, D Greene, M Conway, J Carthy, P Cunningham,(2013) 'An analysis of interactions within and between extreme right communities in social media'. Available here: http://arxiv.org/abs/1206.7050

[192] R Xiang, J Neville & M Rogati (2010) 'Relationship Strength In Online Social Networks', *International Conference on the World Wide Web*.

[193] RE Wilson, SD Gosling, & LT Graham (2012), 'A review of Facebook research in the social sciences', *Perspectives on Psychological Science*, 7(3), 203-220. doi: 10.1177/1745691612442904; M Tsagkias (2012) *Mining Social Media: Tracking Content and Predicting Behavior*, Ph.D. thesis, University of Amsterdam, p10.

[194] Positive and Negative Links in online Social Networks, International Conference on the World Wide Web, Jure Leskovec, Daniel Huttenlocher, Jon Kleinberg, 2010.

[195] All blogs were hosted on xanga.com.

[196] M Chau & J Xu, 'Mining communities and their relationships in blogs: A study of online hate groups' *International Journal of Human-Computer Studies*, p67.

[197] This coverage on the BBC *Newsnight* program provides a tangible example of the way SNA can be used to gain an overview of interactions via social media. http://bit.ly/NNnetVid

[198] Ali Fisher, Nico Prucha, 'The Call-up: The Roots of a Resilient and Persistent Jihadist Presence on Twitter', *CTX* (2014, August): 4(3), 73-8.

[199] The 66 accounts were recommended on a jihadist forum rather than selected by the authors.

[200] Rob Cross, Andrew Parker and Stephen P. Borgatti, 'A bird's-eye view: Using social network analysis to improve knowledge creation and sharing', *IBM Institute for Business Value*, http://www.analytictech.com/borgatti/papers/cross,%20parker%20and%20borgatti%20-%20A_birds_eye_view.pdf

[201] Otto N Larsen, and Richard J. Hill, 'Mass media and interpersonal communication in the diffusion of a news event', *American Sociological Review*, 19, 1954, pp. 426-433.

[202] Full discussion of this content distribution can be found on Jihadica: http://bit.ly/19KMPMX

[203] Node centrality in weighted networks: Generalizing degree and shortest paths: http://toreopsahl.com/2010/04/21/article-node-centrality-in-weighted-networksgeneralizing-degree-and-shortest-paths/

[204]Nico Prucha and Ali Fisher, 'Tweeting for the caliphate: Twitter as the new frontier for jihadist propaganda.' *CTC Sentinel* 6.6 (2013): 19-23.

[205] Ali Fisher, 'How Jihadist Networks Maintain a Persistent Online Presence.'*Perspectives on Terrorism* 9.3 (2015).

[206]TW Va;emte et al., 'How correlated are network centrality measures?', *Connections* (Toronto, Ont.), 28(1), 2008, pp.16–26.

[207] For a full description of this method see: Alistair Willis, Ali Fisher, and Ilia Lvov, 'Mapping networks of influence: Tracking Twitter conversations through time and space.' (2015).

[208] Fisher, A *Everybody's getting hooked up; building innovative strategies in the era of big data*, 2012 p.49

[209] Interview, Alberto Nardelli, founder *Tweetminster*

[210] Manos Tsagkias, 'Mining Social Media: Tracking Content and Predicting Behavior', Ph.D. thesis, University of Amsterdam, 2012.

[211] Ibid., p.54

[212] J Bartlett et al (2013, forthcoming) Politics by Twitter? *Understanding public attitudes using social media*, Demos

[213] M Sparrow, 'Network vulnerabilities and strategic intelligence in law enforcement' *International Journal of Intelligence and Counter Intelligence* 5.3 (1991): 255- 274.

[214] P Klerks (2001) 'The network paradigm applied to criminal organizations: Theoretical nitpicking or a relevant doctrine for investigators? Recent developments in the Netherlands' *Connections* 24.3 (2001): 53-65.

[215] Stephen P. Borgatti, Daniel S. Halgin, *On Network Theory* (2011) http://www.steveborgatti.com/papers/orsc.1110.0641.pdf

[216] T Hegghammer, 'Interpersonal Trust on Jihadi Internet Forums', *Norwegian Defence Research Establishment,* 19 February 2014, p.1

[217] http://www.fastcompany.com/3041479/drugs-guns-and-selfies-gangs-on-social-media

[218] http://news.gc.ca/web/article-en.do?mthd=index&crtr.page=1&nid=852589

[219] http://www.bbc.co.uk/newsbeat/article/29010110/tube-bomb-threat-is-a-social-media-hoax-say-police ; http://www.shropshirestar.com/news/2015/03/17/police-dismiss-online-telford-crime-rumour-over-men-in-van/

[220] *Using Social Media in Emergencies: Smart Practices, smart tips for category 1 responders using social media in emergency management* (DSTL: 2012)

[221] http://www.forbes.com/sites/tarunwadhwa/2013/04/22/lessons-from-crowdsourcing-the-boston-marathon-bombings-investigation/

[222] http://www.newstatesman.com/world-affairs/2013/04/4chan-plays-racist-wheres-wally-find-boston-bomber

[223] R McCreadie, C Macdonald, I Ounis, 'Crowdsourced Rumour Identification During Emergencies', International World Wide Web Conference Committee, 2015

224 Merve Nazliogli, '5 examples of how brands are using co-creation' (October 2013) http://www.visioncritical.com/blog/5-examples-how-brands-are-using-co-creation (accessed 26/11/14). Taken from Ipsos-Mori, (2015) 'Unlocking the Value of Social Media Research'.

[225] Criminal Code, section 183, http://laws.justice.gc.ca/eng/acts/C-46/page-87.html#h-61 (last accessed 27.03.2013). According to section 183 of the Code, 'private communications' are defined as: "any oral communication, or any telecommunication, that is made by an originator who is in Canada or is intended by the originator to be received by a person who is in Canada and that is made under circumstances in which it is reasonable for the originator to expect that it will not be intercepted by any person other than the person intended by the originator to receive it, and includes any radio-based telephone communication that is treated electronically or otherwise for the purpose of preventing intelligible reception by any person other than the person intended by the originator to receive it."

[226] See *Katz v. United States* - 389 US 347 (1967), http://supreme.justia.com/cases/federal/us/389/347/case.html (last accessed 27.03.2013)

[227] Supreme Court of the United States, *Opinion on United States v. Antoine Jones*, 2011, http://www.supremecourt.gov/opinions/11pdf/10-1259.pdf (last accessed 27.03.2013)

[228] C Borchet, F Pinguelo, D Thaw, 'Reasonable expectations of privacy settings: social media and the Stored Communications Act', *Duke Law and Technology Review,* Vol. 13, No. 1, 2015. (http://scholarship.law.duke.edu/cgi/viewcontent.cgi?article=1270&context=dltr)

[229] McCann Truth Central, *The Truth about Privacy: Canada and Beyond*, 2012, http://www.adstandards.com/en/MediaAndEvents/TruthAboutPrivacy.pdf (last accessed 27.03.2013); Toronto Sun, *Canadians oppose internet spy law: poll*, 2011, http://www.torontosun.com/2011/08/25/canadians-oppose-internet-spy-law-poll (last accessed 28.03.2013); J Bartlett, *The Data Dialogue*, Demos, 2010. This is based on a representative population level poll of circa 5,000 people. Also see J Bartlett & C Miller, *Demos CASM submission to the Joint Committee on the Draft Communications Data Bill*, Demos, 2012.

[230] https://www.cigionline.org/internet-survey

[231] McCann Truth Central, *The Truth about Privacy: Canada and Beyond*, 2012, http://www.adstandards.com/en/MediaAndEvents/TruthAboutPrivacy.pdf (last accessed 27.03.2013).

[232] https://www.priv.gc.ca/information/por-rop/2013/por_2013_01_e.pdf

[233] Toronto Sun*, Canadians oppose internet spy law: poll*, 2011, http://www.torontosun.com/2011/08/25/canadians-oppose-internet-spy-law-poll (last accessed 28.03.2013)

[234]http://www.thestar.com/news/canada/2014/11/24/canadians_growing_concerned_over_internet_privacy_poll_shows.html

[235] Accenture (2012) *Are Police Maximizing Technology to Fight Crime and Engage Citizens?* http://www.accenture.com/policecitizensurvey (Accessed 27.03.2013)

[236] Pheoniz Strategic Perspective, Survey of Canadians on Privacy-Related Issues, January 2013, p. 20. https://www.priv.gc.ca/information/por-rop/2013/por_2013_01_e.pdf

[237] See http://www.pre.ethics.gc.ca/eng/policy-politique/initiatives/tcps2-eptc2/chapter1-chapitre1/ (accessed 7 July 2015).

[238] AoIR, 'Ethical Decision-Making and Internet Research: Version 2.0' (September 2012) https://cms.bsu.edu/sitecore/shell//-/media/WWW/DepartmentalContent/ResearchIntegrity/Files/Education/Active/AoIR%20Social%20Media%20Working%20Committee.pdf (accessed 26/11/14).

[239] ESOMAR, 'Guideline on social media research' (July 2011) http://www.esomar.org/uploads/public/knowledge-and-standards/codes-and-guidelines/ESOMAR-Guideline-on-Social-Media-Research.pdf (accessed 26/11/14).

[240] CASRO, 'Social Media Research Guidelines' (October 2011) http://c.ymcdn.com/sites/www.casro.org/resource/resmgr/docs/social_media_research_guidel.pdf (accessed 26/11/14).

[241] ESOMAR, 'Guideline on social media research' (July 2011) http://www.esomar.org/uploads/public/knowledge-and-standards/codes-and-guidelines/ESOMAR-Guideline-on-Social-Media-Research.pdf (accessed 26/11/14) and AAPOR , 'Social Media in Public Opinion Research (May 2014), https://www.aapor.org/AAPORKentico/AAPOR_Main/media/MainSiteFiles/AAPOR_Social_Media_Report_FNL.pdf (accessed 28/05/15),

[242] AoIR, *Ethical Decision-Making and internet Research: Recommendations from the AoIR Ethics Working Committee (Version 2.0)*, 2012, p2

[243] *Ibid*.

[244] British Psychological Society, *Conducting Research on the internet: Guidelines for ethical practice in psychological research online*, 2007.

[245] ESRC, Framework for Research Ethics, latest version: September 2012, http://www.esrc.ac.uk/_images/Framework-for-Research-Ethics_tcm8-4586.pdf (Accessed 27.03.2013). This is no longer the latest version. There are numerous small changes in the 2015 version, but also a significant one - "Harm to research participants and researchers must be avoided in all instances" (2012) becomes "Research should be worthwhile and provide value that outweighs any risk or harm. Researchers should aim to maximise the benefit of the research and minimise potential risk of harm to participants and researchers. All potential risk and harm should be mitigated by robust precautions" (2015). (http://www.esrc.ac.uk/_images/framework-for-research-ethics_tcm8-33470.pdf).

[246] MRS, ‘Code of Conduct 2014’ (September 2014), p.13 .https://www.mrs.org.uk/pdf/mrs%20code%20of%20conduct%202014.pdf (accessed 26/11/14).

Demos – Licence to Publish

The work (as defined below) is provided under the terms of this licence ('licence'). The work is protected by copyright and/or other applicable law. Any use of the work other than as authorized under this licence is prohibited. By exercising any rights to the work provided here, you accept and agree to be bound by the terms of this licence. Demos grants you the rights contained here in consideration of your acceptance of such terms and conditions.

1 Definitions

a 'Collective Work' means a work, such as a periodical issue, anthology or encyclopedia, in which the Work in its entirety in unmodified form, along with a number of other contributions, constituting separate and independent works in themselves, are assembled into a collective whole. A work that constitutes a Collective Work will not be considered a Derivative Work (as defined below) for the purposes of this Licence.

b 'Derivative Work' means a work based upon the Work or upon the Work and other pre-existing works, such as a musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which the Work may be recast, transformed, or adapted, except that a work that constitutes a Collective Work or a translation from English into another language will not be considered a Derivative Work for the purpose of this Licence.

c 'Licensor' means the individual or entity that offers the Work under the terms of this Licence.

d 'Original Author' means the individual or entity who created the Work.

e 'Work' means the copyrightable work of authorship offered under the terms of this Licence.

f 'You' means an individual or entity exercising rights under this Licence who has not previously violated the terms of this Licence with respect to the Work,or who has received express permission from Demos to exercise rights under this Licence despite a previous violation.

2 Fair Use Rights

Nothing in this licence is intended to reduce, limit, or restrict any rights arising from fair use, first sale or other limitations on the exclusive rights of the copyright owner under copyright law or other applicable laws.

3 Licence Grant

Subject to the terms and conditions of this Licence, Licensor hereby grants You a worldwide, royalty-free, non-exclusive,perpetual (for the duration of the applicable copyright) licence to exercise the rights in the Work as stated below:

a to reproduce the Work, to incorporate the Work into one or more Collective Works, and to reproduce the Work as incorporated in the Collective Works;

b to distribute copies or phonorecords of, display publicly,perform publicly, and perform publicly by means of a digital audio transmission the Work including as incorporated in Collective Works; The above rights may be exercised in all media and formats whether now known or hereafter devised.The above rights include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. All rights not expressly granted by Licensor are hereby reserved.

4 Restrictions

The licence granted in Section 3 above is expressly made subject to and limited by the following restrictions:

a You may distribute,publicly display, publicly perform, or publicly digitally perform the Work only under the terms of this Licence, and You must include a copy of, or the Uniform Resource Identifier for, this Licence with every copy or phonorecord of the Work You distribute, publicly display,publicly perform, or publicly digitally perform.You may not offer or impose any terms on the Work that alter or restrict the terms of this Licence or the recipients' exercise of the rights granted hereunder.You may not sublicence the Work.You must keep intact all notices that refer to this Licence and to the disclaimer of warranties.You may not distribute, publicly display, publicly perform, or publicly digitally perform the Work with any technological measures that control access or use of the Work in a manner inconsistent with the terms of this Licence Agreement.The above applies to the Work as incorporated in a Collective Work, but this does not require the Collective Work apart from the Work itself to be made subject to the terms of this Licence. If You create a Collective Work, upon notice from any Licencor You must, to the extent practicable, remove from the Collective Work any reference to such Licensor or the Original Author, as requested.

b You may not exercise any of the rights granted to You in Section 3 above in any manner that is primarily intended for or directed toward commercial advantage or private monetary compensation.The exchange of the Work for other copyrighted works by means of digital filesharing or otherwise shall not be considered to be intended for or directed toward commercial advantage or private monetary compensation, provided there is no payment of any monetary compensation in connection with the exchange of copyrighted works.

C If you distribute, publicly display, publicly perform, or publicly digitally perform the Work or any Collective Works,You must keep intact all copyright notices for the Work and give the Original Author credit reasonable to the medium or means You are utilizing by conveying the name (or pseudonym if applicable) of the Original Author if supplied; the title of the Work if supplied. Such credit may be implemented in any reasonable manner; provided, however, that in the case of a Collective Work, at a minimum such credit will appear where any other comparable authorship credit appears and in a manner at least as prominent as such other comparable authorship credit.

5 Representations, Warranties and Disclaimer

A By offering the Work for public release under this Licence, Licensor represents and warrants that, to the best of Licensor's knowledge after reasonable inquiry:

i Licensor has secured all rights in the Work necessary to grant the licence rights hereunder and to permit the lawful exercise of the rights granted hereunder without You having any obligation to pay any royalties, compulsory licence fees, residuals or any other payments;

ii The Work does not infringe the copyright, trademark, publicity rights, common law rights or any other right of any third party or constitute defamation, invasion of privacy or other tortious injury to any third party.

B except as expressly stated in this licence or otherwise agreed in writing or required by applicable law,the work is licenced on an 'as is'basis,without warranties of any kind, either express or implied including,without limitation,any warranties regarding the contents or accuracy of the work.

6 Limitation on Liability

Except to the extent required by applicable law, and except for damages arising from liability to a third party resulting from breach of the warranties in section 5, in no event will licensor be liable to you on any legal theory for any special, incidental,consequential, punitive or exemplary damages arising out of this licence or the use of the work, even if licensor has been advised of the possibility of such damages.

7 Termination

A This Licence and the rights granted hereunder will terminate automatically upon any breach by You of the terms of this Licence. Individuals or entities who have received Collective Works from You under this Licence,however, will not have their licences terminated provided such individuals or entities remain in full compliance with those licences. Sections 1, 2, 5, 6, 7, and 8 will survive any termination of this Licence.

B Subject to the above terms and conditions, the licence granted here is perpetual (for the duration of the applicable copyright in the Work). Notwithstanding the above, Licensor reserves the right to release the Work under different licence terms or to stop distributing the Work at any time; provided, however that any such election will not serve to withdraw this Licence (or any other licence that has been, or is required to be, granted under the terms of this Licence), and this Licence will continue in full force and effect unless terminated as stated above.

8 Miscellaneous

A Each time You distribute or publicly digitally perform the Work or a Collective Work, Demos offers to the recipient a licence to the Work on the same terms and conditions as the licence granted to You under this Licence.

B If any provision of this Licence is invalid or unenforceable under applicable law, it shall not affect the validity or enforceability of the remainder of the terms of this Licence, and without further action by the parties to this agreement, such provision shall be reformed to the minimum extent necessary to make such provision valid and enforceable.

C No term or provision of this Licence shall be deemed waived and no breach consented to unless such waiver or consent shall be in writing and signed by the party to be charged with such waiver or consent.

D This Licence constitutes the entire agreement between the parties with respect to the Work licensed here.There are no understandings, agreements or representations with respect to the Work not specified here. Licensor shall not be bound by any additional provisions that may appear in any communication from You.This Licence may not be modified without the mutual written agreement of Demos and You.

Jamie Bartlett is Director of the Centre for the Analysis of Social Media at Demos. Louis Reynolds is a researcher at Demos.

ISBN 978-1-909037-92-2