Privacy in Public Spaces: What Expectations of Privacy do we have in Social Media Intelligence?

Lilian Edwards and Lachlan Urquhart[1]

1. Introduction

Since Edward Snowden's revelations burst upon the world in 2013[2], the general population has become sensitised to the fact that our private communications – whether made by phone, text, email or social network – may quite likely be under covert surveillance by law enforcement and intelligence agencies such as the US's NSA, the UK's GCHQ, et al. This is not yet another article about the Snowden revelations and the legality or otherwise of covert state surveillance of private electronic communications. Instead it discusses a connected and worrying area which has received relatively little public attention or legal analysis. Herein we look at the growing use by law enforcement agencies (LEAs) of *publicly* available social media communications (or social media "intelligence", known as SOCMINT) for investigating, prosecuting and perhaps most significantly, attempting to predict and prevent crime and social unrest[3].

SOCMINT embraces a vast amount of material including posts made public on Facebook, tweets sent into the world, videos hosted on YouTube, comments on online public newspaper or TV news sites. Much of this material will qualify as "personal data" within the framework of the EU Data Protection Directive[4] and will contribute to "private life" in terms of art 8 of the European Convention on Human Rights (ECHR)[5]. Much of it may also be "sensitive" in the terminology of the DPD as referring to intimate matters such as a person's sexuality, race, colour or health status[6]. The volume and significance of this material disclosed to the world in today;s digital world is almost incomprehensible. According to the Intelligence and Security Committee of Parliament's 2015 report[7]:

---

[1] Professor of Internet Law, Centre for Internet Law and Policy, University of Strathclyde: lilian.edwards@strath.ac.uk and Doctoral Researcher, Mixed Reality Lab and Horizon Centre for Doctoral Training, University of Nottingham: lachlan.urquhart@nottingham.ac.uk . All links unless otherwise noted were checked as of 20 November 2015.

[2] This is not the place for a bibliography of the Snowden/PRISM revelations or their international political, legal and technical fallout. A good starting point for UK legal ramifications might be the pleadings of Privacy International (PI) in its suit against the UK at the Investigatory Powers Tribunal (IPT) for breaches of RIPA: the claim was disallowed: see [2014] UKIPTrib 13_77-H.

[3] Including, possibly, non-criminal social unrest (see for example, the UK controversy around criminalising "extremist" ideas : see BBC News report "Cameron unveils strategy to tackle Islamist extremism", 20 July 2015.)

[4] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, OJ L 281/31, 23.11.95 (hereafter "DPD") transposed into the UK Data Protection Act 1998 (DPA 1998).

[5] Rome, 4.XI.195 as amended.

[6] See DPD art 8.

[7] Intelligence and Security Committee, *Privacy and security: a modern and transparent legal framework,* HC 1075 2014/15, 12 March 2015 (hereafter the "ISC report") para 55, drawing on "What happens in an internet minute?" Intel Corporation, 5 December 2014.

> "*The internet carries the communications of 2.4 billion internet users. In one minute, those 2.4 billion transfer 1,572,877 gigabytes of data, including 204 million emails, 4.1 million Google searches, 6.9 million messages sent via Facebook, 347,222 posts to Twitter and 138,889 hours of video watched on YouTube*."

The net effect of this deluge of material on to the world's screens is that where once LEAs had to spend enormous amounts of resources, time and effort in acquiring details of the lives of those under suspicion, actually or potentially, by secret or covert intelligence gathering, much simpler, cheaper and easier technological means now exist to monitor much of what we all say, do, think and feel - simply by listening to what we say "publicly" on Facebook, Twitter et al.

SOCMINT, though dabbled with earlier, was first used in anger as a valuable tool for UK policing during the London riots of 2011[8], which simultaneously highlighted both the enormous potential value of digital communications in controlling crime and disturbance, and the UK police's ill-preparedness for such analysis. Since 2011, LEAs have worked to expand and improve their use of digital communications – including calls, texts, social media posts, instant messaging and tweets - as intelligence. These intelligence-led practices are now routinely assisted by "big data" predictive analytics software, allowing law enforcement agencies to anticipate, rather than merely react to, crime and antisocial behaviour[9]. According to the RUSI report of 2015, the majority of intelligence – up to 95% – gathered by intelligence agencies originates from open, not closed, sources, unsurprising given its cheapness to collect and accessibility compared to "secret" intelligence[10]. While not all this "open source" intelligence (known as OSINT – see discussion below) derives from social media, much does: in October 2014, James Clapper, the Director of National Intelligence, described social media as "huge for intelligence purposes"[11] and the extensively researched Anderson Report makes plain the increasing central reliance on SOCMINT[12].

---

[8] See for a general overview and analysis of the London riots, see P Lewis et al *Reading the Riots: Investigating England's Summer of Disorder* (London School of Economics and The Guardian, London 2012) and discussion below.

[9] For critique of applying "big data" profiling to social media data , especially tweets, see C Miller , S Ginnis et al *The road to representivity* (Demos/IPSOS Mori, September 2015) at http://www.demos.co.uk/project/the-road-to-representivity/ .

[10] RUSI *A Democratic Licence to Operate: Report of the Independent Surveillance Review* Panel of the Independent Surveillance Review, Whitehall Reports, 13 July 2015 at https://rusi.org/publication/whitehall-reports/democratic-licence-operate-report-independent-surveillance-review (hereafter, the "RUSI report"), para 3.16.

[11] *A Question of Trust – Report of the Investigatory Powers Review* , June 11 2015 (at https://terrorismlegislationreviewer.independent.gov.uk/a-question-of-trust-report-of-the-investigatory-powers-review/)(hereafter, the "Anderson report"), para 4.27.

[12] "A former head of the bin Laden Unit of the Central Intelligence Agency in the United States noted that '90% of what you need to know' comes from OSINT. According to a report in 2010, 'in the aftermath of 9/11, intelligence failures - particularly a deficient consideration of OSINT … - have been identified as major reasons for the inability to anticipate and prevent these attacks.'" Anderson report, ibid. See also Anderson report, para 4.29: "the extent of that use [of OSINT] is not publicly known." (citing C Hobbs et al (eds), *Open Source Intelligence in the Twenty-First Century: New Approaches and Opportunities*, (2014), p 24.

As discussed below, the assumption commonly made is that "open" SOCMINT - herein defined as social media communications accessible to the world to the world without the protection of "Friends-locks", passwords or encryption - are "fair game" for surveillance, devoid of any expectation of privacy protection. Hence the lack, in the main, of a legal debate around the acquisition of SOCMINT without the consent and usually, knowledge, of the sender .

However, given the vast amount of personal and often intimate information disclosed to the world via social media - and the increasing use of this material to profile and target citizens for special attention or suspicion by LEAs and intelligence organisations – it is necessary to ask if this assumption is now (or ever was) true. Is the collection and processing of SOCMINT without consent or knowledge ever a breach of individual privacy? Or, to put it another way, do we have any reasonable expectation of privacy in personal information even though publicly disseminated? If so, then laws must determine if the public interest in policing proportionately outweighs the privacy interest of those monitored; and a consensus must be reached on what type of authorisation, if any, is needed to gather such information and what redress for misuse is available to the public.[13].

Covert interception of private emails, texts and phone calls by LEAs has been controlled by explicit, if baroquely complex[14], laws since the Regulation of Investigatory Powers Act 2000 (RIPA) and, to some extent, before that, in the Interception of Communications Act 1985. RIPA is the obvious first place to look, if rather in vain, for control of acquisition of SOCMINT (see further below p xx). The Conservative government elected in summer 2015 promised in its first Queens Speech a reconsideration and recasting of both RIPA and DRIPA (the Data Retention and Investigatory Powers Act 2014) and this reformulation is under current debate in the Investigatory Powers Bill 2015[15] (see further below). The new Bill seems unlikely however to do much to clarify the legal issues posed by police acquisition and use of SOCMINT (and OSINT in general) which remain vague and underexplored, even after (or perhaps because of) the ongoing period of legal turmoil in online privacy since the Snowden revelations[16].

[13] Social media communications are also acquired by non-policing authorities eg marketing and research companies. Twitter, eg, makes a substantial part of its revenue by selling access to the "Twitter firehose" (see description at https://www.echosec.net/twitter-api-vs-firehose/ . In this article we are concerned only with policing use, given the impact on personal liberty of police power and attention: however the private aspect of this issue should not be ignored, especially given the private/public nature of much data profiling (see section 2 below).

[14] The ISC report described it as ""absurdly complicated" and the Anderson report as "obscure since its inception …has been patched up so many times as to make it incomprehensible to all but a tiny band of initiates" (para 35).

[15] Introduced as a draft Bill on 4 November 2015 at https://www.gov.uk/government/publications/draft-investigatory-powers-bill . Many related documents are available at https://www.gov.uk/government/collections/draft-investigatory-powers-bill .

[16] Again, it is not the place of this article to explore the full impact of the Snowden revelations on the UK's and EU's data privacy and data retention laws. However, in brief, in 2014, the European Court of Justice ruled that the long controversial EC Data Retention Directive 2006/24/EC was unconstitutional due to its endorsement of Europe-wide blanket surveillance without proper controls over (inter alia) how relevant data retained was to particular, serious crimes, and minimum safeguards for access to data retained (*Digital Rights Ireland Ltd (C-293/12) v Minister for Communications*, Joined Cases C-293/12 and C-594/12). The UK's response to this judgement was to pass emergency legislation, the Data Retention and Investigatory Powers Act 2014 (DRIPA), to ensure continued retention of data by communications service providers. In July 2015, this Act was itself declared invalid as in breach of EU

Below, we therefore

(i) Give a basic introduction to the transition in contemporary surveillance from top down traditional police surveillance to profiling and "pre-crime" methods;

(ii) Review in more detail the rise of open source (OSINT) and social media (SOCMINT) intelligence and its use by law enforcement and security authorities;

(iii) Consider what if any privacy protection is currently given in UK law to SOCMINT;

(iv) Given the largely negative response to the above question, we analyse what reasonable expectations of privacy there may be for users of public social media, with reference to existing case law on art 8 of the ECHR. Two factors are in particular argued to be supportive of a reasonable expectation of privacy in open public social media communications : first, the failure of many social network users to perceive the environment where they communicate as "public"; and secondly, the impact of search engines (and other automated analytics) on traditional conceptions of structured dossiers as most problematic for state surveillance

(v) Conclude that existing law does not provide adequate protection for open SOCMINT and that this will be increasingly significant as more and more personal data is disclosed and collected in public without well-defined expectations of privacy.

2. Contemporary surveillance: from the panoptic Big Brother to public/private profiling

Surveillance has become the guiding organisational principle for social control, particularly by managing populations through collection, sorting, management and risk assessment of data (so called 'dataveillance').[17] Tellingly, a UK Select Committee on Home Affairs report stated, as far back as 2008, "*the foundation for all new surveillance technologies is the database*".[18] Such data collection practices have become extensive, normalised and routine[19], conducted by both state and non-state entities.[20] Critically,

law in an action brought by Liberty alongside MPs David Davis and Tom Watson (*Davis and others v Secretary of State for the Home Department* [2015] EWHC 2092 (Admin). An appeal is likely to follow. Several of the major reports quoted herein which were delivered in summer 2015 (Anderson report, n 11; RUSI report, n 10; ISC report, n 7) arose from the combination of the Snowden revelations and the cracks exposed in UK legal control over state surveillance after public awareness of mass surveillance was raised. A number of other important CJEU privacy decisions raised below including the *Google Spain* and *Schrems* cases (both n 150 infra) are also clearly affected by the fallout from Snowden.

[17]R Clarke, "Information Technology and Dataveillance" in C Dunlop and R Kling (eds), *Computerization and Controversy: Value Conflicts and Social Choices* (Academic Press, Inc. Waltham 1991).

[18] Home Affairs Select Committee, *A Surveillance Society?* , Fifth Report of Session 2007-2008 (HC 2008-2009 58-I).

[19]GT Marx, "What's New about the 'New Surveillance'? Classifying for Change and Continuity" (2002) 1 Surveillance and Society 9.

[20]W Webster et al *Deliverable 2.1 – The Social Perspective: A Report Presenting a Review of the Key Features Raised by the Social Perspectives of Surveillance and Democracy'* (Increasing Resilience in Surveillance Societies, 2013).

contemporary surveillance is understood[21] to have moved beyond the simplistic, traditional notions of a top-down 'Big Brother' or Panoptic state, exercising centralised, institutionalised, disciplinary power.[22] Instead, in '*societies of control*'[23] information is gathered to surveille individuals not just via public powers of information gathering , but from a multitude of private actors – including off- and on-line retailers, employers, insurers, tax authorities, health providers and most notably "information intermediaries" such as social networks, search engines, ISPs, and fixed line and mobile telecommunications operators. Individuals are surveilled, for commercial and entertainments well as security and governance purposes[24], categorised and 'socially sorted', with the state no longer acting as the primary collator of data.[25] Cumulatively these factors can be understood as representing a 'surveillant assemblage' where actors develop strategies of governance and control, often via 'centres of calculation' such as police stations, forensic labs and statistical institutions.[26] As Trottier notes, social media policing is thus part of this model of contemporary surveillance, incorporating a range of citizens, devices and personal software into an 'assemblage' that increases the visibility of everyday life.[27]

Surveillance processes in an "assemblage" world are increasingly hard for individuals to perceive, challenge and resist. As Murakami Wood et al assert, a "*surveillance system is infrastructural, and when its workings are shrouded in technical mystique, it is very hard indeed to make a significant difference…individuals are seriously at a disadvantage in controlling the effects of surveillance*".[28] Similarly, a 2009 House of Lords Select Committee Report noted that, "*surveillance practices are often surreptitious, non-transparent, and unaccountable. The aims, motives and procedures of those who collect and use personal information are often unclear, and therefore difficult to regulate, even when they fall within the scope of the law*".[29]

Despite the decentralised, public/private nature of modern online surveillance, the state still plays a vital role in coordinating and consuming surveillance practices. In the PRISM operations, much of the data was collected by Google, Facebook et al but only when made available to state intelligence agencies (possibly not always with the cooperation or knowledge of the data hosts) could deeply coercive sanctions such as imprisonment or interrogation be a consequence. Trottier again states, in the context of

---

[21] K Haggerty "Tearing Down the Walls: On Demolishing the Panopticon" in D Lyon, *Theorizing Surveillance: The Panopticon and Beyond* (Willan, Cullompton 2006); R Jones "Digital Rule: Punishment, Control and Technology" (2000) 2 *Punishment and Society* 5; R Boyne, "Post Panopticism" (2000) 29 *Economy and Society* 285 ; W Bogard, "Welcome to the Society of Control" in K D Haggerty and RV Ericson eds, *The New Politics of Surveillance Visibility* (University of Toronto Press, 2006)

[22] M Foucault *Discipline and Punish: The Birth of the Prison* (Penguin Books, 1979).

[23] G Deleuze, "Postscripts on the Societies of Control" (1992) 59 October Winter Ed 3.

[24] K Haggerty and R Ericson "Surveillant Assemblage' (2001) 51 *British Journal of Sociology* 605.

[25] D Lyon *Surveillance as Social Sorting: Privacy, Risk and Digital Discrimination*, (Routledge, 2003) 13.

[26] Haggerty and Ericson supra n 24 at 613; B Latour *Science in Action: How to Follow Scientists and Engineers through Society* (Harvard University Press, 2001)

[27] D Trottier *Social Media as Surveillance: Rethinking Visibility in a Converging World* (Ashgate Publishing, 2012).

[28] D Murakami Wood et al, *A Report on the Surveillance Society*, (Surveillance Studies Network, UK Information Commissioner Office 2006).

[29] Select Committee on the Constitution, *Surveillance: Citizens and the State Vol I: Report* (HL, 2008-2009, 18-I)

social media, "*surveillance implies an overview. This refers to the vantage point of the guard tower, but also the overview provided by digital technologies. Domesticated technologies augment this vantage point. Whereas the guard tower and even the CCTV are a top-down attempt to envision social life, social media are an – often unwilling – collaboration between top-down and bottom-up efforts*".[30]

Having outlined the shifting nature of contemporary surveillance, we now outline how policing has changed, and the rise of social media datamining within this.

3. Intelligence lead policing, SOCMINT and OSINT

Open source ("OSINT") and social media ("SOCMINT") intelligence are new categories [31] of intelligence used by police that have joined already familiar intelligence sources such as SIGINT (signals intelligence), interception of communications from electronic sources (ELINT, or between people, COMINT); and HUMINT (human intelligence sources). To understand the significance of the rise of SOCMINT and OSINT, we first need to discuss the emergence of general "intelligence led policing" in the UK.

Tilley defines intelligence led policing as a way of doing practical police business by "*more smartly, incorporating modern information technology and modern methods*".[32] This involves "*developing and maintaining a detailed and up-to date picture of patterns of crime and criminality in order to intervene most effectively to disrupt networks and remove prolific offenders*".[33] Newburn, Willliamson and Wright highlight that the shift from reactive to proactive intelligence-led investigations means "*although police officers are still certainly involved, there are now many non-police actors – from civilian surveillance operatives to highly skilled analysts, without whom the system could not work".*[34]

The growth of 'managerialism' in UK policing has been a key driver of intelligence led policing.[35] Managerialism incorporates a strategic planning approach guided by monitoring of statistical data and a reliance on inter-agency cooperation; a focus on 'standards of service' as a means of measuring efficient uses of resources; and a reliance on 'actuarialism'.[36] Actuarialism promotes risk based approaches to managing crime, as opposed to merely preventing or stopping it, and considers the population in terms of

---

[30] See D Trottier, supra n 27.
[31] D Omand, J Bartlett, and C Miller, "Introducing Social Media Intelligence" (2012) 27 Intelligence and National Security Review 1.
[32] N Tilley "Modern Approaches to Policing: Community, Problem Orientated and Intelligence Led" in T Newburn *Handbook of Policing* (Willan Publishing, 2008) p 373 at 383.
[33] Ibid at 384.
[34] T Newburn, T Williamson and A Wright *Handbook of Criminal Investigation* (2nd edn, Willan Publishing, 2008) at p 653.
[35] T Jones and T Newburn, "The Transformation of Policing: Understanding Current Trends in Policing Systems" (2002) 42 British Journal of Criminology 129 at 136; J Ratcliffe, *Intelligence Led Policing* (Willan Publishing, 2008).
[36] A F Bottoms, "The Philosophy and Politics of Punishment and Sentencing" in C. Clarkson, and R. Morgan, eds *The Politics of Sentencing Reform* (Clarendon Press, 1995) at p 25.

their statistical likelihood of deviance[37]. Further drivers like the influential reports *Helping with Enquiries: Tackling Crime Effectively*[38] and *Policing with Intelligence*[39] coalesced into the emergence of the National Intelligence Model in 1999 which became the roadmap for how police should collect and use intelligence[40]. Thus, when online social media arrive, capturing intelligence from these sources was a useful and natural progression for police.

OSINT is generally defined as the collection, analysis and use of data from openly available sources, for intelligence purposes.[41] This includes the mining of social media intelligence (SOCMINT).[42] Some legal commentators have argued that OSINT data is data on a public website that is accessible without further authorisation or controls [43]. The precise categories of SOCMINT are discussed in depth at p XX below.

Generally, SOCMINT involves the analysis of social media to understand and measure the '*visage of millions of people digitally arguing, talking, joking, condemning and applauding'* online, in order to '*identify criminal activity, indicate early warning of outbreaks of disorder, provide information and intelligence about groups and individuals, and help understand and respond to public concerns'*.[44] Bartlett and others in an important report for the thinktank Demos (hereinafter the "Demos report") highlight the importance of aggregated social media data for conducting sentiment analysis or trend analysis[45] and note the value of crowd-sourcing information from individuals and how listening to social media via "*powerful 'big data' acquisition and analytics tools can help the police spot emerging events, piece together networks and groups, discern public attitudes and improve situational awareness*" [46]. Examples of operational OSINT and SOCMINT systems include the EU Virtuoso OSINT platform[47], which provides a toolkit to pull together OSINT for law enforcement strategic decision making, and Raytheon's RIOT (Rapid Information Overlay Technology) OSINT big data analytics system[48] which tracks a subject over different social networking sites using past location

---

[37] M Feeley and J Simon, "The New Penology: Notes on the Emerging Strategies of Corrections and Its Implications" (1992) 30 Criminology 449 at pp 452-454; K Haggerty and R Ericson, *Policing the Risk Society* (Clarendon, 1997).

[38] Audit Commission, *Helping with Enquiries: Tackling Crime Effectively* (Audit Commission 1993).

[39] HMI Constabulary *Policing with Intelligence* (HMIC 1997).

[40] A. James, *Examining Intelligence Led Policing: Developments in Research, Policy and Practice* (Palgrave Macmillan, 2013) pp 81-86.

[41] BJ Koops, J Hoepman and R Leenes, "Open source intelligence and privacy by design" (2013) 29 Computer Law and Security Review 676.

[42] BJ Koops, "Police Investigations in Internet Open Sources: Procedural Law Issues" (2013) 29 Computer Law and Security Review 654.

[43] I Walden, *Computer Crimes and Digital Investigations* (Oxford University Press, 2007); N Seitz "Transborder search: A new perspective in law enforcement?"(2005) 7 Yale Journal of Law and Technology 24.

[44] D Omand, J Bartlett and C Miller *#Intelligence* (Demos, 2012) at http://www.demos.co.uk/files/ Intelligence - web.pdf?1335197327 .

[45] J Bartlett et al *Policing in an Information Age* (Demos, 2013). Hereafter the "Demos report". The legal (and ethical) analysis in this report was it appears partly updated in Part 4 of J Bartlett and L Reynolds *The state of the art 2015: a literature review of social media intelligence capabilities for counter- terrorism* (Demos, September 2015, hereafter "Demos 2015").

[46] Ibid at 6.

[47] Koops, supra n 42.

[48] S Vaughan-Nicols, "Raytheon Riot: Defense Spying is Coming to Social Networks" , *ZDNet*, 12 February 2013 at http://www.zdnet.com/raytheon-riot-defense-spying-is-coming-to-social-networks-

(e.g. from 'check ins' on services like Foursquare or metadata in photographs published online) and network associations/relationships[49].

Clearly such data can be invaluable to the police, but the Demos report also sounds a note of caution over police adoption of these new resources without proper consideration of privacy, trust and public confidence. Under the British National Security Strategy, they point out, "*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*". [50] Accordingly the report concluded that police use of SOCMINT must be "*grounded in respect for human rights and the associated principles of accountability, proportionality and necessity*"[51].

4. Contemporary police use of social media: the London Riots, 2011, and beyond

The US has made extensive use of SOCMINT for some while : eg, in a 2012 survey of 1,221 federal, state, and local law enforcement agencies, four out of five law enforcement professionals stated they used social media for investigations[52] . The US has also been a leader in the use of such data for predictive profiling: solutions like IBM Predictive analytics use crime statistics, statistical analysis/modelling, and GIS mapping to predict hotspots where police can target resources. Famously, the Memphis Police Department in the US claim this system has reduced serious crime overall by 30% since its introduction.[53] Crowd-sourcing intelligence, via social media (or bespoke apps, as with Facewatch below) has however proved to have pitfalls in the US experience. Twitter and Reddit were used extensively to assist in identification of suspects in the investigation of the Boston Marathon Bombings of 2013[54] but notoriously, this public involvement lead to the false identification and victimisation of an innocent individual. [55]

7000011191/ ; R Gallagher "Software that Tracks People on Social Media Created by Defense Firm" *The Guardian* , 10 February 2013 at http://www.theguardian.com/world/2013/feb/10/software-tracks-social-media-defence.

[49] A broad overview of the various SOCMINT techniques was conducted by Bartlett and Miller for Demos; see *The State of the Art: A Literature Review of Social Media Intelligence Capabilities for Counter-Terrorism* (Demos, 2013)

[50] D Omand, J Bartlett and C Miller, supra n 31 at 7; D Omand, J Bartlett and C Miller for Cabinet Office, *A Strong Britain in an Age of Uncertainty: The National Security Strategy* (HMSO 2010), p 5.

[51] Ibid at 23.

[52] LexisNexis "Law Enforcement Personnel Use of Social Media in Investigations: Summary of Findings" (LexisNexis Risk Solutions Government, 2012)

[53] IBM, 'Memphis PD: Keeping ahead of criminals by finding the "hot spots"' *IBM Press Release* Feb 2011 at http://www.ibm.com/smarterplanet/us/en/leadership/memphispd/assets/pdf/IBM_MemphisPD.pdf .

[54] J O'Mahony "Boston Marathon Bombs: How Investigators use Technology to Identify subjects' *The Telegraph*, 19 April 2013 at http://www.telegraph.co.uk/technology/social-media/10005569/Boston-Marathon-bombs-how-investigators-use-technology-to-identify-suspects.html ; T Simonite, "How Facial Recognition Tech Could Help Trace Terrorism Suspects" *MIT Technology Review,* 18 April 2013 at http://www.technologyreview.com/news/513901/how-facial-recognition-tech-could-help-trace-terrorism-suspects/.

[55] R Sanchez "Boston Marathon Bombings: How Social Media Identified Wrong Suspects" *Telegraph*, 19 April 2013 at http://www.telegraph.co.uk/news/worldnews/northamerica/usa/10006028/Boston-marathon-bombings-how-social-media-identified-wrong-suspects.html.

In the UK and EU, partly due to lack of resources, usage has emerged more slowly. Trottier in 2015 described use of SOCMINT by police in 13 countries in the EU as well as the UK, as still "*in a formative stage"*[56]. The pivotal incident in the UK where social media came to the fore in police was the so-called "London riots" of August 2011, driven initially by anti-austerity protestors but also involving widespread looting and criminality in London and beyond. Police in London intercepted and made use of encrypted Blackberry instant messages;[57] cooperation was ongoing between Blackberry owner RIM and law enforcement authorities to assist in apprehension of suspects[58] and activity on Twitter was tracked via hash tags such as #ukuncut[59]. In the wake of the London riots, London police created a Flickr account, and uploaded photos of suspects, asking members of the public to identify anyone they recognised. 770 people were arrested, and 167 charged as a result.[60] 2,800 CCTV images taken during the riots were also uploaded to the mobile app, Facewatch, which individuals could use to identify suspects for the police.[61] In Manchester[62], where rioting also occurred, information was crowd-sourced across both physical and online space (mainly Facebook and Flickr) in "Operation Shop a Looter" Campaigns through local broadcast media in conjunction with posters and electronic displays asked the public to name and shame suspects around the city (particularly at train stations, cinemas and bar complexes)[63].

Subsequent to the London riots, a number of reviews (discussed below) considered what mistakes had been made during the riots in relation to the use of social media. Lessons learned were eagerly applied during the London Olympics the following year. In 2012, ahead of the London Olympics, some 2,565 intelligence reports were created, following analysis of 31 million items across 56,000 social-media platforms. More recently, photographs and geotags posted by foreign fighters in Syria have been used extensively to identify their likely locations and travel routes, and to build material and evidence for investigations[64].

The UK has also begun, like the US, to specifically use predictive profiling, drawing on social media as one among several sources, as a tool of criminal investigation. The

---

[56] D Trottier "Open Source Intelligence, Social Media and Law Enforcement: Visions, Constraints and Critiques" (2015) 18 (4-5) European Journal of Cultural Studies 542.

[57] V Dodd, 'Police accessed BlackBerry messages to thwart planned riots' *The Guardian* (London August 16 2011); K Wynn, and K Blyth, "Predicting a riot: at what price privacy?" *Practical Law Company* at http://uk.practicallaw.com/9-507-6354

[58] C Williams, "London Riots: BlackBerry Manufacturer offers to help police", *Telegraph*, August 8 2011 at http://www.telegraph.co.uk/technology/blackberry/8689313/London-riots-BlackBerry-manufacturer- offers-to-help-police-in-any-way-we-can.html.

[59] J Ball and P Lewis "Twitter and the Riots: How the News Spread", *Guardian*, 7 Dec 2011 at http://www.theguardian.com/uk/2011/dec/07/twitter-riots-how-news-spread.

[60] Demos report , supra n 45, p 2.

[61] The system was used in the West Midlands in 2013 too, with 200 images uploaded, and 9 arrests made : see BBC News, "Crowd-sourcing used to trace London Riot suspects", 26 June 2012 at http://www.bbc.co.uk/news/uk-england-london-18589273; BBC News "Facewatch app shares CCTV Images of West Midlands Suspects", 24 May 2013 at http://www.bbc.co.uk/news/uk-england-22656258.

[62] E Pieri, "Emergent policing Practice: Operation Shop a Looter and Urban Space Securitisation in the aftermath of the Manchester 2011 Riots" (2014) 12 *Surveillance and Society* 1, 38.

[63] Interestingly, during this operation, police also targeted specific known offenders, by using Automatic Number Plate Recognition (ANPR) to track number plates of cars belonging to convicted criminals, to stop them from entering central Manchester during the unrest; see Pieri supra at 44

[64] See RUSI report, supra n 10, para 3.65.

US developed IBM Predictive system, noted above, was trialed by the Ministry of Justice and various UK police departments in 2010[65] . More recently, the EMOTIVE project used an OSINT approach to plot the mood of the nation via Twitter. The experimental software scans up to 2000 tweets a second and rates them on a scale of eight emotional expressions, to assist in identifying geo-specific civil unrest, track potential criminal behaviour or early threats to public safety [66]. PredPol, a software system designed to aid in predictive profiling, has been adopted by a number of UK police forces including Kent since early 2013. Prospective Mapping West Yorkshire Police, is used to forecast hotspots for burglary and theft from vehicles[67]. Predictive profiling has also been explicitly embraced as a valuable tool in an era of austerity[68]. In response to this expanded uptake, the College of Policing issued explicit national guidance on collection and use of open intelligence in 2015[69]. A large number of research initiatives at UK and EU level continue to explore the use of SOCMINT for predicting crime, its location and occurrence patterns: a recent major project is the ePOOLICE system , which links police data to social media data to identify new crime trends in cybercrime, human trafficking and drug trafficking[70]. With the creation of the Metropolitan Police special open source intelligence unit in 2013, OSINT and SOCMINT seem on the agenda for the foreseeable future[71].

4.1 Problems with SOCMINT uptake by UK police

A number of reports following the London Riots considered the use of OSINT and SOCINT intelligence by the police. The Metropolitan Police Service *5 Days in August: Strategic Review into the Disorder of August 2011* report found that social media was a primary source of information during the riots, with 19% of total information reports assessed by their intelligence body relating to social networking sites, and 14% relating to BBM (337 and 249 out of 1554 reports respectively). Nevertheless, they highlighted that there were insufficient resources to manage the volume of data in real time, including open source data, with the lack of automated search tools slowing progress. Available tools were optimised for business intelligence gathering, not policing.[72]

[65] T Thomson "Crime Software may help police predict violent offences" *Guardian*, 25 July 2010 at http://www.guardian.co.uk/uk/2010/jul/25/police-software-crime-prediction ; IBM, 'Ministry of Justice Chooses IBM Predictive Analytics to Make Streets Safe'a*IBM Press Release*, 16 March 2010.
[66] BBC News "Computer Program uses Twitter to "map mood of nation" September 7 2013 at http://www.bbc.co.uk/news/technology-24001692 .
[67] Both described in M Fielding and V Jones "Disrupting the optimal forager': predictive risk mapping and domestic burglary reduction in Trafford, Greater Manchester" (2011) 14 International Journal of Police Science and Management 30.
[68] See HMIC *Policing in austerity: Rising to the challenge* (2014) and discussion of Kent use of PredPol at p 63.
[69] See College of Policing *Intelligence collection, development and dissemination* (2015) at https://www.app.college.police.uk/app-content/intelligence-management/intelligence-cycle/ . Interestingly the guidance notes that open intelligence may not be "accurate, reliable or valid", is not subject to the same quality standards as closed sources and should be corroborated by supporting evidence.
[70] See POSTnote 470 "Big Data, Crime and Security" 2014.
[71] See P Wright "Meet PRISM's Little Brother: SOCMINT" *Wired*, 26 June 2013; *Social media intelligence technology: Home Office funding*, FOI release, 22 August 2013.
[72] Metropolitan Police Service, *Four Days in August: A Strategic Review into the Disorder of August* 2011 (Met Police Service 2012) 105.

Meanwhile a Review by Her Majesty's Inspectorate Constabulary proposed a New National Framework of "Rules of Engagement" for resolving public disorder in the wake of the 2011 London Riots.[73] Beyond broader police planning and tactical strategies for mobilisation, this framework proposed a "*central information hub to help them [police] anticipate disorder by drawing together all available information, including from direct contact with members of the community and social media monitoring*".[74] Significantly, the report noted that police often felt overwhelmed by the volume of social media data available. One interviewed officer compared the process of finding useful material to "*searching the British Library for a page in a book, without an index to refer to*".[75]

These reports highlighted the need for police to improve technical and organisational capabilities to extract and analyse SOCMINT. A third, non-state report, the *LSE/Guardian "Reading the Riots" Report*[76] helped highlight *methodological* issues in use of SOCMINT data for policing. For example, one prominent question was whether social media had helped incite the riots[77] or alternately helped combat them and restore normality afterwards. The report found evidence that Facebook and Twitter were not as heavily used as Blackberry Messenger texts during the London Riots; and Twitter was more heavily used to coordinate the subsequent clean-up campaign than to organise illegal behaviour. However, Omand, Bartlett and Miller have argued this could not be proven because the 2.6 million tweets analysed in the project were collected via 150 hashtag clouds. Tweets lacking a specific hashtag e.g. #londonriots would systematically be excluded from the dataset. Users co-ordinating criminal looting via Twitter might not wish to broadcast this using hashtags, hence rendering dubious the empirical basis of the assertion.[78]

The reliability and quality of OSINT and SOCMINT, and of data mining and profiling in a policing context generally, has been recognised as an increasingly serious issue since 2011. Certainly the recent rash of social media prosecutions has demonstrated that what people say on line will often not reflect their real intentions – Paul Chambers, for example, did not really mean to blow up Doncaster Airport,[79] nor did most the abusers of Criado-Perez[80], one assumes, really intend to rape her. This does not (in any way) mean that social media abuse is not a vice which should not be investigated or prosecuted by police under appropriate laws,[81] but it does raise worries, in an era characterised (as discussed above) by automated profiling and predictive policing, that SOCMINT profiling, combined with the well-known online "disinhibition effect", will produce bad data, which will lead to bad arrests, bad prosecutions and possibly even

[73] Her Majesty's Inspectorate of Constabulary, *Review into the Disorder of August* (hereafter "HMIC, 2011") .
[74] HMIC , 2011 at 6.
[75] HMIC, 2011 at 31.
[76] Lewis et al, supra n 8.
[77] At one point David Cameron actually threatened powers to close down social media in times of emergency. See BBC News, "David Cameron considers banning suspected rioters from social media", 11 August 2011, at http://www.theguardian.com/media/2011/aug/11/david-cameron-rioters-social-media .
[78] Omand et al, supra n 31 at 9.
[79] See *Chambers v DPP* [2012] EWHC 2157 (Admin).
[80] Criado-Perez story, *The Guardian* (London, 16 December 2013) at http://www.theguardian.com/uk-news/2013/dec/16/two-charged-caroline-criado-perez-tweets.
[81] Dominic McGoldrick "The Limits of Freedom of Expression on Facebook and Social networking Sites: a UK perspective" (2013) 13 (1) *Human Rights Law Review* 125.

bad convictions. Social science researchers in the area also increasingly express dismay that data mined from corpora of (say) tweets fail in meeting basic social science methodological standard. Issues such as sampling, standardisation of populations and exclusion or under representation of certain populations, such as the poor, sub-literate or technophobic are often ignored. While important, this issue is beyond the main focus of this paper.

5. Legal regulation of SOCMINT

Above, we have already classified SOCMINT as either open (accessible to the world) or closed (restricted by Friends locks, passwords, encryption etc). For the purposes of this section we extend this taxonomy further to four classes of social media data.

1. Information (open or closed) which does not (at least *prima facie*) relate to a living person who is identified or identifiable ("personal data[82]" in the terminology of the DPD) eg, aggregated data about views on local transport, bus timetables posted online. This material is irrelevant to the concerns of this paper so long as it remains anonymous or anonymised[83].

2. Information which is open and does relate to a living person who is identified or identifiable. This is the main focus of this paper.

3. Information which is open *or* closed, but is accessed by police via *deceptive, covert or misleading tactics* : eg acquiring access to Friends-locked posts on Facebook via befriending as an invented person with a fake profile; "listening in" on a protected group after joining with an anonymous profile; leveraging certain public responses on Twitter via presentation of provocative messages not truly held by the police observer; searching for content via Google or other engine despite the content provider indicating via the robots.txt standard that they wished this information not to be spidered.

4. Information which is closed *but is accessed via technical "back door" access*, eg email or direct message interception by wiretapping technologies; acquiring traffic logs of communications from ISPs[84].

[82] See Art 2 (h), DPD 1995 transposed into the DPA 1998 s1(1).

[83] This phrase can be misleading given the potential reidentifiability of much anonymous data in a world of data mining and "big data". In the examples given above it is not at all unlikely that views could be reidentified to a particular speaker. Again this important topic cannot be examined in depth in this paper. See the seminal discussion in P Ohm 'Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization', (2010) 57 UCLA Law Review 1701.

[84] Into this category also arguably falls access to social media by law enforcement authorities and security services via extralegal, or at least legally unclear and secret programmes. As disclosed by Edward Snowden, such surveillance and intelligence acquisition programmes are now known to be run by the US NSA, the UK's GCHQ and probably the intelligence agencies of many other countries. The discovery of PRISM, TEMPORA etc. is clearly of considerable importance in the whole debate about regulation of intelligence gathering from social media, but it is not, as discussed above, the main focus of this paper which is devoted to how conventional police and law enforcement authorities acquire intelligence in what must be seen to be justifiable and transparent manners.

The primary instrument which currently regulates acquisition of communications and records by law enforcement authorities in the UK is the Regulation of Investigatory Powers Act 2000 (RIPA) [85]. It should be noted that police acquisition and processing of personal data is exempt from much of DP law[86] under the exemptions for the detection and prevention of crime[87]. Notably, there is no requirement for UK data subjects to consent to the police collecting or otherwise processing their data, nor do they have rights to request access to what data is held about them by the police ("subject access rights", or SARs)[88]. Data processing by the police will also not always fall under the supervision of the UK DP regulator, the Information Commissioner, since other regulators have a specific role eg the Interception of Communications Commissioner[89]. Retention of data, the security of its storage, and other matters addressed by the Second to Eighth Data Protection Principles[90] are however key areas where DP rules are still applicable to policing.

Overarching the domestic laws of the UK Data Protection Act 1998 (DPA) is art 8 of the ECHR, which guarantees the right to respect for private life as a fundamental right. A number of challenges to the legality of RIPA in terms of art 8 have already been taken to the European Court of Human Rights (ECtHR); however the basic legality of the scheme, at least before the Snowden revelations, was confirmed in *Kennedy v UK*.[91]

As a key *a priori* question, we will now consider whether *personal data* disclosed in *public* on *social media* should attract privacy protection, and if so, how much.

6. Does open SOCMINT have any protection in law as "private"?

6.1 *Norms and law : DP, RIPA and the new IP Bill*

[85] Internet related matters and telecoms are generally dealt with as reserved matters to the UK Parliament. We shall not in this paper consider the legal specialities of Scotland which has its own criminal justice and evidence system and thus amends Regulation of Investigatory Powers Act 2000 (RIPA 2000) to some extent within the Regulation of Investigatory Powers (Sc) Act 2000.

[86] See however the potential introduction of a draft directive on data protection in police and criminal matters: Proposal for a Directive of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data by competent authorities for the purposes of prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and the free movement of such data /* COM/2012/010 Final - 2012/0010 (COD) at http://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX:52012PC0010>.

[87] DPD 1995, art 3(2); DPA 1998 ss 28, 29.

[88] DPA 1998, s 7 and s 29(2) .

[89] During the London riots the Information Commissioner made it plain that he claimed no jurisdiction over police acquisition of social media, texts etc and would not comment on its legality. See Information Commissioner Office Disclosure Log, (14 October 2011) at http://ico.org.uk/~/media/documents/disclosure_log/IRQ0417298.ashx. RIPA does however come under the supervision of the Investigatory Powers Tribunal. The complicated supervisory arrangements in RIPA and adjunct legislation are likely to be overhauled and simplified with the creation of one "super regulator" for investigatory powers in the IP Bill.

[90] DPA 1998, Sched 1.

[91] (2010) 50 EHRR 45. Though see now *Davis* case, supra n 16 and further challenges to RIPA after Snowden are apparently pending in both ECtHR and the ECJ. It should also be noted that the EC Charter of Rights contains separate rights both to data protection and to respect for privacy; and these different concepts have been deployed in recent ECJ case law eg *Google Spain*, infra n 150 and are thus significant to the UK even though the UK has not signed up to the Charter.

First reactions to the idea that there should be some kind of legal control of the collection of open SOCMINT tend to run into *a priori* objections of the following kind: "*Of course* police can read public posts - they're public!"[92]. Legal experts, if concerned at all, have mostly responded in a similar way. For example, Gillespie, a leading commenter on online criminal law, asserts that:

> "[*W]hen postings are public and available for all to see it is unlikely that it could be concluded that the viewing of the information is covert in that there must be an awareness that those in authority could look at the postings".*[93]

DP law may also support this position. Art 8(2)(e) on special categories of data (sensitive personal data in UK terminology) normally requires explicit consent for the processing of such data but exempts "*data which are manifestly made public by the data subject"*. It is not clear if this means that data which is *not* sensitive should be exempted from the requirement for a legal ground of processing under art 7 - but it would seem to follow. Furthermore, as pointed out above, the police are exempted from DP rules requiring a lawful ground for the acquisition of personal data.

But not all writers see things as this uncomplicated. In the US, as an interesting comparison from a very different legal system, Bartow, drawing on Semitsu, comments that: "*Facebook [is] a giant surveillance tool, no warrant required, which the government can use in a mind bogglingly creative range of ways with almost no practical constraints from existing laws*". [94] For Bartow, social media surveillance is a covert device to evade the generally strong US Fourth Amendment protections against warrantless search of private material. "[*W]e barely notice that Facebook leaves us almost completely vulnerable to searches and seizures triggered by invasive but mostly invisible government surveillance*", she protests.[95]

We note three points rebutting the "obviousness" of the notion that SOCMINT should have no protection as private matter because someone has chosen to voluntarily disclose their personal and private life to the world.

First, what can be gathered from open SOCMINT is not just the most obvious part, the substantive content, eg text, pictures, videos or links posted[96]. For intelligence

---

[92] See for interesting sentiments along this line, a police forum online *PoliceServices.com* at http://www.policespecials.com/forum/index.php/topic/116579-ripa-requirement-required-for-social-networking-sites/ eg "Sailor": "You don't need a RIPA authorisation to monitor what people post publicly on the internet", 23 February 2011.

[93] A Gillespie, 'Regulation of Internet Surveillance" (2009) 4 *European Human Rights Law Review* 552.

[94] J Semitsu, 'From Facebook to Mug Shot: How the Death of Social Networking Privacy Rights Revolutionized Online Government Surveillance' (2011) 31 *Pace Law Review* 1, 291-381; Ann Bartow 'Facebook and the Fourth Amendment: Expecting Any Privacy May be Unreasonable" *Jotwell* (18 April 2011) at http://cyber.jotwell.com/2011/04/.

[95] To make matters worse in the US, police scrutiny even of closed social media content also does not breach privacy according to Semitsu.

[96] d boyd, *It's Complicated: the Social Lives of Networked Teens* (Yale University Press, New Haven 2014)

purposes, one of the most useful pieces of data that can be extracted from a social media profile is the network of friends or "social graph". Interestingly, on Facebook, the world's most popular social network, it has until recently been impossible, and is still very hard, to shield the Friends list from public exposure[97]. Similarly, Twitter now by default sets tweets to reveal the location from which they were tweeted[98]. Although these settings can be reversed, many users will not know how or bother. Thus valuable location data becomes available to those with access to the Twitter Application Program Interfaces (APIs) [99].

A second obvious but significant point is that social media content about an identified or identifiable person is often not posted by that data subject. On closed Facebook profiles, a photo might be "tagged" with the name of a person who might not even themselves have a Facebook account, and so have no access or notice to remove the tag. Information may quite conceivably be "cut and pasted" from Friends locked profiles eg by hackers, or ex-es[100].

Finally, privacy settings vary from platform to platform and also change constantly over time in a way that requires constant vigilance of users to maintain a privacy status quo. Different privacy settings, and different changes, apply to different types of content eg posts, comments, groups, photos, friends list etc. On most sites, as with Facebook, the overwhelming motivation is to make as much material as possible public to maximise growth of audience and collection of data for marketing revenue. Hence it is well known that many users are deluded in their belief they have adequately protected their privacy via code controls.[101] Indeed Madejski, Johnson and Bellovin found that in a small study of 65 university students, every one had incorrectly managed some of their Facebook privacy settings, thus displaying some personal data to unwanted eyes.[102]

Put together, all these points show that, contrary to popular belief, control of what data about you is made public on social media is not simply a matter of easy voluntary choice. Accordingly the common retort – if you didn't want people (like the police) to read it, why did you make it public? – is not in fact a sensible question to ask. We would argue this contributes strongly to an argument that material placed on

---

[97] Kurt Wagner "Your Private Facebook Friends List Isn't Actually That Private", *Mashable* (New York 2 June 2014) at http://mashable.com/2014/06/02/facebook-friends-list-privacy/.

[98] Note the Twitter privacy policy as of 3 November 2014: "*Our default is almost always to make the information you provide public for as long as you do not delete it from Twitter, but we generally give you settings to make the information more private if you want.*" At https://twitter.com/privacy.

[99] APIs are bespoke packages of tools provided to developers in an accessible interface that they can use to build applications.

[100] Information posted privately is increasingly often hacked and shared "for the lulz" (fun) or for other purposes such as revenge or blackmail: a recent high profile case involved the hacking from the iCloud of celebrity nude pictures including Jennifer Lawrence. See Samuel Gibbs "Google removes results linking to stolen photos of Jennifer Lawrence nude", *The Guardian* , 20 October 2014 at http://www.theguardian.com/technology/2014/oct/20/google-search-results-linking-stolen-nude-photos-jennifer-lawrence .

[101] L Edwards, 'Privacy, Law, Code and Social Networking Sites' in I. Brown, *Research Handbook on Governance of the Internet* (Edward Elgar, 2013) at 321.

[102] M Madejski, M Johnson and S Bellovin "The Failure of Online Social Network Privacy Settings", (Tech Report CUCS-010-11, Columbia University, 2011)

"open" social media can still carry with it reasonable expectations of privacy. (Or in DP terms, may not have been "manifestly" made public.)

Turning to our major case study of the UK, the obvious place to look for any protection in law against police acquisition and use of open SOCMINT is, as noted above, in RIPA, which is supposed to be the main instrument governing the police's investigatory and interception powers[103]. However there is little evidence that the police regard RIPA as constraining their activities in our paradigm case, that of gathering large amounts of open SOCMINT and data mining it for profiling purposes. While there is no authoritative guidance or case law on this point, the Demos reports of 2013 and 2015[104] have been influential in the field, as has the academic work of O'Floinn and Ormerod[105]. The Demos reports, especially the more recent one, claim as their normative basis that there are neither reasonable *personal* expectations of privacy protection, nor such expectations by *society*, in respect of SOCMINT collection. The 2015 report emphasises this as still true even after the public reaction to the Snowden revelations on social media surveillance. They justify the first assertion on the basis that where users disclose open SOCMINT, they do so knowing that the terms and conditions of the site almost invariably state their data may be shared with others. Furthermore, the Demos reports also assert that if the site allows acquisition of data by API, this also means the user should have expected their posts to be open to public acquisition. We strongly reject these assertions and note a number of reasons below p xx why we do not feel this approach does or should reflect the reality of social media use[106].

The Demos reports do suggest that, in some circumstances, the acquisition of SOCMINT will require some kind of authorisation for police - but not under Pt 1 of RIPA, which was supposed by most to be the portion of the Act dedicated to digital investigations. Instead Demos found applicable restrictions under Pt II or RIPA, which covers rules relating to police collection of evidence via directed surveillance and covert intelligence tactics eg such real world activities as following suspects, befriending their relatives or girlfriends and even in some extreme cases entering relationships with these kind of sources[107].

---

[103] Though note that much digital evidence is still eg reportedly acquired as "production" evidence under the rules of PACE in England and Wales and this is a legal alternative to RIPA as confirmed in *R (NTL Group Ltd) v Ipswich Crown Court* [2002] EWHC 1585 (Admin). One likely outcome of the new IP Bill may be to reduce this regulatory overlap.

[104] Supra n 49.

[105] M O'Floinn and D Ormerod "Social networking sites, RIPA and Criminal Investigations" (2011) 10 Criminal Law Review 766.

[106] Demos state that a number of types of SOCMINT data do not carry obvious risks to the personal privacy of users – eg anonymous or aggregated SOCMINT – or where the user has clearly explicitly given permission to use or share personal data – eg responses to crowdsourced police appeals for information on social media. We do not take issue with these as not carrying any reasonable expectation of privacy.

[107] Such undercover operations have been sometimes found to be both inadequately unauthorised and actionable in damages – see eg. "Police apologise to women who had relationships with undercover officers*", Guardian*, 20 November 2015 at http://www.theguardian.com/uk-news/2015/nov/20/met-police-apologise-women-had-relationships-with-undercover-officers .

Firstly, on some occasions, "*directed covert surveillance*" authorisation under Pt II of RIPA may be required. Directed surveillance[108] is surveillance[109], which is *covert*, conducted in such a manner as is likely to result in the obtaining of *private* information about a person, and for the purposes of a specific investigation and not as an immediate response to events (eg following a suspect from a mugging). A directed surveillance authorisation is obtained not from a judge or politician but internally from a senior police officer[110]. "Private information" is defined widely in s 26(10) to include "*any information relating to private or family life*". [111] Hence it is quite possible to imagine "private" matter being gathered by covert surveillance of SOCMINT.

Useful examples of what Demos suggest need authorisation under this category include:

(a) where a *detailed profile* is built of a *named individual* from openly available sources, especially where this requires either automated or manual name recognition techniques (see further discussion of *Rotaru* below, n 125).

(b) where data is collected from a public platform which is technically open but where a reasonable expectation of privacy still exists eg a Facebook page set up for members of a school hobby club or local choir.

O'Floinn and Ormerod argue however that where a police officer calculates[112] the target will be unaware of surveillance - which may be often true in cases of social media scraping, web search engine "spidering" etc – case law points towards a need for covert surveillance authority. *Contra* Gillespie above p 92, we agree. This would indicate a much wider obligation for police to obtain directed surveillance authorisations in respect of SOCMINT than appears to be current practice.

Secondly, the Demos report argues that some acts of SOCMINT collection require authorisation as surveillance involving "*covert human intelligence sources*" (CHIS). The most obvious case here is where police gain access to closed posts, profiles or groups via the construction of fake profiles. It is well known that the UK police have routinely created fake profiles for some investigations eg posing as children on platforms such as Bebo to try to initiate grooming advances. A slightly less obviously invasive technique is to "hang out" using a false identity or at least, not the identity of a policeman, on any forum, open *or* closed, in order to elicit statements or information, possibly by "astroturfing" – expressing fake "grassroots" or provocative views.

Finally in some circumstances both Demos and O'Floinn and Ormerod propose that the rules of Pt1, Ch1 of RIPA on interception of communications may be relevant. By analogy with known mechanisms such as email interception, it seems tempting to say that acquiring access to a Direct Message (private one-to-one communication) on say

[108] RIPA 2000, s 26(2).
[109] RIPA 2000, s 48(2), where it is defined widely to include monitoring, observing or listening to persons, their movements, conversations, activities etc as well as recording such movements etc, and using "surveillance devices".
[110] RIPA s 28.
[111] Home Office Revised Code of Practice for Covert Surveillance and Property Interference (HMSO 2014), supra nxx.
[112] Supra n XX.

Facebook or Twitter, or to a Friends-locked private post, by technical "backdoor" methods, should require an interception warrant and Demos support this. The case for a reasonable expectation in such circumstances seems undeniable. In fact however it is very hard to analogise social media interception to phonecall or email interception.

The main issue is that interception can only be authorised under RIPA when the communication is "*in the course of transmission*".[113] A DM or Friends-locked comment written to Facebook (say) has arguably been conclusively transmitted and so cannot then be intercepted. However s 2(7) of RIPA extends the period of transmission to: *"any time when the system ... is used for storing it in a manner that enables the intended recipient to collect it or otherwise have access to it".*

This would arguably extend the period of "transmission" until sometime after a post or comment had been posted, but while the "intended recipient" still needed to *"otherwise have access"* to it ie until they actually read it. In *R v Coulson*[114] however (a phone-hacking case) the judge concluded that voicemail messages stored for future access remained "in transmission" even after they had arrived, and *even* after they had been listened to. This was partly on the basis that these messages were tied to the mobile operator's proprietary system and could not be easily moved to "offline" storage to be referred to at leisure. Whether this reasoning is also true of social media posts – they are *also* tied to proprietary non-interoperable platforms - as text, they are also much more easily "cut and pasted" than voicemails. It remains an open problem if an interception warrant could appropriately be issued even after a social media post had been read and perhaps replied to[115]. In effect this would mean the interception regime would be appropriate for an indefinite length of time which seems the opposite of the "realtime" interception RIPA intends.

The new regime proposed to replace RIPA's rules on interception in the IP Bill (as introduced in November 2015) would arguably place such "back door" police access under the new concept of "*interference with computer equipment*", or legalised hacking (Part V of the Bill)[116]. In the new proposed scheme, such interference requires, like interception, a dual authorisation by the relevant Secretary of State and a judge, except in emergency situations. This seems to recognize hacking as equivalent in level of severity of interference with "conventional" interception in the current RIPA framework. Although the entire application of both RIPA and quite possibly the IP Bill to SOCMINT is deeply unhelpful, this particular alteration to requiring the so-called "double lock" of judicial and political authorisation, might be helpful.

6.1.1 Will any of this be clearer under the new proposed IP Bill?

---

[113] RIPA s 2(2).
[114] [2013] All ER (D) 287.
[115] The matter is complicated even further by the fact that, after criticism from data protection authorities, Facebook now allow users to download their entire Facebook page to another host or storage medium (but not, say, merely their entire photo album).
[116] Demos supra n 45 also argue that the use of "keylogger" or Trojan viruses to gain passwords to locked accounts should be viewed as "interception". While the UK lacks case law on this topic, in Germany, use of such keyloggers by police has already been held by their Constitutional Court to infringe a new privacy sub-right of integrity of computer systems – see W Abel "Agents, Trojans and Tags: The Next Generation of Investigators" (2009) 23 (1-2)International Review of Law, Computers and Technology 99.

It is difficult to make any firm statements about a Bill in passage and especially one so complex and controversial that its meaning is likely still to be hotly disputed when it has passed. However the obvious answer is no, for at least three reasons.

First, the Bill, perhaps unsurprisingly, takes no steps towards an overt recognition of expectations of privacy in SOCMINT (whether open or closed).

Secondly, despite the Bill's bold claims that "*it will bring together all of the powers already available to law enforcement and the security and intelligence agencies to obtain communications and data about communications*" and that "*it will make sure powers are fit for the digital age*" in fact it still maintains a bifurcation between powers regulating "digital" investigations in the IP Bill and "real world" powers which apparently remain operative in RIPA Pt II. Accordingly, the protection of SOCMINT, such as it is, remains divided uncomfortably between these regimes.

Thirdly, the new Bill's powers are so wide ranging and in many cases so little understood by the legal community, that it will be quite easy for new aspects of the Bill to be pressed into service as needed at any point when there is felt to be a need to defend the routine acquisition of bulk SOCMINT. We have already noted that the new regime on interference with computer equipment, might provide an answer to the difficulty of stuffing backdoor access to closed SOCMINT into the badly shaped "interception" hole. A similar possible opportunity may lie in the new rules on acquisition of "*bulk personal datasets*", which are defined in the Bill's introductory guidance as "*sets of personal information about a large number of individuals, the majority of whom will not be of any interest to the security and intelligence agencies*.." although the examples of these datasets given namely "the telephone directory or the electoral roll" seem very far from SOCMINT. We hope the application of the Bill's new provisions to SOCMINT may be clarified in time through official guidance as well as case law.

Our preliminary conclusion at the end of this section is that the principal organ of UK law regulating police powers, RIPA, does not appear to offer adequate and consistent protection in relation to the acquisition of open SOCMINT. There is no need for political let alone judicial warrant, nor even for internal police authorisation, except in the very limited and uncertain circumstances described above. Protection is not based in the "digital" part of RIPA but scattered across its parts , and this distinction is maintained even if the IP Bill passes in current form. In the next section we consider whether, notwithstanding this strong UK view that open SOCMINT on the whole does not require legal authorisation for acquisition and use, support can be found from European human rights law for privacy rights in SOCMINT.

6.1 The European Convention on Human Rights (ECHR)

The leading international instrument relevant here[117] is the ECHR, art 8, which guarantees the right to respect for private life, but with extensive exceptions for "*national*

[117] It is worth noting parenthetically that the Council of Europe ETS No 185 Convention on cybercrime, Budapest, 23.XI.2001, which is not a human rights instrument but takes account of such in

security, public safety or the economic well-being of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others". The UK Human Rights Act 1998 (HRA) incorporates the ECHR as part of UK domestic law. Under general principles of the ECHR, any restrictions on rights must be based on published, clear and specific legal rules; serve a legitimate aim in a democratic society; and be 'necessary' and 'proportionate' to that aim. These restrictions persist even where incursions into privacy rights have been justified by states on grounds of national security or prevention or detection of crime: as evidenced by a number of prominent ECtHR cases including *Weber v Germany*[118] and *Klass v Germany*[119]. It is clearly the threat of recourse to the ECtHR, as well as the desire to engender public trust and co-operation, that leads to some public statements that police use of SOCMINT is legitimate, proportionate and necessary.

Significantly for current purposes, in recent years ECtHR jurisprudence has clearly endorsed the notion that expectations of privacy may persist even in public places. The leading case is *von Hannover v Germany*[120]. Here, Princess Caroline of Monaco attempted to prevent pictures being published of her in the German press which did not feature public engagements but merely the Princess going about her ordinary life in public, eg, going shopping, out for a walk. The ECtHR found that even though there was a public interest in the reporting of the activities of public figures, it had to be balanced against her rights of privacy to go about in public "off duty". The decisive factor was whether the pictures contributed to a public debate of public interest: these did not and accordingly her legitimate expectation of privacy prevailed. In the UK, the *Hannover* doctrine has been partially but perhaps not entirely adopted by the courts in a series of cases influenced by the Human Rights Act and the leading House of Lords decision in *Campbell v MGN*[121].

While most "privacy in public" cases have involved the famous, it is clear that non-celebrities have even stronger claims since there is usually no countervailing public interest. In *Peck v UK*[122], a mentally ill man successfully sued the UK at the ECtHR

---

building a cross-border framework for digital police investigation, states in art 32 that "publicly available (open source) stored computer data" can be acquired by police authorities across borders without any authorisation of a foreign state. This provision relates to mutual assistance by foreign states, and the circumstances and age of this provision (drafted before the rise of social media) make it questionable if it has any utility in the context of domestic policing. See on updating art 32, the note by M O'Floinn "It wasn't all white light before Prism: Law enforcement practices in gathering data abroad, and proposals for further transnational access at the Council of Europe" (2013) 5 *Computer Law and Security Review* 610 at http://www.sciencedirect.com/science/article/pii/S0267364913001428.

[118] (2006) 46 EHRR SE5.

[119] (1978-80) 2 EHRR 21 .

[120] (2005) 40 EHRR 1.

[121] [2004] UKHL 22. *Campbell* involved the supermodel Naomi Campbell, who successfully argued that photos taken of her on the street outside Narcotics Anonymous, thus showing she had lied about being free from associations with illegal drugs, were a breach of her privacy. See also *Murray v Big Pictures (UK) Ltd* [2008] EWCA Civ 446 (pictures of the author J K Rowling with infant child taken without her permission on the streets of her home town) which was another successful claim (on behalf of the baby alone) for privacy rights of celebrities in everyday public activity, but interestingly, unlike *Campbell*, had no medical treatment issues (indicating a less strong claim for privacy) but did involve an infant plaintiff (indicating a stronger claim, as the infant had not chosen stardom and had the right to grow up without constant press intrusion).

[122] (2003) 36 EHRR 41.

under art 8 for failing to prevent the transmission on TV of footage of his failed suicide attempt in a public place. What remains unclear is if (or when) art 8 restrains mere *observation* of activity in public, as opposed to its recording, storage, use or dissemination. As Koops notes, the fact that the ECtHR has recognised the existence of privacy rights in public does not mean that *all* processing of personal data gathered in public interferes with privacy.[123] The main ECtHR cases in this area have dealt with storage, and in some cases, subsequent use or dissemination of data, rather than its mere collection, "*leaving open the question whether the mere searching for and consultation of data, without storing or using them, constitutes an interference*".[124]

The key dictum most often quoted is from *Rotaru v Romania*[125] where it was held that:

> "*public information can fall within the scope of private life where it is systematically collected and stored in files held by the authorities*"

Koops asserts this means that collection of information *not* so systematically processed may therefore not constitute infringement of private life. In UK case law, the House of Lords has so far only been willing to go as far as saying that the *publication* of data collected in public (eg a photograph snapped in the street) may, exceptionally, be objectionable in all the circumstances of the case, but not its mere collection[126].

6.1.1 Collection of personal data in public by policing authorities

Most of the leading UK cases on a right to privacy in public to date involve infringements of privacy by media organisations. Is there a greater or lesser right as against the *police*? On the one hand, the police clearly have a special role and (limited) authorisation to interfere in private spheres to protect the public. On the other hand where police overstep their powers, the result is arguably a totalitarian state – a worse consequence than where the media over reach. In *Wood v Metropolitan Police Commissioner*[127], a protester against the arms trade argued that his privacy was breached when a police photographer took pictures of him at the AGM of Reed Elsevier, and later attempted to identify him via these and transport documentation. Wood plead that his art 8 rights had been infringed. The police argued that there was no permanent file on Wood and the pictures were to be kept only for sight of police officers to prevent future crimes. Mr. Justice McCombe in the High Court noted:

> "*The majority of the recent high authorities, here and in Strasbourg, are concerned with Article 8 in the context of media intrusion and publication of material relating to celebrity figures with high public profiles…There are few recent cases addressing intrusions on privacy by the state, but .. it is perhaps intrusion by the state with which the draftsmen of the Convention would have been particularly concerned in 1949 and I felt throughout the case the im-*

[123] Koops, supra n 42.
[124] Ibid at 656.
[125] (2000) 8 BHRC. 449 .
[126] *Wood v Commissioner of Police for the Metropolis* [2008] EWHC 1105 (Admin) paras 34 and 43, citing *Campbell , Murray* and *von Hannover.*
[127] [2008] EWHC 1105 (Admin).

> *portance that the courts should attach to vigilance in this area, while recognising the difficulties of police forces in democratic societies in protecting us all from criminal behaviour."*

Interestingly the judge went on to recall the activities of the Staatssicherheitsdienst ("Stasi") in the former German Democratic Republic who were renowned for employing thousands of informers as well as conducting near universal surveillance as far as was possible in a pre-digital world[128]. He continued:

> "*One would hope, for example, that such extreme intrusions would be protected under the ECHR…The allegedly intrusive activity here is, of course, at a far lower level than that, but, in my judgment, it is the development of such state activity against which one has to be vigilant*." (paras 27-28)

Despite this opening assertion, the judge found no invasion of privacy had occurred. Drawing on previous ECtHR cases[129], he found the English courts had "*adopted a very robust approach to questions of interference with rights under Article 8(1) in relation to the taking of photographs in public places... in assisting in the detection of crime*". Before art 8 could be engaged, the infringement of privacy had to reach "a certain level of seriousness"; secondly there had to be a reasonable expectation of privacy; third, the justifications available to the state in art 8(2) had to be examined. The judge found the plaintiff was a known activist attending a public meeting and thus had very little expectation of privacy. Significantly, the images were to be retained, without general disclosure, for very limited purposes, and not as part of a specific dossier on the data subject. Accordingly, there was no interference with the claimant's rights under art 8(1). Even if there had been, it would have been justified in context as in accordance with the law and proportionate.

Most recently, and most appositely for this paper, in *Application for Judicial Review (Northern Ireland)*[130] a boy aged 14 was caught on CCTV involved in serious rioting in Derry. As part of an active police campaign, "Operation Exposure", his image was published and distributed in leaflets put together by the police. He argued his rights under art 8 of the ECHR had been violated. The Supreme Court judges found against the plaintiff by various routes : either art 8 was not engaged (eg because rioting was not part of protected "private life") or that it was, but the interference by the police was justified. Part of what backed this up was that the police had followed a "painstaking" process and only circulated the boy's image as a "last resort"[131].

The most interesting dicta come from the speech of Lord Kerr, who found that art 8 was indeed engaged because of the boy's age and the effect the publication of the photographs may have on him. He made it clear however that this finding did not rest

[128] See discussion comparing Stasi activities to modern ubiquitous surveillance of digital communications in a post-PRISM world in talk by Judith Rauhofer, Staff Seminar Series Edinburgh, (Edinburgh, 18 March 2014) ; slides no longer available online.
[129] See *X v UK* (1973) Decision 5877/72, *Friedl v Austria* (1995) 21 EHRR 83, *PG & JH v UK* (2001) Appl. 44787/98 and *Perry v UK* (2004) 39 EHRR 3.
[130] [2015] UKSC 42.
[131] Ibid, para 76-77.

entirely on whether the boy had a "reasonable expectation of privacy" but was dependent on a myriad "contextual" factors[132]. He also stressed that the fact that the boy was suspected of criminal activity did not *alone* remove any possibility of interference under art 8[133]. However in this case, the advantages to the boy of being diverted from criminal activity, as well as the interest of his community in the prevention of crime and apprehension of offenders, outweighed his interest in privacy[134].

From the above, we can discern *a priori* support in the ECHR for reasonable expectations of privacy in public space, especially when personal data is not only collected but further used, eg published, arranged in dossiers for potential future access (see further 7.3 below), search and use. However, we can also see that the English courts in their interpretations are reluctant to let such expectations escalate to the point where a "public domain" of activity is diminished by privacy rights. Furthermore, despite the potential threat of unfettered police invasion of privacy, the courts are extremely reluctant to confer privacy protection on criminal activity, even where there is only suspicion, and will favour police interference where clear procedural steps have been taken to recognise privacy interests. (Interestingly, in the Demos 2015 report which updates their legal analysis of SOCMINT[135], reasonable expectations of privacy play a much more prominent role than in the 2013 version.)

Below we will argue that support for a reasonable expectation of privacy in SOCMINT has been further increased by (i) the arrival of social media as a mass phenomenon , especially among the young, and (ii) the growth and ease of search and data mining of unstructured free text data online.

7. Moving expectations of privacy in public into the social media and onl;ine search era

7.1 Reconsidering expectations of SOCMINT as "public" by virtue of site terms and conditions

boyd, in her seminal book on young people online, asks challengingly, is everything *done* in public in the new digital world "public"?[136] The 2013 Demos report on legal protections for SOCMINT and its update in 2015[137], in the main takes this stance as read. These reports accept explicitly that a person may have reasonable expectations of privacy in SOCMINT data if (i) he/she thinks it is private, or (ii) if society accepts as objectively reasonable that it is private (with reference to changed expectations after Snowden). Yet an expectation of privacy in the first case is still taken as signaled only where a person has made an "*explicit effort or decision.. to ensure that third parties cannot access this information*"(pp70-71). Examples given are closing accounts to Friends or password protecting them; or placing robot.txt restrictions on search bots.[138]. In relation to (ii), the report asserts, as in 2013, that material is still assumed

[132] Ibid, paras 56. 62.
[133] Ibid, para 41.
[134] Ibid, paras 79-80.
[135] See n 45 supra.
[136] boyd, supra n 96 at 203.
[137] See n 45 supra.
[138] Significantly, these ideas are drawn from a US commentators work : Susan Brenner; see Demos 2015 report at 70.

to be "*reasonably considered [by society] to be open*" in a variety of circumstances, but most notably where a user "*understand[s] this content is likely to be shared and used*" *and* where there is no indication the "*terms of agreement establish that content uploaded is private*" ; *and/or* content is made available to third parties via an API (p 73).

We argue that this is not a valid way to delimit a user's reasonable expectations of privacy in the mass social media age. Even in the simplest case, where the data subject themselves has knowingly disclosed substantive, personal data about themselves, it is simply not credible to argue that accepting standard terms and conditions on a social media site, typically buried under an obscurely placed hyperlink, negates any kind of expectation of privacy. First, as Edwards has noted elsewhere[139] , such "consent" to fixed unilaterally imposed terms and conditions is neither free, specific nor informed (as DP law requires[140]) and has become effectively illusory. Privacy policies have become too bloated for any normal person to read, as well as being largely written in incomprehensible legal language.[141] At the same time, privacy controls on sites have also become labyrinthine, almost invariably are set to defaults which favour disclosure and are changed unilaterally from time to time invariably to disclose more types of data to more classes of audience. [142] Finally there is no "marketplace of choices" where a user could shop to protect their privacy more adequately; virtually all social media sites with any mass audience apply standard term non-negotiable contracts allowing monetization and sharing of user content. While Facebook have made efforts, after severe criticism, to improve the presentation of privacy controls – eg by allowing users to see what their site looks like to a non-Friend third party – these criticisms are still profound.

In particular, even if these criticisms are rejected, while contract law may be taken formalistically to bind the user to terms they do not understand or read *in their relationship with the network*, it is hard to see how such terms could operate to remove their expectations of privacy *as against third parties* such as the police.

7.2 *Reconsidering expectations of privacy for young people on social media sites*

Expectations of privacy are particularly problematic in relation to young adults and minors. boyd , in a decade long research programme, has evolved a convincing theory that teenagers need to socialise and communicate to develop an identity, but that historic venues for this – shopping malls, each other's houses , etc. – are failing them because of the trend towards modern parenting involving intense surveillance and control. As a result teens socialise instead online, but often in public without Friends lock protections – why? boyd argues that teens want to socialise with their peers, including unknown peers, but do not imagine the other audiences – parents and teachers, let alone the police – who are also invisibly able to watch. Teens therefore choose to control the dissemination of their personal data, *not* by using privacy controls

[139] Edwards supra n 132.
[140] See art 2, DPD.
[141] N Bilton, 'Price of Facebook Privacy? Start Clicking'r*New York Times*, (New York, 12 May 2010) at http://www.nytimes.com/2010/05/13/technology/personaltech/13basics.html .
[142] K Opsahl, 'Facebook's Eroding Privacy Policy: A Timeline'a*Electronic Frontier Foundation*, (San Francisco, 28 April 2010) at https://www.eff.org/deeplinks/2010/04/facebook-timeline .

which are arduous to master and frequently change to defeat their efforts anyway ("public by default, private by effort")[143] but by using certain private or shared cryptic language or "social steganography".[144] boyd summarises that:

> "*the dynamics of mediated social situations – including invisible audiences, collapsed contexts and persistent content – further complicate things, making it incredibly difficult for teens to imagine the boundaries of these mediated social situations*" [145]

Yet the results of these practices can be disastrous for teens and young adults. Many social media sites by default retain communications as a persistent and searchable archive. When police get involved, posts or comments may be wrenched from their context and fed into data-mined profiles, leading to mangled meaning and harmful and unexpected consequences. For example, a post on a games or hacker site which refers to "rape" may mean something quite different from the usual sexual offence use of the word[146]. Nissenbaum has argued powerfully that privacy expectations are contextual: we willingly disclose information in certain contexts (eg to our doctor or our friend) which we would not if others were listening (advertisers, parents, the police).[147] On social media sites, such contextual confusions might have lastingly serious effects for users and their police records and profiles.

7.3 Reconsidering expectations of privacy in unstructured vs structured data

The third crucial point that needs reconsidered in the SOCMINT age is the distinction between structured and unstructured data. We noted above the influence of the dictum from *Rotaru* , that mere sporadic monitoring of a data subject may not infringe art 8 rights of privacy, but systematic collection, use and storage of data about him quite possibly will. This notion was reinforced in *Segerstedt-Wiberg v Sweden,*[148] a case involving files kept by the Swedish security police on the applicant drawn from public sources such as newspapers and open meetings, where the ECtHR held:

> "*the information about the applicants that was stored on the Security Police register and was released to them clearly constituted data pertaining to their "private life". Indeed, this embraces even those parts of the information that were public, since the information had been systematically collected and stored in files held by the authorities*.[149]"

Historically, then, what has separated police-state-like ubiquitous surveillance from legitimate police observation has been the compiling and keeping of systematic dossiers. In *Wood* above, the judge referred to the fear of Stasi-style surveillance and this

---

[143] boyd supra n 96 at 61-65.
[144] A Marwick "The Public Domain: Social Surveillance in Everyday Life" (2012) 9 Surveillance and Society 378.
[145] boyd supra n 96 at 61.
[146] C Miller, "This is your Brain Online: How Twitter Changed The Word "Rape" *Politics.co.uk* , 20 January 2014 at http://www.politics.co.uk/comment-analysis/2014/01/20/this-is-your-brain-online-how-twitter-changed-the-word-rape.
[147] H Nissenbaum 'Privacy as Contextual Integrity" (2004) 79 *Washington Law Review* 119-158.
[148](2007) 44 EHRR 2.
[149] (2007) 44 EHRR 2 Para 71. Note that the ECtHR also found that in some cases covert surveillance of political dissidence might also infer a breach of art 10 (freedom of expression).

threat is clearly lurking behind the debate in *Rotaru*, *Segerstedt-Wilburg* and other cases. Drawing on *Segerstedt-Wilburg* , it seems to be argued that publicly available SOCMINT, even if it concerns "private life", still can be observed and collected so long as it is not turned into a detailed dossier on a particular data subject.

In computer science language, as opposed to Strasbourg jurisprudence, what is being posited here is a fundamental difference in terms of privacy-invasive potential between structured and unstructured data. Structured data can be queried and data mined; once, unstructured data could not. But with the arrival of the Internet, automated processing and search algorithms , this distinction has now collapsed. With Google (or similar engine), anyone can construct a dossier instantly of a named person from unstructured materials created across decades. The police can do more still, using analytical profiling tools dedicated to creating predictive models of what people might do or say. There is no need for a Stasi to painstakingly assemble dossiers on everyone in their country in case at some point that person becomes of interest. Now search and data mining tools make the whole of the unrestricted Internet a dossier waiting to happen, at minimum effort and cost, and in very little time.

The European Court of Justice has recently eloquently recognised the impact of automated search on privacy rights in *Google Spain* (the "right to be forgotten" case). They observed:

> "It must be pointed out at the outset that… processing of personal data.. carried out by the operator of a search engine is liable to affect significantly the fundamental rights to privacy and to the protection of personal data when the search by means of that engine is carried out on the basis of an individual's name, since that processing enables any internet user to obtain through the list of results *a structured overview of the information relating to that individual that can be found on the internet* — information which potentially concerns a vast number of aspects of his private life and which, without the search engine, could not have been interconnected or could have been only with great difficulty — and thereby to establish a more or less detailed profile of him. Furthermore, the effect of the interference with those rights of the data subject is heightened on account of the important role played by the internet and search engines in modern society, which render the information contained in such a list of results ubiquitous"[150].

From this perspective, SOCMINT can be seen as the ultimate loophole to avoid the ECtHR constraints on systematic collection of dossiers as engaging art 8.

9. Conclusions

[150] *Google Spain v Costeja Gonzales*, ECJ, Case C‑131/12, para 80. The ECJ may not in general have jurisdiction over nation state police surveillance as a matter currently excluded from (most of) EC data protection law. However their views on nature of modern online surveillance are certainly relevant. See also *Schrems v DPC of Ireland* CJEU Judgment in case C-362/14 para 94. While the *Schrems* case is not strictly relevant to discussion of open SOCMINT , since it concerns interception of "closed" content by extra EU authorities, para 94 is interesting contextual evidence of the unhappiness of the CJEU with general state access to unstructured content in electronic communications.

In January 2015, in the wake of the Charlie Hebdo attacks in Paris, David Cameron argued, in a widely reported speech, that in the interests of the "war against terror", effective encryption of public communications should be banned:

> "*In extremis, it has been possible to read someone's letter, to listen to someone's call, to listen in on mobile communications.. The question remains: are we going to allow a means of communications where it simply is not possible to do that? My answer to that question is: no, we must not.*[151]"

In such a climate – which has since only got worse not better, given apparently insuperable political problems in the Arab world and elsewhere – it is a hard time for academics to argue for more privacy protection against police surveillance, not less, and make any inroads. This has been the story in the UK to date of the persistent criticisms of RIPA, of the failed Communications Data Bill and of the new IP Bill rising from its ashes. It is doubly hard to so argue when we are talking about open posts on social media which many people intuitively think of as without legal significance and which the police assert make their investigations immeasurably easier[152]. Yet this may turn out to be as important a debate as the one still raging over interception of closed communications by security agencies.

Our lives are written out in full and glorious (if sometimes misleading) technicolour detail in those millions of tweets, Facebook posts, and YouTube videoes we post every day. To say that we implicitly give up all expectations of privacy when we join a platform used by millions because of terms and conditions we have not read, did not understand and could not alter seems surreal. To say that only those aware enough of modern methods of data mining and intelligence-lead policing, and techno-literate enough to know how to protect themselves, should be entitled to retain a shred of private life while remaining in public discourse, seems a statement of despair. To say that the young cannot find and create their own identity online while young without storing up problems and police surveillance later is culpable.

Taking all the points above together, we ask: are we writing a blank cheque for mass, automated, Panoptic surveillance via social media by regarding the collection of open SOCMINT as untroublesome to privacy rights? In the 1950s, the Stasi used a staff of c 100,000 secret policemen plus the help of around 500,000 informants to help surveille a population of 16-19 million people in the GDR – a ratio of c. 35 observed per each Stasi agent.[153] Nowadays, the 17-strong staff[154] at the Metropolitan Police Open Source Intelligence Unit can conceivably surveille with greater ease the 8 million strong population of Greater London – a ratio of 470,588 observed to 1 policeman.

---

[151] See "How has David Cameron caused a storm over encryption?", Guardian, 15 January 2015 at http://www.theguardian.com/technology/2015/jan/15/david-cameron-encryption-anti-terror-laws . Rules about encryption , somewhat watered down after public furore, have duely appeared in the new IP Bill.

[152] Of a sample of US police surveyed in February 2014, 73% believed social media helped solve crimes faster, See *Social Media use in Law Enforcement*, LexisNexis, November 2014 at http://www.lexisnexis.com/risk/downloads/whitepaper/2014-social-media-use-in-law-enforcement.pdf.

[153] Figures from Rauhofer supra n 128.

[154] From Wright supra n 71.

What realistic hopes are there for a shift in legal thinking that would take account of the threat to privacy inherent in largescale acquisition of open social media content? We are not talking here of a ban on acquisition of open SOCMINT; merely for a recognition that this is a category of data whose collection must be properly and appropriately authorized under investigatory powers law. There is no sign of this way of thinking in the new IP Bill. It is possible however that the courts, probably not of the UK but more likely of Strasbourg or Brussels[155], may lead the way in imposing privacy protection on domestic law. The ECJ already seem to have taken it upon themselves in the post-Snowden era to show their teeth and create a stronger sense of privacy and the rule of law in the open public spaces of the Internet. We have already mentioned the political significance of the *Google Spain* case, *Schrems* and *Digital Rights Ireland* cases in this space.

Fundamentally, police surveillance of personal data on social media is part of a wider debate about who has the right to appropriate and profit from the digital footprints we now leave in the public corridors of the Internet, and what legal and ethical safeguards should protect us, the users, given the now near impossibility of abstaining from the virtual world. Whether we are talking about our Facebook posts and Likes being harvested to build advertising profiles, our tweets being combed by the well-meaning to determine if we are suicidal[156], our mobile calls and photos in the Cloud being read by the NSA to find out if we are part of a terrorist cell, or our social media posts being mined by local police to see if we are involved in anything from littering to jihadist propaganda, in the end the question is the same: do we have any rights of control over our personal data in public?

Furthermore the debate is moving from being merely about collection of our data as we communicate "online" to our whole lives in non-digital contexts – ie, in the real world. Increasingly, our movements are tracked by sensors in the smart cards we use to access transport and make payments, by our footfall on smart roads and by Internet-connected cars or "autonomous vehicle" we use to get from A to B. Our bodies themselves may be tracked eg if we wear fitness monitors or have implanted medical devices such as heart pacemakers which communicate data even after surgery. This data collected by the "Internet of Things" may be collected in private, typically private environments – eg our energy consumption in our home equipped with smart meters – but will also just as likely be collected in public as a kind of "data discharge" in smart cities, transport systems and shopping malls. Will this data be seen as naturally "public" and therefore lacking reasonable expectations of privacy, as has been the case with open SOCMINT? This kind of future , while beyond the scope of this already long paper[157], is one of the reasons why the authors believe this to be a crucial topic. If we have no privacy in data disclosed in public in future, we may have no privacy at all.

---

[155] Assuming the UK remains in both the EU and the ECHR of course.

[156] See the recent Samaritans radar affair, Kirsty Marrins "Samaritans Radar: 'Charity deserves round of applause for putting mission front and centre' *The Guardian* (London 3 November 2014) <http://www.theguardian.com/voluntary-sector-network/2014/nov/03/samaritans-radar-twitter-mission-charity>. See also Jon Baines 'ICO: Samaritans Radar Failed to Comply with Data Protection Act' *Information Rights and Wrongs* (25 April 2015) at http://informationrightsandwrongs.com/2015/04/25/ico-samaritans-radar-failed-to-comply-with-data-protection-act/ .

[157] See further L Edwards "Privacy, security and data protection in smart cities: a critical EU law perspective", forthcoming 2016 European Journal of Data Protection.