The Law Enforcement Directive (EU) 2016/680 (LED) defines its basic principles, such as purpose limitation and data minimisation, differently than the General Data Protection Regulation (EU) 2016/679 (GDPR). This contribution is exploring the influence of those differences on new policing methods, in particular on the big data policing. After describing the data protection regulatory framework for law enforcement authorities in the EU, we explain our understanding of the notion of big data policing. We then critically interpret the purpose limitation and the data minimisation principle in the GDPR and the LED, thereby busting some myths about the LED, created by other academics. Finally, we explore the boundaries of the abovementioned basic LED principles, in an attempt to measure their success in finding the delicate balance between the high level of personal data protection and the contemporary law enforcement needs.

Keywords: data protection, Law Enforcement Directive, criminal justice, big data, purpose limitation, data minimisation
1. Introduction

The regulation of personal data processing by the police and other law enforcement authorities is an under-researched area of law. It is often perceived as technical, too fragmented, in constant motion, difficult to follow or simply not interesting. However, in recent years this topic has significantly gained importance due to several reasons: data collection and data exchanges by police are a key element in the fight against cybercrime and cyber-enabled crime; police forces are becoming increasingly tech-savvy, investing in the capacity of their IT departments to use state-of-the-art personal data analytical tools, including the big data analytics; and the legislative framework for personal data processing for law enforcement purposes is undergoing consolidation, at least at the level of the European Union (EU).

This development is not only highlighting the importance of data protection law in policing, but also amplifying the impact of its interactions and clashes with the innovative ways of police work. In this chapter, we will explore how the EU data protection law regulates the so called ‘intelligence-led policing’, in particular the police use of big data analytics. We will start by briefly presenting the data protection rules applicable to law enforcement authorities in the EU, and what we consider ‘big data policing’. We will subsequently dissect the construction of two basic principles – purpose limitation and data minimisation – in the EU’s main data protection instrument applicable to law enforcement – the Law Enforcement Directive (EU) 2016/680 (LED)\(^2\). After some considerations about the alleged incompatibility of big data analytics with basic data protection principles, we will demonstrate how the LED, through its peculiar architecture of the two basic data protection principles, manages to strike the delicate balance between competing privacy and security interests in the contemporary world. However, we are fully aware of the speed of technology developments and the fragility of the LED balance, which the EU has a golden opportunity to preserve.

2. Brief presentation of the applicable data protection law

Data protection law is the set of legal provisions that apply to any ‘processing of personal data wholly or partly by automated means’. These provisions are anchored in the OECD data protection guidelines (1980),\(^3\) the Council of Europe Convention 108 (1981)\(^4\), or the UN Guidelines concerning Computerized Personal Data Files (1990) and in the EU regulatory framework on data protection. All of these regimes contain some definitions and a list of principles for those that process personal data.

---


In particular, European data protection law is broad and generic, be it defined at the level of the Council of Europe, or at the level of the European Union. It applies to both private and public sector, including police, national security or defence. Its objective is to offer protection to individuals against potential abuses and to regulate the trans-border flows of personal data. It lays down the basic principles of data protection, in particular, the principles of fairness and lawfulness of data processing, the **purpose limitation principle** (data should be collected for specified, legitimate and explicit purposes and not further processed in a way incompatible with those purposes), the **data minimisation principle** (collection of data should not be excessive or should be limited to what is necessary in order to achieve the purpose), the storage limitation principle (data should not be kept longer than necessary). More recently recognized principles (mainly at the EU level) are the principles of transparency, accountability and privacy/data protection by design. The European data protection law also lays down stricter requirements for the processing of sensitive categories of personal data (data relating to race, politics, health, religion, sexual life or criminal records), grants basic data subject rights to access to his or her personal data and to rectify inaccurate data and establishes specialised oversight quasi-judicial bodies, called supervisory authorities or data protection authorities (DPAs).

In Europe’s regulatory landscape, we will focus on instrument with the highest relevance for big data policing, the LED.

Like the famous General Data Protection Regulation (EU) 2016/679 (the GDPR), the lesser known LED was adopted in May 2016, constituting a major step forward in establishing a comprehensive EU data protection regime, as the first horizontal and legally binding instrument laying down the rules for national and cross-border processing of personal data in the area of law enforcement. Two main objectives of the LED are slightly different and more specific than the objectives of the GDPR. Firstly, the LED seeks to establish an increased level of fundamental rights protection in the area of police and criminal justice. Secondly, the LED is supposed to improve sharing of personal data between the Member States, as they will be able to rely on uniform data protection rules (Article 1(2)).

The rules of the Directive, which are by now transposed by all 27 EU Member States, the UK (who transposed the LED while still an EU Member State) and the four Schengen Area States (Norway, Iceland, Switzerland and Lichtenstein), benefited from the major construction site of the GDPR, as a number of

---

5 Convention 108, Article 3(1): The Parties undertake to apply this Convention to automated personal data files and automatic processing of personal data in the public and private sectors.
7 Citizen have the right to be informed that their data is being used in a processing operation. They also have the right to access their data when these have been processed, e.g., they can investigate how the processing operation is carried out, whether databases exist, what their purpose is, and who is responsible for the processing. Furthermore, in case the data appear to be incomplete, inaccurate, or processed in a manner that is incompatible with the other data protection principles, the data subject has the right to ask for the rectification, or even the erasure of his data. Data subjects are also entitled to object to the processing of their personal data provided there are “compelling legitimate grounds”. Finally, data subjects have the right “not to be subject to a decision which produces legal effects concerning him or significantly affects him and which is based solely on automated processing of data”, which means that important decisions concerning them cannot be taken solely on the automated processing of data, and that they have a right to actively participate in those very decisions.
solutions simply spilled over into the LED text. We would like to focus on some specific LED features of particular importance for big data policing. In that context, there are two basic questions about the LED, asked by two different camps. On the one hand, the law enforcement community (the authorities applying criminal substantive and procedural law on the ground every day) wonder whether the LED is flexible enough in order not to interfere with the efficient progress of criminal investigations and proceedings. On the other hand, the data protection community questions the quality of the LED and wonders whether its provisions ensure a comprehensive and high level of protection of personal data. In the following, we aim to alleviate concerns of both camps. We will demonstrate that the contemporary policing methods are regulated by the LED in a way that strikes the delicate balance between the competing privacy and security interests.

3. Big Data policing

We understand big data as the analysis of large data sets in order to find new correlations -for example, business or political trends or to prevent crime-, and to extract valuable information from large quantities of data.11 Big data is also an umbrella term for technological and societal developments that are already taking place (use of profiles, algorithms, cloud computing, machine learning, commodification of data, open access to governmental data, datafication, securitization and risk society).12 Amongst its promises, big data gives us the ability to predict (generating results that say something about the future of an organization or the result of a concrete action); it gives actionable results and opportunities for direct actions on results found without human intervention and all this in an adaptive, scalable and real-time manner (new speed standards ensure immediate reaction capacity to new situations).13 Big data entered the life of police forces, not only in intelligence headquarters, but also on the street, supporting the work of the individual police officer.14 Ferguson includes under the term big data police technologies predictive systems that identify people or places suspected of crime, surveillance systems to monitor at-risk areas and search systems to mine data for investigative clues or to develop intelligence

---

11 Nikolaus Forgo, Stefanie Hänel & Benjamin Schütze, ‘The Principle of Purpose Limitation and Big Data’, in M. Corrales et al. (eds.), New Technology, Big Data and the Law, Perspectives in Law, Business and Innovation, Singapore: Springer Nature Singapore, 2017, 17-42. Comp. big data is the collection and aggregation of large masses of (publicly, commercially, proprietarily, and/or illicitly) available data from a wide variety of different sources and its analysis, largely in the form of correlation, pattern-recognition, and predictive analysis. So big data is about massive collection and analysis, often secondary analysis or analysis for new purposes, other than those for which the data were originally produced and collected (A. Saetnan, I. Schneider & N. Green (eds.), The Politics and Policies of Big Data: Big Data, Big Brother? Oxon: Routledge, 2018, (1-18), 6. This contribution also contains also a critical discussion of existing definitions, including the commonly used definition proposed by Laney (2001) of big data as great data volume, data velocity, and data variety (and veracity). For more on definitions, see Bart van der Sloop & Sascha van Schendel, ‘Ten Questions for Future Regulation of Big Data: A Comparative and Empirical Legal Study’, JIPITEC, 2016, vol. 7, (110-145), 112-114
12 Bart van der Sloop & Sascha van Schendel, 115.
14 “Predictive analysis and real-time access to intelligence and tasking in the field will be available on modern mobile devices. Officers and staff will be provided with intelligence that is easy to use and relevant to their role, location and local tasking” (quote from the UK National Policing Vision (2016) taken from R. Van Brakel, ‘Pre-Emptive Big Data Surveillance and its (Dis)Empowering Consequences: The Case of Predictive Policing’, in Bart van der Sloop, Dennis Broeders, Erik Schrijvers (eds.), Exploring the Boundaries of Big Data, Amsterdam: Amsterdam University Press, 2016, (117-141), 118
nets of helpful data for groups or across communities. Van Brakel sees big data applications used by police as instrumental in building up a police strategy of pre-emption strategy that directs the police beyond (simply) statistically based decision-making and intelligence led-policing, and allows better predictive mapping (i.e. application of predictive analytics to predict when and where a crime may take place at an aggregate level of analysis), but also predictive identification, where the analysis is at the individual or group level (i.e. application of predictive analytics to predict potential offenders or victims of crime).

An example of predictive mapping is crowd management policing, i.e. the monitoring of language on and use of Twitter and Facebook to make predictions about hotspots for potential crimes. An example of the latter are algorithms to predict recidivism when determining prison sanctions and assessing applications for early release from prison or technologies to profile suspect airplane and train passengers based on their passenger data, data from their records and inferences based on profiles. The disruptive potential of this trend cannot be enough emphasized: you are refused entrance to a train not because you have shown suspected behavior, but because the software decides on the basis of available real-time data and real-time analyses that you might be a suspect. The idea of predicting the future – or, more accurately, predicting a possible future with a certain degree of probability – is the underlying rationale for Google sending us personalized adds, but is now applied to sending us the police.

Broeders et al. insist that the foregoing should be understood as an emerging trend, helped by analytical tools getting more and more effective, and by the more than substantial recent growth of available data in size and variety, due to new processes of data collection: in addition to a directed collection of data (the intentional capture of data on people or objects), comes an ‘automated’ collection (data that are created as an inherent feature of a device or a system), or a ‘volunteered’ collection (data that are created by the voluntary use of systems, devices). Essential is the following observation: “The nature and origins of data that are available for security purposes, therefore, are changing. Public and private data are getting mixed. Relatively hard data (financial data and all kinds of registries) can be linked to softer, more social data. The wealth of data also renders the difference between personal and non-personal data potentially meaningless as it is now relatively easy to ‘construct’ a person on the basis of a limited set of data points that do not directly reference a person. This means there is a limit to anonymisation and pseudonymisation methods and, more importantly, that the legal difference between different types of (personal) data and the level of protection they are awarded is being hollowed out”.

Most academic authors take some time to praise the benefits of big data policing, and then run straightforwardly to a long list of critical observations, primarily about privacy, data protection and non-discrimination. Van Brakel broadens this discussion further by insisting that big data policing, although in theory a tool to improve policy, potentially disempowers certain individuals, groups, and society. Apart from

---

16 “Big Data can be seen as a socio-technological phenomenon within the larger context of the emergence of pre-emptive surveillance, offering the possibility of conducting predictive or real-time analyses of huge amounts of data – structured or unstructured, in different formats, and taken from different types of sources – which may lead to original new insights and knowledge and, as a result, change practices significantly” (R. Van Brakel, 118).
19 See for instance Broeders et al., 309-323
privacy and data protection issues, there are less visible consequences, she argues, such as algorithmic discrimination as a result of several possible types of bias,\(^{20}\) an increasing reliance on decisions made by technology about interventions, chilling effects on citizen’s use of civil liberties because of growing mass surveillance possibilities that might end up in increased distrust in police and disengagement of certain groups in society who are in need of help will cease to turn to the government.\(^{21}\) An almost identical analysis is proposed by Virginia Eubank who writes about big data mechanism used for **poverty profiling**.\(^{22}\)

“Targeting **high-risk** families might lead them to withdraw from networks that provide services, support, and community”.\(^{23}\)

Is the police aware of this? Not always apparently. Ferguson underlines how many amongst the police view these new technologies as race-neutral and objective and are adopted by police departments, hoping to distance themselves from claims of racial bias and unconstitutional practices.\(^{24}\)

When applying this high-level, but well-reasoned criticism on the data protection law on the books, it becomes apparent at the outset that big data policing jeopardises the basic data protection principles of purpose limitation and data minimization. The frontal attack on those principles is inherent to the nature of big data analytics.

Hence, many authors consider data protection as broken and not effective any more to protect privacy and personal data in the age of big data. For example, Paul Ohm in his 2014 dissection of the principles.

To understand his position, it is useful to underline one of the more important methodological innovations brought about by big data technologies: data scientist complement or just replace traditional hypothesis-driven scientific approach focuses (you pick a hypothesis and then turn to selected data to test it out), with a new data-driven or data-intensive approach that begins with data, and uses statistical analyses and machine learning tools and techniques as the data increase in size and complexity, to examine correlations among variables of system response with potential drivers of that response. No preconceived relationships are derived from a theory and many possible relationships are examined.\(^{25}\)

---

\(^{20}\) R. Van Brakel, 124-17: 1) bias that unintentionally creeps into the labelling of examples or the rules that are coded into the algorithm; 2) bias in the data used for the analysis as a result of biased assumptions that are baked into the data by the way it was collected (for instance, arrest rates); 3) bias creep due to technical defects, faults, and bugs in the system, which may lead to more false positives (rating somebody as high risk when little risk actually exists) that meet certain criteria and false negatives. See for a complementary analysis discussing bias in the choice of outcome variable (what you measure to indicate the phenomenon you are trying to predict), the choice of predictive variables (data within a data set that are correlated with the outcome variables) and the choice of validation data, see Virginia Eubank, *Automating Inequality How High-Tech Tools Profile, Police, and Punish the Poor*, NY: St. Martins Press, 2018, 143-147. On input problems (data bias, data error and data incompleteness) and output problems (false positives and negatives, accurate and reliable predictions) see A. Ferguson, *The Rise of Big Data Policing. Surveillance, Race, and the Future of Law Enforcement*, NY: New York University Press, 2017, 191-193 See also H. Lammerant & P. De Hert, ‘Predictive Profiling and its Legal Limits: Effectiveness Gone Forever?’ in B. van der Sloop, D. Broeders & E. Schrijvers (eds), *Exploring the Boundaries of Big Data*, Amsterdam: Amsterdam University Press, Amsterdam 2016, 145-173; H. Lammerant, P. Blok & P. De Hert, ‘Big data besluitvormingsprocessen en sluipwegen van discriminatie,’ *NJC`M bulletin. Nederlands tijdschrift voor de mensenrechten*, 2018, vol. 43, 1, 3-24; Alvaro M. Bedoya, ‘Algorithmic Discrimination vs. Privacy Law’, in E. Selinger, J. Polonetsky & O. Tene (eds.), *Cambridge Handbook of Consumer Privacy*, Cambridge: New York: Cambridge University Press, 2018, 232-240.

\(^{21}\) R. Van Brakel, 125-138

\(^{22}\) Virginia Eubank, *Automating Inequality*, 158. “We might call this poverty profiling. Lie racial profiling, poverty profiling targets individuals for extra scrutiny based not on their behavior but rather on a personal characteristic: living in poverty. Because the model confuses parenting while poor with poor parenting, the AFTS views parents who reach out to public programs as risk to their children.”


\(^{24}\) “After a series of high-profile police shootings and federal investigations into systemic police misconduct, and in an era of law enforcement budget cutbacks, data-driven policing has been billed as a way to turn the page on racial bias. But behind the data are real people, and difficult questions remain about racial discrimination and the potential to distort constitutional protections” (Andrew Guthrie Ferguson, *The Rise of Big Data Policing. Surveillance, Race, and the Future of Law Enforcement*, NY: New York University Press, 2017, 272p., [quote from the bookcover]) See equally on the preemption of politics and political questions by big data, Virginia Eubank, *Automating Inequality*, 168 & 197.

Consequently, big data is not only about creating and using large datasets and exploiting them with less effort and cost, but also about created completely new classes of information and finding unexpected correlations that might be beneficial in one way or another with vague starting points such as ‘learn something about our customers’, or to ‘find the patterns hiding in the data’. 26 The real value of big data is not just a more developed technology, but that it enables to reap positive network effects from combining and re-using data sources.

‘Big data empowers through surprise’ Paul Ohm observes, and on this basis, he concludes that important ideas of data protection such as the notice and choice27 and a core principle as purpose limitation, are not compatible with big data.28 The observation is now regularly heard in legal analysis. Two Dutch authors contrast the benefits of big data, with the myths of data protection (principles that became myths) and identify the following five myths: (1) the myth of anonymous big data; (2) the myth of purpose bonding; (3) the myth of data minimization; (4) the myth of consent and the illusion of transparency; and (5) the myth of security.29

Not much is left of data protection’s ambitions when these myths are demasked. MacCarthy therefore argues for a re-evaluation of these regulatory principles,30 i.e. amending data protection to make a certain big data possible. Without going as far as giving up every and all regulatory restraints, his new big data friendly data protection has not much of the old. Gone are the ideas of data minimization and controlled secondary use/purpose limitation. Via risk assessment alternative controls are proposed when there is potential risk of some degree.31 Equally, problems with anonymization, fairness and discrimination are fixed without even the slightest ambition to prevent them. A transparent framework of responsible use, based on all kind of testing of algorithms and audits by all actors involved will safeguard our values in this big data benefits all society.32

We admit that it is difficult to square purpose limitation (even a mere purpose specification obligation) and data minimisation with big data. A big data analyst often cannot specify a purpose except at a very high level of abstraction - to ‘learn something about our customers’, or to ‘find the patterns hiding in the data’. 'This is why so many big data practitioners are loathed to delete old data; you never know what use we will find for it tomorrow!'33

The LED is a modern data protection instrument, adopted in 2016, transposed to national laws (mostly) in 2018 and 2019. It had to be mindful of the needs of the 21st century intelligence-led policing on one hand,
and of the threats to individuals’ fundamental rights inherent to the digital criminal investigations and proceedings on the other. Hence, the LED via soft law (its recitals) and hard law (its provisions) demonstrates a clear willingness to adapt data protection obligations of law enforcement authorities, compared to those that the authorities bound by the GDPR have to adhere to.34

4. The purpose limitation principle within the LED – bending, not breaking

Most of the fundamental data protection principles were introduced in the European data protection law with the 1981 Council of Europe Data Protection Convention 108. Most of them survived the test of time. The EU embedded those principles in its own data protection legislation, in particular the Directive 95/46/EC35, the Regulation (EC) No 45/200136 and the Framework Decision 2008/977/JHA37. However, we believe that the EU 2016 data protection reform brought about the evolution of the basic principles. The latter still exist in the GDPR and the LED, but in new forms.

In particular, the purpose limitation principle and the data minimisation principle are formulated differently in the GDPR than in the LED. Unlike some authors, we believe these differences are significant and have important consequences for data protection in the Digital Era38. In particular, the LED assigned a double role to these two basic principles: they are the maintaining a high level of personal data protection, while ensuring the appropriate size of the data protection regulatory sandbox for law enforcement authorities. The purpose limitation principle is often considered as the cornerstone of data protection and the prerequisite for most other fundamental requirements.39 Its primordiality is evident from its constitutiveness in Article 8 of the Charter of Fundamental Rights of the EU – the fundamental right to personal data protection. The principle has two main building blocks: purpose specification (specified, explicit and legitimate purposes) and compatible use (no further processing in a manner which is incompatible with the data collection purposes)40.

The principle’s emanation in the LED (Article 4(1)(b), 4(2) and 4(3)) is far simpler and more flexible than its GDPR counterpart. Unlike the GDPR, the LED does not even mention the term further processing41. More importantly, the LED avoids complicated constructions of further processing, such as Article 6(4) GDPR. Instead, the LED contains a number of specific rules about the change of purpose, both within or outside of the LED realm, which we call subsequent processing.

35 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, Official Journal L 281, 23/11/1995 P. 0031 – 0050. This Directive was repealed and replaced by the GDPR.
37 Council Framework Decision 2008/977/JHA of 27 November 2008 on the protection of personal data processed in the framework of police and judicial cooperation in criminal matters, OJ L 350, 30.12.2008, p. 60–71 (BG, ES, CS, DA, DE, ET, EL, EN, FR, IT, LV, LT, HU, MT, NL, PL, PT, RO, SK, SL, FI, SV), Special edition in Croatian: Chapter 16 Volume 002 P. 118 – 129. This Framework Decision was in the meantime repealed and replaced by the LED.
38 For example, Cecile de Terwagne notes the difference in data minimisation principle in the GDPR and LED, but finds that ‘this difference of terms should not have a substantial effect on the scope of the data minimisation principle’. In Christopher Kuner, Lee A. Bygrace, Christopher Docksey (eds), The EU General Data Protection Regulation (GDPR), A Commentary, Oxford University Press, 2020, p. 317.
39 Christopher Kuner, Lee A. Bygrace, Christopher Docksey (eds), The EU General Data Protection Regulation (GDPR), A Commentary, Oxford University Press, 2020, p. 315.
41 Compare Article 5(1)(b) of the GDPR and Article 4(1)(b) of the LED.
In both the GDPR and the LED, the definition of the purpose limitation principle is almost identical. The notable exception is the omission of the word ‘further’ in the Directive. However, both texts prohibit the processing of personal data (be it further or subsequent) for purposes that are incompatible with the original purposes at the time of the collection of data. We will therefore zoom in on the second building block of the purpose limitation principle – the compatible use.

The GDPR has an extremely wide material and territorial scope. It regulates all sorts of processing operations, pursuing both commercial and public interest objectives. In that context, the first striking novelty, which is not so evident from its enacting terms, is that the GDPR considers further processing only possible on the same legal basis as the original data collection, by the same controller. Furthermore, it envisages further processing possible if based on Union or Member State law or the data subject’s consent. For other types of legal basis, the controller will have to carry out a compatibility test under Article 6(4) GDPR. This compatibility test will, therefore, have to be carried out when the initial data collection is based on three out of six legal grounds from Article 6(1) GDPR: (b) contract; (d) vital interest; and (f) legitimate interest. As regards three other legal grounds: (a) consent - needs to be obtained for all purposes, including the ones of further processing; (c) legal obligation and (e) task carried out in the public interest or exercise of official authority - both (c) and (e) legal basis have to be laid down in Union or Member State law, which may provide for further processing (recital 50, 4th sentence). Finally, the GDPR presumes that further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall be deemed compatible with the original collection purposes.

On the other hand, the scope of the LED is narrow, allowing simpler solutions. The LED covers only the processing by ‘competent authorities’ (which is, in the context of the Directive, pretty much an interchangeable term with the ‘controller’) and only for the purposes listed in its Article 1(1). The competent authority is defined as a public authority (Article 3(7)(a)) competent for prevention, investigation, detection and prosecution of criminal offences or the execution of criminal penalties (a law enforcement authority). Another, very limited and circumscribed option, is to consider another entity, which is not a public authority, as the competent authority under the Directive (Article 3(7)(b)). This definition can only apply to entities vested both with public authority and public powers, which means that these actors have to have some coercive powers, e.g. privately run prisons or privatised parts of police forces.

42 ‘collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes’ in the GDPR, as opposed to ‘collected for specified, explicit and legitimate purposes and not processed in a manner that is incompatible with those purposes’ in the Directive.
43 See in particular recital 50 of the GDPR and Article 60(4).
44 The controller shall, in order to ascertain whether processing for another purpose is compatible with the purpose for which the personal data are initially collected, take into account, inter alia:
   (a) Any link between the purposes for which the personal data have been collected, in particular regarding the relationship between data subjects and the controller;
   (b) The context in which the personal data have been collected, in particular regarding the relationship between data subjects and the controller;
   (c) The nature of the personal data, in particular whether special categories of personal data are processed, pursuant to Article 9, or whether personal data related to criminal convictions and offences are processed, pursuant to Article 10;
   (d) The possible consequences of the intended further processing for data subjects;
   (e) The existence of appropriate safeguards, which may include encryption and pseudonymisation.
45 These are so called privileged purposes in Article 4(1)(b) of the GDPR, second sentence. The LED does the same in Article 4(3).
46 The purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, including safeguarding against and prevention of threats to public security.
In this context, the Directive decided to put the flesh on the bones of the principle of purpose limitation in a different way than the GDPR - not through the notion of further processing, but through the specific rules on the change of purpose. This is not a novelty in EU law, since a similar concept can be found in, now defunct, Data Protection Regulation (EC) 45/2001 – another instrument with a narrow scope and specifically designed for public authorities.  

Firstly, the purpose limitation principle is present, with both building blocks alive and well, in Article 4(1)(b) of the LED. Secondly, Article 4(2) develops the principle and the change of purpose further, by explaining the concept: subsequent processing by the same or another controller is permitted if authorized by law and if necessary and proportionate to the new purpose, as long as the new purpose remains within the scope of the Directive. Subsequent processing can also be done for privileged purposes (archiving in the public interest, scientific, statistical or historical use), if carried out within the LED scope, i.e. for broader law enforcement purposes by competent authorities (paragraph 3 of Article 4).

As regards the grounds for lawfulness of processing, the Directive lays down only one legal ground in Article 8 (if necessary for the performance of a task carried out by a competent authority for the purposes of the Directive and based on Union or Member State law), while Article 6 of the GDPR, as we explained above, provides for six different legal grounds. Evidently, the legislator recognized that law enforcement authorities may only carry out tasks permitted by law, and not process personal data for the purposes of the Directive on the basis of consent, a contract or their own legitimate interest.

Hence, we have to disagree with Jasserand and Caruana when they criticize the Directive for ‘missing adequate data subject safeguards’ or ‘derogating from the purpose limitation principle as enshrined in primary law’ (the abovementioned Article 8 of the Charter). In particular, Jasserand argues that the Directive derogates from the principle of purpose limitation because it does not provide for a compatibility test when the personal data originally collected by private parties (e.g. service providers) are subsequently accessed by law enforcement authorities. We argue that the Directive does not introduce any such derogation. The principal flaws in Jasserand’s theory are twofold. Firstly, she does not distinguish between the notion of further processing in the GDPR, as opposed to subsequent processing in the LED. Secondly, her convoluted idea of further processing is a processing operation that could be carried out by a different controller and on a legal basis other than the one used for the original collection of data. While such line of reasoning may transpire from the abovementioned 2013 Opinion of Article 29 Working Party on purpose limitation, it become outdated once the data protection reform package was adopted in 2016. As we already said, further processing, in the GDPR meaning of the term, can only be done by the same controller on the same legal basis. Once data are transmitted to another controller, or the same controller starts using a different legal basis, the processing begins ab novo, with the initial processing - collection of data, followed by informing the data subject pursuant to provisions on the right of information (Articles 13 and 14 GDPR) etc.

In practical terms, if a service provider collected personal data from a customer for a provision of service, it can only use the same legal basis (contract) for further processing, if it manages to pass the compatibility test in Article 6(4) GDPR. We could imagine a scenario in which such further processing will be the transmission or disclosure of personal data to law enforcement authorities. However, such transmission would be on the GDPR side of the coin. Once the police accesses or receives such data, the LED side of the coin would trigger a new processing operation, governed by the LED, i.e. all the rules of the Directive would apply. Every subsequent use of such data by the police or other controllers within the scope of the LED, such as prosecutors or criminal trial courts, is then specifically covered by Article 4(2) of the Directive, coupled with the ‘compatible use’ building block of the purpose limitation principle in Article 4(1)(b). Hence, not only does the Directive not derogate from the principle of purpose limitation – it has more stringent requirements for the change of purpose within the scope of the Directive than the GDPR requirements for further processing when processing is based on Union or Member State law.

5. Launching the data outside of the LED – the other side of purpose limitation

The second LED novelty are provisions on data transmissions of operational law enforcement data from competent authorities to other authorities or private parties within the EU: Article 9(1) and (2). We will call these two paragraphs the launchpad provisions, as they regulate data transmissions from the LED to the GDPR realm.

Article 9(1) provides that the GDPR applies to subsequent processing of personal data originally collected under the Directive, when such data leave the scope of the Directive. An additional safeguard provides that only Union or Member State law can allow such data to be used for purposes other than law enforcement. Moreover, Article 9(2) makes it very clear that law enforcement authorities may have tasks other than the ones covered by the scope of the Directive and have to use the GDPR for carrying out such tasks.

These provisions show that the Directive is opening the door for transmitting operational data to private recipients (e.g. joint controllers in public-private partnerships), and other public actors, such as customs or tax authorities. It even allows the same authority to launch the data out of the LED and to process it under the GDPR for a non-law enforcement purpose, such as archiving. Some may argue that the launchpad provisions are yet another attack on the purpose limitation, another deficiency of the LED. However, we stick to our above reasoning. We argue that, mutatis mutandis, the respect for the purpose limitation in launchpad provisions is ensured in a twofold manner. Firstly, launching the data outside of the LED scope and into the GDPR will result in new processing governed by the rules of the GDPR. This means that the law enforcement authority or the recipient, as the case may be, has to treat the ‘arrival of the launched data’ in the same way as original data collection under the GDPR. Data subjects will have to be informed about the data collection, they will have their GDPR rights, the principle of transparency (which does not exist in the LED) becomes applicable etc. Secondly, launching can take place only if authorized by law, just like under the GDPR, further processing can take place if legal basis for processing is laid down in law. Admittedly, such laws can be (and often are) deficient: not clear, foreseeable or precise. But the quality of law can always be challenged before domestic and European courts, as well as the necessity and proportionality of such legislative measures.

In light of the above, we find that the LED’s architecture of the purpose limitation principle allows law enforcement authorities to collect personal data from a number of legally available sources without exces-
sive red tape. They can access the data originally collected by other actors (private companies) or collect the data themselves (e.g. open source), as they can carry out any other processing operation under the LED, always based on law. The LED also allows compatible, necessary and proportionate re-purposing of personal data within the material scope and the actors of the criminal justice system. In our view, this is enough to allow big data policing, while prohibiting dodgy policing practices, such as arbitrary data collection (without legal basis), obscure, clandestine and unspecified purposes of data processing, or cooperation with private companies who clearly violate the GDPR\(^52\).

6. Data minimization under the LED

Data minimisation principle requires data controllers to limit themselves, from the original data collection and throughout the processing, only to the data required for accomplishing the purposes of processing. This principle is sometimes designated as a ‘direct consequence’ of the purpose limitation principle\(^53\). At the outset, it might seem that nothing new happened on this front ever since the 1981 Convention 108. However, we argue the contrary - note how the data minimisation principle is defined in Article 5(1)(c) GDPR! It requires that personal data be ‘adequate, relevant and \textit{limited to what is necessary} in relation to the purposes for which they are processed’. On the other hand, the LED, in Article 4(1)(c) keeps the old, familiar language: data should be ‘adequate, relevant and \textit{not excessive} in relation to the purposes for which they are processed’.

This difference is striking. The data minimisation principle has three building blocks: the LED blocks are the same as in all the older abovementioned instruments. In the GDPR, first two are the same. In both instruments, once the purposes of processing are specified, only personal data that are adequate and relevant for those purposes can be processed. However, due to the GDPR novelty in the third building block, the controllers also have to demonstrate that the data collected is absolutely necessary, by showing concrete measures that were taken in order to minimise the amount of data used to serve a given purpose\(^54\).

The European Data Protection Board argues that the GDPR data minimisation principle is therefore \textit{substantiating and operationalising} the principle of necessity. Consequently, in the GDPR-concept of further processing, the controller has to periodically consider whether processed personal data is still adequate, relevant and \textit{necessary}, or if the data shall be deleted or anonymised\(^55\).

In the LED, the third building block implies a much easier burden of proof for the controller. There is no need to demonstrate the strict necessity of the data by limiting oneself to the necessary minimum. The controllers under the LED can operate with less precision, they can grab and hold on to data in a rougher manner. They just have to make sure not to process \textit{excessive} datasets. Linking this finding to our above discussion on big data policing, one cannot but conclude that the LED language affects both the legal and computational definition of the data minimisation principle, which is different than the one in the GDPR.


As a consequence, the LED facilitates law enforcement data collection and subsequent processing immanent to the intelligence-led, big data policing. At the same time, it prevents law enforcement authorities from mass surveillance-type of collecting and processing personal data ‘just in case’. After all, the LED data minimisation principle does ensure the same level of protection as the Convention 108 and the pre-2016 EU data protection legislation.

7. Conclusion

In this contribution we looked at the LED’s basic principles as framing big data policing developments. We discussed data protection, for the simple reason that if personal data is processed, big data providers have to comply; and although big data is sometimes about non-personal data, it is more often about or with personal data.

We saw that the Directive recalibrates the general data protection principles but hesitates to enter into the details about the processing work done by contemporary police. Big data relevant processing practices (web crawling, data mining, data matching, etc.) are not mentioned in the LED. Moreover, ideas such as predictive policing are launched in the recitals and provisions of the Directive without any elaboration apart from the requirement that such processing operations need to be envisaged by law. However, contrary to what we previously argued, we now agree, with Purtova, that the absence in either the GDPR or the LED of a binding provision pertaining specifically to data transfers from private parties to the law enforcement authorities is irrelevant, since transmissions or disclosures are just another instance of data processing. The message is that its general rules and principles suffice in order to provide an adequate level of data protection, and, unlike Caruana and Jasserand, we do not believe that certain safeguards are missing from the Directive.

How long will the LED-specific solutions be able to preserve their delicate balance? After all, the technology-led policing is substantially different to general law enforcement personal data processing. It is potentially more harmful to individuals, it usually takes place unnoticed and ultimately involves different technical specifications. In light of the pace of the technological development, our gut feeling is that, in this area, the EU will quickly have to complement and specify the LED; the opportunity should not be missed in the upcoming AI Act.

57 N. Purtova, 64
58 Cited above
The Brussels Privacy Hub Working Papers series

N°1 “The data protection regime applying to the inter-agency cooperation and future architecture of the EU criminal justice and law enforcement area” (November 2014) by Paul De Hert and Vagelis Papakonstantinou (35 pages)

N°2 “The new cloud computing ISO/IEC 27018 standard through the lens of the EU legislation on data protection” (November 2014) by Paul de Hert, Vagelis Papakonstantinou, Irene Kamara (25 pages)

N°3 “Towards efficient cooperation between supervisory authorities in the area of data privacy law” (October 2015) by Dariusz Kloza, Antonella Galetta (24 pages)

N°4 “The data protection regime in China” (November 2015) by Paul De Hert and Vagelis Papakonstantinou (30 pages)

N°5 “The right to privacy and personal data protection in Brazil: time for internet privacy rights?” (February 2016) by Vinicius Borges Fortes (23 pages)

N°6 “Permissions and Prohibitions in Data Protection Jurisdiction” (May 2016) by Mistale Taylor (25 pages)

N°7 “Structure and Enforcement of Data Privacy Law in South Korea” (October 2016) by Haksoo Ko, John Leitner, Eunsoo Kim and Jong-Gu Jung (20 pages)

N°8 “The “Right to be Forgotten” and Search Engine Liability” (December 2016) by Hiroshi Miyashita (15 pages)

N°9 “European Human Rights, Criminal Surveillance, and Intelligence Surveillance: Towards “Good Enough” Oversight, Preferably but Not Necessarily by Judges” (March 2017) by Gianclaudio Malgieri and Paul De Hert (25 pages)

N°10 “Does Technology Drive Law? The Dilemma of Technological Exceptionalism in Cyberlaw” (July 2017) by Meg Leta Jones, JD, PhD (31 pages)

N°11 “The Microsoft Ireland case and the cyberspace sovereignty trilemma. Post-territorial technologies and companies question territorial state sovereignty and regulatory state monopolies” (July 2018) by Paul De Hert and Johannes Thumfart (27 pages)

N°12 “Understanding the balancing act behind the legitimate interest of the controller ground: a pragmatic approach” (August 2018) by Irene Kamara and Paul De Hert (35 pages)

N°13 “Big data analytics by telecommunications operators and the draft ePrivacy Regulation” (September 2018) by Vagelis Papakonstantinou and Paul de Hert (13 pages)

N°14 “Enforcement in Indonesia Data Privacy Laws: The 2018 Facebook-Cambridge Analytica scandal as a case study” (October 2018) by Anbar Jayadi (21 pages)

N°16 Big data analytics in electronic communications: A reality in need of granular regulation (even if this includes an interim period of no regulation at all) (June 2019) by Vagelis Papakonstantinou and Paul de Hert (25 pages)

N°17 Data Localisation: Deconstructing myths and suggesting a workable model for the future. The cases of China and the EU (September 2019) by Author: Yanqing Hong, Senior Fellow, Law and Development Institute, Peking University of China, Edited by Vagelis Papakonstantinou, Brussels Privacy Hub (31 pages)

N°18 Challenging algorithmic profiling: The limits of data protection and anti-discrimination in responding to emergent discrimination (January 2020) by Dr Monique Mann and Professor Tobias Matzner (18 pages)

N°19 Access to the Internet in the EU: a policy priority, a fundamental, a human right or a concern for eGovernment? (February 2020) by Lina Jasmontaite and Paul de Hert (23 pages)

N°20 The Proposed ePrivacy Regulation: The Commission’s and the Parliament’s Drafts at a Crossroads? (March 2020) by Elena Gil González, Paul De Hert & Vagelis Papakonstantinou (31 pages)

N°21 Article 8 ECHR compliant and foreseeable surveillance: the ECtHR’s expanded legality requirement copied by the CJEU. A discussion of European surveillance case law (April 2020) by Paul De Hert & Gianclaudio Malgieri (42 pages)

N°22 The “Ethification” of Privacy and Data Protection Law in the European Union. The Case of Artificial Intelligence (May 2020) by Niels van Dijk and Simone Casiraghi (23 pages)

N°23 Logic and Key Points of China’s Cybersecurity Review Measures (June 2020) by Yanqing Hong, Senior Fellow, Law and Development Institute, Peking University of China, Edited by Vagelis Papakonstantinou, Brussels Privacy Hub (9 pages)

N°24 Individuation: re-imagining data privacy laws to protect against digital harms (July 2020) by Anna Johnston (22 pages)


N°26 European Law Enforcement and US Data Companies: A Decade of Cooperation Free from Law (September 2020) by Angela Aguinaldo and Paul De Hert (16 pages)

N°27 Fashion ID and Decisively Influencing Facebook Plugins: A Fair Approach to Single and Joint Controllership (June 2021) by Paul De Hert and Georgios Bouchagiar (24 pages)
The Brussels Privacy Hub Working Papers series

N°28 Adding and removing elements of the proportionality and necessity test to achieve desired outcomes. Breyer and the necessity to end anonymity of cell phone users (September 2021) by Paul De Hert and Georgios Bouchagiar (26 pages)

N°29 Facial recognition, visual and biometric data in the US. Recent, promising developments to regulate intrusive technologies (October 2021) by Paul De Hert and Georgios Bouchagiar (46 pages)

N°30 Necessity knows no law in contaminated times: the rule of law under pandemic police and pandemic legislation’ (“Nood breekt wet in besmette tijden: de rechtsstatelijkheid van de pandemiepolitie en pandemiewetgeving”) (November 2021) by Paul De Hert (33 pages)

N°31 The fundamental right to personal data protection in criminal investigations and proceedings: framing big data policing through the purpose limitation and data minimisation principles of the Directive (EU) 2016/680 (December 2021) by Paul De Hert and Juraj Sajfert (17 pages)

The Brussels Privacy Hub Working Papers series

The Brussels Privacy Hub Working Papers are intended to circulate research in progress for comment and discussion. The Working Papers focus on all areas of data protection and privacy research and can contain empirical research on privacy issues and analytical work on privacy governance and regulation in the EU; global flows of data; reconciling law enforcement and privacy interests; privacy challenges posed by new technologies; comparative research on privacy in different regions; jurisprudential issues of privacy protection; and many others

Available at www.brusselsprivacyhub.org/publications.html

Editorial Board: Paul De Hert and Christopher Kuner

Contact: info@brusselsprivacyhub.eu