scripted

Volume 17, Issue 2, August 2020

Processing Data to Protect Data: Resolving the Breach Detection Paradox

*Andrew Cormack**

DOI: 10.2966/scrip.170220.197

Abstract

Most privacy laws contain two obligations: that processing of personal data must be minimised, and that security breaches must be detected and mitigated as quickly as possible. These two requirements appear to conflict, since detecting breaches requires additional processing of logfiles and other personal data to determine what went wrong. Fortunately Europe's General Data Protection Regulation (GDPR) – considered the strictest such law – recognises this paradox and suggests how both requirements can be satisfied. This paper assesses security breach detection in the light of the principles of purpose limitation and necessity, finding that properly-conducted breach detection should satisfy both principles. Indeed the same safeguards that are required by data protection law are essential in practice for breach detection to achieve its purpose. The increasing use of automated breach detection is then examined, finding opportunities to further strengthen these safeguards as well as those that might be required by the GDPR provisions on profiling and automated decision-making. Finally we consider how processing for breach detection relates to the context of providing and using on-line services concluding that, far from being paradoxical, it should be expected and welcomed by regulators and all those whose data may be stored in networked computers.

Keywords
Data protection; breach detection; incident response

* Chief Regulatory Adviser, Jisc, Didcot, UK, andrew.cormack@jisc.ac.uk

1 Introduction: the need for breach detection

One of the core principles of data protection – whether expressed in the European General Data Protection Regulation (GDPR),[1] the Council of Europe Convention 108,[2] or the FTC Fair Information Practice Principles[3] - is that the processing of personal data should be minimised. However all of those documents also demand that personal data must be protected, including more or less explicit requirements to be able to detect, investigate and mitigate the impact of security breaches. The benefits of a data controller reducing harm in its own processing will quickly be lost if a malicious intruder can gain undetected access to the data and cause havoc.

Even for information on paper, detection and investigation requires the collection and retention of additional personal data, such as records of who was authorised to access files and who (including unauthorised persons) actually did. For information stored in digital form on networked computers, the corresponding records of accesses and attempted accesses may involve very large collections of data. Without such logs it will be much harder to detect breaches and impossible to analyse and contain their impact. This creates a paradox: that protecting personal data against security breaches requires data controllers to collect and process more, not less, personal data. This paper demonstrates not only that these breach detection activities can be done fully in accordance with

[1] Regulation (EU) 2016/679 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 and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (hereinafter 'GDPR').

[2] Council of Europe, Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data, Strasbourg, 28 Jan 1981.

[3] Federal Trade Commission, Fair Information Practice Principles (25 June 2007), available at https://web.archive.org/web/20090331134113/http://www.ftc.gov/reports/privacy3/fairinfo.shtm (accessed 19 August 2019).

the strict requirements of the GDPR but that they should be seen as both necessary and reassuring by data subjects, data controllers, and regulators. Organisations that do not process data to detect and mitigate breaches should be a much greater concern than those that do.

The growing importance of protecting digital information and the systems that contain it was stressed at the 2017 launch of the EU Cybersecurity Act:

> With recent ransomware attacks, a dramatic rise in cyber-criminal activity, the increasing use of cyber tools by state actors to meet their geopolitical goals and the diversification of cybersecurity incidents, the EU needs to build a stronger resilience to cyber-attacks.[4]

Resilience has two main components: reducing the number of attacks that succeed (prevention) and reducing the impact of those that do (detection and recovery): thus Recital 25 of the Act seeks to help Member States and Union institutions "to prevent, detect and respond to cyber threats and incidents".[5] Likewise, while Article 32 of the General Data Protection Regulation (GDPR) requires that anyone processing personal data must take "appropriate technical and organisational measures" to prevent security breaches, the parallel requirement in Article 33 to notify breaches recognises that prevention alone is not enough. Recital 85 is explicit that to avoid "physical, material or non-material damage" to individuals, organisations must also be able to respond to breaches

[4] European Commission, "State of the Union 2017 – Cybersecurity: Commission scales up EU's response to cyber-attacks" (Brussels, 19 September 2017), available at http://europa.eu/rapid/press-release_IP-17-3193_en.htm (accessed 19 August 2019).

[5] Regulation (EU) 2019/881 of the European Parliament and of the Council of 17 April 2019 on ENISA (the European Agency for Cybersecurity) and on information and communications technology cybersecurity certification and repealing Regulation (EU) No 526/2013 (Cybersecurity Act), Recital 25.

"in an appropriate and timely manner" when they occur.[6] The Article 29 Working Party's guidance on Breach Notification, endorsed by the European Data Protection Board,[7] confirms that "the ability to detect, address, and report a breach in a timely manner" is an "essential element" of the Article 32 duty.[8] This dual requirement is now a common pattern in European legislation: sectors where breaches may cause disruption to society, rather than directly affecting personal data, are also required to have detection and response measures alongside their preventive ones, for example by Chapters IV and V of the Network and Information Security Directive covering energy, transport, banking, financial markets, health, water and digital infrastructures;[9] Article 19 of the eIDAS Regulation[10] covering electronic identification and trust services; and Article 4(3) of the amended ePrivacy Directive[11] covering electronic communications. Detection and response are as important as prevention.

The Article 29 Working Party outlines what is involved in detecting and analysing breaches: "For example, for finding some irregularities in data processing the controller or processor may use certain technical measures such as data flow and log analysers, from which [it] is possible to define events and

[6] GDPR, *supra* n. 1, Recital 85.

[7] European Data Protection Board, "Personal Data Breach Notifications" (25 May 2018), available at https://edpb.europa.eu/node/67 (accessed 19 August 2019).

[8] Article 29 Working Party, "Guidelines on Personal data breach notification under Regulation 2016/679" (18/EN WP250rev.01) (hereinafter "Breach Notification"), p. 13.

[9] Directive (EU) 2016/1148 of the European Parliament and of the Council of 6 July 2016 concerning measures for a high common level of security of network and information systems across the Union (NIS Directive).

[10] Regulation (EU) 910/2014 of the European Parliament and of the Council of 23 July 2014 on electronic identification and trust services for electronic transactions in the internal market and repealing Directive 1999/93/EC (EIDAS Regulation).

[11] Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications), as amended by Article 2(4)(c) of Directive 2009/136/EC of the European Parliament and of the Council of 25 November 2009 (ePrivacy Directive).

alerts by correlating any log data".[12] The mention of irregularities and correlations indicates a need to consider both historical and contextual information: breaches will often be detected as a divergence from normal behaviour or as a group of events happening around the same time. Such activities therefore involve additional processing beyond that required to service individual transactions: for example in *Breyer v Germany* the European Court of justice recognised that "aiming to ensure the general operability of those [web] services" might require retaining and using logs after the completion of the transactions to which they referred.[13] More specifically, the European Network and Information Security Agency's (ENISA) 2011 report identified the "must-have tools" for detection of network security breaches as "firewalls, antivirus (alerts), IDS/IPS and NetFlow";[14] for analysing security breaches, the Forum of Incident Response and Security Teams (FIRST) consider that relevant sources may include "Netflow data, Router logs, Proxy server logs, Web application logs, Mail server logs, DHCP server logs, Authentication server logs, Referring databases, Security equipment, such as firewall or intrusion detection logs".[15]

Both the example of website logs in *Breyer* and the longer lists from ENISA and FIRST indicate that the information needed to detect and investigate breaches is likely to be already held – if only briefly – by the organisation that operates the online service. To send and receive packets, a networked computer must process the Internet Protocol (IP) header data from which netflow, router,

[12] Article 29 Working Party, "Breach Notification", *supra* n. 8, p. 13.

[13] *Patrick Breyer v Bundesrepublik Deutschland*, Case C-582/14 [2016] ECLI:EU:C:2016:779 (hereinafter *Breyer*), para. 64.

[14] ENISA, "Proactive detection of network security incidents" (2012), available at https://www.enisa.europa.eu/publications/proactive-detection-report/ (accessed 19 August 2019), p. 105.

[15] Forum of Incident Response and Security Teams, "Establishing a CSIRT" (version 1.2, November 2017), available at https://www.first.org/resources/guides/Establishing-CSIRT-v1.2.pdf (accessed 19 August 2019), p. 28.

firewall and Intrusion Detection/Protection System logs are derived; to deliver a web or email service, it must process the application headers that are recorded in logfiles; to safely connect local user devices it must provide DHCP, anti-virus and proxy services; to provide authentication it must maintain user accounts. A key feature of these information sources – in legal terms – was noted as long ago as 2003: that, to be useful, they must contain Internet Protocol (IP) addresses and timestamps.[16] Under European law, at least, this means they are likely to constitute personal data, so breach detection and analysis – involving recording, retaining and processing these data sources – will itself be subject to the GDPR.[17] Fortunately both legislation and case law are aware of this: Recital 49 of the Regulation recognises a need for "processing of personal data to the extent strictly necessary and proportionate for the purposes of ensuring network and information security"; *Breyer* recognised that retaining and processing logs to detect and investigate attacks might be lawful.[18]

Thus a resolution of the breach detection paradox should be possible: where additional personal data processing is necessary to protect personal data, the law both requires and permits this.

Breach detection involves processing, for a second purpose, of personal data that the service or network operator must already process for the primary purpose of providing their service. Under the GDPR, the first two principles to consider in such situations are the relationship between the two purposes,

[16] Moira West-Brown et al., "Handbook for Computer Security Incident Response Teams" (Software Engineering Institute, April 2003), available at https://resources.sei.cmu.edu/library/asset-view.cfm?assetid=6305 (accessed 19 August 2019), p. 84.

[17] Article 29 Working Party, "Opinion 4/2007 on the concept of personal data" 01248/07/EN WP 136 (hereinafter 'personal data'), pp.15-16.

[18] *Breyer*, *supra* n. 13, para 64.

"purpose limitation",[19] and the necessity of the additional processing for the second purpose, "data minimisation".[20] The results of that inquiry will then guide compliance with the remaining principles. This paper therefore examines breach detection from the perspectives of purpose and necessity. Since automation is increasingly needed to handle the growing volume of data relevant to breach detection, we then investigate how this may affect the purpose and necessity analysis, and what additional requirements may result from the GDPR's specific provisions on profiling and automated decision making. Finally we conclude that breach detection is not only compatible with the GDPR, but should be welcomed and expected by regulators, operators and data subjects as a key part of the provision of any internet-connected system or service.

2 Purpose

The Article 29 Working Party considers that

> Specification of purpose is an essential first step in applying data protection laws and designing data protection safeguards for any processing operation. ... The principle of purpose limitation is designed to establish the boundaries within which personal data collected for a given purpose may be processed and may be put to further use [21]

This should, for example "prevent the use of individuals' personal data in a way (or for further purposes) that they might find unexpected, inappropriate or otherwise objectionable".[22]

[19] GDPR, *supra* n. 1, Article 5(1)(b).

[20] GDPR, *supra* n. 1, Article 5(1)(c).

[21] Article 29 Working Party, "Opinion 03/2013 on purpose limitation" 00569/13/EN WP 203 (hereinafter "Purpose Limitation"), p.4.

[22] *Ibid*., p. 11.

For breach detection the purpose is both clear and set out in law. GDPR Recital 49 concerns:

> ensuring network and information security, i.e. the ability of a network or an information system to resist, at a given level of confidence, accidental events or unlawful or malicious actions that compromise the availability, authenticity, integrity and confidentiality of stored or transmitted personal data, and the security of the related services offered by, or accessible via, those networks and systems[23]

Most of the data processed by breach detection systems will therefore serve two purposes: providing a networked service, and keeping that service secure. Both purposes are known, specified and legitimate (according to Recital 49) at the time when data are collected.

The Working Party recognises that "[p]ersonal data can be collected for more than one purpose. In some cases, these purposes, while distinct, are nevertheless related to some degree. In other cases the purposes may be unrelated".[24] The two cases require different safeguards to protect individuals' interests: the following analysis suggests that properly-conducted breach detection should have no difficulty in satisfying the requirements of both.

2.1 Breach Detection as a Compatible Purpose

The first option, covered by GDPR Article 6(4), is that a group of purposes may be "compatible". The Working Party explain that this requires an assessment of "the relationship between the purposes...; the context … and reasonable expectations of data subjects…; the nature of the personal data and the impact of

[23] GDPR, *supra* n. 1, Recital 49.

[24] Article 29 Working Party, "Purpose Limitation", *supra* n. 21, p. 16.

the further processing…; the safeguards adopted by the controller".[25]

The close relationship between operating a service and securing it has been increasingly recognised by legislation, case law, and regulators' guidance. Both GDPR Recital 49[26] and *Breyer*[27] link breach detection and response to the provision of networked services; the Working Party's Guidelines on Breach Notification encourage all data controllers and processors to "put in place processes to be able to detect and promptly contain a breach".[28] These bases in law[29] together with widespread reporting of the harm caused by on-line security incidents and regulators' criticisms,[30] should mean data subjects very "reasonably expect" that those providing services will also do what is necessary to secure them and the data they contain.[31]

Concerning nature and impact, the kinds of data used for breach detection will normally be the same as those involved in providing the service. The Working Party note that additional processing with a negative or uncertain impact is unlikely to be compatible:[32] the purpose of breach detection in fact demands that the impact on users be positive.

Finally, security teams involved in breach detection have at least as strong an interest as their users in applying organisational and procedural safeguards to their information and processing secure.[33] Logfiles and information derived

[25] *Ibid.*, p. 3.

[26] GDPR, *supra* n. 1, Recital 49.

[27] *Breyer*, *supra* n. 13, para. 64.

[28] Article 29 Working Party, "Breach Notification", *supra* n. 8, p. 6.

[29] Article 29 Working Party, "Purpose Limitation", *supra* n. 21, p. 25.

[30] For example BBC, "British Airways faces record £183m fine for data breach" (8 July 2019), available at https://www.bbc.co.uk/news/business-48905907 (accessed 19 August 2019); BBC "UK watchdog plans to fine Marriott £99m" (9 July 2019), available at https://www.bbc.co.uk/news/technology-48928163 (accessed 19 August 2019).

[31] Article 29 Working Party, "Purpose Limitation", *supra* n. 21, p. 13.

[32] *Ibid.*, p. 26.

[33] Andrew Cormack, "Incident Response: Protecting Individual Rights Under the General Data Protection Regulation", (2016) 13(3) *SCRIPTed* 258-282, p. 276.

from them are likely to contain information that would help an attacker find weaknesses in a system;[34] they can also reveal to an attacker whether or not their activities have been detected. Both undermine the defenders' purpose. These files and processing will therefore normally be kept separate from the operation of the service and subject to additional technical and organisational controls. For example security data and systems will normally implement strong access controls and those with access to them will be under contractual obligations of confidentiality. The technical safeguards that can and should be used during breach detection and investigation are described in section 4.

Six years ago – before Recital 49 and *Breyer* had explicitly recognised the link between service provision and service security – the Working Party nonetheless cited as compatible purposes "preventing fraud and abuse of the financial system"[35] and a smart grid operator that "wishes to implement an intelligent system, including an analytics tool, to detect anomalies in usage patterns, which may give reasonable suspicion of fraudulent use". In particular the latter

> stems from, and is in furtherance of, the initial purposes of providing energy to the customers and charging them for the energy they use. Customers could reasonably expect that their provider will take reasonable and proportionate measures to prevent fraudulent use of the energy, in the interest not only of the energy company, but also those customers that are paying their bills correctly[36]

Provided appropriate safeguards are applied, such processing is considered

[34] Bernie Lantz, Rob Hall, and Jason Couraud, "Locking Down Log Files: Enhancing Network Security by Protecting Log Files", (2006) VII(2) *Issues in Information Systems* 43-47, p. 44.

[35] Article 29 Working Party, "Purpose Limitation", *supra* n. 21, p. 53.

[36] *Ibid*., pp 69-70.

compatible. The same should apply to the processes organisations use to detect misuse of online systems and data: this, too, is in the interests of both organisations and their customers.

Note that this would not extend to other uses of the data generated by use of computers and networks – for example to enforce policy or investigate crime, including attempts to identify attackers. These would constitute additional purposes, requiring their own assessment and safeguards. Where organisations use the same data and systems for multiple purposes, they must ensure these are kept distinct by appropriate organisational and technical safeguards, appropriate to each purpose and the risks it involves.

2.2 Breach Detection as a Separate Purpose

Treating breach detection as a "compatible purpose" to the operation of an on-line service means both activities have the same legal basis (probably "necessary for contract" under Art.6(1)(b)) and the same obligations apply to both sets of processing. This may be helpful when existing service data are later discovered to have value for breach detection;[37] however users may gain additional protections if breach detection is treated as a separate purpose, necessary for a legitimate interest of the service operator, as suggested by both Recital 49 and the *Breyer* judgment.

Under this approach, the separate purpose must be "specified, explicit [and] legitimate"[38] and the processing must fully satisfy the requirements of the appropriate legal basis. In addition to the common requirement (under both GDPR Articles 6(1)(b) and 6(1)(f)) that it must be necessary, legitimate interest processing must satisfy the balancing test that the interest is not overridden by

[37] *Ibid.*, p. 21.

[38] *Ibid.*, p. 12.

the data subject's fundamental rights and freedoms. Individuals also have a right to seek a review of that balancing test against their own particular circumstances, under the Article 21 right to object.

None of these requirements should cause significant difficulties for the operator of an online service who wishes to use service data to detect breaches. As discussed above, the purpose is specified when data are first collected; both *Breyer* and Recital 49 indicate that it is legitimate. It can therefore be made explicit to users. The practical issue of how to inform users about processing of data that is observed, rather than provided directly by the user, is common to both the "compatible purpose" and "separate purpose" approaches. Regulators' practice on their own websites[39] indicates that including breach detection and response in a privacy notice is an appropriate mechanism. Cormack explains how the balancing test will generally be satisfied by the existing practice of incident response and security teams.[40] In particular, unlike examples of secondary processing considered by Balboni et al, detecting and remedying breaches is not an action whose benefit to the controller "is considered to prevail over the protection of personal data",[41] but a shared interest that enhances that protection, actively supporting users' rights and freedoms.

Security teams can therefore provide additional reassurance to their users, beyond what the law requires, by meeting the requirements of both approaches to purpose. Respecting purpose compatibility ensures that security activities are closely related to the operation of the service, stay within the expectations of

[39] Information Commissioner's Office, "Visitors to Our Website" (ICO, undated), available at https://ico.org.uk/global/privacy-notice/visitors-to-our-website/#sec (accessed 23 August 2019).

[40] Cormack, *supra* n. 33, p. 274.

[41] Paolo Balboni et al., "Legitimate Interest of the Data Controller. New Data Protection Paradigm: Legitimacy Grounded on Appropriate Protection" (2013) 3(4) *International Data Privacy Law* 244-261, p. 247.

users and are subject to appropriate safeguards. Treating security, in addition, as a separate purpose further ensures that it is always explicitly declared to service users and that their rights and freedoms, not just the security of the individual service, are taken into account.

Finally, the Article 29 Working Party mentions "surprise" as an indicator of non-compatible processing.[42,43] Given the stress in legislation, case law and guidance on the importance of protecting personal data it seems likely that the Working Party – and service users – would actually be more surprised by a service provider that does not process data to protect its systems than by one that does.

3 Necessity

Having concluded that breach detection satisfies the purpose requirements of the GDPR, the next question is what processing is necessary to achieve it.

3.1 When is processing "necessary"?

Recital 39 of the GDPR states that "personal data should be adequate, relevant and limited to what is necessary for the purposes for which they are processed".[44] The Article 29 Working Party has explained this use of "necessary", in both the Regulation and its preceding Directive, as meaning "any processing of personal data involved is the minimum amount required to fulfil its aim", [45] noting also that "if other effective and less invasive means to achieve the same goal exist,

[42] Article 29 Working Party, "Purpose Limitation", *supra* n. 21, p. 24.

[43] Article 29 Working Party, "Guidelines on transparency under Regulation 2016/679" 17/EN WP260 rev.01 (hereinafter "Transparency"), p. 24.

[44] GDPR, *supra* n. 1, Recital 39.

[45] Article 29 Working Party, "Opinion 01/2014 on the application of necessity and proportionality concepts and data protection within the law enforcement sector" 536/14/EN WP 211 (hereinafter "Necessity and Proportionality"), p. 18.

then it would not be 'necessary'".[46] This definition seems to exclude any possibility of further qualifying the word "necessary", since any processing less than the minimum required cannot, by definition, fulfil the aim, so fails the requirement that processing be "adequate".

It is therefore puzzling to find both Recital 49 of the GDPR and Article 5(3) of the *ePrivacy Directive* requiring that processing – for network and information security, and of cookies, respectively – must be "***strictly*** necessary" (emphasis added).[47,48] The explanation appears to be that these phrases derive from a different source: the requirement in Articles 7 and 8 of the Charter of Fundamental Rights that any interference with rights must be "necessary in a democratic society".[49] In this context, "necessary" has been ruled by the European Court of Justice to be "not synonymous with indispensable" and "[n]or should it be interpreted too literally, as this would set too high a bar and make it unduly difficult for otherwise legitimate activities which may justifiably interfere with fundamental rights to take place".[50] Thus the *e-Privacy Directive*'s requirement for cookies to be "strictly necessary" – which the Working Party interpret as "if cookies are disabled, the functionality will not be available"[51] – narrows the wider Charter sense of "necessary" down to that contained in data protection law.

This section will therefore interpret any qualified use of "necessary" as deriving from the Charter sense and follow the Working Party's approach that,

[46] Article 29 Working Party, "Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679" 17/EN WP251rev.01 (hereinafter "Profiling Guidelines"), p. 17.

[47] GDPR, *supra* n. 1, Recital 49.

[48] ePrivacy Directive, *supra* n. 11, Article 5(3).

[49] Charter of Fundamental Rights of the European Union 2012/C 326/02, Articles 7 and 8.

[50] Article 29 Working Party, "Necessity and Proportionality", *supra* n. 45, p. 6.

[51] Article 29 Working Party, "Opinion 04/2012 on Cookie Consent Exemption" 00879/12/EN WP 194, p. 4.

in data protection law, "necessary" (whether qualified or not) involves a requirement that processing must be reduced to the minimum possible that will still achieve the objective. Furthermore, if two approaches are "equally effective", then the less intrusive should be adopted.[52]

3.2 What processing is necessary for breach detection?

Very rarely, a security breach may involve only a single event: far more often there will be several preparatory steps involved. The key to early detection is to spot these sequences of events, ideally before the critical point of infection or compromise. For example detecting an attacker scanning for vulnerabilities involves recognising the same test being run against several internal addresses; malware infections are often detected by linking a local machine's visit to an infected website with its subsequent "call-home" connection to a machine controlled by the attacker;[53] a phishing incident will often be revealed by the same account logging in from a rapid sequence of geographically implausible places. In each case linking the individual events into the sequence that reveals them to be abnormal and needing further investigation requires them to be stored in association with a relevant identifier, such as the IP address or account name. Processing those identifiers is thus "necessary", in the narrow data protection sense, as there is no less intrusive way to recognise the critical sequences of events.

[52] Article 29 Working Party, "Necessity and Proportionality", *supra* n. 45, p. 21.

[53] Guofei Gu et al., "BotHunter: Detecting Malware Infection Through IDS-Driven Dialog Correlation" (2007) *16th Usenix Security Symposium* 167-182.

Although random malicious traffic on the Internet is so prevalent[54] that every user is at risk of becoming a victim – hence likely to benefit directly from early detection and mitigation of breaches – in any given period some users will avoid this fate. This raises the question whether processing personal data of those fortunate individuals is also "necessary". Not recording data for a particular machine or account obviously means that individual user will not benefit from detection and response when the worst does happen. But, since attackers commonly use their initial success to attack others within the system or organisation, any gaps in recording will also put others at risk. Finally, detecting unusual traffic depends on comparison with a normal baseline: problems will often be detected when behaviour varies from that of uncompromised computers and accounts. Thus comprehensive logging and processing of data is necessary, in the narrow data protection sense, for breach detection, analysis and response.

The legal position of breach detection is therefore different, in context as well as scale, from government powers to retain data for law enforcement purposes that were analysed, and found not "strictly necessary", by the European Court of Justice in *Digital Rights Ireland*. In that case most of the retained data related to individuals who were not "even indirectly, in a situation which is liable to give rise to criminal prosecutions".[55] Woods notes the subsequent case of *Tele2/Watson* describing this retention as "indiscriminate" because "there is no link between the data retention and the threat posed by a

[54] SANS, "Survival Time" (Internet Storm Centre, undated), available at https://isc.sans.edu/survivaltime.html (accessed 19 August 2019) suggests every Internet-connected computer receives hostile traffic several times a day.

[55] *Digital Rights Ireland Ltd v Minister for Communications, Marine and Natural Resources, Minister for Justice, Equality and Law Reform, Commissioner of the Garda Síochána, Ireland, The Attorney General*, Case C-293/12 [2014] ECLI:EU:C:2014:238 (hereinafter *Digital Rights Ireland*), para. 58.

specific individual"; this "goes beyond what is 'strictly necessary'".[56] In particular, collecting data in case individuals commit criminal acts "transform[s] [them] into potential suspects".[57]

By contrast, when detecting security breaches, all those whose data are retained are likely to be victims; indeed, according to Eurobarometer, 42% of them already have been.[58] This activity is not "indiscriminate" and does not transform their status. Logging and processing information to detect breaches and provide help supports the rights and freedoms of individual users, not just an "objective of general interest" such as "the fight against serious crime".[59,60] Finally, *Tele2/Watson* saw government data retention as an exception to the privacy protections in the *ePrivacy Directive*,[61] whereas "ensuring network and information security" is explicitly recognised as contributing to those protections by both the amended *ePrivacy Directive* and the GDPR.[62]

According to Woods, even where data retention may be necessary, "stringent safeguards to prevent abuse would be of central importance in

[56] Lorna Woods, "Automated Number Plate Recognition: Data Retention and the Protection of Privacy in Public Places" (2017) 2(1) Journal of Information Rights, Policy and Practice 1-21, p. 18.

[57] *Ibid*., p. 12.

[58] Eurobarometer, "Special Report 464a: Europeans' Attitudes Toward Cyber Security" (European Commission, 2017), p. 66, available at http://ec.europa.eu/commfrontoffice/publicopinion/index.cfm/ResultDoc/download/DocumentKy/79734 (accessed 23 August 2019).

[59] *Digital Rights Ireland*, *supra* n. 55, para. 44.

[60] *Ibid*., para. 41.

[61] *Tele2 Sverige AB v Post-och telestyrelsen and Secretary of State for the Home Department v Tom Watson, Peter Brice, Geoffrey Lewis*, Cases C-203/15 and C-698/15 [2016] ECLI:EU:C:2016:970 (hereinafter Tele2), para. 88.

[62] Directive 2009/136/EC of the European Parliament and of the Council of 25 November 2009 amending Directive 2002/22/EC on universal service and users' rights relating to electronic communications networks and services, Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector and Regulation (EC) No 2006/2004 on cooperation between national authorities responsible for the enforcement of consumer protection laws, Recital 53.

determining whether such powers were proportionate".[63] Not only is breach detection compatible with such safeguards, omitting them is likely to make it significantly less effective.

4 Safeguards

Both necessity and purpose limitation principles therefore consider the safeguards that can be applied to the processing as a relevant factor. As noted in the earlier purpose limitation discussion, the organisational safeguards needed to ensure the effectiveness of breach detection and investigation are strongly aligned with those required to ensure privacy is protected. Here we consider the technical safeguards that can, and should, be used.

Nearly all the identifiers used for breach detection – including IP, MAC and email addresses – have the technical characteristics of pseudonyms, defined in the GDPR as:

> the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person.[64]

As noted in *Breyer*,[65] for website logs – and others that record the activity of external users – the "additional information" is not just held separately, but by an entirely different organisation: police powers are likely to be required to

[63] Woods, *supra* n. 56, p. 18.

[64] GDPR, *supra* n. 1, Article 4(5).

[65] *Breyer*, *supra* n. 13, paras 47-8.

obtain it. Even where logs relate to users within the organisation, the additional information is normally generated by separate systems – those concerned with authentication and address allocation – from the network flows and application logs that are the main resource for breach detection.

Furthermore most breach detection can be done without the attribution step. As described above, the first stage in detection is to link several events – each associated with a pseudonymous identifier such as an IP address – into a sequence that may indicate a security breach. Analysis to determine whether a breach is the most likely cause of such an alert can also normally done using just the pseudonymised data. Only when this investigation concludes that a breach probably has occurred is it necessary to identify the individuals involved: to contact them, confirm what has happened and provide assistance. Events that do not correlate into alerts, and alerts whose investigation reveals them to have an innocent explanation, can left as unattributed pseudonyms. Breach detection can therefore be done within a framework recognised by the GDPR both as a safeguard that "can reduce the risks to the data subjects concerned and help controllers and processors to meet their data-protection obligations",[66] and as an "appropriate technical measure" for implementing data protection principles including minimisation[67] and security.[68]

Bolognini and Bistolfi consider that in situations where the purpose may require identification of a subset of individuals, the GDPR's approach to pseudonyms in fact provides the best protection since – unlike anonymisation, which takes data outside the scope of data protection law – pseudonymisation provides both technical safeguards and continuing regulation: it "is able to

[66] GDPR, *supra* n. 1, Recital 28.

[67] GDPR, *supra* n. 1, Article 25(1).

[68] GDPR, *supra* n. 1, Article 32(1)(a).

mitigate the risks of a data subject's direct identification, guaranteeing that the data controller uses the data in compliance with norms governing data protection".[69] Treating security event and alert data as pseudonyms ensures that data protection law regulates both the data and processes used for breach detection and the data and processes for linking breaches to individual victims.

Where pseudonyms are used, GDPR Article 11 relaxes the normal rule that individuals must be informed in advance of processing, recognising that identifying individuals to inform them that their pseudonymised data are being processed would remove the benefit of the safeguard.[70] The purpose of breach detection and response thus encourages organisations to do at least as much as the legal requirement, by providing general information to all users of a system that data will be processed for breach detection and response, and informing specific users who do need to be identified, immediately after that linking takes place.

Again, there is a contrast with law enforcement data retention where, according to Spina, "the fact that data are retained and subsequently used without the subscriber or registered user being informed is likely to generate in the minds of the persons concerned the feeling that their private lives are the subject of constant surveillance".[71] Effective breach detection and response require the user to be informed, and action to be taken, as soon as possible after the event. Not informing a user when it appears likely that they are a victim of a security breach would defeat our purpose.

[69] Luca Bolognini and Camilla Bistolfi, "Pseudonymisation and Impacts of Big (personal/anonymous) Data Processing in the Transition from the Directive 95/46/EC to the new EU General Data Protection Regulation" (2017) 33(2) *Computer Law and Security Review* 171-181, p. 178.

[70] GDPR, *supra* n. 1, Article 11.

[71] Alessandro Spina, "Risk Regulation of Big Data: Has the Time Arrived for a Paradigm Shift in EU Data Protection Law?" (2014) 5(2) *European Journal of Risk Regulation* 248-252, p. 251.

Breach detection and response can, and should, therefore follow the "differentiated approach" recommended by Mantelero for data and processing minimisation.[72] Analysis to detect problems is done using pseudonyms, affected users are identified at the last stage of response when offering them assistance. Any lessons learned can be shared to help others using either anonymised data or pseudonyms (such as remote IP addresses) that are only meaningful to the recipient organisation.[73] This approach also contributes to the security and privacy of data, users and systems, since it minimises the risk of analysts inadvertently discovering or disclosing information that is not relevant to the investigation. If unusual activity on a network is analysed and found not to be malicious, the analyst can ignore it without ever knowing which individual users were involved.

Finally, Bolognini and Bistolfi note that using linked pseudonyms to identify and assist individual victims may well involve less, and more predictable, limitation of rights and freedoms than imposing preventive measures on a larger, anonymous, group.[74] The specific question of whether this may constitute "profiling", and how it can be done in accordance with the GDPR, will be considered after examining the general issues raised by the increasing use of automation in breach detection.

5 Automated Breach Detection

Techniques for breach detection have been developed continuously over more than twenty years.[75] Originally these involved manual inspection of logfiles and

[72] Alessandro Mantelero, "Data Protection, e-ticketing, and Intelligent Systems for Public Transport" (2015) 5(4) *International Data Privacy Law* 309-320, p. 312.

[73] Cormack, *supra* n. 33, p. 281.

[74] Bolognini and Bistolfi, *supra* n. 69, p. 180.

[75] West-Brown et al., *supra* n. 16.

network flows; visualisation and investigation tools were then developed to help analysts perform these procedures.[76] Over the same period, our use of networked computers has expanded massively in both scale and complexity, generating security data in much greater volumes than human analysts can handle. In 2014 a national research computing service generated less than 10 Gigabytes of logs a day:[77] in 2018 a single medium-sized university, over 200GB.[78]

Humans can no longer look at every event on a network or system: indeed looking at individual events is unlikely to be sufficient to reveal most security breaches. As the Article 29 Working Party notes, breaches generally appear as anomalies within the normal patterns of activity and detecting them requires correlating events occurring at different times, in different locations or, indeed, reported by entirely different systems.[79] For example Huang, Kalbarczyk and Nicol describe a hybrid breach detection system that combines information about network flows with logs from applications and content inspection systems.[80] Automation is therefore an essential first stage in most breach detection processes: typically software will be used to analyse the events recorded in flows and logfiles, to identify groups of events that may indicate security breaches – either because they match known patterns of unwanted activity, or because they

[76] Peter Haag, "Watch Your Flows with NfSen and NFDUMP" (2005), available at https://meetings.ripe.net/ripe-50/presentations/ripe50-plenary-tue-nfsen-nfdump.pdf (accessed 23 August 2019).

[77] Jingwei Huang, Zbigniew Kalbarczyk, and David M Nicol, "Knowledge Discovery from Big Data for Intrusion Detection Using LDA" (2014) *IEEE International Congress on Big Data* 760-761, p. 760.

[78] Arthur Clune, University of York, personal communication. 17th August 2018.

[79] Article 29 Working Party, "Breach Notification", *supra* n. 8, p. 13.

[80] Huang et al, *supra* n. 77, p. 760.

do not match normal patterns – and to alert human analysts of the need to investigate these correlated groups.[81]

This section reviews how automation affects the earlier discussion of purpose, necessity and safeguards; the next considers it in the light of the GDPR's provisions on profiling and automated decision making.

5.1 Automation and Purpose

Introducing automation does not change the purpose of breach detection – "ensuring network and information security"[82] – it merely changes the means by which part of that purpose is achieved. In fact, automation should guarantee adherence to that purpose, since event reduction and correlation programs can be written to specifically target groups of events likely to indicate security breaches. Unlike human analysts, their focus is hard-coded and cannot wander onto other implications of the data they may see.

Automation may even allow the same breach detection purpose to be achieved through fundamentally different, and less intrusive, techniques: not just a faster version of what was previously done by a human. Zeuch et al describe how an algorithm needed to examine fewer log fields than a human to detect attacks,[83] Anderson et al suggest how malware infections can be detected from the encryption parameters used, rather than having to decrypt all traffic.[84]

[81] Tyler Wall, "SIEM Implementation Strategies" (Tripwire, March 13 2018), available at https://www.tripwire.com/state-of-security/incident-detection/log-management-siem/siem-implementation-strategies/ (accessed 23 August 2019).

[82] GDPR, *supra* n. 1, Recital 49.

[83] Richard Zeuch, Taghi Khoshgoftaar, and Randall Wald, "Intrusion Detection and Big Heterogenous Data: A Survey" (2015) 2:3 *Journal of Big Data* 1-41, p. 34.

[84] Blake Anderson, Subharthi Paul, and David McGrew, "Deciphering Malware's use of TLS (without Decryption)" (2016) 14(3) *Journal of Computer Virology and Hacking Techniques* 195-211.

Even where programs implement the same method as humans, they can be written to ensure compliance with requirements such as the legitimate interests balancing test, discussed in section 2.2. For example, in accordance with privacy by design principles,[85] automated Denial of Service detection processes benefit from the structure of their input data: inspecting low-risk headers first and passing most legitimate traffic based on these few fields, then performing more detailed inspection of higher-risk data only for flows whose headers raise concern.[86] Parts of messages that contain insufficient information about attacks to justify examining them can be ignored. This ensures that actions involving a slightly greater (but still low) risk to individuals' privacy are only taken where this is justified by a greater risk of individuals being harmed a security breach.

5.2 Automation as a Safeguard

Considering automation as the first stage in Mantelero's differentiated approach to pseudonyms suggests that it is likely to act as a safeguard of individuals' rights. One of the main aims of automation is to eliminate the "noise" (from a breach detection perspective) represented by the majority of legitimate and non-threatening activity, allowing human analysts to concentrate on threatening and unusual patterns. Events with a harmless explanation are not only protected by pseudonymisation, as discussed in the previous section: with an automated event reduction system they are unlikely to be seen by human analysts at all. Inspection by computer – for example when checking emails for malware – has been treated by the Article 29 Working Party as less privacy intrusive than the

[85] GDPR, *supra* n. 1, Article 25.

[86] E.g. F5 Networks, "The F5 DDoS Protection Reference Architecture" (19 December 2014), available at https://f5.com/resources/white-papers/the-f5-ddos-protection-reference-architecture (accessed 23 August 2019).

same check being done by a human.[87] Automation of such tasks should be considered as a positive safeguard.

As well as making the breach detection process more privacy-respecting, automation may make it faster and more effective. With early detection of breaches recognised as an important way to reduce their impact,[88] approaches such as those described by Zeuch et al – which identified 784 security incidents in a dataset where traditional techniques found only eight[89] – may make a major contribution to data protection.

Kuner et al suggest that automation may also act as a safeguard against discrimination and bias: "human decision-making is often influenced by bias, both conscious and unconscious, and even by metabolism … intriguing possibility that it may in future be feasible to use an algorithmic process to demonstrate the lawfulness, fairness and transparency of a decision made by either a human or a machine to a greater extent than is possible via any human review of the decision in question". [90] Tired, hungry incident responders and users of their systems should welcome the consistency and respectfulness of automated decisions.

5.3 Automation and Necessity?

This analysis suggests that, for types of breach where it is known to be effective, automation can reduce both the data protection risks from security breaches – because they should be detected and resolved more quickly – and the risks

[87] Article 29 Working Party, "Working Document: Privacy on the Internet – An integrated EU Approach to On-line Data Protection" 5063/00/EN/FINAL WP37, p. 40.

[88] GDPR, *supra* n. 1, Recital 85.

[89] Zeuch et al, *supra* n. 83, p. 2.

[90] Christopher Kuner et al, "The Challenge of Big Data for Data Protection" (2012) 2(2) *International Data Privacy Law* 47-49, p. 47.

arising out of the detection process itself – because there is less human inspection of users' activities and safeguards can be built in. In particular, the human intrusion into legitimate online activities will be much less, as these will be classified as non-malicious by machines, rather than human eyes. For more than a decade, automation has been recognised as a way to defend against malware and spam "without prejudice to confidentiality of the communications":[91] it may well be appropriate to view more modern automated breach detection techniques in the same light.

Indeed, under the "necessity" principle that processing should choose the least intrusive among a number of different ways of achieving its purpose,[92] it might be argued that the law should positively encourage the greater use of automation, where it can replace human inspection. This is likely, however, to require consideration of GDPR Article 22, which applies specifically to "Automated individual decision-making, including profiling". The next section considers how this might affect the use of automation in detecting breaches.

[91] Article 29 Working Party, "Opinion 2/2006 on privacy issues related to the provision of email screening services" 00451/06/EN WP 118 (hereinafter "Email Screening Services"), p.40.

[92] Article 29 Working Party, "Profiling Guidelines", *supra* n. 46, p. 17.

6 Profiling and Automated Decision Making

"Profiling" is defined in Article 4(4) of the GDPR as:

> any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements[93]

Authors vary in their assessment of the Regulation's attitude to profiling. De Hert and Papakostantinou consider the Regulation, like its predecessor Directive, treats it as a potentially beneficial activity whose risks can be mitigated by regulatory controls. Therefore "the new rules do allow profiling operations to take place even based on sensitive data under the general, but not always applicable, condition that special measures for the protection of individuals have also been implemented".[94] Rubinstein agrees that automating decision-making can "substantially improve its accuracy and scope",[95] but is less optimistic about the law's power to ensure that improvement is used to benefit individuals.

This section considers whether breach detection will involve either "profiling" or "automated decision making" within the GDPR definitions and, if so, how it can comply with the law's requirements.

[93] GDPR, *supra* n. 1, Article 4(4).

[94] Paul De Hert and Vagelis Papakostantinou, "The New General Data Protection Regulation: Still a Sound System for the Protection of Individuals?" (2016) 32 *Computer Law & Security Review* 179-194, p. 189.

[95] Ira Rubinstein, "Big Data: The end of privacy or a new beginning?" (2013) 3(2) *International Data Privacy Law* 74-87, pp. 77-78.

6.1 Profiling

Breach detection will involve processing information about the use of networks and systems, to identify attacks and those who have been affected by them. It could be argued that this falls outside the Regulation's definition of Profiling, as the purpose is to identify insecure machines and accounts, not to "evaluate … personal aspects" of their users. However such hair-splitting should be unnecessary, as the "special measures for the protection of individuals" set out in Recital 71[96] are, in any case, things that strongly support the aims of automated breach detection systems and their operators. Those developing such systems already strive to identify "appropriate mathematical or statistical procedures". False positives (alerts when there is no security breach) and false negatives (failure to detect an actual breach) both undermine the effectiveness of systems and waste operators' and users' time, so developers and operators are keen to "ensure … that factors which result in inaccuracies in personal data are corrected and the risk of errors minimised". Much of the information processed would help an attacker – not least by informing him whether his activities have been detected and recognised – so there is a strong incentive to use both technical and organisational measures to keep it secure. In 2013 the Article 29 Working Party recommended pseudonyms (discussed in section 4) as a specific safeguard for profiling.[97] Discriminatory algorithms – where protected characteristics of the attacker or victim affect the likelihood of an attack being detected – would constitute false positives, false negatives, or both, so should quickly be rejected.

[96] GDPR, *supra* n. 1, Recital 71.

[97] Article 29 Working Party, "Advice paper on essential elements of a definition and a provision on profiling within the EU General Data Protection Regulation" (13 May 2013) (hereinafter "Profiling Advice"), p. 4.

If profiling involves "systematic and extensive evaluation of personal aspects" then Article 35(3)(a) requires a data protection impact assessment (DPIA).[98] Since breach detection systems are not intended to "evaluate personal aspects" at all, it seems unlikely that they would reach this threshold. However, as their legal basis is the legitimate interests of the organisation, they will in any case be subject to the data minimisation and rights balancing tests required by Article 6(1)(f). Before committing to a large-scale, expensive and resource-intensive deployment, organisations are likely to perform a detailed assessment of the shared risks and benefits for both the organisation and its users, in terms very similar to a formal DPIA.

Whether or not breach detection systems involve profiling in the GDPR sense, they will therefore benefit greatly from being developed and used in accordance with the Regulation's wishes. In fact the Regulation does not impose any requirements merely because an activity falls within the definition of "profiling". Instead GDPR Article 22 places requirements on "automated individual decision-making", which is considered in the next section.

6.2 Automated Decision Making

In most cases automated breach detection will be used to raise alerts when sequences of events require further investigation by human analysts. Any subsequent action will normally be based on the conclusions reached by those analysts, taking into account their previous experience and the context surrounding the particular sequence of events. For example an analyst should quickly identify when a spike in network traffic is due to a new release of a

[98] GDPR, *supra* n. 1, Article 35(3)(a).

popular operating system, rather than an attack.[99] According to the Article 29 Working Party's Guidelines on Automated Individual Decision-Making and Profiling, this involvement of "someone who has the authority and competence to change the decision … consider[ing] all the relevant data" will take the activity outside the scope of Article 22.[100]

In a few situations, however, the threat to a system, data or users will be sufficiently clear and urgent that operators will choose to have an alert trigger an immediate automated response. Such responses are commonly used to block senders of virus-infected or spam emails;[101] in some countries to quarantine ISPs' customers whose systems appear to have been compromised;[102] and increasingly to re-direct distributed denial of service (DDoS) attacks away from their targets. These systems – sometimes referred to as Incident Prevention Systems (IPS) – make decisions without prior human review, so might constitute "automated individual decision-making", regulated by GDPR Article 22.

Article 22(1) states that:

> The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her

Applying this to automated incident prevention therefore raises three questions: what constitutes a solely automated decision? What right may be created? Does

[99] Alex Hern, "iOS7 update doubles UK and German net traffic and may have reached 100m" (The Guardian, 19 September 2013), available at https://www.theguardian.com/technology/2013/sep/19/ios-7-update-traffic-100-million (accessed 23 August 2019).

[100] Article 29 Working Party, "Profiling Guidelines", *supra* n. 46, p. 21.

[101] Article 29 Working Party, "Email Screening Services", *supra* n. 91, p. 2.

[102] Jeroen Pijpker and Harald Vranken, "The Role of Internet Service Providers in Botnet Mitigation" (2016) *Proceedings of the 23rd European Intelligence and Security Informatics Conference* 24-31, p. 26.

the decision sufficiently affect an individual to create that right?

Kuner et al note that Article 22(1) is an expansion – from profiling to any kind of automated decision-making – of Article 15(1) of the 1995 *Data Protection Directive*.[103] Analysing that Article, Bygrave concluded that "a response on the part of computer software … to particular constellations of data and data input" probably is a decision.[104] He also considers that a decision is solely automated "if a decision … originates from an automated data-processing operation the result of which is not actively assessed by [any person] before being formalised as a decision".[105] Unsurprisingly, automated incident prevention does therefore involve solely-automated decisions.

Bygrave considered the Article 15 right "one of the most difficult to construe properly" in the Directive.[106] He noted that it "does not take the form of a direct prohibition on a particular type of decision making":[107] a Member State could comply either by creating an individual right of human review after an automated decision was made, or by proactively banning such decisions.[108] A right to review is stronger – since it concerns the decision reached, not just the risks involved in the processing that led up to it – than the right to object that already exists (under Art.14 DPD/Art.21 GDPR) whenever processing is based on legitimate interests.[109] Rubinstein interprets Article 15 as a right to "resist automated decisions and seek human intervention";[110] in 2013 the Article 29

[103] Christopher Kuner et al., "Machine Learning with Personal Data: is Data Protection Law Smart Enough to Meet the Challenge?" (2017) 7(1) *International Data Privacy Law* 1-2, p. 1.

[104] Lee Bygrave, "Minding the Machine: Article 15 of the EC Data Protection Directive and Automated Profiling" (2001) 17(1) *Computer Law and Security Report* 17-24, p. 19.

[105] *Ibid.*, p. 20.

[106] *Ibid.*, p. 17.

[107] *Ibid.*

[108] *Ibid.*, p. 18.

[109] Ibid.

[110] Rubinstein, supra n. 95, p. 79.

Working Party also appear to have intended this interpretation of Article 22 as an individual, retrospective right:

> Data subjects should also have the right to access, to modify or to delete the profile information attributed to them and to refuse any measure or decision based on it or have any measure or decision reconsidered with the safeguard of human intervention.[111]

However five years later the Working Party concluded that the Article 22 "right" was on the contrary "[a] general prohibition on this type of processing […] to reflect the potential risks to individuals' rights and freedoms".[112] Whether automated incident prevention continues to be permitted therefore depends, since none of the Art.22(2) exemptions applies, on whether it "similarly significantly affects" data subjects.

An automated action that prevents an individual becoming a victim of crime might seem to significantly affect them however, from the context, Bygrave considers that Article 15 in fact requires a decision that is "significantly **adverse** in its consequences" (emphasis added),[113] suggesting that "it is extremely doubtful that Art. 15(1) may apply when a decision has purely beneficial effects for the data subject."[114] In their 2013 analysis the Article 29 Working Party expected the future Regulation to provide "a reasonable degree of discretion to assess the actual effects – positive and negative";[115] in 2018 that "only serious impactful events will be covered" by Article 22.[116] Although not explicit, this seems to confirm that only adverse effects and impacts are of concern (the

[111] Article 29 Working Party, "Profiling Advice", supra n. 97, p. 3.

[112] Article 29 Working Party, "Profiling Guidelines", supra n. 46, p. 9.

[113] Bygrave, supra n. 104, p. 20.

[114] *Ibid.*, p. 19.

[115] Article 29 Working Party, "Profiling Advice", *supra* n. 97, p. 4.

[116] Article 29 Working Party, "Profiling Guidelines", *supra* n. 46, p. 21.

requirement for an "adverse legal effect" is made explicit in s.49(2)(a) of the UK's *Data Protection Act 2018*). Since the effect of automated incident prevention should be positive: to remove, or at least reduce, the impact of the crime on the victim, this "seriously impactful" test should ensure it does not fall within the Working Party's Article 22 ban.

The attacker whose aims – such as the installation of profitable ransomware – are thwarted might wish to argue that this does constitute a "significantly adverse" outcome for them. However automated blocking of such an attack will not create "legal effects" of the kind discussed by the Working Party (cancellation of a contract, denial of a social benefit granted by law, refusal of entry to a country).[117] Any subsequent process that did lead to legal effects such as fines or imprisonment would be the result of considerable human decision-making within the prosecution system, so would fall outside Article 22.

In the past the Article 29 Working Party has strongly supported automated scanning and blocking of virus-infected e-mails.[118] Automated systems are now used to protect against ransomware and Denial of Service attacks that can shut down even global service providers[119] and user organisations.[120] This new interpretation of Article 22 as a prohibition makes it essential that its threshold is set well above the level of actions required to defend users, networks and systems. In particular, Regulators must be cautious in

[117] *Ibid*.

[118] Article 29 Working Party, "Email Screening Services", *supra* n. 91, p. 6.

[119] Nicky Woolf, "DDoS attack that disrupted internet was largest of its kind in history, experts say" (The Guardian, 26 October 2016), available at https://www.theguardian.com/technology/2016/oct/26/ddos-attack-dyn-mirai-botnet (accessed 19 August 2019).

[120] Alex Hern and Samuel Gibbs, "What is WannaCry ransomware and why is it attacking global computers?" (The Guardian, 12 May 2017), available at https://www.theguardian.com/technology/2017/may/12/nhs-ransomware-cyber-attack-what-is-wanacrypt0r-20 (accessed 19 August 2019).

interpreting "similarly significant" non-legal effects to ensure that automatically depriving criminals of financial opportunities does not fall within the ban.

6.3 Automation in practice

By interpreting Article 22 as a prohibition, and therefore having to apply a high threshold, the Working Party has removed lower-impact automated decisions from both the additional information provisions in Article 15(1)(h) and the safeguards in Article 22(3). Even if this means that organisations are not legally required to operate their automated defences in accordance with these Articles, it may well benefit their purpose to do so.

Users of online services should be reassured to know that their providers are using automated technologies to detect activity that is abnormal or has the characteristics of known attacks and, if appropriate, to block it. As discussed in the previous chapter, automated alerts reduce the quantity of personal data that needs to be inspected by human incident responders, thus providing greater privacy for legitimate use. The Working Party has recognised that having a program, rather than a human, check for malicious content is more privacy-protecting;[121] it should also be faster and more effective. A public notice of the presence of automated defences might even discourage some attackers who conclude that the benefits of attacking that organisation's systems and users are not worth the risk. Such transparency should not, of course, go as far as telling an attacker how to circumvent the defences and evade detection, but the level of explanation proposed by the Article 29 Working Party in 2018 should not create these risks.[122]

[121] Article 29 Working Party, "Email Screening Services", *supra* n. 91, p. 40.

[122] Article 29 Working Party, "Profiling Guidelines", *supra* n. 46, p. 25.

Even for high-impact decisions, the law does not appear to require data controllers to notify data subjects when an automated rule has been triggered.[123] The purpose of breach detection will, nonetheless, often encourage operators to do so. Where a user has been placed in quarantine, the security team will want to assist them in removing the malicious software or other cause. Blocking of a DDoS attack is likely to be of interest to an organisation's managers and IT staff, but not to the majority of users who benefit from the silent protection. The volume of automatically-blocked e-mails (48% of all messages in 2018, according to Symantec[124]) is likely to mean recipients will not want to be interrupted every this happens, but systems will normally offer the option to periodically review such messages and provide feedback if algorithms are mis-classifying them.

These opportunities to review and tune algorithms based on user feedback again mean that security teams are likely to want to do more than the law requires. Recital 71 applies only to high-impact decisions and suggests only that "the controller should … ensure, in particular, that factors which result in inaccuracies in personal data are corrected and the risk of errors minimised".[125] However incident responders whose algorithms have failed to accurately detect a breach – even if using entirely accurate personal data and with minimal impact, this time – will have a strong incentive to improve them.[126]

It therefore appears that even automated incident prevention can be done in compliance with the GDPR. However there is sufficient concern about large-scale automated data processing that mere legal compliance may not be sufficient

[123] *Ibid*.

[124] Symantec, "Internet Security Threat Report Vol.24" (Symantec, 2019), p. 1, available at https://www.symantec.com/security-center/threat-report (accessed 12 August 2019).

[125] GDPR, *supra* n. 1, Recital 71.

[126] Wall, *supra* n. 81.

to ensure public confidence and support. The final section considers how breach detection, including automation, can achieve that.

7 Beyond Compliance: Avoiding "Creepiness"

Concerns about large-scale processing of personal data are widespread, crossing even the boundaries between traditionally different privacy cultures. In the USA, Leonard finds a "perception that business data analytics principally involved hidden and deliberatively secretive identification and targeting of individual consumers for 'one to one' marketing":[127] in Europe the concern is "the personal dignity and integrity of individuals compromised by decisions made by automated processes, when contrasted with decisions made by humans having regard to individual circumstances and constrained by human rights laws and also, perhaps, human empathy?".[128]

Leonard notes, however, the "highly contextual way in which 'creepiness' concerns arise":[129] machines are not always bad, humans not always good. As Nissenbaum[130] would predict, in some contexts automation is perceived as a benefit, in others a threat. Doubts are widespread whether compliance – even with strict European privacy laws – will be sufficient to avoid these concerns. Rubinstein worries that the GDPR, while recognising "issues associated with targeting, profiling, and consumer mistrust, relies too heavily on the discredited informed choice model, and therefore fails to fully engage with the impending

[127] Peter Leonard, "Customer Data Analytics: Privacy Settings for 'Big Data' Businesses" (2014) 4(1) *International Data Privacy Law* 53-68, p. 54.

[128] *Ibid.*, p. 55.

[129] *Ibid.*, p. 54.

[130] Helen Nissenbaum, *Privacy in Context* (Stanford: Stanford University Press, 2010), p. 195.

Big Data tsunami".[131] De Hert and Papakonstantinou note that drafting of this law started before big data began to "challenge the limits of legislation".[132]

This final section therefore summarises, first, how breach detection contributes to, rather than conflicting with, expectations of online service use; then how, in each of the areas discussed, the purpose of breach detection is best served by doing more than the law requires.

7.1 Contributing to the Online Context

Nissenbaum suggests that, whether or not a use of data complies with the applicable law, individuals are likely to perceive it as breaching privacy if it conflicts with their expectations for the context in which it was provided.[133] This goes beyond the Article 29 Working Party's use of surprise as an indicator of incompatible processing,[134,135] since a fully disclosed secondary use may still conflict with contextual expectations. However, security – including breach detection and remediation – should be a basic expectation whenever we go online. Legislators and regulators are making this expectation more explicit; the media frequently remind us of the risks posed by services whose security measures are insufficient.[136] The fact that personal data are processed for these purposes should neither surprise users, nor breach their contextual expectations.

In contrast to the systems that concern Leonard, breach detection is the opposite of a secret transfer of value from individual to provider. As discussed in section 2.1, its primary purpose, which requires it to be done openly, is to

[131] Rubinstein, *supra* n. 95, p. 74.

[132] De Hert and Papakonstantinou, *supra* n. 94, p. 180.

[133] Nissenbaum, *supra* n. 130, p. 140.

[134] Article 29 Working Party, "Purpose Limitation", *supra* n. 21, p. 24.

[135] Article 29 Working Party, "Transparency", *supra* n. 43, p. 24.

[136] Rory Cellan-Jones, "Dixons Carphone Admits Huge Databreach" (BBC News, 13 June 2018) available at https://www.bbc.co.uk/news/business-44465331 (accessed 19 August 2019).

protect those whose data may be at risk. The benefits that accrue to service operators are a secondary result of achieving that primary purpose: sales are not lost, reputations are protected (or even enhanced), fines and compensation do not need to be paid. Here the interests of individuals and providers are strongly aligned, not conflicting. As discussed in section 4, Individuals may be 'targeted', in the sense of receiving personalised attention, but this will only happen when they appear to have been victims of a security breach and need help. Not informing victims would defeat the purpose of the processing and leave both individual and provider exposed to continuing harm.

As discussed in section 6.2, only a few breach detection processes will involve fully-automated decision-making. In most cases, humans will check the results and recommendations of automated systems against their own experience and knowledge of context – precisely to ensure that a breach, rather than an unexpected but legitimate activity, is the most likely cause – before taking any action. Even where defensive actions are fully automated, a rapid human feedback process is essential to achieve the desired goal of permitting legitimate traffic while blocking hostile activity. Automation should, in fact, free up human resource to make these context-dependent decisions: in breach detection, machines and humans have highly complementary roles.

7.2 More than Compliance

Recital 49 of the GDPR sets a high standard for network and information security activities. Most of the information processed will be subject to data protection law; in addition to the normal requirements of necessity, proportionality, fairness, etc., as a legitimate interest processing must, unlike any other legal basis, be explicitly tested against the risk to individuals' rights and freedoms. For breach detection, these are not just legal requirements: they are essential to delivering the objective of improving the security of users, data systems and networks. Indeed, as this paper has shown, that objective often encourages security teams to implement more safeguards than the law requires. They should not be worried by – or seek to avoid – falling within regulation's scope.

Section 2 discussed purpose limitation. Unlike many large-scale data processing activities, breach detection is focussed, from the start, on a single, well-defined purpose. Furthermore, that purpose now has strong support from regulators, and publicising it can provide direct benefits by discouraging casual attackers. Whereas the law establishes two types of secondary purpose – those that are compatible and those that are separately declared – breach detection will be most effective if done in accordance with both sets of obligations. Its activities should be closely linked to the continuing, secure, delivery of online services and they should always be designed to enhance, rather than put at risk, the rights and freedoms of individuals using those services.

Section 3 examined necessity: the requirement that processing be done in the least intrusive way that will achieve the objective. Although breach detection does require processing of large quantities of information about the use of networked services, this reflects the universal risk of becoming a victim of an on-line attack and the relevance of that information for detection and mitigation.

Section 4 examined technical and procedural safeguards. Breach detection can largely be done using pseudonyms, with individuals only being identified when there is a high likelihood that they have become a victim and need help. This should be done as soon as possible: keeping information long after a breach is discovered is directly contrary to the purpose of the processing.

Sections 5 and 6 considered the use of automation in breach detection: how this may affect the issues raised in purpose and necessity, and how it may be affected by the new Article 22 rules on profiling and automated decision-making. Section 5 identified multiple benefits of automation: allowing purpose and safeguards to be written into code, rather than just policy and practice;

reducing the need for human inspection of legitimate activities, and allowing harmful ones to be identified and addressed more quickly. Section 6 concluded that even fully automated responses to attacks are unlikely to involve the "serious impactful effects"[137] to which Article 22 applies. Nonetheless, the requirements on processing that does exceed that threshold are beneficial to breach detection and response, so security teams using automation should follow them anyway: accuracy of algorithms is essential and data protection impact assessments likely to be beneficial for large-scale deployments; informing those subject to decisions is a key part of helping them recover from being victimised. A notice that automated breach detection is being used to improve security should reassure legitimate users and may discourage attackers. Security teams will be delighted if the latter wish to exercise their right to object to this processing!

Finally, this section has shown that there is, in fact, no paradox. Processing personal data and protecting it from security breaches should be inextricably linked in expectation, law and practice. Where protecting personal data requires further processing, not only can this be done in accordance with the law, the protection is likely to be ineffective if it is not. In many cases, effective breach detection requires security teams to take even more care than the law requires. Well-conducted breach detection activities should be a source of reassurance to data subjects, data controllers and regulators alike.

[137] Article 29 Working Party, "Profiling Guidelines", *supra* n. 46, p. 21.