“A balance between security and privacy online must be struck…”

#INTELLIGENCE

Sir David Omand
Jamie Bartlett
Carl Miller

Demos is a think-tank focused on power and politics. Our unique approach challenges the traditional, 'ivory tower' model of policy making by giving a voice to people and communities. We work together with the groups and individuals who are the focus of our research, including them in citizens' juries, deliberative workshops, focus groups and ethnographic research. Through our high quality and socially responsible research, Demos has established itself as the leading independent think-tank in British politics.

In 2012, our work is focused on four programmes: Family and Society; Public Services and Welfare; Violence and Extremism; and Citizens. Alongside and connected with our research programes, Demos has political projects focused on the burning issues in current political thinking, including the Progressive Conservatism Project, the Centre for London at Demos and Demos Collections, bringing together topical essays by leading thinkers and commentators.

Our work is driven by the goal of a society populated by free, capable, secure and powerful citizens. Find out more at www.demos.co.uk.

First published in 2012

Magdalen House, 136 Tooley Street,
London, SE1 2TU, UK

ISBN 978 1 906693 08 3
Series design by modernactivity
Typeset by Chat Noir Design, Charente
Printed by Lecturis, Eindhoven

Set in Gotham Rounded
and Baskerville 10
Cover paper: Flora Gardenia
Text paper: Munken Premium White

#INTELLIGENCE

Sir David Omand
Jamie Bartlett
Carl Miller

DEMOS

Contents

Foreword

This report provides a balanced and sobering commentary on a subject that has recently attracted elsewhere more wild and unconsidered comment than common sense. There is no doubt that the changing risk landscape and the changed operational environment have presented a challenge to those who hunt down terrorism and other very serious crime. The growing use of social media now has to be examined as an important part of this much bigger intelligence picture.

The authors make serious recommendations as to how social media and social networks can and should become part of the everyday intelligence effort. As they argue, this must be done accountably, proportionately and in a fair, balanced and reviewable way. They are correct that the current legislation, including the Regulation of Investigatory Powers Act 2000, should be re-examined and rewritten to fit the current situation. There should be an expert and independent advisory panel to oversee the process. The boundaries between the public digital space and the private digital space should be clarified. The public should be given an informative narrative, so that they understand what is being done and that it is safe and in their interests.

Policy makers would be well advised to read the paper carefully, and use it as a template.

Lord Carlile of Berriew CBE, QC
The Independent Reviewer of Terrorism Legislation (2001–11)
April 2012

Executive summary

Social media is transforming society. We are transferring more and more of our lives onto vast digital social commons. The emergence of these increasingly significant public spaces poses a dilemma for government.

On the one hand, the emergence of these social spaces holds a great opportunity for more effective, agile and responsive government and wider social and economic gain. In particular social media intelligence – which we term 'SOCMINT' – could contribute decisively to public safety: identifying criminal activity; giving early warning of disorder and threats to the public; or building situational awareness in rapidly changing situations. As society develops and adopts new ways to communicate and organise, it is vital that public bodies, including law enforcement and the intelligence community, keep up with these changes.

On the other, at the heart of national security is public understanding and support for the steps being taken to keep us safe. Democratic legitimacy demands that where new methods of intelligence gathering and use are to be introduced they should be on a firm legal basis and rest on parliamentary and public understanding of what is involved, even if the operational details of the sources and methods used must sometimes remain secret. As the public debate surrounding the proposed UK Communications Capability Development Programme (CCDP) demonstrated, any new proposals for intelligence gathering in an internet age will raise issues over access to personal data and their use by the state, as well as broader concerns about the effect surveillance work might have on the economic and social value of the internet as a place of free exchange of ideas and information.

In respect of SOCMINT these conditions of democratic legitimacy are presently absent. Social media does not fit easily

into the policy and legal frameworks that guarantee to the public that intelligence activity in general is proportionate, accountable, and balances various public goods, such as security and the right to privacy. People now share vastly more personal information about themselves, their friends and their networks in new and varied ways: what is 'public' and what is 'private' is not always obvious, and differs greatly across social media platforms and even within social media platforms. Moreover, new and emerging technology potentially allows more invisible and widespread intrusive surveillance than ever before. Consequently, ways of managing the possible harms associated with the state accessing and using social media data have to be understood. This is why SOCMINT cannot be readily fitted into the current framework that manages the state's intrusion into people's private lives. The Regulation of Investigatory Powers Act, for example, was passed in 2000, long before social media use was widespread. New harms may also need to be considered, such as the risk surveillance might entail to the economic and social benefit of the internet as a site of the free exchange of ideas. Ensuring intelligence and security work is proportionate, legitimate and based on public consent depends on measuring and managing the possible harms it might entail; for SOCMINT how this is to be done is still unclear.

Intelligence work is also only justified in any circumstances if it is able to improve the quality of decision making. But as the summer 2011 riots revealed, SOCMINT is not yet capable of making a decisive contribution to public security and safety. There are new forms of online behaviour, norms and language that make analysis and verification difficult. Translating often unprecedentedly large, complex and conflicting bodies of information into actionable, robust insight is a significant challenge that has not been overcome. SOCMINT does not fit easily into the existing systems we have developed to ensure intelligence collected can be confidently acted on.

The way forward

We believe that SOCMINT could potentially make a dramatic, legitimate and publicly acceptable contribution to public security and safety. But two conditions must be met which respond to the challenges outlined above.

First, SOCMINT must be based on a publicly argued and sound legal footing, with clarity and transparency over use, storage, purpose, regulation and accountability. This means the harms it entails must be identified and managed, and legislation amended or created accordingly. This not only is important for the public, but also provides an enabling environment in which agencies of the state feel confident and able to act.

Second, SOCMINT must be able to produce reliable, powerful insight that can be acted on. This means there needs to be greater investment in human and technology capabilities, and the creation of a new inter-disciplinary approach fusing technological capability and humanistic understanding together as social media science.

This report lays out a template for how both of these challenges, in the immediate and longer terms, could be approached.

We believe any use of SOCMINT by the state – including the CCDP – should be based on the following six principles:

- principle 1: there must be sufficient, sustainable cause
- principle 2: there must be integrity of motive
- principle 3: the methods used must be proportionate and necessary
- principle 4: there must be right authority, validated by external oversight
- principle 5: recourse to secret intelligence must be a last resort if more open sources can be used
- principle 6: there must be reasonable prospect of success

We believe the principles provide a secure framework within which Britain's responses to new technology challenges can be consistent, and be demonstrated to be consistent, with Britain's approach to civil liberties and information rights.

Government should take a two-route approach to the use of SOCMINT, making a clear distinction between open source non-intrusive SOCMINT and intrusive or surveillance SOCMINT.

Route one would be open source non-intrusive SOCMINT, which can be conducted on a similar basis to non-state actors, such as universities and commercial companies. This should be tightly bound with conditions relating to anonymity, data protection or based on the full consent of the producers of that information. This might include such activity as openly crowd sourcing information through Twitter or Facebook to gain situational awareness in the event of public disorder, or gauging general levels of community tension. This type of activity would not be used to identify individuals, or as a means of criminal investigation and should not puncture the privacy wishes of any user. As such, this would not fall under existing legislation that governs intrusions into people's privacy: individual departments and agencies would be responsible for how to undertake this type of activity. Inevitably it is possible that, while undertaking route one SOCMINT, criminal or possible criminal activity is found. In the event, this should be then transitioned into the second route, set out below.

Route two SOCMINT is the exercise of state-specific powers of access intended to result in the identification of individuals and access to private information. This is SOCMINT as intrusive surveillance and interception. Accessing social media could range from relatively minor intrusions (such as collecting publicly available data about specific individuals) to more significant intrusions, such as intercepting and reading personal communications. Such access needs to be governed by a series of ethical principles which we set out below, and animated through a legal framework that maintains an association between harm, privacy, authorisation, agency and cause, such as limits on the number of agencies permitted to undertake it. In the immediate term, this type of activity could be governed by relevant legislation contained in parts I and II of RIPA 2000, although we believe an interdepartmental review and a Green Paper are needed to reach a sustainable settlement based on public consent and acceptance.

The Government should undertake an interdepartmental review of current legislation – notably RIPA 2000 – and existing systems of oversight to determine what applies to SOCMINT now.

Following that, there needs to be public and parliamentary debate about the use of SOCMINT. However, it is important to ensure there is some form of oversight and regulation governing its use. We believe RIPA 2000 is the most appropriate legislation currently available. An interdepartmental review must review what types of SOCMINT might fall under RIPA 2000 parts I and II, and the relevant degrees and type of authorisation required. Existing mechanisms of oversight for all intelligence and policing work, including the Parliamentary Intelligence and Security Committee and the independent police commissioners, need to determine how SOCMINT should relate to their current procedures and operations. We recommend that as far as possible the association of privacy, authorisation, agency and cause as exists under RIPA 2000 is maintained for SOCMINT.

In the long term, the Government should publish a green paper subject to public consultation about how it plans to use and manage social media analysis in the public interest, including for the purposes of public security.

This must include a position on how to define and measure the possible harm entailed by SOCMINT access, and how it can therefore be balanced against other public goods. This requires the provision of information about the circumstances legitimising the use SOCMINT, the bodies capable of conducting it, the system of authorisation and oversight that will exist, and how abuses to this system will be prevented and redressed.

Government needs to evolve and strengthen SOCMINT capabilities. An independent expert scientific and industrial advisory panel and SOCMINT centre of excellence should be established.

A single, networked hub of excellence should coordinate SOCMINT development across different branches of

government, and structures of engagement and funding must be created to involve extra-governmental actors, especially industrial and academic actors, in the process. Strengthening SOCMINT capability also includes the creation of a 'SOCMINT culture', where SOCMINT practitioners and users understand the cultural, linguistic and technological underpinnings of the platform.

1 The dilemma of social media intelligence as a public good

We live in the age of social media. Facebook, Twitter, Google+ and LinkedIn are all examples of the rapid transfer of people's lives – interactions, identities, arguments and views – onto a new kind of public and private sphere; a vast digital social commons.[1] This transfer is happening on an unprecedented scale. On Facebook alone, 250 million photos are added per day,[2] as are 200 million tweets on Twitter.[3] There are 4 billion video views per day on YouTube.[4]

Data of this size are known as 'big data'. Big data are both more – quintillions of bytes and growing at around 50 per cent a year[5] – and linked: aggregated webs of information rather than discrete databases.[6] Social media is an extremely important class of big data, and are increasingly subject to collection and analysis. Measuring and understanding the visage of millions of people digitally arguing, talking, joking, condemning and applauding is of wide and tremendous value. Unsurprisingly, big data are valuable; they are already described as a new class of economic asset, like currency or gold.[7]

The family of big data approaches applied to make sense of social media is currently known as 'social media analytics' (SMA). SMA is a broad church, ranging from the general aggregation of social media content to mapping 'social graphs' of relationships between people, to drawing 'wisdom of the crowd' solutions to emergency situations, to conducting linguistic analysis of forum posts and network analysis of Twitter users. Hundreds of companies offer software and services to measure the 'buzz' emanating from social media.[8] Advertisers listen to this buzz to track attitudes surrounding their brands, and companies monitor their reputation and spot emerging crises concerning their products. Fledgling academic efforts have used social media to inform investments into hedge funds.[9]

As people transfer more of their lives onto social media platforms, they become an increasingly significant public space, and therefore of interest to, and used by, public bodies. Understanding the content of social media presents an opportunity for public bodies better to understand, and respond to, the public they serve.[10] Public health experts are learning to scan tweets and search requests to identify pandemics earlier than traditional methods.[11] US psychologists believe Facebook contains valuable indicators of mental health, and indeed the social media profiles of a number of participants in school shootings, such as the suspect in the Ohio School Shooting, TJ Lane, seem to show some indicative content.[12] The United Nations believes that tapping into social media can help tackle global unemployment and food insecurity.[13] Political parties are also starting to explore the use of these technologies for electoral advantage. Even small-scale uses have delivered large returns. Highly customised Facebook adverts helped deliver the rank underdog Sam Kooiker an astonishing victory in Rapid City's mayoral elections in South Dakota.[14]

On a larger scale, the 2012 US presidential election – dubbed 'the first Facebook election' – sees President Obama's re-election campaign team using automated social media collection to both organise and directly message prospective voters on an unprecedented scale.[15] Similar tactics are becoming increasingly common in UK politics, especially following the recent successes of the Scottish National Party and its sophisticated online database 'Activate'.[16]

'SOCMINT': intelligence and insight from social media

Social media is now significantly relevant to security and public safety. Facebook, for example, has been used to coordinate contract killings, boast about serious animal abuse, conduct cyber-stalking, plan sexual assaults, breach court orders and cause distress through anti-social 'trolling'.[17] In late 2010, it was reported that the police received 7,545 calls from the public that year concerned with Facebook.[18]

There are many ways social media is likely to affect policing and security work. They could facilitate direct engagement with the public. For example, Greater Manchester Police have developed a social media application to share information – including a newsfeed, missing persons and police appeals – with the public.[19] They might also bring new risks, too, such as leaking of confidential information, the identification of undercover agents, or the reputational risks involved with not responding to social media complaints or concerns. New guidelines and policies are likely to be needed on how to manage these new opportunities and threats.

This paper focuses on one specific way in which social media can be used by police and intelligence agencies: the opportunity to generate and use social media intelligence – 'SOCMINT' – in the interests of safety and security.

The explosion of social media, together with the rapid development of SMA capabilities, now provides an opportunity to generate intelligence that could help identify criminal activity, indicate early warning of outbreaks of disorder, provide information and intelligence about groups and individuals, or help understand and respond to public concerns.

This is already happening. Cases show that everyone from international criminal fugitives to bigamists have been caught using social media.[20] A number of police forces in the UK and elsewhere are believed to be trialling various types of automated social media collection and analysis to collect information to help criminal investigations and gauge the 'temperature' of communities they are working with.[21] Police constabularies have used Flickr to crowd source identifications of suspects from photographs. Underlying this has been significant public investment in the capabilities to generate SOCMINT. In the UK, the Ministry of Defence's Cyber and Influence Science and Technology Centre has released a number of calls for research to develop capabilities including 'cyber situational awareness', 'influence through cyber' and 'social media monitoring and analysis: crowd sourcing'.[22]

The dilemma of SOCMINT as a public good

Government faces a dilemma over when, where and how it collects and uses SOCMINT. On the one hand, in a modern society characterised by widespread social media use, SOCMINT is likely to be an increasingly relevant component of intelligence work in support of public safety and security. There will be public pressure to use it, and an imperative on the agencies tasked with protecting society to become expert and effective practitioners in the collection and analysis of social media using the most powerful means available.

On the other hand, the methods employed to protect society rest ultimately on some form of public acceptability and involvement. Britain's National Security Strategy recognises that security and intelligence work in general is predicated not only on the public's consent and understanding, but also on the active partnership and participation of people and communities. Serious and recognised damage to security occurs when the state's efforts are not accepted or trusted.[23]

Public acceptability can be secured and maintained through two important public demonstrations. First, that SOCMINT is able to make an effective and necessary contribution toward safety and security; second, that this contribution is being proportionately and appropriately balanced against other desirable public goods – such as the right to private life. In sum, intelligence activity must effectively contribute to a public good but not detract from or threaten any others in ways that are not recognised and appropriately managed.

In general terms, the law enforcement and intelligence communities maintain public confidence through a delicate settlement of practices, procedures, laws and regulations. Statutory provisions, including the incorporation of the provisions of the European Human Rights directive into domestic law through the Human Rights Act 2000, ensure respect for human rights (such as the right to a private life) is upheld, if necessary through the courts. The Regulation of Investigatory Powers Act 2000 (RIPA 2000) part I governs interception, and in part II surveillance (such as eavesdropping) and covert human intelligence sources. Statutory monitoring and oversight bodies, both judicial and parliamentary, provide means

of investigation and redress. Within the intelligence profession and the bodies that use intelligence, various procedures, doctrines, expertise and processes exist that embody the principles of proportionality and accountability, and help to establish and ensure validity and trustworthiness. Taken together, this settlement helps to ensure that intelligence activity is legal, ethical and effective.

SOCMINT potentially disrupts this equilibrium. Social attitudes towards what is private or public, and therefore what counts as intrusive or not, are blurred and changing. It is unclear whether social media platforms are public spaces, private spaces, or something else altogether. The legislation that covers the processing of personal data – the Data Protection Act 1998 – and the intrusive collection of intelligence about suspects by government agencies and the police – RIPA 2000 – were both written into law before the advent of social media. And even if SOCMINT could gather information in ways that manage its potential harm, it is unclear today how it would be effectively employed. Many of the methods to generate SOCMINT are inadequate in their current form, and do not fit within the existing organisational boundaries and processes that translate good information into more informed action. One of the reasons that social media sources were not used by the police when responding to the 2011 summer riots was that it did not fit their existing process of intelligence collection, validation and reaction.

Therefore, for SOCMINT to contribute towards the public good, and become a justified part of intelligence and policing work, its existence and use need to be demonstrated both to be ethically and legally sound, and capable of producing value-adding intelligence in practice. We discuss each in turn.

In chapter 2 we examine the ethical challenges involved in the collection of SOCMINT, in particular how it affects the balancing of three public goods constant to all security and intelligence work: public safety, the right to privacy and the economic and social wellbeing of the nation. Given that the great economic and social power of the internet – and social media – lies in its openness as a free, broadly unfettered space, and given the context of fundamentally changing and blurring social

norms and a system of regulation that predates the arrival of social media, how can the delicate balance between the government's need to access certain information for purposes of national security, protecting the right to privacy, and improving the economic and social wellbeing of the nation and its citizens be maintained? We conclude that SOCMINT will become an important part of future intelligence efforts, but needs to be balanced against these other public goods.

In chapter 3, we set out six specific ethical tests that can be applied to ensure that SOCMINT can be undertaken according to a sound moral and legal basis. We recommend that a long-term settlement about SOCMINT should be based on public and parliamentary debate, and according to these principles. In the immediate term, we propose that SOCMINT activity could be governed by existing legislation covering privacy intrusions (RIPA 2000 parts I and part II) and set out a framework for relating social media privacy breaches to existing definitions of privacy.

In chapter 4, we examine methodological, interpretative and practical challenges to SOCMINT's contribution to public safety: how can huge, confusing, conflicting and overwhelming corpora of social media information be turned into powerful, reliable and useable insight? How can SOCMINT be accurately and robustly accessed, analysed, interpreted, validated, disseminated and used? We conclude that the provision of capabilities that can yield actionable evidence requires an underlying transition from the current suite of social media analytics tools into a new discipline – social media science – involving a much more meaningful and intensive fusion of the computational, technological and humanities approaches.

In each of these chapters, it is not our aim to provide an exhaustive set of concrete solutions to these challenges – the scope of the topic is far too varied and fast-paced. Rather, the paper sketches out how each challenge should be understood and then approached. Overall it suggests a strategic approach to build experience and public understanding on which more permanent solutions (including potentially legislation and international agreements) can be based.

2 Balancing security with other public goods

The problem

From Hobbes' *Leviathan* onward, the political theory of the state understands security to be the state's first duty.[24] Modern approaches continue this tradition but define national security broadly to cover the major threats and hazards facing the citizen, and see a state of security as resting on public confidence that those risks are being effectively managed in relation to other public priorities. Security is therefore sustained public protection through the delivery of a balanced set of public goods.

All security and intelligence work rests on a delicate balance between three classes of public goods: the maintenance of national security including public order and public safety; citizens' right to the rule of law, liberty and privacy; and the overall economic and social wellbeing of the nation and its citizens. In her 2011 Reith Lecture, the former Director General of the Security Services Dame Eliza Manningham-Buller emphasised the importance of these values in maintaining security, and explicitly placed intelligence work within the framework of articles 2, 5 and 8 of the European Convention on Human Rights (ECHR) and the Human Rights Act: the rights to life, security and liberty, and a private life.[25] Public opinion polling tentatively agrees with regulated, restricted access. In a 2011 Eurobarometer poll, when considering police access to online personal data, around 40 per cent of UK respondents felt that police access to data on social networking sites should be allowed but within the framework of an investigation, around 20 per cent with the authorisation of a judge, and 37 per cent for 'general crime prevention activities'.[26]

In most circumstances these three classes of public goods should be mutually reinforcing: security promotes inward investment and market confidence promoting economic

wellbeing and social harmony that in turn supports the maintenance of security. There are times however when choices have to be made. Within a rights-based approach, the only justification for one public good to be hazarded is the provision of another.

The UK's approach to managing and defining appropriate trade-offs is to adhere to a legal framework defined by statute, common law and regulatory codes of practice that both define and limit the powers of the state to create new law, restrict liberty and intrude on privacy, for example, in the name of public order and security. The practice of this approach is scrutinised by a number of independent oversight bodies, some drawn from the judiciary to ensure compliance with the law, others drawn from (or accountable to) Parliament to examine policy and cases.

The Human Rights Act (especially article 8, the right to privacy) and the Data Protection Act (especially 'schedule 1')[27] lay out the circumstances under which personal information can be processed by public authorities and private organisations. At the European level, the European Directive on Data Protection (95/45/EC) requires the 'unambiguous consent' of the subject before 'personal data' can be shared.[28] RIPA 2000 established the overall principles and procedures under which nominated public bodies including the police may breach normal privacy in the form of intrusive interception and surveillance operations. The Security Service Acts 1989 and 1996 and the Intelligence Services Act 1994 also lay down restrictions on the purposes for which the national intelligence agencies may collect intelligence and insist on safeguards for that information.

Social media is a potentially disruptive phenomenon that is already affecting and in some cases redefining how these three classes of public goods can be attained: security and public safety, privacy and consent, and the economic and social wellbeing of the nation.

We discuss each below, and their implications for the collection and use of SOCMINT.

SOCMINT's possible contribution to public goods: safety and security

The justification for the state engaging in intelligence work is that intelligence can help achieve better decision making in the public interest by reducing ignorance on the part of the decision taker, whether a police officer, military commander or policy maker. A number of trends and examples now suggest that social media is already and will increasingly be an essential source of intelligence and insight for the proper maintenance of security and public safety on the part of most national governments.

Social media's increasingly central role in how society interacts is important. Worldwide there are 845 million Facebook users, of whom 483 million access the website every day,[29] while in February 2012 the number of Twitter users grew to over 500 million.[30] In June 2011 the number of UK Facebook users was measured at 29.8 million people, or 58 per cent of people online.[31] It is an increasingly important space.

What is conducted in this space now has clear consequences for security and safety. On Thursday 4 August, Mark Duggan was shot and killed by a police officer in Tottenham. By the morning of Saturday 6 August social media channels showed increasing hostility, including explicit threats, against the police. From 7 August, social media information indicated the possible spread of disorder to other parts of London, then England. Over the next few days, content indicating criminal intent or action ratcheted in huge numbers through both open source social networking, such as Twitter, and closed system networks, such as the BlackBerry Messaging service and closed groups such as chat forums. Similarly, huge numbers of messages appeared trying to provide information to the police, either about an outbreak of disorder or the identities of the people behind it.[32]

Following the August 2011 riots the police acknowledged that they had been insufficiently equipped to deal with intelligence gathering via social media. One intelligence professional said it was like 'searching the British Library for a page in a book without an index to refer to'.[33] Social media did not fit into their systems of receiving, corroborating, prioritising and disseminating information, and therefore was not properly acted on. Her Majesty's Chief Inspector of Constabulary noted,

'With some notable individual exceptions, the power of this kind of media (both for sending out and receiving information) is not well understood and less well managed.'[34] He concluded that '[t]he police have much to learn about social media, and the quickly shifting modern communications of today'.[35]

The summer 2011 riots are just one example among many. When society develops and adopts new methods of communication and organisation – such as social media public institutions, including the police and intelligence services, have a responsibility to react and adapt. Groups like the English Defence League use sites like Facebook to plan and organise their demonstrations, and access to such data could be a vital source of information to help more effective policing.[36] In the UK, thousands of crimes have been linked to Facebook.[37]

Looking at the current technologies now on the horizon – as well as the threats we now face – the following SOCMINT capabilities could contribute decisively in the future to public security. This includes understanding social resentments, grievances and radicalisation, and the identification of specific criminal intent or individuals:

- *Crowd-sourced information*. This could help ensure a better flow of information between citizens and the government, especially in times of emergency.[38] With access to social media, passive bystanders can become active citizen journalists, providing and relaying information from the ground. The report by Her Majesty's Inspectorate of Constabulary (HMIC) into the riots notes, for example, a messaging service on West Midlands Police's website, which allowed citizens to post messages and questions, allowing the police to build up a picture of the situation on the ground in real-time, as well as allowing people to identify pictures of suspects uploaded to the site.[39] Tapping into the 'wisdom of the crowds' is already of great, demonstrated value. For example, the open-source platform Ushahidi has allowed large groups of people to provide collective testimony on everything from the earthquake in Haiti to blocked roads in Washington DC.[40] These applications, impressive as they are, are only the beginning, and the stronger the techniques to make

sense of information of this kind, scale and dynamism, the more effective the responses, from providing snow ploughs to drinking water, that can be made.

- *Real-time situational awareness.* This is the ability to collect and cluster social media and output in a way that indicates and describes unfolding events. Analysis of Twitter has shown that, while the majority of Twitter traffic occurred after an event had been reported by a mainstream news outlet, 'bursts' of tweets indicating a significant event often pre-empt conventional reporting.[41] Social media traffic analysis could allow for a more rapid identification of events than traditional reporting mechanisms. With the application of geo-location techniques this could lead, for example, to a constantly evolving map showing spikes in possible violence-related tweets, facilitating a faster, more effective, and more agile emergency response.
- *Insight into groups.* This would include the ability to better understand activities and behaviour of certain groups already of interest to police or intelligence agencies. SOCMINT could spot new, rapidly emerging 'hot topics' that spring up within group-specific conversations and how the group reacts to a specific, perhaps volatile, event. Through these and other techniques, SOCMINT might indicate the overall levels of anger within a group, and their key concerns and themes that animate intra-group discussions. At a higher level of specificity, information can also be identified and extracted regarding when a group is planning demonstrations or flashmobs, which could lead to violence or increasing community tensions; football fans planning 'meets', which could cause major economic disruption; groups planning counter-demonstrations, which could change the kind of policing required to maintain public order.
- *Research and understanding.* Research based on social media could contribute to our understanding of a number of phenomena. This could include the thresholds, indicators and permissive conditions of violence; pathways into radicalisation; an analysis of how ideas form and change; and investigation of the socio-technical intersections between online and offline personae. Beneath the tactical and operational level, a background of more generic and distanced understanding is important for security

work. For instance, the British counter-terrorism strategy aims to reduce the threat from terrorism so that people can go about their normal lives, freely and with confidence, and it is understood that the long-term way to do this is through tackling the underlying social, ideational and political causes of terrorism.

- *Identification of criminal intent or criminal elements in the course of an enquiry both for the prevention and prosecution of crime.* This could include the surveillance of social media use by individuals suspected of involvement in a crime or criminal conspiracy, the cross-referencing of such individuals' accounts, the identification of accomplices, the uncovering of assumed identities, the identification of criminal networks that operate through social media sites, and the provision of social media content suspected of being evidence of a crime to the Crown Prosecution Service.

This list is by no means exhaustive, and does not capture the full range of possibilities. Indeed, the technology for potentially far more intrusive surveillance also exists. As the technology continues to evolve, new applications and opportunities will doubtless emerge. While SOCMINT capabilities could contribute to security, they could also potentially entail hazard to other public goods, especially privacy and consent. We turn next to examine this downside.

SOCMINT's possible harm to public goods: privacy and consent

Privacy itself is an elusive concept. Article 8 of the ECHR (echoed in the UK Human Rights Act 2000) enshrines the right to respect for 'a person's private and family life, his home and correspondence', but privacy has no formal definition within UK law. Respecting privacy can mean that data are kept confidentially, gathered anonymously, used in a self-determined way (the principle of 'informed consent'), and that people are able to see them and correct errors, or, of course, that no data are gathered at all.

Many broad and fundamental changes in society are transforming what privacy means to people. Social media

challenges clear-cut distinctions of what is private and what is not. McKinsey Global Institute has calculated that 30 billion pieces of content are shared on Facebook each month, many of them personal.[42] This sharing of such a large amount of voluntarily uploaded personal data, and the number of people and institutions to whom these data are accessible, is unprecedented; depending on the user-selected privacy settings employed, personal information added to Facebook can be viewed by all of Facebook's 845 million other users. Far from being incidental, this move towards the widespread dissemination of personal information is fundamental to the ethos of social networking sites. Facebook's privacy settings inform users that the ability to share information 'allows us to provide Facebook as it exists today', while Twitter states more explicitly that '[m]ost of the information you provide to us is information you are asking us to make public'.[43]

Encouraging users to share their personal information is central to these companies' business plans and lies at the heart of the commercial competition between tech giants like Google and Facebook.[44] The practice of gathering vast amounts of personal information and selling it to third parties, in particular advertisers, is highly lucrative and consequently the quantity of personal data held by some sites about their users is huge. For instance, in 2011 an Austrian student, Max Schrems, made a request to access the information held on him by Facebook and subsequently received several CDs containing over 1,200 PDF pages chronicling in minute detail his actions on the site since 2008.[45] The privacy implications of this were detailed by Schrems himself when he said that Facebook knew, or had the ability to know, about the most intimate details of his life, including, 'every demonstration I've been to, my political thoughts, intimate conversations, discussion of illnesses'.[46]

Indeed as a result of these changing behaviours, Mark Zuckerberg, Facebook's CEO, declared that privacy is 'no longer a social norm'.[47] Attitudes towards privacy – especially broad, generic and in-principle attitudes – are notoriously hard to measure objectively. Broad behavioural norms, such as the amount of information we now share, suggest the concept of

privacy is certainly changing. Most of us accept that both private and public bodies – from Tesco through its Clubcards to Amazon, Oyster and Google – learn and record a vast amount about us daily. In a Eurobarometer poll, a bare majority of UK respondents considered photos of themselves to be personal data, less than half considered 'who your friends are' to be personal data, 41 per cent thought that details of the websites they visit were personal data, and only 32 per cent thought their tastes and opinions were personal data, yet in contrast, large majorities regard financial data as personal.[48]

However, although research suggests that users recognise disclosing personal information is an increasingly important part of modern life, the majority have concerns about what this means.[49] In a 2008 European Commission Poll, around 80 per cent of people agreed that 'people's awareness about personal data protection in the UK is low'.[50] The majority of us barely or never read the terms and conditions when downloading apps or uploading information.[51]

Indeed, there is a profound tension between the privacy of consumers' information on the one hand, and (often commercial) data sharing on the other. The economics of the internet requires revenue-earning data sharing. Facebook's chief of engineering Lars Rasmussen described striking a balance between users' control of their data and the practice of data sharing as something which is 'always going to be our main challenge'. These tensions are not yet resolved.[52]

The concept of privacy seems therefore both important and relevant to people, but also in a state of flux, where its boundaries of definition are being fundamentally redrawn. The debate will continue to rage about where these redrawn boundaries on the possession, sharing and use of personal information, now lie – indeed what privacy is.[53] The crucial implication for the collection of SOCMINT is that the framework for recognising and managing incursions into privacy is struggling to keep pace.

The quasi-public sphere of many social media platforms (indeed a function of social media is its openness and networked design) does not fit easily into existing legislation that governs

privacy. This is because much of it was written into law at a time when what was private and public was more distinguishable. In 2010 two US state courts, for example, returned conflicting rulings in two very similar cases, based on two entirely different readings of statute. In the first case, brought in California, a reading of electronic communications legislation exempted a defendant from turning over his 'private' Facebook and MySpace messages because of his friends-only privacy settings. A New York judge, under the same legislation, in the same year, returned the opposite verdict. The plaintiff was told to hand over all of her Facebook postings and images because, in creating and using social media accounts, 'she consented to the fact that her personal information would be shared with others, notwithstanding her privacy settings'.[54]

UK law faces similar challenges. The European Directive on Data Protection 95/46/EC controls how personal data can be used and shared. Sharing any identifying information requires the notice and 'unambiguous consent' of the individual in question. Enacted before the rise of social media, any individual sharing information without the express consent of the individual to whom the data pertains may, under this directive, be seen as a 'data processor' transgressing European law. In January 2012 it was announced that a replacement data protection regulation was being drafted, but this process is not yet complete.[55] Indeed, Google's attempts to rationalise its privacy policies into a single, unitary one is being opposed by the EU data authorities.[56]

These issues of privacy, intrusion, and ownership of data are especially significant for the state's use of SOCMINT.

There are many ways the state can collect and use information about people, while different systems exist for carrying this out. Each system identifies and limits potential harm from accessing private information. When the state conducts clinical research, for example, consent is requested and anonymity often guaranteed; when the state draws on the research of others we apply fair usage and ascribe credit to the individuals concerned. When the state obtains biometric information such as a DNA sample from a suspect, consent is

not required but restrictions are applied to the retention of material.

When the state carries out surveillance, the activity is usually covert and individual consent is irrelevant. Instead, to deal with concerns of privacy and intrusiveness, society's consent is needed, expressed through enabling legislation that provides safeguards, for example through requiring warrants for the activity to be obtained in advance. European Court of Human Rights case law upholds the right of states to conduct intrusive investigation of its citizens by police, security and intelligence agencies, for example to counter terrorism and for the prevention and detection of serious crime. However, this case law insists there must be a statutory basis for such activity, legislation governing the bodies that conduct it, and the ability of the citizen to have independent investigation of allegations of abuse of these powers and effective redress if proven.

As already mentioned, RIPA 2000 parts I and II provide in the UK the legal basis for the use of intrusive methods of gathering intelligence (and a schedule to the Act lists the public bodies that may engage in different classes of intrusion), and the more intrusive the activity (for example, the need to make lawful covert entry into a suspect's house to plant a listening device) the higher the degree of prior authorisation that is mandated under the Act. Until recently, such activity mostly involved interception of postal and telephone communications, and directed surveillance and eavesdropping. The widespread use of mobile telephony and now the internet has made the monitoring of communications meta-data (who called whom, when and for how long) a fruitful additional source of intelligence, quite apart from the content of communications.

The advent of SMA brings yet another potential source of intelligence, one that did not exist when RIPA 2000 was conceived and, as we shall discuss below, one that is hard simply to slot into the existing categories of authorisation (warrant) provided for in the Act or that might be covered by data protection legislation. The properties and capabilities of the techniques and technologies that can be used to produce SOCMINT make it hard to draw an exact parallel with any one

of these existing categories. This uniqueness is caused by a number of things. Shifting public attitudes, new types of technologies and surveillance, and new security challenges are all intrinsic difficulties in measuring and managing potential harm:

- *Fungibility and heterogeneity*. SOCMINT cuts across several categories and can be in more than one category at a time. The broad scanning of tweets has similarities to mass surveillance such as the deployment of CCTV in crowded places. The close following of an individual's Facebook page during the course of an investigation has similarities to *de visu* surveillance as 'authorisable' by a senior police officer. Accessing encrypted BlackBerry messaging by cracking the security PIN is an interception of communications under RIPA 2000 for which a warrant would be required.
- *Expectations of privacy*. There is a lack of clarity over the boundaries of what should be considered 'private' as against 'public' space when using social media compared with our existing understanding of these terms, and lack of clarity over the relationship between internet service providers (ISPs) and social network providers and governments.
- *Generality*. Unlike other forms of intrusive investigation there may be no named suspect or telephone number to target and the output may be general rather than specific to an individual (such as noting an increase in social media communications in a specific area where demonstrations are taking place).
- *Scalability*. Many of the automated techniques can be scaled up from the collection of hundreds to millions of pieces of social media data easily and cheaply. The scale is difficult to fix in advance as part of the authorisation process.
- *Flexibility*. The scope of many 'scraping' technologies (for instance, the keywords they scan for) can be changed easily. This means they can easily be redirected away from their original mission and function, which may be justified operationally by tactical changes on the ground but would pose problems for any warranting system.
- *Invisibility*. Like other forms of covert surveillance, the operation of SMA techniques will normally not be visible to the social

media users themselves and will override what they may assume are their privacy settings.

To these SMA specific issues can be added a number of wider concerns about digital surveillance:

- the general rise of information systems that have vast capacities to capture, stockpile, retrieve, analyse, distribute, visualise and disseminate information
- the general decrease in public understanding of the extent and type of 'surveillance' processes being operated by the state and by the private sector (for example through collection of browsing history); the Information Commissioner, in a report to Parliament in 2010, flagged this general trend as posing new challenges to the management of the possible harm of surveillance[57]
- collateral intrusion: the inevitable intrusion into the lives of citizens whom no agency has any reason to suspect of wrongdoing but who may be in innocent contact with suspects[58]
- the general implication of suspicion of wrongdoing from widespread collection of information on specific communities outside the course of a specific investigation; this was a concern also outlined by the Information Commissioner in 2010[59]
- the proliferation of surveillance capabilities – especially free and cheap software – could make it easier for criminals and other non-authorised persons to access information[60]
- the possibility of more, and more damaging, leaks or unauthorised releases of information because there are larger, more centralised information repositories

The lack of legal and conceptual clarity for the use of SOCMINT by government has led to accusations by privacy campaign groups that governments are routinely misusing or abusing their powers. The reaction to the CCDP, described by Privacy International as giving the government 'enormous scope to monitor and control the Internet', is the most recent example.[61] In Canada, the proposed online surveillance bill (Protecting Children from Internet Predators Act) has been

fiercely criticised.[62] In the UK, privacy groups such as the Open Rights Group have also launched campaigns against the Government's alleged plans to 'snoop' on emails, Facebook and other web traffic through the proposed CCDP.[63] Privacy International has launched a series of freedom of information requests about the Metropolitan Police's use of SMA, which have not been answered.[64] Big Brother Watch has raised concerns about plans like Westminster Council's 'Your Choice' programme, which it worries could breach citizens' privacy by accessing their communications via social networking sites.[65] The same group also reported on the alleged monitoring of BlackBerry's BBM service following the arrest of two young men in Essex after they tried to organise a mass waterfight, condemning any such monitoring as 'a gross violation of privacy'.[66]

There are many serious difficulties that make these controversies intractable. Indeed, lying at their heart are contested questions of jurisdiction based on the nationality of users, their physical location, the physical location of servers hosting or visited by software, the physical location of 'victims' (personal, corporate or state), and the physical location or nationality of other involved parties, such as ISPs and search engine providers.

While some of these issues are some way from resolution, there needs to be clarity on the extent to which certain types of social media data could be admissible in court, what would be the evidential requirements and where international jurisdictions are relevant (accessing UK citizens' data if they are hosted on a separate server, or posted on a site that is hosted in a country without such stringent use conditions).

Managing SOCMINT's possible harm to the economic and social wellbeing of the nation

The internet as a free and open space – of course within reasonable limits – provides an immense economic and social benefit to the UK. Intelligence and security work is intended to protect our prosperity, not undermine it. Indeed, as Foreign Secretary William Hague explained in early 2011, 'nothing would

be more fatal or self-defeating than the heavy hand of state control on the internet, which only thrives because of the talent of individuals and of industry within an open market for ideas and innovation'.[67] This sentiment is echoed by Wikipedia founder, Jimmy Wales, who stated that 'the biggest threat to the Internet is not cybercriminals, but misguided or overreaching government policy'.[68]

The risk must be recognised that the unregulated large-scale collection and analysis of social media data will undermine confidence in, and therefore the value of, this space. The idea that the economic and social benefit of the internet is premised on its openness and freedom of government control is not new. From the early 1990s, a powerful argument and vision has existed about what the internet is for and what it should be like: an opportunity to evolve past the nation-state system into post-territorial, self-governing communities who operate under their own floating social contracts of consent-through-use. John Perry Barlow's famous *Declaration of Cyberspace Independence* declared to the 'weary giants of flesh and steel' that cyberspace was developing its own sovereignty and 'you are not welcome among us'.[69]

The foundations of the internet itself are based on these self-consciously revolutionary beliefs. In designing the internet's universal language – the TCP/IP protocol – the 'fathers of the internet' embraced an open architecture that distrusted centralised control, did not make judgements on content, and allowed any computer or network to join.[70] Many of the most important infrastructural and technological developments of the internet have been driven in a way that is consistent with this ethos. Today, many of the most powerful and influential members of the digital world still embody many of its tenets. Mark Zuckerberg, in an open letter to potential investors on Facebook's prospective initial public offering, writes about 'the Hacker Way' as 'testing the boundaries of what can be done'.[71]

The potential ability of government to collect, collate and analyse large amounts of data from social media represents at least the possibility of taming the wilder reaches of this chaotic and open space. One of the justifications for security and

intelligence work is to maintain the economic and social welfare of the nation and its citizens. On introducing the RIPA 2000 bill to Parliament, the then Home Secretary Jack Straw MP said 'we are actively trying to ensure that our system protects individuals' Convention rights, while recognising how vital such investigatory powers are to the protection of society as a whole'.[72] Any consideration of SMA use must therefore be based on understanding the risk of diminishing the value of this space.

3 An ethical and legal approach to SOCMINT

At the heart of any framework that can legitimise the use of any kind of SOCMINT, there must be a clear-cut distinction between SOCMINT activity that is a form of intrusive investigation, and SOCMINT that is not.

This recognises that there are times when we can legitimately control what information we give away and how it is used, but there are also times when individual control must be over-ridden. The circumstances where this can happen are based on collective decisions and assent about the state's authority.

We believe this can be achieved through the creation of two routes for government bodies to manage the safe exploitation of social media data.

The first route is non-intrusive, open source SOCMINT, which can be conducted on a similar basis to academic institutions and commercial companies, with conditions relating to anonymity and data protection.

The second route would be for the state to use specific powers of access intended to result in the identification of individuals, either by personal characteristics or the URLs associated with their social media use. This is 'SOCMINT as interception and surveillance'. Such access and use is intrusive, and needs to be governed by a series of ethical principles and a legal framework that maintains an association between harm, privacy, authorisation, agency and cause.

The existence of the two routes preserves the essentially dualistic nature of government. Under route one, it operates on the same footing as any private or academic body but under route two (investigating said information) it acts as an entity that has unique powers and sovereignty and therefore requires a specific and unique structure, animated by the six principles, in order to maintain public support.

Of course, the boundary separating route one and route two might sometimes be porous. A government official might, for example, encounter freely available information under route one, such as a manifesto similar to Anders Bering Breivik's, which requires action. In such instances, the case might then become a matter of surveillance for security rather than research for understanding, and be considered a route two activity (which, as it would be directed surveillance of open source material, would require relatively low levels of authorisation).

This approach would both avoid the possibility of officials having their 'hands tied' when encountering material requiring further action, but also ensures that any Government collection of personally identifiable SOCMINT without consent is subject to the balances and guarantees provided by the six principles and legal framework.

Route 1: Open SOCMINT – consent, control and reasonable expectation

The first route is 'open SOCMINT'. Open SOCMINT is not intrusive because its access and use is controlled by the user through the vehicle of consent. Without this user consent, open SOCMINT will not be able to:

- identify individuals
- be used as a means of criminal investigation
- puncture the privacy wishes of any user – any information accessed under route 1 is freely accessible to anyone.

In understanding what public content and private content are, we suggest the best way forward at present is to draw an analogy between the digital public domain and the digital private domain. Where social media activity is taking place in the digital 'public domain', accessing it is not in principle intrusive. Content that can be found by anyone who wishes to search for it, because it is freely and openly available (such as tweets) is, in an important sense, public. There are wider issues relating to public understanding of privacy settings, which we

believe should be an issue for the education system, but is beyond the scope of this paper.

This puts government use of open SOCMINT on the same footing as the commercial exploitation of SMA and academic research. We see such use expanding rapidly and it is in the public interest to improve government's understanding of public attitudes and responses to public policy. This is, broadly, SOCMINT for understanding.

Consent is usually therefore the answer to this loss of control over personal data. However, the precise mechanism through which consent might be expressed is not always obvious, especially when different social networking sites have a wide variety of privacy settings and consent policies. Indeed, what constitutes consent in respect of social media data is still subject to a number of intense debates in academia, and among private companies.

For open SOCMINT, then, harm is conceived not as intrusion into someone's private space, nor the wider issues of trust and implied suspicion (since neither of these would happen within open SOCMINT), but by the loss of control over the information.

Wherever possible, government ought to look at how academia is dealing with these issues. While approaches are being developed, it might be useful for government agencies tasked with these decisions to consider a test of 'reasonable expectation': would the producer of the information in question reasonably expect that it be used in this way?

Reasonable expectation can be protected through a characteristic openness of how or when this kind of SOCMINT is conducted. Since this form of SOCMINT is not secret, and does not need to be in order to be effective, it should be as transparent as possible that:

- all such collection, retention, and sharing policies are publicised and justified
- the reason why the information is collected is publicised
- the information commissioner should monitor the development of this form of information processing to ensure that it conforms

with the principles of the Data Protection Act, including being fairly and lawfully processed in accordance with individuals' rights

For this type of usage, authorisation would be for the departments and public bodies themselves, operating within the policy laid down centrally and subject to data protection legislation.

Route 2: SOCMINT as intrusive interception and surveillance, six principles and RIPA 2000

The second route would be for the state to use specific powers of access intended to result in the identification of individuals, either by personal characteristics or the URLs associated with their social media use. This is SOCMINT as interception and surveillance. Such access and use is intrusive.

How this type of SOCMINT should be conducted and regulated needs to be resolved at a legal and operational level, which in the long term needs both public and parliamentary debate. We believe that at the heart of an enduring, effective settlement exist six principles, adapted from those earlier suggested by Sir David Omand for the intelligence community in his book, *Securing the State*.[73] These principles draw on the 'just war' tradition of ethical thinking as a tested way of reconciling opposing concerns of the public good with the necessity of harming others, and thus provide a framework for testing individual cases and instances.[74]

These principles were applied by the independent Chief Inspector of Constabulary (HMIC) in his report to the Home Secretary on the recent use of police officers under deep cover posing as members of radical groups.[75] Towards a challenge similar to the use of SOCMINT – the need to have a technique available for the prevention of serious crime yet avoid a loss of public confidence in the impartiality and ethical sense of the police service – the HMIC pointed to the dangers of trying to be too precise on a subject that is liable to morph, and instead recommended a framework of principles

for the guidance of those involved in conducting and overseeing such operations.

These are the six principles that we believe could be used to guide and inform decisions by any public body using or wishing to create the capability to exploit social media covertly as a way of generating SOCMINT intelligence through route two:

- principle 1: there must be sufficient, sustainable cause
- principle 2: there must be integrity of motive
- principle 3: the methods used must be proportionate and necessary
- principle 4: there must be right authority, validated by external oversight
- principle 5: recourse to secret intelligence must be a last resort if more open sources can be used
- principle 6: there must be reasonable prospect of success

Principle 1: There must be sufficient, sustainable cause

This first and overarching principle forces the big picture to be taken into account: the overall purposes that could justify the acquisition by a public body of capabilities to gather, understand and use social media data. SOCMINT capabilities will become increasingly easy to acquire and use and affordable as technology advances. Just because it can be done does not mean that it should be done.

Indeed, experience shows that constant attention is needed to prevent the availability and cheapness of technology tempting use beyond what is reasonable and proportionate. RIPA 2000 powers have been used to control dog fouling on beaches and preventing school catchment areas being manipulated by the use of false addresses by parents. This speaks to the need for safeguards to prevent any possibility that a series of SOCMINT measures – each in themselves justifiable – together creep towards an undesirable end point: a publicly unacceptable level of overall surveillance; the possession of a dangerous capability; and the overall damage to a medium that is of obvious intrinsic value beyond security.

As we have noted already, SOCMINT can contribute towards the public good of safety and security. The increasing centrality of social media as a form of communication, the increasing capabilities available to tap into and interpret it, and the changing nature of the security threats faced all argue for law enforcement and some other public bodies to develop some kind of SOCMINT capability.

Application of the principle of requiring sufficient, sustainable cause is therefore necessary to ensure that SOCMINT remains within the boundaries required to deliver an intelligence benefit to the public, resisting bureaucratic empire building, finding ways to employ spare capacity or simply the *banalisation* of the technology available from commercial suppliers.

A schedule needs to be drawn up by the Home Office listing the public bodies with sufficient cause to be authorised to develop and use SOCMINT techniques. We would expect this to be similar to that currently in schedule 1 of RIPA 2000.

As the intrusiveness of the kind of SOCMINT increases, the causes for which it can be legitimately used must decrease. For the gathering and use of publicly available tweets, analogous to 'directed surveillance' under RIPA 2000, a broad number of causes could be considered legitimate, being:

- in the interests of national security
- for the purpose of preventing or detecting crime or of preventing disorder
- in the interests of the economic wellbeing of the UK
- in the interests of public safety
- for the purpose of protecting public health
- for the purpose of assessing or collecting any tax, duty, levy or other imposition, contribution or charge payable to a government department
- for any purpose (not falling within the above) which is specified for this purpose by an order made by the Secretary of State

However, for the interception of private messages on (for example) Twitter, analogous to the interception of private communications under RIPA 2000 part I, the legitimate causes

as provided for in RIPA part II must be national security, preventing or detecting serious crime or disorder, and safeguarding the economic wellbeing of the UK.

The national intelligence community should be able to exploit SOCMINT in support of its missions. As UK legislation at present limits the work of the intelligence agencies to national security, the detection and prevention of serious crime, and the economic wellbeing of the nation, we believe this narrower 'sufficient and sustainable cause' restriction should apply to their use of SOCMINT as well.

Principle 2: There must be integrity of motive

This principle refers to the need for integrity throughout the whole intelligence system, from the statement of justification for access, accessing the information itself, through to the objective analysis, assessment and honest presentation of the resulting intelligence. The justification for seeking SOCMINT in individual cases must be clear and not mask other motives on the part of the investigating officers. Intelligence is by its nature usually incomplete and fragmentary, and can be wrong or subject to deception. In presenting intelligence to end-users the limitations and caveats must be made clear. Nor must the decision to use (or not to use) SOCMINT, or the conclusions drawn from it, be influenced by local or national political considerations or media pressure.

Principle 3: The methods used must be proportionate and necessary

There is a well-established principle in law enforcement that the extent of harm likely to arise from any specific action being taken should be proportionate to the harm that it is being sought to prevent. This principle is well known from its common law application in the doctrine of minimum necessary force (only using such force as is necessary to achieve the lawful objective). The late Professor RV Jones coined the term 'minimum necessary intrusion' as the equivalent rule governing domestic surveillance.

In the UK, a check on the principles of proportionality and necessity are built in to the system of authorisation of intrusive intelligence gathering. The level of authorisation (from police superintendent to senior judge, from security service officer to the secretary of state) depends on the intrusiveness of any intelligence method – the extent to which each intervention might be held to puncture article 8 of the Human Rights Act 2000, the right to respect for one's private and family life, home and correspondence.

In assessing proportionality, the extent of intrusion has to be assessed. That would mean a lower threshold for covertly monitoring material that the user has not restricted than in cases where they had only a limited list of friends who had access, or where they used a system such as BlackBerry that required a PIN. Published guidance notes under RIPA 2000 could be a vehicle for providing a consistent approach across government bodies and law enforcement agencies to these judgements.

Principle 4: There must be right authority, validated by external oversight

Having an adequate paper trail (or its electronic equivalent) for key decisions is essential for confidence of staff and ministers, and for the operation of investigations and redress in any cases of suspected abuse of powers. The warranting and certification processes laid down in statute ensure that commands are lawful. Self-regulation must, however, be recognised as the main guidance system for everyday operations based on an ethos that has internalised ethical principles and respects the need to operate at all times within the law and thus to seek the necessary approvals.

There is therefore a general principle that there must be an audit trail for the authorisation of actions that may carry moral hazard with an unambiguously accountable authority within a clear chain of command. We believe this principle should apply to any intelligence operations, including SOCMINT. This is an important way in which proportionality and accountability is realised in practice.

Given the novelty of social media as a mode of communication and possible form of intelligence, the extent to which existing bodies would take responsibility for oversight of the SOCMINT area is not yet clear. It is therefore important that the Home Office and Ministry of Justice establish how far existing regulators and oversight bodies, including the Information Commissioner, Independent Safeguarding Authority (ISA) Commissioner, Surveillance (RIPA 2000) Commissioner, Independent Police Complaints Authority, HMIC and policing and crime commissioners in the future, should be given responsibilities and a set of consistent policies to deal with the issues SOCMINT poses.

Principle 5: Recourse to secret intelligence must be a last resort if more open sources can be used

Because of the moral hazards of all intrusive secret intelligence gathering methods, those authorising such operations should ask whether the information could reasonably be expected to be obtained through other means, ranging from fully open sources to information freely volunteered from within the community.

Applying this principle to SOCMINT, less intrusive forms should be preferred to more intrusive covert forms. SOCMINT should be based wherever possible on the clearest possible mandate of informed consent. The most preferred route is to access explicitly 'consenting' information from the online community, for example crowd-sourced information that has been explicitly volunteered on a particular Facebook wall or hashtag. Recourse to covert intelligence gathering, including via exploitation of social media, should be confined to cases where the information is necessary for the lawful purpose of the operation and cannot reasonably be expected to be gained by other means.

Principle 6: There must be reasonable prospect of success

Even if the purpose is valid (principle 1) and the methods to be used are proportionate to the issue (principle 3) there needs to be a hard-headed assessment of the consequences of failure. This includes risk of harm to suspects who turn out to be innocent, risks of collateral damage to others and those to whom a duty of care is owed, and, not least, the risk to future operations and institutional reputations if the operation were to go awry.

The 'success' of intelligence is not the information or even secrets that it collects, but the value it adds to decision making. Indeed, the justification for creating SOCMINT capabilities and applying them, with all the recognised hazards this entails, is that it contributes to the public good of safety and security. It is therefore morally imperative that SOCMINT operations present a reasonable chance that they will yield actionable, useable intelligence that contribute to consequential decisions, such as deploying emergency services to the right place at the right time.

The capabilities to capture, analyse and understand social media data – those crucially required for SOCMINT – are nascent, emerging fields. As with any developing area of knowledge and research, methodological flaws exist, from how data are accessed and analysed, to how they are interpreted and used. These shortcomings in understanding need to be rectified so that Government can use social media information to contribute meaningfully to insight and therefore reduce ignorance.

Failure to develop and use SOCMINT properly would jeopardise security, lead to bad decisions and strain the public's trust in the state to use its powers in support of public goods.[76] The next chapter considers this principle in further detail.

Short term: SOCMINT and RIPA 2000

In the short term it is important that intrusive SOCMINT is governed by some form of regulatory oversight. This is important for legal and ethical reasons, as well as ensuring there is public confidence in the way SOCMINT is conducted. For this reason, we believe the most appropriate legislation at present is RIPA 2000. Below we consider how SOCMINT might slot into the existing framework set out by the Act.

Broadly speaking, the legislation that governs the state's access to people's personal lives is based on a distinction between public and private space. There is an outer public space where the expectation of privacy is low (for example justifying the ubiquity of CCTV) and increasing privacy is understood spatially: bugging someone's home or car is considered more intrusive, and has greater limitations imposed on it, than bugging a public space.

As the type of surveillance becomes increasingly intrusive, RIPA 2000 imposes three vital and increasingly narrow conditions onto it: the agencies that can conduct it, who must authorise it, and the reasons why the surveillance can be legitimately conducted.

At the less intrusive end of the RIPA 2000 scale, the Act lists in a schedule the public bodies that can employ directed intelligence methods. These bodies include, for example, local government organisations where less intrusive techniques are used for public health and child protection. At the most intrusive end of the scale, in part I of the Act only the intelligence agencies and law enforcement can use warranted interception to gather intelligence.

Maintaining the connection between the scale of intrusion on the one hand, and cause, legitimate agency and authority on the other, is vital to balance the possible good of the collection and use of SOCMINT with the possible harm.

However, the place of SOCMINT methods within the range of techniques covered by RIPA 2000 is not immediately apparent, because a spatial understanding of privacy cannot be applied to social media, with its characteristic of sometimes being considered a public, sometimes a private, and more often a quasi-private space. People often share what would usually be considered private things about their lives in ways that are unprecedentedly public, sometimes with unexpected or unrealised consequences.

Although some forms of SOCMINT can be placed at the low end of the scale in comparison with other techniques open to intelligence agencies under RIPA 2000, the scale of intrusion entailed by SOCMINT access varies greatly, and SOCMINT as

we are defining it does not fit into just part I or part II of RIPA 2000. To gather and analyse a suspect's open tweets looks similar to the 'directed surveillance' of a conversation occurring in a public space. However, to gather and analyse someone's Facebook messages seems closer to reading their private correspondence.

SOCMINT should follow the same general principle: that the greater level of intrusion and corresponding likelihood of harm, the greater the justification, authorisation and limits on use. How to determine what types of SOCMINT would fall under which parts of RIPA 2000 requires a conceptual framework of what constitutes privacy on social media and the sorts of harms associated with breaching that privacy (and thus the degree of intrusion entailed by a SOCMINT intervention and where it might fall under RIPA 2000).

This is not simple. One way to approach this subject is to draw a series of analogies between different kinds of digital space and physical public and private spaces, and relate intruding into these spaces with standing typologies. Taking this analogy, it is possible to define a series of different degrees of SOCMINT intrusions.

Where social media activity is taking place in the digital 'public domain', the access and use could be considered a very low level of intrusion. Content that can be found by anyone who wishes to search for it, because it is freely and openly available (such as tweets), is in an important sense public. The ability to collect and analyse a named individual's tweets in this way is similar to 'directed surveillance' – carrying out surveillance on an individual in a public space. This is surveillance of a specific individual or individuals, and falls under the authorisation stipulations of RIPA 2000 part II to manage its potential harm, but is not technically intrusive and not comparable to interception as regulated under part I. Under current guidelines, directed surveillance of people in a public space requires a relatively low level of authorisation and can be undertaken by a number of agencies. A similar approach might be taken for SOCMINT of this type.

However, other social media activity takes place in a more private space, where communication and data sharing is not publicly available, and is only accessible to people selected by the user. Knowing the existence (but not the content) of private communications is similar to the current 'access to communications data' provisions in RIPA 2000. This form of surveillance requires higher standards of necessity, although it can still be authorised internally.

More intrusive would be to actually *enter* that private space to access information. Examples include people using social network sites to send messages directly to other individuals (or indeed everyone in their network), Direct Messaging (DM) on Twitter, or sharing personal photos and details on friendship networks. This is more analogous to a private space, or the site of a private conversation. Accessing these types of data might be considered a greater intrusion into privacy, and therefore should be considered as interception and require higher levels of authorisation under RIPA 2000 part I, with narrower grounds for justification and necessity, with a smaller number of agencies able to conduct it.

Finally – and more intrusive still – would be the covert collection of information, such as an undercover agent joining a digital group to monitor activities. Although technically very easy to do, this might be viewed as analogous to an undercover agent joining an offline group or organisation, which is regulated for the police under section 93 of the 1997 Police Act. Indeed, according to the HMIC report into undercover policing (which recommended Sir David Omand's six principles), the use of undercover agents by the police should require higher levels of authorisation than are currently in operation.

These examples are far from comprehensive, of course, and whether in all cases the use could strictly fall under the present terms of RIPA 2000 remains to be determined. However, they are designed to illustrate a way of approaching the difficulty of mapping SOCMINT activity onto existing RIPA 2000 and other relevant legislation until a more considered set of decisions can be taken.

4 Turning data into insight

As principle six states, the underlying justification for intelligence work – both open source and secret – is that it helps improve decision making through the 'the reduction of ignorance', by using timely, sufficiently accurate, reliable, useable information.

For information to be considered 'intelligence' it needs to meet certain thresholds of how it is gathered, evidenced, corroborated, verified, understood and applied. Different sources and kinds of information have developed signature ways of meeting this challenge. For example, open-source intelligence (OSINT) triangulates reliable sources; human intelligence (HUMINT) might consider the track record of the agent; imagery intelligence (IMINT) needs to pay attention to the technical characteristics of the collection platform; and signals intelligence (SIGINT) would need to understand the context of the language used. All source intelligence assessments try to get an overall picture on the basis of these different types and reliability of contribution.

Whenever a new form of technology emerges, it takes some time before rigorous and robust systems of capture, analysis and interpretation are developed. Typically, new forms of data collection should first help with establishing situational awareness: 'what' is happening, 'when', 'where' and involving 'whom'. The second step is to use the intelligence to examine explanatory hypotheses and to select the most convincing explanation consistent with the available data of 'why' the situation is as it is and 'what for'. This can ideally then lead to the final level of being confident enough to make predictive judgements of 'what next' or 'where next', and modelling of 'what if' the authorities were to react in different ways.

The key challenge is that there is not yet the human and technological infrastructure to turn the large volumes of social

media data into timely, reliable, decisive SOCMINT. The consequence is that, during the August 2011 riots for example, the surging social media output did not fit into the police's formal process to evaluate reliability, did not therefore count as intelligence, and consequently was not acted on.[77]

This chapter cannot cover all the many issues that obtain, and the issues it does cover certainly are not necessarily relevant for all the different branches and kinds of SOCMINT that might be produced. However, some of the most significant and general difficulties in generating insight from social media can be seen by examining how far (or not) SOCMINT fits into the traditional intelligence cycle, in particular the functional steps of collection, processing and analysis, dissemination and feedback to influence future collection.

Collection: data access

One of the difficulties of SOCMINT is not a paucity of data, often the key problem in the collection of secret intelligence, but a deluge. The term 'access' is preferred over 'collection' to indicate that we are dealing with a very different process from that of traditional intelligence gathering. During the week of the August 2011 riots, for example, millions of riot-related tweets were sent, consisting of news stories, rumours, reactions and indications of criminal intent.[78] Knowing what data ought to have been accessed and analysed – sorting the wheat from the chaff – is the critical consideration.

One useful way to approach this challenge is to draw on how traditional methods of data collection deal with large data sets. In statistics, researchers collect samples, which are a manageable amount of data that represent the population being researched. The reliability and validity of inferences or extrapolations made depend on the quality, especially representativeness, of the sample collected. Over the past century, statisticians have developed techniques to use small data sets to make general conclusions, particularly through the use of randomised sampling. Simply put, inferences and conclusions can only be reasonably drawn if one knows how the sample was constructed

and the data collected, and what this means about the inferences that are then made.

The broad problem is that social sciences have not developed an approach to robustly sample social media data sets. Only a very limited amount of work has been done to develop different types of sampling for automated systems of data collection.[79] More attention has been paid to methodologies that produce a large sample (something that computational approaches are good at delivering), rather than methodologies that produce a representative one (something computational methods are less apt at delivering). Moreover, the emerging, highly technical and computer science-driven developments in social media sampling practices, including 'forest-fire' (wherein links and nodes 'burn' outward from a random seed to create the sample), user-activity and location-attribute-based techniques, have had very little uptake within the social science community.

Very little research based on social media data sets acknowledges the sampling frame applied, and how this might limit or bias the results that are drawn. Typical data acquisition strategies remain 'samples of convenience' or 'incidental sampling', which means the most readily available or easily accessible – rather than the most representative – are collected.[80] For obvious reasons this type of sampling limits the strength of conclusions drawn.

One prominent example is the recent *Reading the Riots* collaboration involving the *Guardian*, a number of British universities, freelance journalists and community researchers. The project gathered 2.6 million tweets about the August 2011 riots and drew a number of conclusions, including that there were very few examples of Twitter being used to express or encourage criminal activity. However, the data set of tweets was collected using around 150 'hashtag clouds', which means only tweets that included an identifying hashtag, such as #londonriots, were collected and analysed. It is possible, however, that people who use a hashtag when tweeting are not representative of all tweets about the riots; for example, they might be less likely to talk about criminal activity because hashtags are usually employed by users to disseminate tweets to

a wider audience. In statistics, this is known as 'missing at non-random data', which means certain data might be systemically absent as a result of the sampling method. This is considered a serious problem when drawing conclusions, because when people are absent from a dataset for a reason other than chance, they share a related and likely important trait (or traits) that could have a substantial impact on the research findings.

Unlike in traditional social science research, technical considerations might also affect the quality of sample. For example, in the case of Twitter, the publicly available application programme interface is limited to 150 requests per hour, going up to 20,000 when it is 'whitelisted'.[81] While this can capture an enormous dataset by traditional social science standards, it can only capture a small amount, and not an automatically representative sample, of the total number of tweets. Researchers can gain access to larger proportions of the tweet-feed. The 'firehose' gives access to all tweets, the 'gardenhose' gives 10 per cent, and the 'spritzer' gives 1 per cent. Unfortunately, precisely how these different access levels affect the quality of the data, and what sorts of systemic bias they might hide, are not fully known – and very rarely stated.

For the interests of public security, what are especially required are data acquisition strategies that allow online information to be related to offline phenomena. Demographic representativeness is key for this. The people who use Twitter and Facebook tend to be younger, richer, more educated and more urban than the population in general.[82] However, when looking at demography, it is not only the general population that is important, but also the community that accounts for the information that is gathered. An enduring rule of online social collaborations is that 80 per cent of the user-generated content for any given site will tend to come from a highly productive 20 per cent of users.[83] A 2010 study of Twitter found this to be broadly true: 22.5 per cent of users accounted for around 90 per cent of all activity.[84]

There is no simple solution to these problems. In the long term, it is important for social and computer science disciplines to jointly develop new forms of sampling techniques. In the

more immediate short term, analysts using SOCMINT must interrogate how the data were collected, especially the sample, whenever drawing conclusions.

Processing and analysis

The analysis of SOCMINT – trying to draw meaning from the data themselves – presents even greater challenges. Because social media data sets are so large, a number of broadly computational approaches have been developed to infer and extract 'meaning' automatically, without the routine presence of a human analyst. The most important approach is a variant of artificial intelligence – 'machine learning' – where algorithms are taught to recognise patterns and therefore meaning within pieces of information that human beings need therefore never see.

Machine learning has a number of important applications, from identifying clusters and anomalies in large data sets to the extraction of semantic information from text. A particularly important application is 'sentiment analysis', where an algorithm looks for certain qualities and properties in a piece of text (a form of 'unstructured data') that it considers to correlate statistically with a certain emotion, or 'sentiment'. Once human input has defined what sentiments are being searched for, and what textual examples of these sentiments are, the algorithm is able, with varying degrees of specificity and accuracy, to classify enormous volumes of data automatically on the same basis. Sentiment analysis has been applied for a number of aims, from measuring Twitter users' feelings towards political parties to predicting the future of box office revenues.[85]

The ability to extract automatic meaning from text opens many research opportunities, and social researchers can now contemplate handling bodies of information on a previously unmanageable scale. However, social media analytics is currently much better at counting examples of online human behaviour than explaining *why* they are and what it might mean.

To make this sort of sense of any form of communication, context is critical. A central tenet of all semiotics and linguistics is that language is textured: the intent, motivation, social

signification, denotation and connotation of any utterance is mutable and dependent on the context of situation and culture. The accuracy of any interpretation depends on a very detailed understanding of the group or context that is being studied. For example, most groups of people use vernacular and group-specific language that a generic or standardised sentiment lexicon or thesaurus would often misinterpret.

However, because automatic data collection is required to process the sheer volume of data now available, many of the contextual cues – the thread of a conversation, information about the speaker, the tone of the utterance and the information about the speaker – are often lacking in analysis of social media data. Therefore utterances have to be abstracted out of the wider situational, contextual and cultural picture – what we would call their 'naturalistic setting'. The act of 'scraping' a social media platform – such as collecting tweets or Facebook posts – usually does not collect the utterance's position in a social network (such as whether they were talking to their friend) or a conversational network (such as whether the utterance was a heated rebuttal in an argument).

Context is also shaped by the norms and mores of the medium we are using. A number of studies are beginning to identify norms, rules and behaviours that dictate communication via social media that differ in significant ways to how people might communicate offline. Some studies for example argue for an 'online disinhibition effect' – that the invisible and anonymous qualities of online interaction lead to disinhibited, more intensive, self-disclosing and aggressive uses of language.[86]

Identification with groups or movements has also changed. Traditional forms of membership to a group or a movement are relatively intense, often involving subscription fees and membership lists. For many online groups, however, a single click of a mouse is sufficient to express some form of affiliation. This is a more ephemeral, looser and possibly less involved form of affiliation. Indeed, one recent study of 1,300 Facebook fans of the English Defence League found that only three-quarters considered themselves 'members' of the group, and only one-quarter of those had ever actually been on a march.[87]

Sometimes even the specific social media platform is important in shaping behaviour. Twitter, for example, limits users to just 140 characters, which has resulted in a new lexicon of shortened and abridged words and entirely new phrases. Hundreds of twitter glossaries exist to introduce new users to the twitter-specific dictionary. Furthermore, it appears that Twitter might be particularly given to the spread of rumour and misinformation. Sharing information – such as retweeting it or forwarding it to friends – is not necessarily an indication of agreement. For example, *Reading the Riots* illustrated how quickly rumours of London Zoo being looted (including a story about an escaped tiger in Primrose Hill) spread across the network.

Taken together, these phenomena constitute the rapid emergence of distinct social media sub-cultures, which are developing new language and new uses of language in clearly distinct and often non-literal ways.[88] Indeed, a new branch of sociology – digital sociology – is devoted to understanding these cultural consequences of the use of internet technologies.

When context is not considered, there can be profound consequences and potential for misinterpretation. In 2010, Paul Chambers declared to his 650 Twitter followers his intention of 'blowing [Robin Hood] airport sky high!!'[89] Undoubtedly in jest, his conviction for the 'menacing use of a public communication system' under the Communications Act 2003 has attracted wide public criticism. Jonathan Bennett, the district judge, noted the 'huge security concerns' within the context of the times in which we live, but perhaps not the Twitter-specific situational and cultural context of the utterance.[90] In a similar case, Leigh Van Bryan and Emily Bunting were denied entry to America after tweeting 'free this week for a quick gossip/prep before I go and destroy America? x'.[91]

Many of the technologies that have been developed in the private sector by online advertising and digital reputation industries have been created for the needs of these industries: to gain a general understanding of attitudes toward a product or whether a new advertising campaign is creating a 'buzz'. But context is especially important for security and intelligence work because of the need for a high degree of confidence in

information, the value of predictive and explanatory analyses, and the consequences of making error.

Although there are no simple solutions to these difficulties, some steps are possible. First, big data computational tools must become more 'human-sized' – sympathetic to the human subject they wish to measure. Sentiment analysis must involve analysts and experts who understand the norms and behaviours of the groups involved. Second, any analysis of social media data sets should always be based on an understanding of the medium itself: the specific online culture, language and behaviour. Project Raynard, for example, stresses the importance of establishing norms in online environments, and then look for deviations from that norm. Any agency using SOCMINT must recognise the analytical and interpretative limitations of the field, how they reflect on the kind of insight that can be drawn, and the kind of decisions that can be made on the basis of those limitations.

Dissemination

The effective use of intelligence from social media data sets also depends on it getting to the right people quickly, securely and presented in a format that makes sense to strategic and operational decision makers. Depending on the purpose of the SOCMINT, this may range from a footnoted, caveated and in-depth strategic analysis paper, to the operational use of single-screen, real-time visualisations of data available on portable devices.[92]

However, several general challenges will need to be addressed. First, SOCMINT dissemination must reflect the general difficulties in using SOCMINT: its complexity, scale, dynamism, and – given the problems outlined above relating to both access and interpretation – any presentation of data needs to be presented with new procedures and caveats.

Second, SOCMINT dissemination must slot into existing intelligence channels – police, emergency service response, the Security Service, the Joint Terrorism Analysis Centre, Cabinet Office Assessments Staff and so on. However, this requires

specific training for gold, silver and bronze commanders and additional training for frontline officers who could benefit from the daily use of such intelligence.

Third, for security purposes, SOCMINT dissemination and retention must be handled as intelligence. Existing controls must be applied to ensure the safekeeping of SOCMINT data and regulating their dissemination, including overseas. The unregulated dissemination of SOCMINT data would risk jeopardising public confidence in this form of intelligence. Whether lost, insecurely held, or subject to hostile access, as government increases the amount of personal information held on its servers, the potential for various data compromises, and the harm it might cause to public confidence, will inevitably grow.

Similarly to the sliding scale of authorisation required under RIPA 2000 for the access of personal data, we need equivalent safety and security settings which regulate the use of SOCMINT, where the greater the risk of breaches of personal data rights, the more secure the levels of storage and access required.

Effective application of data visualisation techniques will be required to render complex and often interlinked intelligence in a way that is intuitively comprehensible, but conserves the networked nature of the information. More experience of using such SOCMINT data analysis techniques is needed in order to draw up detailed rules and regulations for its safe management.

Validation and use

The way SOCMINT can add value relates to how the operators – such as frontline police officers – will actually use the information, and how they ought to interpret and act on it (such as deploying reserve forces in the build up to a march). In addition to the many methodological hurdles that stand in the way of the responsible interpretation of data, the social media being monitored is itself prone to contain misleading information of a polemical nature, which may involve the recirculation of selective half-truths, mistakes and outright distortions.

Validation of SOCMINT intelligence is therefore an important function for the SMA analyst.

One risk that must be accounted for when considering validating SOCMINT data is the risk of engineering the 'observation effect': the tendency of individuals to change their behaviour if they believe they are being observed. In 2009, the LSE's Policy Engagement Network warned of this 'effect' in a briefing paper responding to the then-government's Interception Modernisation Programme. The report feared that when the public became aware that communications data were being collected and collated, there would be a risk that 'it will generate a chilling effect on the individual's right to free expression, association and might dissuade people from participating in communications transactions'.[93] Evidently, this would limit the effectiveness of social media and online communications data as sources of intelligence.

Related to this issue is the problem of 'gaming' – the deliberate use of this media as a means of misleading or confusing an observer, in this case the law enforcement agencies. In the context of intelligence work, this problem is not new, and as early as the Second World War German intelligence operatives in occupied France had developed the tactic of 'Funkspiel' as a means of transmitting misinformation to the enemy. Yet while gaming has long represented a problem in intelligence work, the nature of SMA intelligence makes deception easier, given the ubiquity of social media, its widespread use and the democratisation of computer coding and technical know-how. In a recent example, a leaked cache of emails allegedly belonging to Bashar al-Assad indicated that an aide, Hadeel al-Assad, posted pro-regime commentary under an assumed Facebook identity that was mistaken as genuine and given international coverage by CNN.[94]

For these reasons, there must be a thorough (yet sufficiently rapid) process to ensure that an item of SOCMINT intelligence can, as far as possible, be validated before it reaches the user. Unlike other forms of single-source intelligence such as HUMINT, the validation of SOCMINT has to take place further up the 'food chain' from the functions of access and processing

of data. Validation of SOCMINT can only be done effectively when all sources of intelligence, including open source material, can be brought to bear.

First, this validation process must take the form of a reporting framework that rates the 'confidence' in any piece of freestanding piece of SOCMINT. By pointing out potential vulnerabilities and biases in the acquisition and analysis of the information, we may gauge the importance of the information collected and caveat the conclusions that may be drawn.

Second, we must be able to relate SOCMINT to other kinds of evidence to produce an overall picture – the equivalent of an 'all-source assessment': the value of SOCMINT relative to other forms of intelligence must be evaluated and the ways in which various types of intelligence can be used in tandem needs to be investigated. The crucial points here are the exact application of SOCMINT in a given circumstance and its 'strength' in relation to other forms of intelligence. To complicate the issue, both will of course vary according to the situation. For example, in identifying broad societal trends, SMA intelligence may well be able to provide reliable and actionable intelligence, whereas in the context of a riot or crowd control, a handful of tweets, for example, are unlikely to be of much operational validity in the absence of corroborative evidence gathered as HUMINT or other types of intelligence. The fundamental point is that methods of evaluating the quality of SOCMINT have to be developed if it is to become an effective source of information, and a framework which in effect ranks various types of data and how they are to be applied to each other will also be necessary.

A number of strategies will be useful to create processes to validate SOCMINT. More methodologically mature forms of offline research can be conducted in parallel to SOCMINT projects to allow the results to be compared. For example, it would be especially useful to establish rules about how online phenomena maps onto offline behaviour. Retrospective analysis can also be used to quantify SOCMINT accuracies and diagnose instances where accuracy was confounded. In addition to the specific validation responsibilities placed on the agency that

collected the intelligence, there needs to be a very general up-skilling of all the branches of government that might be involved in this work. It will be impossible to use this medium without analysts, police officers or judges who understand its norms and mores.

It also requires an accompanying change in the skills of analysts and others using it, to ensure they are able to make valid and reasonable inferences from the data, based on an understanding of how offline and online norms and behaviour can differ. Ultimately, the value of SOCMINT can only really be understood through using it. Understanding will slowly emerge as SOCMINT is trialled– we must expect varying degrees of success in different contexts.

Conclusion and recommendations

SOCMINT presents both challenges and opportunities for policing and intelligence agencies. As social media becomes an increasingly important form of communication, SOCMINT can help these agencies better serve the public and protect society. However, social media – and the tools potentially available to government to access and interpret both open and secret information – is changing quickly. Public norms surrounding privacy and consent are also being transformed. For SOCMINT to become an important and valuable part of intelligence work – capable of producing high quality information and on a sound footing that is publicly accepted – there are a number of legal, ethical, interpretive and analytical challenges that need to be addressed.

Social media science

UK policing and intelligence agencies must become world leaders in the ethical, effective and legitimate collection and analysis of social media data. This will require new relationships with industry and academia, and concerted, long-term investment to build technological, methodological and presentational capabilities. The opportunity that the explosion of social media use offers is remarkable. While the continued development of automated systems of access and analysis is vital, computers alone are not the answer.

The origin of the main risks in using information from social media arises from a lack of interaction between the humanities and the statistical and computational disciplines. Those disciplines best equipped to understand and explain human behaviour – the social and behavioural sciences, political science, psephology, anthropology and social psychology – have

not kept pace in relating this insight to the big data approaches necessary to understand social media. Conversely, these very same big data approaches that form the backbone of current SOCMINT capabilities have not used sociology to employ the measurements and statistics they use to the task of meaningfully interpreting human behaviour.

SOCMINT must evolve to become a new, applied, academic discipline: social media science. The transition from social media analytics to social media science requires a much more meaningful and intensive fusion of the computational, technological and humanities approaches. Only through this fusion can data-led explanations of human behaviour also be humanistic explanations of human behaviour.

Ethics, law and public acceptability

Technology and capability is only half the picture. Consistent with Britain's modern approach towards national security, the work of the intelligence agencies and law enforcement is made easier when the public broadly understands and accepts why, when and with what restrictions intelligence and policing work is undertaken.

Without an explicitly articulated approach towards generating SOCMINT based on respect for human rights and the associated principles of accountability, proportionality and necessity, there is a serious risk that this vital confidence and trust will be undermined. Indeed, a number of polls suggest that the British public are concerned about this issue: 60 per cent of British adults for example feel that police access to data on social networking sites should have some form of restriction or be in some way regulated.[95]

Government faces an awkward paradox in establishing the conditions and structures to promote public understanding and acceptance of the actions of law enforcement and the intelligence agencies. To be effective, the sources and methods of the police and intelligence community must be kept secret. In intelligence and policing, the UK's recent approach to this paradox has been to debate, and ultimately legislate in

Parliament, the legal framework to govern this work, and to create independent (but sometimes quite necessarily closed) bodies that monitor and review their activities, such as the Parliamentary Intelligence and Security Committee, a number of senior judicial commissioners, and the independent police commissioner. It is through these mechanisms that the necessary work of police and intelligence agencies sometimes remains secret (and effective) while being accountable.

Retaining that balance is crucial to satisfy perhaps the single most important measure when making decisions about intelligence work, the 'public acceptability test': would the secretary of state or senior officer or official authorising any intrusive intervention be happy to stand in front of the public and justify his or her operational decision in the event of it being made public? SOCMINT that is carried out consistently according to the same principles that govern the UK's intelligence and policing agencies now (of proportionality, necessity, operating within the law, and accountability) maximises the likelihood this test is satisfied, and that the significant capability of SOCMINT will be harnessed for public good, economic, social and security.

Overall, we believe that Government must construct and articulate a strategic, systematic and comprehensive 'big picture' of its use of SOCMINT, rather than allow a tactical and piecemeal one to implicitly emerge. This will require ministerial and political engagement. Public understanding and acceptance of any settlement on the use of SOCMINT relies on the processes of political and public debate to raise the important issues, manage competing interests and perspectives within society and, where necessary, criticise proposed solutions. Consistency of application will be important: we doubt that this is an area where individual policing and crime commissioners should be able to limit (or demand) the application of SOCMINT by their force, for whose use chief constables should be operationally independent within national policy. The potential use of SOCMINT for such serious crime as terrorism suggests that the national policy should cover all of the UK, as does RIPA 2000 with appropriate delegations to the devolved administrations.

Recommendations

We offer these general recommendations to help guide any agency – from central government to local police forces – to make these difficult decisions.

We recommend the police and intelligence agencies use social media as a form of intelligence and insight, but it must be based on a legal footing, transparency over use and purpose, regulation, accountability and public understanding.

Social media now potentially allow government agencies to better serve the public and protect society. Any modern intelligence and policing services should be expected to use social media to help them fulfil their responsibilities. However, security and intelligence work in general is predicated not only on public consent and understanding, but also on people's and communities' partnerships and active participation. There is also the danger that SOCMINT could result in a chilling effect on the use of social media itself, which would have negative economic and social consequences for the country as a whole. Therefore, the Government needs to articulate its approach to SOCMINT and place it on the same footing as other types of intelligence and security work.

The use of SOCMINT needs to be based on a distinction between the public digital space and the private digital space. The greater the degree of intrusion into the private space requires greater cause, oversight, legitimate agency and authority.

Not all privacy breaches into social media are the same. While some content is clearly public, and can be undertaken according to existing data-protection law, other content is more analogous to private communication, and should be subject to similar restrictions. Any decisions should be guided by the following six principles:

- principle 1: there must be sufficient, sustainable cause
- principle 2: there must be integrity of motive
- principle 3: the methods used must be proportionate and necessary

- principle 4: there must be right authority, validated by external oversight
- principle 5: recourse to secret intelligence must be a last resort if more open sources can be used
- principle 6: there must be reasonable prospect of success

Government should take a two-route approach to the use of SOCMINT, making a clear distinction between open source non-intrusive SOCMINT and intrusive or surveillance SOCMINT.

Route one would be open source non-intrusive SOCMINT, which can be conducted on a similar basis to non-state actors, such as universities and commercial companies. This should be tightly bound with conditions relating to anonymity, data protection or based on the full consent of the producers of that information. This might include such activity as openly crowd sourcing information through Twitter or Facebook to gain situational awareness in the event of public disorder, or gauging general levels of community tension. This type of activity would not be used to identify individuals, or as a means of criminal investigation and should not puncture the privacy wishes of any user. As such, this would not fall under existing legislation that governs intrusions into people's privacy: individual departments and agencies would be responsible for how to undertake this type of activity. Inevitably it is possible that, while undertaking route one SOCMINT, criminal or possible criminal activity is found. In the event, this should be then transitioned into the second route, set out below.

Route two SOCMINT is the exercise of state-specific powers of access intended to result in the identification of individuals and access to private information. This is SOCMINT as intrusive surveillance and interception. Accessing social media could range from relatively minor intrusions (such as collecting publicly available data about specific individuals) to more significant intrusions, such as intercepting and reading personal communications. Such access needs to be governed by a series of ethical principles which we set out below, and animated through a legal framework that maintains an association between harm, privacy, authorisation,

agency and cause, such as limits on the number of agencies permitted to undertake it. In the immediate term, this type of activity could be governed by relevant legislation contained in Parts I and II of RIPA 2000, although we believe an interdepartmental review and a Green Paper are needed to reach a sustainable settlement based on public consent and acceptance.

This general approach leads to a series of more specific recommendations.

The Government should undertake an interdepartmental review of current legislation – notably RIPA 2000 – and existing systems of oversight to determine what applies to SOCMINT now.

There needs to be public and parliamentary debate about the use of SOCMINT. However, in the immediate term it is important to ensure there is some form of oversight and regulation governing its use. We believe RIPA 2000 is the most appropriate legislation currently available. An interdepartmental review must review what types of SOCMINT might fall under RIPA 2000 parts I and II, and the relevant degrees and type of authorisation required. Existing mechanisms of oversight for all intelligence and policing work, including the Parliamentary Intelligence and Security Committee and the independent police commissioners, need to determine how SOCMINT should relate to their current procedures and operations. We recommend that as far as possible the association of privacy, authorisation, agency and cause as exists under RIPA 2000 be maintained for SOCMINT.

The Government should, in the light of its emerging policy approach, conduct research and consultations on legislation governing social media data access, analysis and usage in order to make systematic recommendations for consideration by Parliament, for example by tasking the Law Commission for England and Wales.

The Government should publish a green paper about how it plans to use and manage social media analysis in the public interest, including for the purposes of public security.

The Government should publish a green paper as soon as possible on how it plans to manage over the next few years the opportunities offered by social media analysis and the moral and legal hazards that the generation and use of SOCMINT raises. This needs to include definition of the potential harms that SOCMINT pose, how harm can be judged and measured, and how these risks can be balanced and managed. It is important that the Government provides a position on the practicalities and specifics involved, including information on the relationship between the Government, ISPs and social network providers, the scope of information collected, the bodies authorised to collect it, who will have access to certain capabilities and with what safeguards. This might include attitudinal polling and behavioural research to try to clarify emerging societal norms on what is public and what is private, and whether these categories are still useful. Given the difficulties in objectively assessing privacy norms in the abstract, specific polling using social-media-based scenarios should be considered. The green paper might include proposals about how to balance necessity and proportionality, the oversight structures that are necessary, how and when SOCMINT should be stored and shared, and how the relationship between government and ISPs should be managed. The green paper should also include plans for public education on SMA. A schedule listing public bodies with sufficient cause to be authorised to develop and use SOCMINT techniques needs to be drawn up by the Home Office.

The Government should create a single, networked hub of SOCMINT excellence.

Without coordination, a number of different agencies might develop technologies and capabilities in isolation, which might be redundant, overlapping or not inter-operable. A single hub of excellence for the police service should coordinate these activities and develop new methods of access, interpretation, verification, presentation and dissemination, working in close cooperation with the UK intelligence community.

The Government should set up an independent expert scientific and industrial advisory panel to look at future developments in SOCMINT use.

Social media platforms and the tools available to analyse them are changing at a remarkable speed. Agencies need to be ready to adapt and change as new opportunities and threats arise. Therefore, it is important that an independent and cross-disciplinary set of experts can advise government agencies on their use of the latest technologies. Such a panel could be established under the auspices of the UK Security and Resilience Industry Suppliers' Community (RISC), which already brings together industrial trade associations (including Intellect) and academia to work with government on issues of national security.

Government needs urgently to coordinate the development of operational capability of SMA for law enforcement and intelligence through contracting with private technology companies and academia.

To maximise the potential of SOCMINT and to become world leaders in it, the Government must commission development work with a wide range of external partners. Many of the latest advances in technological analytics capabilities are from the private sector, yet much of the available and best humanities expertise is situated in universities. The Government should bring these groups together with the public sector, through a formal structure of engagement to provide early wins. A pre-procurement structure similar to the 'Niteworks' programmes as currently run by the Ministry of Defence would be a good start.

Government needs to produce a robust approach to the safe storage and access to the data it collects, and communicate that approach with the public.

The collection of large sets of data about citizens brings the risk that the data are not stored safely and could be leaked or accessed by third parties. It is vital that the levels of security and access are proportionate to the possible harm to citizens' right to

privacy in the event of a breach. Clear policies need to be put in place to protect against data theft, and these policies must be made known to the public.

There should be development of training and doctrine for a SOCMINT culture for practitioners.

Technology will only get you so far; analysts are still essential. People who understand online culture, behaviour and norms, and who can apply context, are essential for ensuring SOCMINT is used carefully and accurately. The potential impact of social media on police and intelligence agencies is far more profound than SOCMINT capability. It could become a valuable way of communicating with the public, but also brings risks such as breaches to data protection law, the identification of individual officers and their families, or reputational issues relating to how to respond to emergency calls or requests made through social media platforms. Guidance to set out some principles and approaches to deal with these challenges and opportunities is badly needed.

Notes

1 For a visual sense of the entire span of social media platforms, see B Solis and JESS3, *The Conversation Prism: The art of listening, learning and sharing*, www.theconversationprism.com/ (accessed 17 Apr 2012).

2 'The value of friendship', *Economist*, 4 Feb 2012, www.economist.com/node/21546020 (accessed 17 Apr 2012).

3 Twitterblog, '200 million tweets a day', 30 Jun 2011, http://blog.twitter.com/2011/06/200-million-tweets-per-day.html (accessed 17 Apr 2012).

4 YouTube, 'Statistics', www.youtube.com/t/press_statistics (accessed 17 Apr 2012).

5 This is an estimate by IDC, a technology research firm. See S Lohr, 'The age of big data', *New York Times*, 12 Feb 2012, www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html?_r=1&scp=1&sq=big%20data&st=cse (accessed 17 Apr 2012).

6 For a powerful, if controversial, description of the 'Petabyte Age', see C Anderson, 'The end of theory: the data deluge makes the scientific method obsolete', *Wired*, 23 June 2008, www.wired.com/science/discoveries/magazine/16-07/pb_theory (accessed 17 Apr 2012).

7 World Economic Forum, *Big Data, Big Impact: New possibilities for development*, 2012, www.weforum.org/reports/big-data-big-impact-new-possibilities-international-development (accessed 17 Apr 2012).

8 For a comprehensive wiki of 'social media monitoring solutions', see 'A wiki of social media monitoring solutions: master list', http://wiki.kenburbary.com/ (accessed 17 Apr 2012).

9 TO Sprenger and IM Welpe, 'Tweets and trades: the information content of stock microblogs', 1 Nov 2010, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1702854 (accessed 17 Apr 2012).

10 M Wind-Cowie and R Lekhi, *The Data Dividend*, London: Demos, 2012, www.demos.co.uk/files/The_Data_Dividend_-_web.pdf (accessed 17 Apr 2012).

11 A Signorini, AM Segre and PM Polgreen, 'The use of Twitter to track levels of disease activity and public concern in the U.S. during the Influenza A H1N1 pandemic', *PLoS ONE* 6, no 5, 2011, www.plosone.org/article/info:doi%2F10.1371%2Fjournal.pone.0019467 (accessed 17 Apr 2012).

12 J Hoffman, 'Trying to find a cry of desperation amid the Facebook drama', *New York Times*, 23 Feb 2012, www.nytimes.com/2012/02/24/us/facebook-posts-can-offer-clues-of-depression.html?_r=3 (accessed 17 Apr 2012); 'T.J. Lane Facebook photos: suspect faces charges in Chardon High School shooting (slideshow)', *Huffington Post*, 28 Feb 2012, www.huffingtonpost.com/2012/02/28/tj-lane-facebook-photos_n_1307836.html#s736080&title=TJ_Lane_Facebook (accessed 17 Apr 2012).

13 For instance the UN Global Pulse Programme, 'UN unveils initial findings on uses of real-time data for development work', UN News Centre, 8 Dec 2011, www.un.org/apps/news/story.asp?NewsID=40667&Cr=global&Cr1=pulse (accessed 17 Apr 2012).

14 E Pilkington, 'How Facebook helped elect the mayor of Rapid City, South Dakota', *Guardian*, 17 Feb 2012, www.guardian.co.uk/world/2012/feb/17/facebook-mayoral-election-south-dakota?intcmp=239 (accessed 17 Apr 2012).

15 E Pilkington and A Michel, 'Obama, Facebook and the power of friendship: the 2012 data election', *Guardian*, 17 Feb 2012, www.guardian.co.uk/world/2012/feb/17/obama-digital-data-machine-facebook-election (accessed 17 Apr 2012).

16 'SNP praised for social media use', *Herald Scotland*, 26 Oct 2011, www.heraldscotland.com/politics/political-news/snp-praised-for-social-media-use.15566873 (accessed 17 Apr 2012).

17 Criminal Justice Degrees Guide, '20 cases solved by Facebook', 2 Mar 2012, www.criminaljusticedegreesguide.com/features/20-cases-solved-by-facebook.html (accessed 17 Apr 2012); for the case of trolling, see 'Internet "trolls" jailed for mocking dead teenagers on Facebook', *Telegraph*, 14 Sep 2011, www.telegraph.co.uk/news/uknews/crime/8760504/Internet-troll-jailed-for-mocking-dead-teenagers-on-Facebook.html (accessed 17 Apr 2012).

18 C Gill, 'The Facebook crimewave hits 100,000 in the last five years', *Daily Mail*, 14 Dec 2010, www.dailymail.co.uk/news/article-1338223/Facebook-crime-rises-540-cent-3-years-police-chiefs-16-forces-reveal.html (accessed 17 Apr 2012).

19 'Police launch facebook crime-fighting tool', *Metro*, [no date], www.metro.co.uk/news/143242-police-launch-facebook-crime-fighting-tool (accessed 17 Apr 2012).

20 'Fugitive found on Facebook', *Independent*, 13 Feb 2012, www.independent.co.uk/news/uk/crime/fugitive-found-on-facebook-6804822.html (accessed 17 Apr 2012); 'Pinoy Convicted of Bigamy in UK', Rappler, 14 Dec 2011, www.rappler.com/nation/1994-pinoy-bigamist-convicted-in-uk (accessed 17 Apr 2012).

21 'Facebook crimes probed by Humberside Police', *Hull Daily Mail*, 24 Aug 2011, www.thisishullandeastriding.co.uk/Facebook-crimes-probed-Humberside-Police/story-13191231-detail/story.html (accessed 17 Apr 2012); Westminster's 'Your Choice' programme: see '"Choose life, not gangs": problem kids told to clean up or face the consequences', City of Westminster, 29 Sep 2011, www.westminster.gov.uk/press-releases/2011-09/choose-life-not-gangs-problem-kids-told-to/ (accessed 17 Apr 2012).

22 Ministry of Defence and Centre for Defence Enterprise, Cyber and Influence Science and Technology Centre, 'CDE call for research proposals', 1 Nov 2011, www.science.mod.uk/controls/getpdf.pdf?603 (accessed 17 Apr 2012).

23 'We need to build a much closer relationship between government, the private sector and the public when it comes to national security', Cabinet Office, *A Strong Britain in an Age of Uncertainty: The National Security Strategy*, Oct 2010, www.cabinetoffice.gov.uk/sites/default/files/resources/national-security-strategy.pdf (accessed 17 Apr 2012).

24 That, at core, we consent to have magistracy held above us in order to move from a pre-political state of war: *bellum omnium contra omnes*. 'Every man divest himself of the right he hath to all things by nature' through the formation of a covenant 'whereby the liberty of performing or not performing is taken away', Thomas Hobbes, *The Elements of Law, Natural and Politic*, London: Elibron Classics, 2005, ch 15.2, p 63; ch 15.9, p 65.

25 E Manningham-Buller, 'Security', Reith lecture, 13 Sep 2011, pp 2–7, http://downloads.bbc.co.uk/rmhttp/radio4/transcripts/2011_reith4.pdf (accessed 17 Apr 2012).

26 European Commission, *Attitudes on Data Protection and Electronic Identity in the European Union*, Special Eurobarometer 359, Jun 2011, http://ec.europa.eu/public_opinion/archives/ebs/ebs_359_en.pdf (accessed 17 Apr 2012).

27 Which defines the circumstances under which personal information can be processed by public authorities and private organisations. These are in summary that the information must be fairly and lawfully processed; processed for limited purposes; adequate, relevant and not excessive; accurate and up to date; not kept for longer than is necessary; processed in line with rights of data subjects under the Act; and secure, and not transferred to other countries without adequate protection.

28 The three agencies of British intelligence, the Security Service, the Secret Intelligence Service (SIS) and Government Communications Headquarters (GCHQ) have moved some way from the cloak of the Secret Vote and Royal Prerogative. For over the last 20 years, they have been avowed parts of government, and their actions have been properly defined and circumscribed constitutionally and legally. This framework is overseen by independent bodies, with ultimate accountability to the national parliament. RIPA 2000 regulated all forms of intrusive surveillance, with a senior judge as commissioner. Separately the UK passed the Security Service Act and then the Intelligence Services Act (for SIS and GCHQ), which provide for regulation of the three UK intelligence agencies by senior judges and oversight by the Parliamentary Intelligence and Security Committee. In addition, the information commissioner oversees electronic communications regulations, the chief surveillance commissioner oversees covert surveillance, the interceptions of communications commissioner oversees the issue and operation of warrants under RIPA 2000 and the intelligence services commissioner reviews warrants under the Intelligence Services Act. The Investigatory Powers Tribunal allows for citizens who believe they are the illegitimate subjects of surveillance to complain and seek restitution.

29 Facebook, 'Fact Sheet', http://newsroom.fb.com/content/default.aspx?NewsAreaId=22 (accessed 17 Apr 2012).

30 E Barnett, 'Twitter "to hit 500 million registered users"', *Telegraph*, 22 Feb 2012, www.telegraph.co.uk/technology/twitter/9098557/Twitter-to-hit-500-million-registered-users.html (accessed 17 Apr 2012).

31 C Arthur, 'Has Facebook peaked? New drop in number of UK users', *Guardian*, 13 Jun 2011, www.guardian.co.uk/technology/2011/jun/13/has-facebook-peaked-drop-uk-users (accessed 17 Apr 2012).

32 HMIC, *The Rules of Engagement: A review of the August 2011 disorders*, Her Majesty's Inspectorate of Constabulary, 2011, www.hmic.gov.uk/media/a-review-of-the-august-2011-disorders-20111220.pdf (accessed 17 Apr 2012).

33 Ibid, p 31.

34 Ibid, p 30.

35 Ibid.

36 J Bartlett and M Littler, *Inside the EDL: Populist politics in a digital age*, London: Demos, 2011, www.demos.co.uk/publications/insidetheedl (accessed 17 Apr 2012).

37 Gill, 'The Facebook crimewave hits 100,000 in the last five years'.

38 Twitter was used by pupils as an *ad hoc* emergency broadcasting system during the Ohio school shooting. See L Dugan, 'Twitter used as an impromptu broadcast system during Ohio school shooting', Media Bistro, 28 Feb 2012, www.mediabistro.com/alltwitter/twitter-used-as-impromptu-emergency-broadcast-system-during-ohio-school-shooting_b19030

39 HMIC, *The Rules of Engagement*.

40 'The rise of crowdsourcing', *Wired*, Jun 2006, www.wired.com/wired/archive/14.06/crowds.html (accessed 17 Apr 2012).

41 'Reading the riots: investigating England's summer of disorder' [interactive], *Guardian*, www.guardian.co.uk/uk/interactive/2011/aug/24/riots-twitter-traffic-interactive (accessed 17 Apr 2012).

42 S Sengupta, 'Zuckerberg's unspoken law: sharing and more sharing', *New York Times*, 23 Sep 2011, http://bits.blogs.nytimes.com/2011/09/23/zuckerbergs-unspoken-law-sharing-and-more-sharing/ (accessed 17 Apr 2012).

43 Twitter, 'Privacy policy', http://twitter.com/privacy/previous/version_2 (accessed 17 Apr 2012); Facebook, 'Data use policy', www.facebook.com/about/privacy/your-info (accessed 17 Apr 2012).

44 J Jarvis, 'Mark Zuckerberg's masterplan for the 'sharing economy', *Guardian*, 2 Feb 2012, www.guardian.co.uk/technology/2012/feb/02/mark-zuckerberg-sharing-economy (accessed 17 Apr 2012)

45 K Hill, 'Max Schrems: The Austrian Thorn in Facebook's Side', Forbes, 7 Feb 2012, www.forbes.com/sites/kashmirhill/2012/02/07/the-austrian-thorn-in-facebooks-side/ (accessed 17 Apr 2012).

46 Ibid.

47 B Johnson, 'Privacy no longer a social norm, says Facebook Founder', *Guardian*, 11 Jan 2011, www.guardian.co.uk/technology/2010/jan/11/facebook-privacy (accessed 17 Apr 2012).

48 European Commission, *Attitudes on Data Protection and Electronic Identity in the European Union*.

49 See D Boyd and E Hargittai, 'Facebook privacy settings: who cares?', *First Monday* 15, no 8, 2010, http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/3086/2589 (accessed 17 Apr 2012); C Edwards and C Fieschi (eds), *UK Confidential*, London: Demos, 2008, p 20, www.demos.co.uk/publications/ukconfidential (accessed 17 Apr 2012); European Commission, *Attitudes on Data Protection and Electronic Identity in the European Union*, pp 23 and 30.

50 European Commission, *Data Protection in the European Union: Citizens' Perceptions*, 2008, http://ec.europa.eu/public_opinion/flash/fl_225_en.pdf (accessed 17 Apr 2012).

51 R Henry and C Flyn, 'Tap, tap, tapping us all up – mobile apps invade privacy', *Sunday Times*, 26 Feb 2012. This is related to the Eurobarometer data that found that 58 per cent of UK respondents read privacy statements on the internet.

52 In November 2011 the USA's Federal Trade Commission (FTC) accused Facebook of 'unfair and deceptive' practices after discovering that the website had allowed advertisers and application developers to access personally-identifiable information about users as well as information held on deleted accounts, all in contravention of guarantees made to its users. In future, the FTC will require Facebook to get the 'affirmative express consent' of users in overriding their privacy preferences and the company has in addition been subjected to biennial privacy audits for the next 20 years. Google, too, has received criticism in the USA and Europe over its proposed handling of users' personal data. See 'Private data, public rules', *Economist*, 28 Jan 2012, www.economist.com/node/21543489 (accessed 17 Apr 2012).

53 For recent deliberative research into people's conceptions of privacy, see P Bradwell, *Private Lives: A people's inquiry into personal information*, London: Demos, 2010, www.demos.co.uk/files/Private_Lives_-_web.pdf (accessed 17 Apr 2012).

54 K Hill, 'Do your social media settings matter if you get sued?', Forbes, 27 Feb 2010, www.forbes.com/sites/kashmirhill/2010/09/27/do-your-social-networking-privacy-settings-matter-if-you-get-sued/ (accessed 17 Apr 2012).

55 The draft legislation is available. See European Commission, *Proposal for a Regulation of the European Parliament and of the Council on the Protection of Individuals with regard to the Processing of Personal Data and on the Free Movement of such Data (General Data Protection Regulation)*, 25 Jan 2012, http://ec.europa.eu/justice/data-protection/document/review2012/com_2012_11_en.pdf (accessed 17 Apr 2012).

56 'Google privacy changes "in breach of EU law"', BBC News, 1 Mar 2012, www.bbc.co.uk/news/technology-17205754 (accessed 17 Apr 2012).

57 Information Commissioner's Office, *Information Commissioner's Report to Parliament on the State of Surveillance*, Nov 2010, p 14.

58 Royal Academy of Engineering, *Dilemmas of Privacy and Surveillance: Challenges of technological change*, 2007, www.raeng.org.uk/news/publications/list/reports/dilemmas_of_privacy_and_surveillance_report.pdf (accessed 17 Apr 2012).

59 Information Commissioner's Office, *Information Commissioner's Report to Parliament on the State of Surveillance*

60 For a prominent recent example see J Holliday, 'Operation Motorman: Guido Fawkes under fire over publication of files', *Guardian*, 10 Apr 2012, www.guardian.co.uk/media/2012/apr/10/operation-motorman-guido-fawkes (accessed 17 Apr 2012).

61 G Hosein, 'FAQ: the Communications Capabilities Development Programme', Privacy International, 3 Apr 2012, https://www.privacyinternational.org/blog/faq-the-communications-capabilities-development-programme (accessed 17 Apr 2012).

62 'Online critics point to foreign experience', CBC News, 21 Feb 2012, www.cbc.ca/news/canada/story/2012/02/21/pol-c30-surveillance-caution.html?cmp=rss (accessed 17 Apr 2012).

63 Open Rights Group, 'Stop the government snooping on every email and Facebook message', http://action. openrightsgroup. org/ea-campaign/clientcampaign.do?ea. client.id=1422 &ea.campaign.id=8227 (accessed 17 Apr 2012).

64 WhatDoTheyKnow.com, 'Social media monitoring policies', 2011 and 2012, www.whatdotheyknow.com/request/social_media_ monitoring_policies (accessed 17 Apr 2012).

65 Big Brother Watch, 'Westminster Council unveils plans for social media monitoring', Sep 2011, www.bigbrotherwatch.org.uk/ home/2011/09/westminster-council-unveils-plans-for-social-media-monitoring.html#.T1eQH4dmKzY (accessed 17 Apr 2012).

66 Big Brother Watch, 'Monitoring BBM to stop a water fight,' Aug 2011, www.bigbrotherwatch.org.uk/home/2011/08/monitoring-bbm-to-stop-a-water-fight.html#.T1eQiIdmKzY (accessed 17 Apr 2012).

67 M Hick, 'Hague: governments must not censor internet', *Huffington Post*, 1 Nov 2011, www.huffingtonpost.co.uk/ 2011/11/01/william-hague-government-internet-censorship_n_1069298.html (accessed 17 Apr 2012).

68 'London hosts cyberspace security conference', BBC News, 1 Nov 2011, www.bbc.co.uk/news/technology-15533786 (accessed 17 Apr 2012).

69 JP Barlow, 'A declaration of the independence of cyberspace', 8 Feb 1996, https://projects.eff.org/~barlow/Declaration-Final.html (accessed 17 Apr 2012).

70 See J Goldsmith and T Wu, *Who Controls the Internet?*, Cambridge: Cambridge University Press, 2006.

71 Facebook's Chief Executive's letter to potential investors', BBC News, 2 Feb 2012, www.bbc.co.uk/news/technology-16859527 (accessed 17 Apr 2012).

72 *Hansard*, 6 Mar 2000, vol 345, col 767.

73 D Omand, *Securing the State*, London: Hurst & Co, 2010. See also M Quinlan, 'The just war tradition and the use of armed force in the twenty-first century', annual lecture of the War Studies Department, King's College London, 25 Jan 2006. *Jus ad intelligentiam* refers to the idea that there are limitations to the [specific] circumstances in which law enforcement and relevant government agencies would be justified in building and using the capacity to undertake secret intelligence work since by its nature such work carries inevitable moral hazard; the methods for accessing secret intelligence even against legitimate targets are also limited *jus in intelligentia* principles such as proportionality and necessity that prevent abuse and the banalisation of such activities by the state. To this we might add *jus post intelligentiam* – the opportunities for citizen recourse, oversight and redress in the event that the powers of intelligence gathering are misused.

74 A line of thinking owed to discussion with the late Sir Michael Quinlan, as set out in David Omand, 'Ethical guidelines in using secret intelligence for public security', *Cambridge Review of International Affairs* 19, no 4, 2006, pp 613–28.

75 HMIC, *A Review of National Police Units which Provide Intelligence on Criminality Associated with Protest*, Her Majesty's Inspectorate of Constabulary, 2012, www.hmic.gov.uk/media/review-of-national-police-units-which-provide-intelligence-on-criminality-associated-with-protest-20120202.pdf (accessed 18 Apr 2012).

76 European Commission, *Attitudes on Data Protection and Electronic Identity in the European Union*: 11 per cent said that they feared that their views and behaviours could be misunderstood.

77 HMIC, *Rules of Engagement*, p 28, paras 2.3 and 2.4

78 R Proctor, F Vis and A Voss, 'Riot rumours: how misinformation spread on Twitter during a time of crisis, *Guardian*, 7 Dec 2011, www.guardian.co.uk/uk/interactive/2011/dec/07/london-riots-twitter (accessed 17 Apr 2012).

79 See for instance J Leskovec, J Kleinberg and C Faloutsos, 'Graph evolution: densification and shrinking diameters', *Data* 1, no 1, Mar 2007, www.cs.cmu.edu/~jure/pubs/powergrowth-tkdd.pdf (accessed 16 Apr 2012); J Leskovec and C Faloutsos, 'Sampling from large graphs' in *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2006, www.stat.cmu.edu/~fienberg/Stat36-835/Leskovec-sampling-kdd06.pdf (accessed 17 Apr 2012); P Rusmevichientong et al, 'Methods for sampling pages uniformly from the world wide web' in *Proceedings of the AAAI Fall Symposium on Using Uncertainty Within Computation*, 2001, pp 121–8.

80 See, for instance, B O'Connor et al, 'From tweets to polls: linking text sentiment to public opinion', Time Series, Pittsburgh PA: Carnegie Mellon, 2010. The authors collected their sample using just a few keyword searches.

81 See M de Choudhury et al, 'How does the data sampling strategy impact the discovery of information diffusion in social media?', Tempe AZ: Arizona: Arizona State University and IBM TJ Watson Research Centre, 2010.

82 For information on Twitter and Facebook demographics, see Digital Buzz Blog, 'Infographic: Facebook vs Twitter demographics', 21 Dec 2010, www.digitalbuzzblog.com/infographic-facebook-vs-twitter-demographics-2010-2011/ (accessed 16 Apr 2012).

83 See C Shirky, *Here Comes Everybody: The power of organizing without organizations*, New York: Penguin, 2008.

84 Sysomos, 'Twitter statistics for 2010', Dec 2010, www.sysomos.com/insidetwitter/twitter-stats-2010/ (accessed 16 Apr 2012).

85 J Weng et al, 'Event detection in Twitter', HP Laboratories, 6 Jul 2011, www.hpl.hp.com/techreports/2011/HPL-2011-98.html (accessed 17 Apr 2012); S Asur and BA Huberman, 'Predicting the future with social media', HP Laboratories, 29 Mar 2010, www.hpl.hp.com/research/scl/papers/socialmedia/socialmedia.pdf (accessed 17 Apr 2012).

86 J Suler, 'The online disinhibition effect', *Journal of Cyberpsychology and Behaviour* 7, no 3, 2004, pp 321–6. See also J Suler, 'The psychology of cyberspace: the online disinhibition effect', http://users.rider.edu/~suler/psycyber/disinhibit.html (accessed 17 Apr 2012).

87 Bartlett and Littler, *Inside the EDL*.

88 'Twitterology high and low', *The Economist*, 31 Oct 2011, www.economist.com/blogs/johnson/2011/10/technology-and-language?fsrc=scn%2Ftw%2Fte%2Fbl%2Ftwitterologyhighandlow (accessed 16 Apr 2012).

89 R Booth, 'Twitter joke case reaches high court', *Guardian*, 8 Feb 2012, www.guardian.co.uk/law/2012/feb/08/twitter-joke-case-court-appeal (accessed 16 Apr 2012).

90 Jack of Kent (David Allen Green), 'Paul Chambers: a disgraceful and illiberal judgment', 11 May 2010, http://jackofkent.blogspot.com/2010/05/paul-chambers-disgraceful-and-illiberal.html (accessed 16 Apr 2012).

91 A Parker, 'US bars friends over Twitter joke', *Sun*, 30 Jan 2012, www.thesun.co.uk/sol/homepage/news/4095372/Twitter-news-US-bars-friends-over-Twitter-joke.html (accessed 16 Apr 2012).

92 For instance, in deprived areas of Berlin, civil servants have increasingly used portable devices connected to database records when visiting care homes for the elderly and hospitals. These devices give constant, mobile access to databases, enabling public servants to understand the needs of individuals and families, track their previous contact and check for problems and underlying issues that may have been recorded by other agencies. See J Millard, 'eGovernance and eParticipation: lessons from Europe in promoting inclusion and empowerment', paper presented to UN Division for Public Administration and Development Management (DPADM) workshop, E-Participation and E-Government: Understanding the Present and Creating the Future, Budapest, Hungary, 27–28 Jul 2006, unpan1.un.org/intradoc/groups/public/documents/UN/UNPAN023685.pdf (accessed 23 Jan 2012).

93 Briefing on the Interception Modernisation Programme, Policy Engagement Network paper 5, Jun 2009, p 56.

94 'Shopping amid a massacre: leaked e-mails from Syria's regime', CNN International, 16 Mar 2012, http://edition.cnn.com/2012/03/15/world/meast/syria-al-assad-e-mails/index.html?iphoneemail (accessed 16 Apr 2012).

95 European Commission, *Attitudes on Data Protection and Electronic Identity in the European Union*.

References

'A wiki of social media monitoring solutions: master list', http://wiki.kenburbary.com/ (accessed 17 Apr 2012).

'"Choose life, not gangs": problem kids told to clean up or face the consequences', City of Westminster, 29 Sep 2011, www.westminster.gov.uk/press-releases/2011-09/choose-life-not-gangs-problem-kids-told-to/ (accessed 17 Apr 2012).

'Facebook crimes probed by Humberside Police', *Hull Daily Mail*, 24 Aug 2011, www.thisishullandeastriding.co.uk/Facebook-crimes-probed-Humberside-Police/story-13191231-detail/story.html (accessed 17 Apr 2012).

'Facebook's Chief Executive's letter to potential investors', BBC News, 2 Feb 2012, www.bbc.co.uk/news/technology-16859527 (accessed 17 Apr 2012).

'Fugitive found on Facebook', *Independent*, 13 Feb 2012, www.independent.co.uk/news/uk/crime/fugitive-found-on-facebook-6804822.html (accessed 17 Apr 2012).

'Google privacy changes "in breach of EU law"', BBC News, 1 Mar 2012, www.bbc.co.uk/news/technology-17205754 (accessed 17 Apr 2012).

'Internet "trolls" jailed for mocking dead teenagers on Facebook', *Telegraph*, 14 Sep 2011, www.telegraph.co.uk/news/uknews/crime/8760504/Internet-troll-jailed-for-mocking-dead-teenagers-on-Facebook.html (accessed 17 Apr 2012).

'London hosts cyberspace security conference', BBC News, 1 Nov 2011, www.bbc.co.uk/news/technology-15533786 (accessed 17 Apr 2012).

'Online critics point to foreign experience', CBC News, 21 Feb 2012, www.cbc.ca/news/canada/story/2012/02/21/pol-c30-surveillance-caution.html?cmp=rss (accessed 17 Apr 2012).

'Pinoy Convicted of Bigamy in UK', Rappler, 14 Dec 2011, www.rappler.com/nation/1994-pinoy-bigamist-convicted-in-uk (accessed 17 Apr 2012).

'Police launch facebook crime-fighting tool', *Metro*, [no date], www.metro.co.uk/news/143242-police-launch-facebook-crime-fighting-tool (accessed 17 Apr 2012).

'Private data, public rules', *Economist*, 28 Jan 2012, www.economist.com/node/21543489 (accessed 17 Apr 2012).

'Reading the riots: investigating England's summer of disorder' [interactive], *Guardian*, www.guardian.co.uk/uk/interactive/2011/aug/24/riots-twitter-traffic-interactive (accessed 17 Apr 2012).

'Shopping amid a massacre: leaked e-mails from Syria's regime', CNN International, 16 Mar 2012, http://edition.cnn.com/2012/03/15/world/meast/syria-al-assad-e-mails/index.html?iphoneemail (accessed 16 Apr 2012).

'SNP praised for social media use', *Herald Scotland*, 26 Oct 2011, www.heraldscotland.com/politics/political-news/snp-praised-for-social-media-use.15566873 (accessed 17 Apr 2012).

'T.J. Lane Facebook photos: suspect faces charges in Chardon High School shooting (slideshow)', *Huffington Post*, 28 Feb 2012, www.huffingtonpost.com/2012/02/28/tj-lane-facebook-photos_n_1307836.html#s736080&title=TJ_Lane_Facebook (accessed 17 Apr 2012).

'The rise of crowdsourcing', *Wired*, Jun 2006, www.wired.com/wired/archive/14.06/crowds.html (accessed 17 Apr 2012).

'The value of friendship', *Economist*, 4 Feb 2012, www.economist.com/node/21546020 (accessed 17 Apr 2012).

'Twitterology high and low', *The Economist*, 31 Oct 2011, www.economist.com/blogs/johnson/2011/10/technology-and-language?fsrc=scn%2Ftw%2Fte%2Fbl%2Ftwitterologyhighandlow (accessed 16 Apr 2012).

'We need to build a much closer relationship between government, the private sector and the public when it comes to national security', Cabinet Office, *A Strong Britain in an Age of Uncertainty: The National Security Strategy*, Oct 2010, www.cabinetoffice.gov.uk/sites/default/files/resources/national-security-strategy.pdf (accessed 17 Apr 2012).

Anderson C, 'The end of theory: the data deluge makes the scientific method obsolete', *Wired*, 23 June 2008, www.wired.com/science/discoveries/magazine/16-07/pb_theory (accessed 17 Apr 2012).

Arthur C, 'Has Facebook peaked? New drop in number of UK users', *Guardian*, 13 Jun 2011, www.guardian.co.uk/technology/2011/jun/13/has-facebook-peaked-drop-uk-users (accessed 17 Apr 2012).

Asur S and Huberman BA, 'Predicting the future with social media', HP Laboratories, 29 Mar 2010, www.hpl.hp.com/research/scl/papers/socialmedia/socialmedia.pdf (accessed 17 Apr 2012).

Barlow JP, 'A declaration of the independence of cyberspace', 8 Feb 1996, https://projects.eff.org/~barlow/Declaration-Final.html (accessed 17 Apr 2012).

Barnett E, 'Twitter "to hit 500 million registered users"', *Telegraph*, 22 Feb 2012, www.telegraph.co.uk/technology/twitter/9098557/Twitter-to-hit-500-million-registered-users.html (accessed 17 Apr 2012).

Bartlett J and Littler M, *Inside the EDL: Populist politics in a digital age*, London: Demos, 2011, www.demos.co.uk/publications/insidetheedl (accessed 17 Apr 2012).

Big Brother Watch, 'Monitoring BBM to stop a water fight,' Aug 2011, www.bigbrotherwatch.org.uk/home/2011/08/monitoring-bbm-to-stop-a-water-fight.html#.T1eQiIdmKzY (accessed 17 Apr 2012).

Big Brother Watch, 'Westminster Council unveils plans for social media monitoring', Sep 2011, www.bigbrotherwatch.org.uk/home/2011/09/westminster-council-unveils-plans-for-social-media-monitoring.html#.T1eQH4dmKzY (accessed 17 Apr 2012).

Booth R, 'Twitter joke case reaches high court', *Guardian*, 8 Feb 2012, www.guardian.co.uk/law/2012/feb/08/twitter-joke-case-court-appeal (accessed 16 Apr 2012).

Boyd D and Hargittai E, 'Facebook privacy settings: who cares?', *First Monday* 15, no 8, 2010, http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/3086/2589 (accessed 17 Apr 2012).

Bradwell P, *Private Lives: A people's inquiry into personal information*, London: Demos, 2010, www.demos.co.uk/files/Private_Lives_-_web.pdf (accessed 17 Apr 2012).

Briefing on the Interception Modernisation Programme, Policy Engagement Network paper 5, Jun 2009.

Criminal Justice Degrees Guide, '20 cases solved by Facebook', 2 Mar 2012, www.criminaljusticedegreesguide.com/features/20-cases-solved-by-facebook.html (accessed 17 Apr 2012).

De Choudhury M et al, 'How does the data sampling strategy impact the discovery of information diffusion in social media?', Tempe AZ: Arizona: Arizona State University and IBM TJ Watson Research Centre, 2010.

Digital Buzz Blog, 'Infographic: Facebook vs Twitter demographics', 21 Dec 2010, www.digitalbuzzblog.com/infographic-facebook-vs-twitter-demographics-2010-2011/ (accessed 16 Apr 2012).

Dugan L, 'Twitter used as an impromptu broadcast system during Ohio school shooting', Media Bistro, 28 Feb 2012, www.mediabistro.com/alltwitter/twitter-used-as-impromptu-emergency-broadcast-system-during-ohio-school-shooting_b19030

Edwards C and Fieschi C (eds), *UK Confidential*, London: Demos, 2008, www.demos.co.uk/publications/ukconfidential (accessed 17 Apr 2012).

European Commission, *Attitudes on Data Protection and Electronic Identity in the European Union*, Special Eurobarometer 359, Jun 2011, http://ec.europa.eu/public_opinion/archives/ebs/ebs_359_en.pdf (accessed 17 Apr 2012).

European Commission, *Data Protection in the European Union: Citizens' Perceptions*, 2008, http://ec.europa.eu/public_opinion/flash/fl_225_en.pdf (accessed 17 Apr 2012).

European Commission, *Proposal for a Regulation of the European Parliament and of the Council on the Protection of Individuals with regard to the Processing of Personal Data and on the Free Movement of such Data (General Data Protection Regulation)*, 25 Jan 2012, http://ec.europa.eu/justice/data-protection/document/review2012/com_2012_11_en.pdf (accessed 17 Apr 2012).

Facebook, 'Data use policy', www.facebook.com/about/privacy/your-info (accessed 17 Apr 2012).

Facebook, 'Fact Sheet', http://newsroom.fb.com/content/default.aspx?NewsAreaId=22 (accessed 17 Apr 2012).

Ford R, 'Beware rise of Big Brother state, warns data watchdog', *The Times*, 16 Aug 2004.

Gill C, 'The Facebook crimewave hits 100,000 in the last five years', *Daily Mail*, 14 Dec 2010, www.dailymail.co.uk/news/article-1338223/Facebook-crime-rises-540-cent-3-years-police-chiefs-16-forces-reveal.html (accessed 17 Apr 2012).

Goldsmith J and Wu T, *Who Controls the Internet?*, Cambridge: Cambridge University Press, 2006.

Henry R and Flyn C, 'Tap, tap, tapping us all up – mobile apps invade privacy', *Sunday Times*, 26 Feb 2012.

Hick M, 'Hague: governments must not censor internet', *Huffington Post*, 1 Nov 2011, www.huffingtonpost.co.uk/2011/11/01/william-hague-government-internet-censorship_n_1069298.html (accessed 17 Apr 2012).

Hill K, 'Do your social media settings matter if you get sued?', Forbes, 27 Feb 2010, www.forbes.com/sites/kashmirhill/2010/09/27/do-your-social-networking-privacy-settings-matter-if-you-get-sued/ (accessed 17 Apr 2012).

Hill K, 'Max Schrems: The Austrian Thorn in Facebook's Side', Forbes, 7 Feb 2012, www.forbes.com/sites/kashmirhill/2012/02/07/the-austrian-thorn-in-facebooks-side/ (accessed 17 Apr 2012).

HMIC, *A Review of National Police Units which Provide Intelligence on Criminality Associated with Protest*, Her Majesty's Inspectorate of Constabulary, 2012, www.hmic.gov.uk/media/review-of-national-

police-units-which-provide-intelligence-on-criminality-associated-with-protest-20120202.pdf (accessed 18 Apr 2012).

HMIC, *The Rules of Engagement: A review of the August 2011 disorders*, Her Majesty's Inspectorate of Constabulary, 2011, www.hmic.gov.uk/media/a-review-of-the-august-2011-disorders-20111220.pdf (accessed 17 Apr 2012).

Hobbes T, *The Elements of Law, Natural and Politic*, London: Elibron Classics, 2005.

Hoffman J, 'Trying to find a cry of desperation amid the facebook drama', *New York Times*, 23 Feb 2012, www.nytimes.com/2012/02/24/us/facebook-posts-can-offer-clues-of-depression.html?_r=3 (accessed 17 Apr 2012).

Holliday J, 'Operation Motorman: Guido Fawkes under fire over publication of files', *Guardian*, 10 Apr 2012, www.guardian.co.uk/media/2012/apr/10/operation-motorman-guido-fawkes (accessed 17 Apr 2012).

Hosein G, 'FAQ: the Communications Capabilities Development Programme', Privacy International, 3 Apr 2012, https://www.privacyinternational.org/blog/faq-the-communications-capabilities-development-programme (accessed 17 Apr 2012).

Information Commissioner's Office, *Information Commissioner's Report to Parliament on the State of Surveillance*, Nov 2010.

Jack of Kent (David Allen Green), 'Paul Chambers: a disgraceful and illiberal judgment', 11 May 2010, http://jackofkent.blogspot.com/2010/05/paul-chambers-disgraceful-and-illiberal.html (accessed 16 Apr 2012).

Jarvis J, 'Mark Zuckerberg's masterplan for the "sharing economy"', *Guardian*, 2 Feb 2012, www.guardian.co.uk/technology/2012/feb/02/mark-zuckerberg-sharing-economy (accessed 17 Apr 2012)

Johnson B, 'Privacy no longer a social norm, says Facebook Founder', *Guardian*, 11 Jan 2011, www.guardian.co.uk/technology/2010/jan/11/facebook-privacy (accessed 17 Apr 2012).

Leskovec J and Faloutsos C, 'Sampling from large graphs' in *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2006, www.stat.cmu.edu/~fienberg/Stat36-835/Leskovec-sampling-kdd06.pdf (accessed 17 Apr 2012).

Leskovec J, Kleinberg J and Faloutsos C, 'Graph evolution: densification and shrinking diameters', *Data* 1, no 1, Mar 2007, www.cs.cmu.edu/~jure/pubs/powergrowth-tkdd.pdf (accessed 16 Apr 2012).

Lohr S, 'The age of big data', *New York Times*, 12 Feb 2012, www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html?_r=1&scp=1&sq=big%20data&st=cse (accessed 17 Apr 2012).

Manningham-Buller E, 'Security', Reith lecture, 13 Sep 2011, http://downloads.bbc.co.uk/rmhttp/radio4/transcripts/2011_reith4.pdf (accessed 17 Apr 2012).

Millard J, 'eGovernance and eParticipation: lessons from Europe in promoting inclusion and empowerment', paper presented to UN Division for Public Administration and Development Management (DPADM) workshop, E-Participation and E-Government: Understanding the Present and Creating the Future, Budapest, Hungary, 27–28 Jul 2006, unpan1.un.org/intradoc/groups/public/documents/UN/UNPAN023685.pdf (accessed 23 Jan 2012).

Ministry of Defence and Centre for Defence Enterprise, Cyber and Influence Science and Technology Centre, 'CDE call for research proposals', 1 Nov 2011, www.science.mod.uk/controls/getpdf.pdf?603 (accessed 17 Apr 2012).

O'Connor B et al, 'From tweets to polls: linking text sentiment to public opinion', Time Series, Pittsburgh PA: Carnegie Mellon, 2010.

Omand D, 'Ethical guidelines in using secret intelligence for public security', *Cambridge Review of International Affairs* 19, no 4, 2006.

Omand D, *Securing the State*, London: Hurst & Co, 2010.

Open Rights Group, 'Stop the government snooping on every email and Facebook message', [2012], http://action.openrightsgroup.org/ea-campaign/clientcampaign.do?ea.client.id=1422&ea.campaign.id=8227 (accessed 17 Apr 2012).

Parker A, 'US bars friends over Twitter joke', *Sun*, 30 Jan 2012, www.thesun.co.uk/sol/homepage/news/4095372/Twitter-news-US-bars-friends-over-Twitter-joke.html (accessed 16 Apr 2012).

Pilkington E, 'How Facebook helped elect the mayor of Rapid City, South Dakota', *Guardian*, 17 Feb 2012, www.guardian.co.uk/world/2012/feb/17/facebook-mayoral-election-south-dakota?intcmp=239 (accessed 17 Apr 2012).

Pilkington E and Michel A, 'Obama, Facebook and the power of friendship: the 2012 data election', *Guardian*, 17 Feb 2012,www.guardian.co.uk/world/2012/feb/17/obama-digital-data-machine-facebook-election (accessed 17 Apr 2012).

Proctor R, Vis F and Voss A, 'Riot rumours: how misinformation spread on Twitter during a time of crisis, *Guardian*, 7 Dec 2011, www.guardian.co.uk/uk/interactive/2011/dec/07/london-riots-twitter (accessed 17 Apr 2012).

Quinlan M, 'The just war tradition and the use of armed force in the twenty-first century', annual lecture of the War Studies Department, King's College London, 25 Jan 2006.

Royal Academy of Engineering, *Dilemmas of Privacy and Surveillance: Challenges of technological change*, 2007, www.raeng.org.uk/news/publications/list/reports/dilemmas_of_privacy_and_surveillance_report.pdf (accessed 17 Apr 2012).

Rusmevichientong P et al, 'Methods for sampling pages uniformly from the world wide web' in *Proceedings of the AAAI Fall Symposium on Using Uncertainty Within Computation*, 2001.

Sengupta S, 'Zuckerberg's unspoken law: sharing and more sharing', *New York Times*, 23 Sep 2011, http://bits.blogs.nytimes.com/2011/09/23/zuckerbergs-unspoken-law-sharing-and-more-sharing/ (accessed 17 Apr 2012).

Shirky C, *Here Comes Everybody: The power of organizing without organizations*, New York: Penguin, 2008.

Signorini A, Segre AM and Polgreen PM, 'The use of Twitter to track levels of disease activity and public concern in the U.S. during the Influenza A H1N1 pandemic', *PLoS ONE* 6, no 5, 2011, www.plosone.org/article/info:doi%2F10.1371%2Fjournal.pone.0019467 (accessed 17 Apr 2012).

Solis B and JESS3, *The Conversation Prism: The art of listening, learning and sharing*, www.theconversationprism.com/ (accessed 17 Apr 2012).

Sprenger TO and Welpe IM, 'Tweets and trades: the information content of stock microblogs', 1 Nov 2010, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1702854 (accessed 17 Apr 2012).

Suler J, 'The online disinhibition effect', *Journal of Cyberpsychology and Behaviour* 7, no 3, 2004.

Suler J, 'The psychology of cyberspace: the online disinhibition effect', http://users.rider.edu/~suler/psycyber/disinhibit.html (accessed 17 Apr 2012).

Sysomos, 'Twitter statistics for 2010', Dec 2010, www.sysomos.com/insidetwitter/twitter-stats-2010/ (accessed 16 Apr 2012).

Twitter, 'Privacy policy', http://twitter.com/privacy/previous/version_2 (accessed 17 Apr 2012).

Twitterblog, '200 million tweets a day', 30 Jun 2011, http://blog.twitter.com/2011/06/200-million-tweets-per-day.html (accessed 17 Apr 2012).

UN Global Pulse Programme, 'UN unveils initial findings on uses of real-time data for development work', UN News Centre, 8 Dec 2011, www.un.org/apps/news/story.asp?NewsID=40667&Cr=global&Cr1=pulse (accessed 17 Apr 2012).

Weng J et al, 'Event detection in Twitter', HP Laboratories, 6 Jul 2011, www.hpl.hp.com/techreports/2011/HPL-2011-98.html (accessed 17 Apr 2012).

WhatDoTheyKnow.com, 'Social media monitoring policies', 2011 and 2012, www.whatdotheyknow.com/request/social_media_monitoring_policies (accessed 17 Apr 2012).

Wind-Cowie M and Lekhi R, *The Data Dividend*, London: Demos, 2012, www.demos.co.uk/files/The_Data_Dividend_-_web.pdf (accessed 17 Apr 2012).

World Economic Forum, *Big Data, Big Impact: New possibilities for development*, 2012, www.weforum.org/reports/big-data-big-impact-new-possibilities-international-development (accessed 17 Apr 2012).

YouTube, 'Statistics', www.youtube.com/t/press_statistics (accessed 17 Apr 2012).

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 authorised 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, dramatisation, fictionalisation, 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 here under. 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 Licensor 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 towards 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 towards 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 utilising 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 licenced 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.

The growth of social media poses a dilemma for security and law enforcement agencies. On the one hand, social media could provide a new form of intelligence – SOCMINT – that could contribute decisively to keeping the public safe. On the other, national security is dependent on public understanding and support for the measures being taken to keep us safe.

Social media challenges current conceptions about privacy, consent and personal data, and new forms of technology allow for more invisible and widespread intrusive surveillance than ever before. Furthermore, analysis of social media for intelligence purposes does not fit easily into the policy and legal frameworks that guarantee that such activity is proportionate, necessary and accountable.

This paper is the first effort to examine the ethical, legal and operational challenges involved in using social media for intelligence and insight purposes. It argues that social media should become a permanent part of the intelligence framework but that it must be based on a publicly argued, legal footing, with clarity and transparency over use, storage, purpose, regulation and accountability. *#Intelligence* lays out six ethical principles that can help government agencies approach these challenges and argues for major changes to the current regulatory and legal framework in the long-term, including a review of the current Regulation of Investigatory Powers Act 2000.

Sir David Omand is a Visiting Professor at the War Studies department at King's College London and the former Director of GCHQ. Jamie Bartlett is Head of the Violence and Extremism Programme at Demos. Carl Miller is a Demos Associate.

ISBN 978-1-909037-08-3 £10