NEW COLLEGE
OXFORD

Reading the Investigators Their Rights: A review of literature on the General Data Protection Regulation and open-source intelligence gathering and analysis
Author: Anjuli Shere
Source: *The New Collection*, Vol. 14 (Aug., 2020), pp. 3-21
Published by: The Middle Common Room (MCR) of New College, Oxford
Stable URL: http://mcr.new.ox.ac.uk/journal/Contents2020.php

 ISSN 1757-2541.

Reading the Investigators their Rights: A review of literature on the General Data Protection Regulation and open-source intelligence gathering and analysis

*Anjuli Shere**
Centre for Doctoral Training in Cyber Security

Abstract

Open-source intelligence gathering and analysis (OSINT) techniques are no longer predominantly the remit of private investigators and journalists. An estimated 80-90% of data analysed by intelligence agencies is also now derived from publicly available material. Additionally, the massive expansion of the internet and, in particular, social media platforms, have made OSINT increasingly accessible to civilians who simply want to trawl the Web for information on a specific individual, organisation or product. In May 2018, the European Union's General Data Protection Regulation (GDPR) was implemented in the UK through the new Data Protection Act, intended to secure personal data against unjustified collection, storage and exploitation. This document presents a preliminary literature review of work related to the GDPR and OSINT, which was collated as the basis for an as-yet-unpublished study evaluating the effects of the GDPR on OSINT capabilities in the UK. The literature reviewed is separated into the following six sections: 'What is OSINT?', 'What are the risks and benefits of OSINT?', 'What is the rationale for data protection legislation?', 'What are the current legislative frameworks in the UK and Europe?', 'What is the potential impact of the GDPR on OSINT?', and 'Have the views of civilian and commercial stakeholders been sought and why is this important?'. As OSINT tools and techniques are accessible to anyone, they have the unique capacity for being used to hold power to account. It is therefore important that new data protection legislation does not impede civilian OSINT capabilities.

Keywords: GDPR (General Data Protection Regulation)● Data Protection Act● Open-Source Investigation● data protection● literature review

* anjuli.shere@new.ox.ac.uk.

INTRODUCTION

The ability of modern technology to swiftly, surreptitiously and easily gather vast amounts of data has been accompanied by a myriad of data breaches, many of which have been high profile and hugely disruptive on an individual, organisational and global scale. Notable examples of breaches include: the 2012–2014 illicit access to the US Office of Personnel Management's databases, which is suspected to have exposed the background checks of more than 22 million current and former federal employees to Chinese agents [1]; and the 2013 Yahoo breach that revealed to attackers the personal information of 3 billion users [2].

To minimise the risk of such data breaches continuing to occur, emphasis has been placed on the creation of data protection legislation, intended to preserve information security and therefore defend individual privacy in an age of increasingly intrusive and pervasive technologies. It is against this backdrop that the European Union's (EU) General Data Protection Regulation (GDPR) was first proposed in 2012 [3]. The United Kingdom's (UK) 2018 Data Protection Act (DPA) virtually mirrors the GDPR and ensures legislative consistency with the rest of the EU, even in the event of Brexit [4].

In 1988, Roger Clarke coined the term "dataveillance", meaning "the systematic use of personal data systems in the investigation or monitoring of the actions or communications of one or more persons" [5]. Clarke argued that mass dataveillance was conducted by both public and private organisations and "involves a generalized suspicion that some (as yet unidentified) members of the group may be of interest" [5]. One "facilitative mechanism" by which "multifactor file analysis" (profiling of personal data) is conducted is "matching", the "expropriation and merger of data held in separate data systems" [5], such as between governmental and corporate organisations.

The GDPR was largely marketed as intending to combat dataveillance of individuals and in particular the matching that is prevalent as a result of popular social media platforms and other companies collecting user data [6]. However, in their article for the *European Data Protection Law Review*, Giurgiu and Larsen argued that "the ambition of the Regulation was to create data protection rules fit for the Digital Single Market that would alleviate problems companies face when doing business in the Union" [7]. This indicates that the GDPR focuses on privileging private corporations when standardising laws across the EU, despite the fact that companies are often the perpetrators of dataveillance and data hoarding that impinge upon individual privacy. Additionally, Giurgiu and Larsen noted that the implementation of the GDPR enhanced the role of data protection authorities as domestic and international "guardians of the respect of EU data protection law" [7]. These data pro-

tection authorities, such as the UK's Information Commissioner's Office (ICO), are ostensibly independent bodies, but may be funded or otherwise reliant on governmental support. For example, although the ICO reports to the UK Parliament, it is sponsored by the governmental Department for Digital, Culture, Media and Sport. It is therefore possible to argue that the GDPR's key beneficiaries include these already-powerful institutions: multinational corporations and often state-sponsored bodies, such as data protection authorities. That companies and such "arm's length bodies" (as the ICO is referred to in its DCMS Management Agreement [8]) gain more power or assistance through the GDPR suggests a stronger pro-company and pro-gov slant to the law than media coverage of the GDPR's motivations would infer. This was supported by Austria's comments as the only EU member state that voted against the GDPR replacing the 1995 EU Data Protection Directive. Austria highlighted that the GDPR's "legitimate interests" justification differs from the 1995 Directive by allowing data controllers to act on the claim that their interests in a given scenario supersede the confidentiality of data subjects. As a supporting reason for their rejection of the GDPR, the country argued that "the mere presence of legitimate interests of the data controller – without a requirement to weigh those interests against the data subject's confidentiality interests – cannot justify data processing" [9]. This assumed increase in protection of individuals' data and yet the inclusion of loopholes that maintain the power held by state-affiliated and corporate institutions reflects an inconsistency of prioritisation that Wachter contends is rife throughout the GDPR and, while it would benefit OSINT analysis by these entities, would likely hinder OSINT investigations that challenge their authority [10].

Ironically, efforts to secure one public good may result in infringement of another [11]; for example, the use of automated live facial recognition cameras to prevent crime in UK city centres [12]. Similarly, following the implementation of the GDPR, there has been a notable imposition of professional constraints on investigators who work with predominantly open-source intelligence (OSINT), i.e. those who analyse publicly available data. In 2013, the then Information Commissioner, Christopher Graham, spoke against the Leveson Report's recommendations for reforming the previous UK Data Protection Act in a way that would allow the Information Commissioner's Office to regulate press activities, thereby having a "chilling effect" on potential investigations [13].

While the majority of the GDPR has the potential to affect OSINT capabilities through the restriction of access to and availability of online personal data, there are two chapters that are directly relevant to OSINT investigators and may have the longer-term effect of curtailing their ability to operate as normal. Chapter 2 (Art. 5-–11) addresses accountability through the documentation of compliance with six fundamental principles that ensure that data is only gathered for explicit and legit-

imate purposes and processed lawfully, fairly, transparently, is accurate, current and held securely at all times. Chapter 5 (Art. 44—50) also enshrines legislatively that the collected data is not stored for longer than is necessary, to ensure that data processing is legal, consensual and carried out for the public good [14]. Chapter 5 is concerned with the international transfer of personal data and dictates that this information may only be shared to non-EU countries if those countries are certified as meeting GDPR standards for data protection [14]. As a large proportion of OSINT practitioners are domiciled in a different location to the subjects of their professional investigations, this may drastically restrict their ability to conduct OSINT on EU subjects. At present, many people are working remotely, but the scope of this study specifically covers the GDPR's effect on the ability to conduct UK-centric OSINT investigations.

However, due to the GDPR's potential to "incite a shift not only in the handling of personal data but in the attitude towards it as well" [15], various aspects of it have been examined at length by diverse parties to whose work and personal lives the GDPR is relevant. These aspects include: its inherent flexibility as a piece of legislation designed to be implemented across a continent [16]; the feasibility of its commitment to defending the "right to withdraw consent and the right to be forgotten" [17]; and the likelihood of its long-term success. This final feature concerns the increasing prevalence of household and wearable Internet of Things devices and their associated infringements of privacy [18]. This study aims to assess the current state of the literature related to the efficacy of legislation like the GDPR and the DPA, given the recent implementation of the GDPR and its potential impact on OSINT investigations in the UK.

METHODOLOGY

The core objective of this research is to survey the existing literature and understandings of the GDPR and OSINT investigations, including how and where they overlap. This literature review utilises diverse sources of information on open-source materials, the challenges and benefits of open-source intelligence gathering (OSINT), and the justifications for data protection legislation such as the GDPR and other European legislative frameworks. This was intended to set a foundational understanding of the existing publicly available research on OSINT and the potential implications of historically relevant regulatory measures.

The literature review and aspects of the discussion of this study and its results followed the PRISMA guidelines for systematic review, drawing from a variety of sources, including ICO reports, news articles and academic papers to identify

as many perspectives on the impact of regulation on OSINT as possible. The use of the PRISMA method allowed research using the most relevant and current literature on the GDPR, OSINT and the surveillance state. Initial literature searches used terms such as "open-source", "open-source investigations", "OSINT", "GDPR", "General Data Protection Regulation", "surveillance state", "open-source intelligence gathering and analysis" and "data protection legislation". Additional keywords, such as "dataveillance" [19], that emerged as significant during the information-gathering process were noted and used for further research across academic literature from the fields of Computer Science, Cyber Security, Political Science and Law. In total, over 50 publications were found and used to build a comprehensive profile of the topic. The questions used as sub-headings throughout the rest of this paper to organise the literature and findings of this study were derived from salient thematic groupings found during the study's literature search.

LITERATURE REVIEW

What is Open-Source Intelligence Gathering and Analysis (OSINT)?

One way in which the state and corporate interests may be held to account is using OSINT to confirm or discredit claims made by powerful actors. OSINT traditionally involved the comprehensive examination of publicly available material such as conventional mass media (e.g. television, radio, domestic and foreign press), specialised journals, conference proceedings, think tank studies, academic and governmental databases and photos, usually found in physical archives [20]. Intelligence gathering may take any number of forms, the majority of which involve collection from or processing by information systems. Intelligence gathering methods can be largely grouped into the following disciplines, which utilise a variety of sources, including the cultivation of insiders (part of HUMINT), interception of signals (SIGINT), and interrogation of geospatial images (GEOINT). The vast and growing expanse of the Internet "as an extension of the human mind" has allowed OSINT to develop into a practice that combines elements from all the aforementioned investigative specialties, thereby transcending the control of a small number of local parties or even linguistic limitations [21]. Not only is individuals' personal data now abundant on social media sites, but it is also possible to link multiple aspects of a person's work and private life solely by searching for their names. Additionally, commercial organisations expanding their role into domains that were formerly reserved for governments has meant that data that would previously have been classified is increasingly shared

publicly, for example geospatial information such as maps and other live and historic imagery products.

What are the risks and benefits of OSINT?

As there is an ongoing shift of data into the public arena, OSINT skills are useful in turning this data into actionable intelligence. However, the danger of doing so is a loss of privacy that may not be reversible. The evolution of regulations such as the GDPR signifies the widespread prioritisation of privacy, rather than facilitation of civilian investigative capabilities. The EU recognises the value of OSINT tools for law enforcement and, from 2010 to 2013, the European Commission even sponsored the 'Versatile InfoRmation Toolkit for end-Users oriented Open Sources explOitation' (VIRTUOSO) project which built "a technical framework for the integration of tools for collection, processing, analysis and communication of open source information" [22]. Fears about VIRTUOSO's compatibility with the EU's core norms and values were driven partly by concern about the privacy infringements inherent in OSINT. Therefore, part of the project's development process was an evaluation of its likely ethical and legal ramifications. The conclusion was that it is the responsibility of the developers of OSINT platforms and networks to include functionalities that maximise end-users' ability to adapt their settings to match their desired level of privacy [23]. This was intended to provide guidance for including "privacy by design" in this OSINT toolkit. Further, in 2014 military and security experts Hribar *et al.* wrote that "most of the data processed by intelligence services is based on public/open sources (speculatively, from 80 to 90%)" [24]. Both UK military and policing organisations are open about their extensive use of OSINT. Studies have been conducted into how these cases overlap in practice, concluding that the main differences relate to: policy regarding chain of custody of data collected; the extent to which third-party suppliers are contracted; willingness to utilise data from the dark web; and counterintelligence approaches [20].

Prior to the GDPR OSINT practices were accessible to almost anyone, but the strict prosecution and much heftier fines associated with GDPR noncompliance mean that smaller and less well-resourced operations are likely to struggle financially if they do not adapt to new data protection regimes. If small private OSINT investigations are lacking funds, they are also less likely to afford legal advice and support that would enable them to exploit legislative loopholes, such as effectively claiming "legitimate interests" to justify data processing to the ICO or other relevant data protection authorities. In comparison, state-affiliated, larger, international agencies, and large private corporations like social media giants that utilise OSINT are much more likely to have an advantage in terms of human, financial, legal and technological re-

sources, meaning that their OSINT investigations may not be hindered to the same extent as smaller OSINT operations. This would affect all investigators who operate without the special capabilities of state sponsorship or mogul protection, including journalists. Since a major normative problem with OSINT data mining and mapping is the inclusion of information that may not have legally been leaked into the public domain, such as insider testimonies, OSINT analysts could be prosecuted for their work. There is typically a strong governmental backlash against any journalists who would use the information, possibly accompanied by abuses of power to aid cover-ups. A prominent example of this out of fear of the consequences of intelligence leaks is President Nixon's pre-Watergate hiring of "The Plumbers", a group of "ex-CIA fixers" whose role was to find and silence the employee who was leaking information to the press [25].

A large part of OSINT is Social Media Intelligence (SOCMINT): the collection of publicly available information shared on social networking sites [26]. Although much of this data was covered under the definition of personal data given in the 1995 European Data Protection Directive, the GDPR's updated and expanded definition includes personal data that may be gleaned from profiles on modern social media sites, such as IP addresses, mobile device identifiers, geolocation and information relating to one's identity (e.g. social, cultural, physical, psychological, genetic, mental, or economic elements of an individual's life) [14]. These features were previously ungoverned but are now covered as part of the GDPR and DPA legal frameworks for regulation. At its most basic, OSINT may be seen as a "form of directed surveillance" of individuals through sites like Twitter, Facebook and Instagram, which effectively act as journals for individuals to fill with intimate details of their lives for public viewing [20]. In 2008, a UK Home Affairs Select Committee report entitled 'A Surveillance Society?' specified that surveillance "encompasses not only the use of monitoring and recording technology but also the creation and use of databases of personal information and the record of our communications in the digital age" [27]. The report stated that these databases are ubiquitous and no longer predominantly created by countries, but rather by decentralised private actors who categorise and use the information collected as a "surveillant assemblage" that automatically makes all aspects of daily life publicly visible, creating an online persona or "data double" [28]. These online spaces are structured to enable and encourage the sharing of information that can then be traded for ephemeral rewards, such as Twitter 'likes'. Indeed, Lyon framed modern dataveillance as a marriage of "computer technologies on the one hand, and consumerism on the other", arguing that "user-generated content" is the lifeblood of these practices [29]. This presents a counterpoint to Bentham's "panopticon", "the idea of a new principle of construction applicable to any sort of establishment, in which persons of any description are to be kept under inspection"

[30], where the perception of constant surveillance was intended as a punitive measure. Similarly, in the Foucauldian model of "Panopticism", which links the original conception of an enclosed surveillance system to capitalist motivations [31], subjects are not willingly complicit as they are in modern practice when divulging personal information on social media sites like Instagram.

Describing the surveillance state as a "laboratory of power" [31], Foucault's inference was later echoed by Clarke who argued that the threat of surveillance is a form of social control, moulding subjects' behaviour to make it more desirable to dominant powers. In the contemporary international system, these powers are no longer limited to states but also include commercial interests. These stakeholders are inextricably financially and socially linked, such that even speculation about regulation of aspects of industry has an impact on the career prospects of political figures globally. It is widely known that states are not solely responsible for contemporary data collection or the processing of this data into workable intelligence. However, they are still often seen as playing "a vital role in coordinating and consuming surveillance practices" of privately acquired information, and "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" [26].

In contrast, civilian OSINT has been used by individuals and public bodies to verifiably hold to account the institutions that are responsible for the construction and maintenance of a surveillance state-industrial complex by using some of the techniques and materials that have historically been used to uphold the surveillance state. Prominent examples of this are independent investigations into states' nuclear disarmament and non-proliferation commitment [32]. OSINT is also valuable for counteracting some of Clarke's stated dangers of dataveillance, both to the individual and to society, such as by challenging "inequitable application of the law" [5].

What is the rationale for data protection legislation?

Former US National Security Agency (NSA) contractor Edward Snowden's revelations in 2013 that large swathes of the world's population are under continual dataveillance by their own domestic law enforcement and intelligence agencies drastically heightened public awareness of the threats to their privacy and identity that resulted from inappropriate collection and processing of personal data [26]. One of Austria's core objections to the GDPR was that it gives "equal value" to the interests of both data controllers and data subjects, thereby justifying data processing without prioritising data subjects' confidentiality [9]. This is in contrast to its predecessor, the 1995 Data Protection Directive, Article 12 of which dictates subjects' right to

promptly and comprehensively access or block their data even during and after processing [33]. The GDPR's chapter 2 does stress consent and compliance regarding data collection, storage and processing [14], and it also affects OSINT in a way that the previous Directive could not because online resources were not so expansive in 1995. However, Wachter noted that, once processing is justified, the GDPR does not create much protection for processed data [10]. This therefore means that, since the GDPR's implementation, the increasingly prevalent use of OSINT to compile profiles on individual subjects remains largely unencumbered by law if the perpetrators are able to cite "legitimate interests". Previously, legislation such as Article 8 of the European Convention on Human Rights [34], exempted publicly available data from ordinary consent requirements, so that the onus was on data subjects to avoid sharing personal information on public platforms.

Despite the GDPR's emphasis on data subjects' consent, it would be simplistic to argue that an individual's creation, curation and maintenance of an online presence are entirely consensual decisions. Multiple aspects of contemporary social activities have been effectively interwoven with features offered by Internet platforms, such that attempting full integration into society requires agreeing to the 'Terms and Conditions' of a social networking site. Examples of this include 'closed' Facebook groups that one can only enter if their account is linked to their academic email address, or else they risk losing access to event invitations and support. In this way, meaningful participation in social life is restricted to those who are willing to experience "the daily renewal of a twenty-first century Faustian compact" [35].

There are many components to this compulsion to be dependent on online fora, not least the lack of correspondingly relevant and efficient alternatives to these platforms. Similarly, "digital connection is now a means to others' commercial ends" [35], with free and yet lucrative search engines like Google indexing and selling data to companies whose feedback subtly modifies user behaviour. Although other search engines that do not track user input exist, such as DuckDuckGo, they are inconvenient as they do not filter responses by popularity and so can cause long time delays for searches that would be swift and simple elsewhere. This effectively comprises a form of peer and institutional pressure that relies on the complexity of the 'Terms and Conditions' and the social benefits of inclusion in this online realm to make "the realities of being tracked, parsed, mined, and modified" seem vague and appealing [35]. Data protection legislation is therefore viewed as a way of mitigating the many possibilities and implications of commercial exploitation of individuals' loss of privacy and manipulation of online identity. The latter involves the generation of behavioural surplus, i.e. "personal data collected for the primary purpose of predicting and changing individual behaviours, rather than for the primary purpose of improving a service to individual users" [36].

The companies that have profited the most from these actions have disproportionately high market influence and so norms are not considered strong enough deterrents. A 2019 Australian Privacy Foundation report submitted to the Australian Competition and Consumer Commission on the dangers of surveillance capitalism argued that "the history of Google and Facebook shows that any voluntary measures will be evaded and defeated, and that the only realistic approach when dealing with these companies is legal compulsion coupled with penalties" [36]. As such, the concerns that lead to the drafting of new data protection legislation around the world affect individuals. They are also increasingly governmental, driven by states' worry that not meeting rising data protection standards will hinder international intelligence sharing and trade.

What are the current legislative frameworks in the UK and Europe relevant to OSINT?

For two decades, the UK's data protection principles were derived from its 1998 Data Protection Act, which had been intended to implement an EU Directive on data processing [37]. However, in 2018 a new Act of the same name came into effect, as a result of the UK government's desire "to support businesses in their use of data, and give consumers the confidence that their data is protected and those who misuse it will be held to account," as well as "prepar[ing] Britain for Brexit" [38], following the enactment of the EU's GDPR.

The prevalence of dataveillance and its resultant deluge of personal data catalysed the creation and implementation of the GDPR as "a modernised consumer-focused toolkit for privacy protection across Europe" [39]. It posits individual privacy infringement as a societal issue to be curtailed by curbing intelligence gathering, storage and processing capabilities. The regime applies to all organisations that conduct these activities regarding the personal data of EU residents. To establish effective regulation and data protection, the GDPR is split into eleven chapters detailing definitions, principles of accountability and specifying remedies and penalties in the event of transgression.

Due to the UK government's decision to invoke Article 50 of the Treaty on the EU and thereby withdraw the UK from the Union, provisions have been made to ensure that data protection legislation in the UK is equivalent to the regime newly in effect across the continent. The UK's determination to adhere to international standards is driven by recognition that countries observing the GDPR are bound to only share data with sources that have similarly stringent regulations. As such, the DPA is "the UK's third generation of data protection law", combining "the applied GDPR" with additional amendments that are specific to the UK [40]. It is therefore applicable to all organisations that function in the UK and process UK citizens' data. In addi-

tion to elucidating the "general functions" and "international role" of the Information Commissioner [41], the DPA includes the following changes: Adjustments to immigration data processing requirements to ensure that they match GDPR standards but also reflect the UK national context; a section that condenses EU Data Protection Directive 2016/680 relating to law enforcement agencies processing data in a "competent" and ethical fashion so that, post-Brexit, law enforcement and criminal justice agencies can collaborate by sharing data; and a requirement similar to Council of Europe Data Protection Convention 108, ensuring the compliance of British intelligence services with international data protection regulations.

Also relevant to this discussion is "Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA", also known as the Police Data Protection Directive (PDPD) [42]. It was part of the consolidation of EU data protection rules, alongside the creation of the GDPR, to better reflect the realities of law enforcement investigations into digital crimes, particularly regarding law enforcement agencies' use of big data analytics [43]. According to the PDPD, "competent authorities may include not only public authorities such as the judicial authorities, the police or other law-enforcement authorities but also any other body or entity entrusted by Member State law to exercise public authority and public powers for the purposes of this Directive" [44]. Today, the GDPR is usually applicable instead of the PDPD, except in cases where data processing "must be carried out by a competent authority […] for the purposes of the prevention, investigation, detection and prosecution of criminal offences or the execution of criminal penalties, including the safeguarding against and the prevention of threats to public security" [43].

Much of the PDPD echoes elements of the GDPR, with PDPD chapters 2, 3 and 5 diverging. Particularly notable is that the PDPD specifically endows these competent authorities with the ability to collect and process personal data so far as they are "not excessive in relation to the purposes for which they are processed" [44], which deviates from the GDPR's mandate that personal data be "limited to what is necessary" [14]. Additionally, Sajfert and Quintel explicitly acknowledged that the PDPD theoretically grants supervisory bodies lots of power, but that it "does not oblige the Member States to vest their national supervisory authorities with any particular corrective powers in respect of police and criminal justice authorities", thereby potentially allowing law enforcement agencies to subvert the PDPD without effective penalty [43]. The stated purposes of the PDPD are only related to law enforcement and national security, not corporate or other interests, and so data subjects' rights (e.g.

to directly access or erase the information held on them) are not enshrined in the PDPD as they are in the GDPR. Sajfert and Quintel point out that Article 11 of the PDPD prohibits automated processing that adversely and/or significantly affects data subjects, but that this does not rule out profiling by non-automated means [43]. This effectively renders non-automated OSINT investigations, including profiling, by law enforcement agencies allowable. Sajfert and Quintel also note that both the GDPR and PDPD focus on individual data subjects' rights, with the latter legislation leaving a "loophole" that diminishes data protections for groups of individuals, except for the PDPD's prohibition of processing of purely sensitive data, which is absent from the GDPR [43].

What is the potential impact of the General Data Protection Regulation (GDPR) on OSINT?

Due to the novelty of the GDPR and its current status as the most intricate and substantial data protection legislation in the world, there is little literature from which to determine the regulation's impact on OSINT capabilities. Although there are extensive publications discussing the utility of the GDPR for safeguarding and preventing the unethical use of data, there are no studies assessing whether civilian and commercial OSINT capabilities are impeded by increased regulation or if innovative techniques are already being utilised to ensure the continued availability of big data.

Rather than centring data protection conversations on the conventional Western idea of privacy as "the concept around which resistance to intensive or extreme surveillance may be mobilized", Lyon claimed that "many of today's surveillance issues are better thought of in social justice terms" [45]. In this vein, although OSINT is part of a data processing patchwork of capabilities that take advantage of the online information sharing culture, it is uniquely able to use this infrastructure against the powerful institutions that instigated such a culture. Therefore, as OSINT is newly regulated by the GDPR and DPA, it may be decreasing in efficacy and as such diminishing its potential for positive societal use. This literature review inspired me to conduct a short-term study of the effects of the GDPR and DPA on OSINT investigations in the UK, which is yet to be published.

Have the views of civilian and commercial stakeholders been sought and why is this important?

Prior to the GDPR, data protection mechanisms were only enshrined on a state-level and so were fragmented and applied inconsistently. The EU's acceptance of the GDPR signals the advent of data protection legislation not only affecting individuals and companies, but also applying to instruments of the state. In effect, this distinc-

tion means that the primary objective of the GDPR's creation was to protect the data itself, rather than to protect control over data, thereby enhancing the privacy of the individual against governments and corporations. However, there exists a tension in that the GDPR emphasises the role of national data protection authorities, independent public bodies that are nonetheless supported by a governmental department of the EU member state in which they exist [46]. As previously mentioned, the integral role of state-sponsored data protection authorities in both disseminating information about and enforcement of the GDPR arguably contributes to public perceptions of the GDPR and the DPA as tools used to legitimise traditional notions of state sovereignty and enforce the surveillance state. Considering that media coverage of the GDPR has typically emphasised its motivation as being to protect the data of those less powerful than state entities (i.e. individuals), this tension could undermine its stated purpose. Rather than OSINT remaining accessible to almost everyone, the discipline could begin to emulate the current balance of power in other spheres, in favour of government actors rather than non-state OSINT investigators.

All data yielded is used to conduct pattern of life analysis and make predictions about individuals and their choices that can be engineered to support existing power structures. This paradigm broadly acts "to identify individuals who may be worth subjecting to personal surveillance, and to constrain the group's behaviour *[sic]*" [5]. The latter purpose is significant in that it only works if subjects are acutely aware of and uncomfortable with the dataveillance. It is possible that the GDPR increased public awareness of the prevalence of mass dataveillance thereby catalysing a popular shift in attitude towards more privacy-enhancing behaviour. This would make data collection more difficult but could also signal a decline in behaviours that are considered socially inappropriate, such as "oversharing" of personal information online, thereby potentially supporting a function of the surveillance state.

CONCLUSION

This literature review is in no way a comprehensive summary of the work available on data protection regulation or OSINT tools and techniques, but rather is intended to provide a foundation for research into legislative and policy decisions that have implications for stifling civilian investigative capabilities. At present, there appears to be no published work examining the effects of the GDPR on UK-based investigations, either public or private. Following this literature review, I conducted a short-term (as yet unpublished) study into this topic and surmised that this is an area that would benefit from further research. It could lead to the creation of well-evidenced

guidelines for decision-makers that would protect both the privacy of individuals' personal data and the ability of members of the public to research and analyse the actions of governments and corporations, to maintain the balance of power and the level of accountability necessary for democracy.

ACKNOWLEDGEMENTS

With thanks to Prof. Ian Loader for his supervision of this study and to Miranda R. Melcher for her support and editing.

Bibliography

[1] E. E. Cummings, *Hearing on 'OPM: Data Breach': Opening Statement*. Washington DC, 2015.

[2] R. McMillan and R. Knutson, 'Yahoo Triples Estimate of Breached Accounts to 3 Billion', *Wall Street Journal*, Oct. 04, 2017.

[3] European Commission, 'Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation)'. European Commission, Jan. 25, 2012, Accessed: May 28, 2019. [Online]. Available: https://web.archive.org/web/20121203024154/http://ec.europa.eu/justice/data-protection/document/review2012/com_2012_11_en.pdf.

[4] M. Burgess, 'What is GDPR? The summary guide to GDPR compliance in the UK', *Wired UK*, Jan. 21, 2019.

[5] R. Clarke, 'Information technology and dataveillance', *Commun. ACM*, vol. 31, no. 5, pp. 498–512, May 1988, doi: 10.1145/42411.42413.

[6] R. Brandom, 'GDPR launches today. Here's what you need to know', *The Verge*, Mar. 28, 2018. https://www.theverge.com/2018/3/28/17172548/gdpr-compliance-requirements-privacy-notice (accessed Apr. 21,2020).

[7] A. Giurgiu and T. A. Larsen, 'Roles and Powers of National Data Protection Authorities', *Eur. Data Prot. Law Rev.*, vol. 2, no. 3, pp. 342–352, 2016, doi: 10.21552/EDPL/2016/3/9.

[8] Information Commissioner's Office and Department for Digital, Culture, Media Sport, 'Information Commissioner's Office: Management Agreement 2018-2021', Information Commissioner's Office, London, UK, 2018. Accessed: Apr. 21, 2020. [Online]. Available: https://ico.org.uk/media/about-the-ico/documents/2259800/management-agreement-2018-2021.pdf.

[9] A. Tanca, *Written Procedure for the adoption of the Council's position at first reading and the statement of the Council's reasons on the Draft Regulation of the European Parliament and of*

the Council on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). Brussels, Belgium: Council of the European Union, 2016, p. 6.

[10] S. Wachter, 'Data protection in the age of big data', *Nat. Electron.*, vol. 2, no. 1, pp. 6–7, Jan. 2019, doi: 10.1038/s41928-018-0193-y.

[11] A. Roosendaal and J. Wright, 'Will data protection eventually kill privacy?', in *TILTing perspectives 2017: regulating a connected world*, Tilburg University, the Netherlands, May 2017, p. 12.

[12] S. Morris, 'Facial recognition tech prevents crime, police tell UK privacy case', *The Guardian*, May 22, 2019.

[13] S. Swinford, 'Leveson could have "chilling effect" on journalism, Information Commissioner warns', *The Telegraph*, Jan. 07, 2013.

[14] European Parliament and Council of the European Union, 'General Data Protection Regulation (GDPR) – Official Legal Text', *General Data Protection Regulation (GDPR)*, May 25, 2018. https://gdpr-info.eu/ (accessed Mar. 03, 2020).

[15] S. Sirur, J. R. C. Nurse, and H. Webb, 'Are We There Yet?: Understanding the Challenges Faced in Complying with the General Data Protection Regulation (GDPR)', in *Proceedings of the 2nd International Workshop on Multimedia Privacy and Security - MPS '18*, Toronto, Canada, 2018, pp. 88–95, doi: 10.1145/3267357.3267368.

[16] G. Buttarelli, 'The EU GDPR as a clarion call for a new global digital gold standard', *Int. Data Priv. Law*, vol. 6, no. 2, pp. 77–78, May 2016, doi: 10.1093/idpl/ipw006.

[17] E. Politou, E. Alepis, and C. Patsakis, 'Forgetting personal data and revoking consent under the GDPR: Challenges and proposed solutions', *J. Cybersecurity*, vol. 4, no. 1, Jan. 2018, doi: 10.1093/cybsec/tyy001.

[18] J. Lindqvist, 'New challenges to personal data processing agreements: is the GDPR fit to deal with contract, accountability and liability in a world of the Internet of Things?', *Int. J. Law Inf. Technol.*, vol. 26, no. 1, pp. 45–63, Mar. 2018, doi: 10.1093/ijlit/eax024.

[19] J. Cobbe, 'Big Data, Surveillance, and the Digital Citizen', *SSRN Electron. J.*, 2018, doi: 10.2139/ssrn.3234984.

[20] D. Wells and H. Gibson, 'OSINT from a UK perspective: Considerations from the law enforcement and military domains', p. 32, 2017.

[21] M. Glassman and M. J. Kang, 'Intelligence in the internet age: The emergence and evolution of Open Source Intelligence (OSINT)', *Comput. Hum. Behav.*, vol. 28, no. 2, pp. 673–682, Mar. 2012, doi: 10.1016/j.chb.2011.11.014.

[22] COMMISSARIAT A L ENERGIE ATOMIQUE ET AUX ENERGIES ALTERNATIVES, 'Final Report Summary - VIRTUOSO (Versatile InfoRmation Toolkit for end-Users oriented Open Sources explOitation)', European Commission, May 2015. Accessed: May 28, 2019. [Online]. Available: https://cordis.europa.eu/project/rcn/94446/reporting/en.

[23] C. Cuijpers, 'Legal aspects of open source intelligence – Results of the VIRTUOSO project', *Comput. Law Secur. Rev.*, vol. 29, no. 6, pp. 642–653, Dec. 2013, doi: 10.1016/j.clsr.2013.09.002.

[24] G. Hribar, I. Podbregar, and T. Ivanuša, 'OSINT: A "Grey Zone"?', *Int. J. Intell. CounterIntelligence*, vol. 27, no. 3, pp. 529–549, Sep. 2014, doi: 10.1080/08850607.2014.900295.

[25] A. S. Hulnick, 'The Downside of Open Source Intelligence', *Int. J. Intell. CounterIntelligence*, vol. 15, no. 4, pp. 565–579, Nov. 2002, doi: 10.1080/08850600290101767.

[26] L. Edwards and L. Urquhart, 'Privacy in public spaces: what expectations of privacy do we have in social media intelligence?', *Int. J. Law Inf. Technol.*, vol. 24, no. 3, pp. 279–310, Sep. 2016, doi: 10.1093/ijlit/eaw007.

[27] House of Commons: Home Affairs Committee, 'HC 58-I: A Surveillance Society?', Fifth Report of Session 2007–08, Jun. 2008. Accessed: May 28, 2019. [Online]. Available: https://publications.parliament.uk/pa/cm200708/cmselect/cmhaff/58/58i.pdf.

[28] K. D. Haggerty, Richard V. Ericson, 'The surveillant assemblage', *Br. J. Sociol.*, vol. 51, no. 4, pp. 605–622, Dec. 2000, doi: 10.1080/00071310020015280.

[29] D. Lyon, *The Culture of Surveillance: Watching as a Way of Life*, 1st ed. Cambridge, UK: Polity Press, 2018.

[30] J. Bentham, *The Works of Jeremy Bentham*, vol. 4, vol. 4. RareBooksClub.com, 2012.

[31] M. Foucault, *DISCIPLINE AND PUNISH: The Birth of the Prison*. UK: Penguin Books Ltd, 1991.

[32] Ridgeway Information, 'Ridgeway Information: Open Source Intelligence', 2019. https://www.ridgeway-information.com/open-source-intelligence (accessed Apr. 21, 2020).

[33] The European Parliament and The Council of the European Union, *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*, vol. OJ L. 1995.

[34] European Court of Human Rights and Council of Europe, 'European Convention on Human Rights', Nov. 1998.

[35] S. Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. London: Profile Books Ltd., 2019.

[36] G. Greenleaf, A. Johnston, B. Arnold, D. Lindsay, and R. Clarke, 'Digital platforms: The need to restrict surveillance capitalism', Australian Privacy Foundation submission to the ACCC, Feb. 2019.

[37] Parliament of the United Kingdom, *Data Protection Act 1998: Chapter 29*. London: The Stationery Office, 1998.

[38] Department for Digital, Culture, Media Sport, 'Press release: Government to strengthen UK data protection law', *GOV.UK*, Aug. 07, 2017. https://www.gov.uk/government/news/government-to-strengthen-uk-data-protection-law (accessed May 28, 2019).

[39] S. Davies, 'The Data Protection Regulation: A Triumph of Pragmatism over Principle?', *Eur. Data Prot. Law Rev.*, vol. 2, no. 3, pp. 290–296, 2016, doi: 10.21552/EDPL/2016/3/5.

[40] Information Commissioner's Office, 'Data Protection Act 2018: For Organisations', May 13, 2019. https://ico.org.uk/for-organisations/data-protection-act-2018/ (accessed May 28, 2019).

[41] Parliament of the United Kingdom, *Data Protection Act 2018: Chapter 12*. London: The Stationery Office, 2018.

[42] M. Hildebrandt, *Law for Computer Scientists and Other Folk*, 1st ed. New York, NY, USA: Oxford University Press, 2020.

[43] Protection Directive (EU) 2016/680 for Police and Criminal Justice Authorities', *Edw. Elgar Publ.*, p. 22, Dec. 2017.

[44] The European Parliament and The Council of the European Union, *Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA*, vol. OJ L. 2016.

[45] D. Lyon, *Surveillance Studies: An Overview*. Cambridge, UK; Malden, MA: Polity Press, 2007.

[46] European Commission, 'What are Data Protection Authorities (DPAs)?', *European Commission - European Commission*, 2018. https://ec.europa.eu/info/law/law-topic/data-protection/reform/what-are-data-protection-authorities-dpas_en (accessed May 28, 2019).