EJLT European Journal of Law and Technology

Between GDPR and Law Enforcement Directive in Security Research: The Use of Personal Data by Law Enforcement Authorities

Stergios Aidinlis,[*] David Barnard-Wills, Leanne Cochrane, Krzysztof Garstka, Agata Gurzawska and Joshua Hughes[**]

Abstract:

European law enforcement agencies (LEAs) increasingly seek to make additional use of the personal data they have gathered, particularly for the purpose of research. This raises practical data protection challenges for these agencies and their research partners. LEAs may be uncertain about which data protection instrument – if any – should govern such processing; a question best answered by disentangling the blurry boundary between operational and research activities. This article takes on that task, by examining the applicability of the General Data Protection Regulation (GDPR) and the Law Enforcement Directive (LED) to LEA research activities, in particular those based on LEA-held personal data. It also considers the practical implications of choosing one instrument over the other and gives guidance on legal choices that follow, with respect to issues such as identifying the legal basis for processing, and setting the data controllership arrangements.

Keywords: data protection, law enforcement, security research, controller, EU

[*] Durham University.
[**] All authors are affiliated with Trilateral Research Ltd. The research leading to these results received funding from the European Community's Horizon 2020 research and innovation programme under grant agreements: No 833276 (INSPECTr); No 883543 (CC-DRIVER); No 786687 (COPKIT); No 883297 (DARLENE); No 833115 (PREVISION); No 833635 (ROXANNE); and No 740558 (TITANIUM).

1. Introduction

EU laws regulating the use and re-use of personal data held by law enforcement authorities, for research purposes such as technology development, lack clarity despite attempts to balance the rights of individuals with the demand for innovation. The confusion around the applicable legal regime engenders risks that data protection law (as well as traditional research ethics safeguards) might be applied incorrectly or not at all, and that researchers might opt to work in areas with a more consistent governance framework and avoid engaging in security technology research that could make law enforcement more effective. Technological tools that have emerged out of EU research projects, shared cost-free with law enforcement authorities (LEAs) under licence with Europol via the Europol Tool Repository,[1] are already being used by European LEAs in investigations into organised crime and human trafficking. Resolving the confusion surrounding this legal regime could accelerate further similar research, improving LEA capabilities to investigate criminal activity. Section 2 below describes this LEA research context in more detail, setting the scene for the legal questions that dominate the paper.

The first and core legal question is when does a research activity involving the processing of personal data (such as in LEA case files), and conducted by or for European LEAs, fall within the core law enforcement activities of 'prevention, investigation, detection, prosecution' and is therefore regulated under the Law Enforcement Directive (LED),[2] and when does it fall outside of these terms, and is therefore regulated by the General Data Protection Regulation (GDPR)?[3] The legal regimes under which LEAs process personal data can be easy to delimit in some cases, for example, processing personal data for operational purposes of prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties would come under the LED, whilst processing of personal data for administrative purposes of human resources would come under the GDPR.[4] Yet, there are grey areas where it is not obvious which regime should apply, and 'research' is one of them. It can broadly contribute to 'prevention' of criminal offences (as

[1] Europol, 'Innovation Lab' (2023) <https://www.europol.europa.eu/operations-services-and-innovation/innovation-lab> accessed 14 August 2024.

[2] Directive (EU) 2016/680 of the European Parliament and of the Council on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offenses or the execution of criminal penalties and on the free movement of such data, and repealing Council Framework Decisions 2008/977/JHA [2016] OJ L119/89 (Law Enforcement Directive, hereafter: LED).

[3] Regulation (EU) 2016/679 of the European Parliament and the of the Council, on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and the repeal of Directive 95/46/EC (General Data Protection Regulation) [2016] OJ 119/1 (hereafter: GDPR).

[4] See Information Commissioner's Office, 'Scope and Key Definitions: What if we are processing for other general purposes?' (Information Commissioner's Office, 2021) <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-law-enforcement-processing/scope-and-key-definitions/> accessed 14 August 2024.

captured by the LED), but there are also research-specific provisions in the GDPR (i.e., Article 9(2)(j), Article 89). This gives rise to complex discussions around the meanings of both prevention and research, and the appropriate legal basis for such activity.[5]

Section 3 of this paper answers the question on applicability through two methodological steps. Firstly, it takes a profound look at the interplay between articles and recitals of both legislative instruments. Secondly, it examines diverse national legal frameworks, both implementing EU data protection laws, as well establishing the function of LEAs. A conclusion on the application issue is then proposed. Section 4 follows the identification of a relevant data protection instrument and offers guidance on subsequent practical choices: how to allocate a suitable legal basis for discussed processing; and how to arrange the controller−processor relationship in case of sharing LEA data with external researchers. Taking a broader perspective, section 5 explores the impact of choices made with respect to issues covered in sections 3 and 4. Specifically, the tangible impact on the data subject depending on whether the GDPR or the LED governs the processing of their data; and the impact on the quality and integrity of security research. The article concludes with a set of recommendations on how the law in this area could be clarified.

Three disclaimers must be made before proceeding. Firstly, whilst predictive policing has attracted significant critical attention,[6] the focus here is on the wider purpose of policing research with personal data, especially where LEA case files are involved. Secondly, this paper addresses good-faith research efforts. There are valid concerns regarding operational deployments of new technologies being incorrectly described as 'pilots' or 'research',[7] but the focus here is on the issues raised within appropriately structured research projects. Finally, it should be noted that this paper focuses on civil security and public safety police data as opposed to national security data and

[5] A 2022 review of the functioning of the LED noted the difficulty in delineating the scope of the GDPR and LED but did not provide further clarity in the context of research. See EC COM 2022) 364 final, 'Communication from the Commission to the European Parliament and the Council: First report on application and functioning of the Data Protection Law Enforcement Directive (EU) 2016/680 ('LED')' (European Commission, 2022).

[6] Sarah Brayne, Alex Rosenblat and Danah Boyd, 'Predictive Policing' (2015) *Data & Society* <https://www.datacivilrights.org/pubs/2015-1027/Predictive_Policing.pdf> accessed 14 August 2024; Albert Meijer and Martijn Wessels, 'Predictive Policing: Review of the Benefits and Drawbacks' (2019) 42(12) *International Journal of Public Administration* 1031; Rosamunde van Brakel, 'Rethinking Predictive Policing: Towards a Holistic Framework of Democratic Algorithmic Surveillance' in Marc Schuilenberg and Ronald Peeters (eds), *Algorithmic Societies: Power, Knowledge and Technology in the Age of Algorithms* (Routledge, 2021), 45; Janet Chan, 'The Future of AI in Policing: Exploring the Sociotechnical Imaginaries' in John McDaniel and Ken Pease (eds), *Predictive Policing and Artificial Intelligence* (Routledge, 2021), 89.

[7] See, for example, references to a 'trial licence' during unlawful use of facial recognition algorithms at Brussels airport:, 'Belgian police illegally used facial recognition software' *The Brussels Times* (11 October 2021) <https://www.brusselstimes.com/news/belgium-all-news/188743/belgian-police-illegally-used-facial-recognition-software> accessed 14 August 2024.

research initiatives, to better focus upon areas where EU law applies (in general, the EU Treaties do not allow EU laws to govern national security matters), and because of the secrecy surrounding national security and intelligence data use.[8]

2. Law Enforcement and the Use of Personal Data for Technology Research and Development

Law enforcement is a data-intensive activity, closely linked to practices such as investigation, administration,[9] surveillance and knowledge management.[10] As LEAs go about their usual activities, they produce significant quantities of data, including, but not limited to, collected evidence, results of investigations, intelligence and records of their own activities, some of which will contain personal data. Personal data has the same definition in the LED as the GDPR and refers to:

> [A]ny information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.[11]

Some of the most relevant personal data types used by LEAs in research activities come from closed LEA case files.[12] These files include personal data of various sorts in police charge sheets, interview statements and crime reports, e.g., names and other information relating to suspects, victims, witnesses, including 'special category data' which can pertain to a person's race, sexual orientation, political views of other more sensitive matter. Closed case files also often contain data on criminal offences or convictions; though it should also be noted that closed case files can also contain data that is not personal, such as relevant legislation and policies, drug analysis reports, case progression or decision information, provided none of the above is linked to a natural person. These case files provide an especially rich dataset for technology research such as data-based machine-learning tools for police work; and, like most

[8] It has to be mentioned that LEA research activities falling with the scope of national security activities could fall outside of GDPR, LED and EU law in general. Article 2(3)(a) LED states that the Directive does not apply to 'processing of personal data (…) in the course of an activity which falls outside the scope of Union law'. Following this reasoning, Art. 4(2) Treaty on European Union (TEU) requires the Union to respect the 'essential State functions, including ensuring the territorial integrity of the State, maintaining law and order and safeguarding national security. In particular, national security remains the sole responsibility of each Member State.'

[9] Mark Neocleous, *A Critical Theory of Police Power: The Fabrication of Social Order* (Verso, 2021), 46.

[10] Richard V. Ericson and Kevin D. Haggerty, *Policing the Risk Society* (OUP, 1997).

[11] Art. 3(1) LED; Art. 4(1) GDPR.

[12] The paper excludes the use of data currently held for ongoing investigations. It is important to note that the commonly used term 'cold case' is an ongoing case where investigative leads have dried up, rather than being closed due to conclusion of legal processes.

contemporary institutions, LEAs are under pressure to use this data to increase their efficiency and effectiveness.[13]

Police use of personal data is socially sensitive with real risks linked to the unique role of the police and their coercive capacity. Empirical studies have shown how police organisations investigating predictive policing are influenced by managerialist approaches to organisation and a drive for actionable intelligence.[14] On the one hand, such usage might be mostly direct and operational; on the other, it could be aimed at identifying trends, patterns, correlations in order to improve the long-term effectiveness of the police forces. The latter type of enquiry comes closer to 'research' and the focus of this paper.

The concept of 'research' as defined by the European Federation of Academies of Sciences and Humanities (ALLEA) refers to the 'quest for knowledge obtained through systematic study, thinking, observation, and experimentation'.[15] ALLEA further notes that 'while different disciplines may use different approaches, they each share the motivation to increase our understanding of ourselves and the world in which we live'.[16] The research context that most underpins the analysis of this article is that of European Commission's Horizon 2020 and (successor) Horizon Europe research and development programmes. These programmes support and facilitate research by security actors, including LEAs, by funding research projects in the realms of forensics, cybercrime and security, radicalisation, supply chain security, financial crime, civil–military cooperation, and efforts against external security threats.[17] The programmes do not include, nor are they responsible for, LEA operational activities in the fight against crime and terrorism. Research projects in this space typically take the form of a consortium of different types of research actors, including LEAs, private industry and SMEs, research institutions and universities, collaboratively writing a research proposal in response to a call from the European Commission (EC). However, the issues covered by this article can still arise when LEAs share data gathered during their operational activities, when LEAs have their own technological research capacity and conduct research projects entirely on their own, or when LEAs commission research from third parties. Third parties such as technology companies are themselves often interested in LEA (and other public sector) data where their work involves research artificial intelligence (AI), big data and machine learning due to the importance of large real-world data sets for the development of machine learning

[13] Henrik Schildt, *The Data Imperative: How Digitalization is Reshaping Management, Organizing and Work* (OUP, 2020).
[14] Sarah Egbert and Matthias Leese, *Criminal Futures: Predictive Policing and Everyday Police Work* (Routledge, 2021).
[15] European Code of Conduct for Research Integrity (ALLEA, 2023) <https://allea.org/wp-content/uploads/2023/06/European-Code-of-Conduct-Revised-Edition-2023.pdf> accessed 14 August 2024, 3.
[16] ALLEA European Code of Conduct (n 15), 5.
[17] See European Commission, 'Innovation and Security Research' (Migration and Home Affairs, 2024) <https://home-affairs.ec.europa.eu/policies/internal-security/innovation-and-security-research_en> accessed 14 August 2024.

models. The drive for efficiency, together with an element of technological solutionism,[18] make the AI-driven tools that might emerge from such research attractive to law enforcement and beyond.

To aid the reader's understanding of the research context, it is helpful to conceive of a research project emerging as follows. An LEA wishes to build a predictive algorithm, or perhaps less controversially,[19] to build tools to automatically filter information. The most relevant and useful data it holds includes personal data in closed case files that include: the identity of the offender, victim, and their contacts; content data and metadata of communications (phone calls, SMSs, social media, emails); photos; electronic identification details of device; possession details of device; location details; professional situation and job applications; bank account details; financial transactions; personal documents; forensic artefacts). At an early stage, the LEA does not know if the predictive algorithm (or the automatic filtering) is technically feasible, or if such an approach can be successful, but it wishes to explore the possibility and joins a consortium with researchers to investigate. The consortium put together a proposal for funding in response to a relevant call, which wins subject to a contract with the EC (a 'Grant Agreement') and an agreement between the consortium members (the 'Consortium Agreement'). The proposal and later the Grant Agreement emphasises the consortium partner's obligations to comply with all relevant human rights and data protection legislation, and best practices in research ethics. The consortium must determine the applicable legislation and the applicable legal basis therein for processing the LEA-held personal data, and the correct data-sharing arrangements should other (non-LEA) partners − such as technology developers − want access to the data for this shared work. That need for a legal basis turns on the question of applicability of the LED and the GDPR to the described context.

3. Applicability of GDPR and LED to Research with LEA Case Files

3.1 Does Research with LEAs Case Files Fall within Article 1(1) LED?

This section of the paper analyses the legal provisions laid out by the LED and GDPR, in order to find out which instrument was intended by the European lawmaker to cover LEA research activities with personal data. Specifically, the LEA-held personal data processed in the described scenario, though many of our findings apply beyond it.

For LEAs to rely on the LED, personal data must be processed according to Article 1(1) for the purposes of the 'prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, including the safeguarding against and the prevention of threats to public security'.[20] Since LEAs may intend to use the results of their research, such as the technologies developed, in the longer

[18] Evgeny Morozov, *To Save Everything Click Here: The Folly of Technological Solutionism* (Public Affairs, 2014).
[19] Brayne, Rosenblat and Boyd (n 6).
[20] LED, Art. 1(1).

term for these Article 1(1) purposes, they can at times be keen to immediately rely on the LED rather than the GDPR as the applicable legal framework for processing personal data in research projects.

This paper argues that processing personal data for research is too conceptually different to be classed as 'investigating, detecting, or prosecuting crimes', or 'safeguarding public security'. It could, however, reasonably be argued to come within the meaning of crime 'prevention', a line that could be followed by LEAs relying on the LED as their legal basis.

Police-commissioned or -engaged research is rarely 'disinterested' pure science, but rather it is conducted to support organisational objectives. That connection to the organisation's mandate and strategic vision, however, can be attenuated, with a long and breakable chain between the research goal and the impact on crime prevention. The legislative framework does not provide clarity with regards to the transition points from processing of personal data for research purposes − traditionally governed by the GDPR − to processing for the purposes of crime prevention − where the LED applies (or for national security purposes, where EU law does not apply[21]).

The meaning of 'prevention' itself is not defined in the LED, and most of its use in that instrument is in terms of preventing 'criminal offences' or 'threats to public safety' in the sense of stopping offences or threats from being realised. It could be argued that this vagueness means that any processing of personal data intended to stop crime or threats to public safety generally could be 'prevention' and therefore within the meaning of the LED; such a reading would seem to include all research activities by LEAs intended to stop reoccurring or regular crimes, (as well as other threats to safety) from manifesting.

Some of the recitals to the LED provide a (persuasive, non-binding) interpretation of 'prevention' that includes a wide range of measures intended to stop crimes and threats to public safety occurring. For example, Recital 12 refers to LEAs engaging in activities such as policing 'demonstrations, major sporting events and riots' and 'maintaining law and order… to safeguard… fundamental interests of the society'. In terms of a wide understanding of 'prevention' that could include research, Recital 27 is the most pertinent and provides that it can be:

> [N]ecessary for competent authorities to process personal data collected in the context of the prevention, investigation, detection or prosecution of specific criminal offences beyond that context in order to develop an understanding of criminal activities and to make links between different criminal offences detected.

Arguably, then, processing personal data about 'specific criminal offences' beyond the immediate context could be read to include research where it develops 'an

[21] Treaty on European Union (Maastricht Treaty) [1992] OJ C191/1, Art. 4 – 'In particular, national security remains the sole responsibility of each Member State.'

understanding of criminal activities' and makes 'links between different criminal offences detected'.

As regards 'research' specifically, the LED mentions processing for 'archiving in the public interest, scientific, statistical or historical' purposes three times. First, Article 4(3) provides that, subject to appropriate safeguards, such processing (paraphrased further as 'research') can be carried out for the purposes under Article 1(1), thus covering, for example, prevention of criminal offences. Second, Article 9(2) states that, where the public mandate of LEAs goes beyond law enforcement purposes (again, as set out under Article 1(1)), then activities that are not for law enforcement purposes, including research, should be regulated under the GDPR.[22] Furthermore, Articles 5(1)(b), 9(2)(j) and 89 GDPR all explicitly refer to the use of personal data for research purposes, and so it can be assumed that the drafters of the GDPR intended it to be the primary legal regime for research activities. Third, Recital 26 suggests that Member States should lay down domestic law to provide safeguards for research purposes. Consequently, the requirements under Article 4(3) regarding safeguards provided in Member State law (supported persuasively by Recital 26) indicate that the processing of personal data for research purposes under the LED should only take place where Member State law provides safeguards for the rights and freedoms of data subjects in LEA research.

Such safeguards can come in variety of ways. They can include general technological and organisation measures (such as security of storage, data accuracy verification, anonymisation/pseudonymisation), but they could also be specific ethical safeguards tied to research with personal data. The challenge here is that while there is ample guidance on law enforcement ethics in general (such as the European Code of Police Ethics (2001)[23]), there is an arguable scarcity of guidance on LEA research. While the Code covers key concepts within the police, such as 'loyalty', 'consent', 'impartiality', 'discretion' and 'professionalism', it does not provide any further guidance on research ethics in the law enforcement context. National codes of ethics for law enforcement at the national level take a similar approach (e.g., Belgium,[24] Bulgaria,[25]

[22] ICO, 'Legitimate Interest', Guide to the General Data Protection Regulation (GDPR), <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/lawful-basis-for-processing/legitimate-interests/> accessed 14 August 2024.

[23] Council of Europe, Committee of Ministers, Recommendation Rec(2001)10 of the Committee of Ministers to member states on the European Code of Police Ethics (Council of Europe, 2001).

[24] Service Public Fédéral, 'Code de déontologie des services de police' (BE, author translation: Federal Public Service Justice, 'Code of Ethics of the Police Services') (2006) 2006A00301, 2006-05-10/33, <http://www.ejustice.just.fgov.be/cgi_loi/change_lg.pl?language=nl&la=N&cn=2006051033&table_name=wet> accessed 14 August 2024.

[25] Министерството На Вътрешните Работи, 'Етичен Кодекс За Поведение На Държавните Служители' (BG, author translation: Ministry of Interior, 'Code of Ethics for Civil Servants' (2014), Order No. 8121h-348 of 25 July 2014, <https://www.mvr.bg/docs/default-source/structura/96de0a6d-etichen_kodeks-pdf.pdf> accessed 14 August 2024.

Latvia[26]). However, there are some emergent examples of specific guidance for LEAs on particular research areas, most notably facial recognition.[27] What remains unclear for Article 4(3) purposes is who should verify the existence and quality of such safeguards. There is certainly a role for research funders to require and periodically review such safeguards – something, e.g., the European Commission embraces on Horizon programmes.[28] Beyond research funders, one could imagine data protection authorities being capable of acting to ascertain the safeguards' adequacy – though this is unlikely to happen in a preventive, proactive manner.

Taking all of this into account, it does seem possible for research with LEA case files to be conducted under the LED, on a broad reading of 'prevention' and provided that adequate safeguards (supported by national laws) are in place. Such research would be falling within the LED most convincingly when used in the context of linking different criminal offences. Where the purpose of processing is too vague or distant from the notion of prevention, or where suitable safeguards are not in place, the processing/research in question should be seen as falling within the GDPR.

However, what is missing from this equation is the issue of choice – made both by the national legislators implementing the Directive with corresponding acts of legislation, and by data controllers deciding on where their research goals and safeguards stand. The remaining sections of the article seek to cover and inform both perspectives.

3.2 Applicability of National Frameworks – Leaning towards LED?

The above sections covered the possibility that both the LED and the GDPR could apply as the legal basis for processing LEA case data in research contexts. The GDPR

---

[26] Iekšlietu Ministrijas, 'Ētikas Kodekss' (LV, author translation: Ministry of Interior, 'Code of Ethics') (2020), <http://www.iem.gov.lv/lat/dokumenti/etikas_kodekss/> accessed 14 August 2024; Latvijas Valsts Policija, 'Valsts Policija Ētikas Kodekss' (LV, author translation: State Police Republic of Latvia, 'Code of Ethics of the State Police') (2020), <https://www.vp.gov.lv/en/ethics> accessed 14 August 2024.

[27] In May 2023, the EDPB adopted the final version of guidelines on the use of facial recognition technology in law enforcement, which include in Annex II 'Practical Guidance for Managing FRT [Facial Recognition Technology] Projects in LEAs', directing LEA researchers on roles and responsibilities, safeguards at inception, procurement/development and deployment stages. European Data Protection Board, 'Guidelines 05/2022 on the use of facial recognition technology in the area of law enforcement, Version 2.0' (2023), 33, <https://www.edpb.europa.eu/system/files/2023-05/edpb_guidelines_202304_frtlawenforcement_v2_en.pdf> accessed 14 August 2024. Non-EU examples include reports by the UK Home Office's Biometrics and Forensics Ethics Group. Nina Hallowell et al, 'Ethical issues arising from the police use of live facial recognition technology' (2019), Biometrics and Forensics Ethics Group Facial Recognition Working Group <https://assets.publishing.service.gov.uk/media/5c755ffc40f0b603d660be32/Facial_Recognition_Briefing_BFEG_February_2019.pdf > accessed 14 August 2024.

[28] See European Commission, 'How to complete your ethics self-assessment' (2021), EU Grants <https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/common/guidance/how-to-complete-your-ethics-self-assessment_en.pdf> accessed 14 August 2024.

should be broadly understood as the default regime for research activities; however, research activities conducted for Article 1(1) LED purposes, most realistically for crime 'prevention', *can* fall within the scope of the LED. What was not discussed was the need for the LED choice of the legal basis to be supported by the national legal framework implementing the LED due to the LED being an EU Directive and not a Regulation. Of this, there are several examples.

There is a relatively high degree of variability among Member States' laws which cover LEA processing of personal data in the context of research. This is true of the domestic data protection legislation both elaborating on the GDPR and transposing the LED, and of the domestic statutes which set out LEA functions (note that in some cases, these combined issues are dealt with in the one legislative instrument). Some statutes establishing LEAs specifically identify research as a police function. In Romania, for example, Article 26(1)(18) of Law No. 218/2002 on the organisation and functioning of the Romanian Police specifically identifies that the police have the function to:

> carry out, independently or in cooperation, scientific assessments, studies and research to improve the methods and means used in police work, in particular forensic technical and scientific, informational analysis, prevent and combat crime or other illegal acts, and to identify new methods and means.[29]

With the above as an explicit LEA public function, it is not surprising that the Romanian law implementing the LED, in Articles 1(1) and 5(2)–(3), also explicitly permits the further processing of LED data for research purposes, provided the research objective is also for an Article 1(1) LED purpose.[30] In this way, the research purpose is compatible with the law enforcement purpose and is presented as though a subset of it, along the lines potentially foreseen by Recital 27 and Article 4(3) LED, as discussed above.

Ireland takes a similar approach to the Romanian law, albeit without the explicit LEA 'research' function.[31] In Ireland, Part 5 Data Protection Act 2018 (DPA 2018), transposing the LED, also specifically conceives of the possibility that LEAs can process

---

[29] Legea nr. 218/2002 privind organizarea și funcționarea Poliției Române (RO, author translation: Law No. 218/2002 on the organisation and functioning of the Romanian Police) <http://legislatie.just.ro/Public/DetaliiDocument/157719> accessed 14 August 2024.

[30] Lege nr. 363/2018 privind protecția persoanelor fizice referitor la prelucrarea datelor cu caracter personal de către autoritățile competente în scopul prevenirii, descoperirii, cercetării, urmăririi penale și combaterii infracțiunilor sau al executării pedepselor, măsurilor educative și de siguranță, precum și privind libera circulație a acestor date (RO, author translation: Law No. 363/2018 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of preventing, detecting, investigating, prosecuting and combating criminal offences or the execution of penalties, educational and security measures, as well as on the free movement of such data) <http://legislatie.just.ro/Public/DetaliiDocument/209627> accessed 14 August 2024.

[31] See the seven explicit functions of the Irish police in s. 7 Garda Síochána Act 2005 (IE), <https://www.irishstatutebook.ie/eli/2005/act/20/enacted/en/html> accessed 14 August 2024.

personal data for scientific research which is for a purpose identified in Article 1(1) LED. Specifically, section 71(6) DPA 2018 provides that:

> A controller may process personal data, whether the data were collected by the controller or another controller, for—
>
> (…)
>
> (b) scientific or historical research purposes (…)
>
> provided that the said processing—
>
> (i) is for a purpose specified in section 70(1)(a), and
>
> (ii) is subject to appropriate safeguards for the rights and freedoms of data subjects.

The purposes described in section 70(1)(a) DPA 2018 are:

> (i) the prevention, investigation, detection or prosecution of criminal offences, including the safeguarding against, and the prevention of, threats to public security, or
>
> (ii) the execution of criminal penalties.

Part 5 of the legislation proceeds to extend these provisions to the processing of special categories of personal data (section 73(1)(b)(viii)).

As mentioned, Irish LEAs do not have a specific 'research' function, contrary to the earlier Romanian example, but Part 5 DPA 2018 clearly conceives that LEAs may engage in research activities for an Article 1(1) LED purpose. That broader law enforcement purpose, in domestic law, is most logically that of 'crime prevention', one of seven functions set out in section 7 Garda Síochána Act. Yet, following on from earlier discussions, it is important to recognise that Part 5 DPA 2018 represents an example of an EU Member State adopting a broad approach to 'prevention' when transposing the LED. In such scenarios, should LEAs say they are doing research for crime prevention, they notably might *not distinguish* between any LED meaning of 'prevention' and the scope of their own 'crime prevention' function. Instead, LEAs might identify that the Article 1(1) scope of the LED, i.e., its application to the processing of personal data for the purposes of 'prevention, investigation, detection or prosecution of criminal offences', are not cumulative activities. In this interpretation, prevention activities do not in fact have to be linked with an investigation, detection or prosecution activity, and so on. This approach undermines any argument that the prevention activities within the LED must be linked with a specific criminal offence.

In other contexts, other LEA-specific legislation (which is not the national implementation of the LED) exists, which also assumes the LEAs may engage in research activity. For example, in Belgium, Article 44/11/10 of the federal Loi sur la fonction de police (Law on the Functioning of the Police) 1992 restricts the processing

of LEA personal data for research purposes to cases that have been explicitly authorised by an 'appropriate authority'.[32] This approach does not follow a clear GDPR or LED framework, but constitutes the main national law covering LEA engagement in research activities. As such, upon receiving permission from the appropriate authorities, all stipulations by that authority together with GDPR safeguards should be assumed to apply.

Yet again, in other national contexts, LEAs engage with research activities under the personal data provisions which apply to persons generally (rather than to the LEA specifically) and so more clearly pursue a GDPR approach.[33] In Estonia, for example, the LEA may process personal data without the consent of the data subject for research purposes provided it is pseudonymised (although exceptions do exist).[34] Where this data includes special category data, the LEA must first attain approval from the relevant ethics committee, or in the absence of a relevant ethics committee, the national DPA.[35] Section 6(1) of the Estonian Isikuandmete kaitse seadus (Personal Data Protection Act 2018) sets out the generally applicable provision:

> (1) Personal data may be processed without the consent of the data subject for the needs of scientific and historical research and official statistic, in particular in a pseudonymised format or a format which provides equivalent level of protection. Prior to transmission of personal data for processing for the needs of scientific and historical research or official statistics, personal data shall be replaced by pseudonymised data or data in a format which provides equivalent level of data protection.

This small set of examples demonstrates the variety of approaches in national law regulating processing of LEA case data for research. Scientific research is not often an explicit function of LEAs under establishing statutes. Only the Romanian example is provided above; but the lack of an explicit research function, does not necessarily prevent LEAs from asserting that they can process LEA case data for research projects on the grounds of the wider public interest function that they carry, and thereby engaging Article 6(1)(e) GDPR (i.e., 'the performance of a task carried out in the public interest or the exercise of official authority vested in the controller'). Indeed, interpreting Irish law, this appears to be a secondary, albeit less preferred route, available to the LEA.[36]

---

[32] Loi sur la fonction de police (BE, author translation: Law on the Functioning of the Police) (1992), 1992000606, 1992-08-05/52), <http://www.ejustice.just.fgov.be/cgi_loi/change_lg.pl?language=fr&la=F&cn=1992080552&table_name=loi> accessed 14 August 2024.

[33] Isikuandmete kaitse seadus (EE, unofficial translation: Personal Data Protection Act) (2018), § 6, <https://www.riigiteataja.ee/en/eli/523012019001/consolide#para16 > accessed on 14 August 2024.

[34] Ibid § 6(1).

[35] Ibid § 6(4).

[36] See especially, Data Protection Act 2018, Pt. 3, s. 55 (IE).

While there is undoubtedly a complex network of national laws for LEAs and legal teams to work through concerning the use of LEA-controlled personal data for research, the array of approaches outlined above suggests that the confusion over the applicability of the GDPR and LED in these contexts has extended into the national legal frameworks and their respective interpretations. While EU Directives deliberately leave discretion to Member States on how to implement the goal – in this case, the safeguarding of natural persons when processing LEA-held personal data for Article 1(1) LED purposes – it may be that this flexible implementation mechanism suggests to Member States an overextended scope when juxtaposed with the inflexible GDPR mechanism which regulates an overlapping space. That is not to suggest that some of the more broadly construed national laws are incompatible with the LED, but rather to call for greater delineation at the EU level between the legislative boundaries.

4. Two Key Aspects of Data Protection-Compliant Research with LEA-Held Personal Data

Research with LEA-held personal data is a complex endeavour in terms of achieving data protection compliance. Activities following the frameworks of both GDPR and LED have to address multiple questions; two of the most pertinent ones are which legal basis to rely on, and how to arrange the controller–processor relationships. The following subsections explore both questions, indicating differences between GDPR and LED where they materialise.

4.1 Which Legal Basis for the LEA?

It is important to note that the use of LEA-collected personal data in research projects typically assumes that the data subjects will not have consented to their data being used in this way, and that processing these data will be repurposing (or secondary processing), rather than initial processing. Within the GDPR, researchers (i.e., technology developers, data scientists, police officers, academics, etc.) could attempt to gain consent from data subjects and process their personal data on the legal basis of consent.[37] However, it could be problematic to gain that consent, and the resulting variation runs the risk of creating high systemic bias.[38] The very appeal of LEA-collected personal data for researchers is that it already exists and holds some purported veracity (even if in practice there are significant distortions due to differences in the historical conduct of policing[39]). Within the LED, the situation is somewhat simpler; as described in section 5.1 below, it does not recognise consent as a legal basis for processing.

[37] GDPR (n 3), Art. 6(1)(a).

[38] This has been the typical research pathway for qualitative criminologists trying to understand, for example, offender pathways with the assistance and participation of former offenders. But this method essentially creates *new* data (often anonymised) with the researcher, not the LEA as the data controller.

[39] Ruha Benjamin, *Race After Technology: Abolitionist Tools for the New Jim Code* (Polity, 2019).

Looking at the GDPR, where repurposing is compatible with the original purpose for which the data were gathered, Recital 50 GDPR persuasively states that 'no legal basis separate from that which allowed the collection of the personal data is required'.[40] Read together with Article 5(1)(b), compatible repurposing can use the same legal basis as the original purpose for which the data were originally processed. Where repurposing is incompatible with the original purpose, a new legal basis is required.

Repurposing for research is regulated in both a specific and a general way. Article 5(1) GDPR states:

> 1. Personal data shall be:
>
> …
>
> (b) collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes; further processing for archiving purposes in the public interest, *scientific or historical research purposes* or statistical purposes shall, in accordance with Article 89(1), *not be considered to be incompatible* with the initial purposes ('purpose limitation').[41]

Therefore, repurposing personal data for research is *prima facie* assumed to be compatible with any original purpose. The general regime for regulating repurposing is found in Article 6(4) GDPR. The European Data Protection Supervisor (EDPS) is clear that even where repurposing for research purposes is *permitted*, the test from the more general regime should still be applied.[42] This requires data controllers to consider: links between the original purpose and the repurposing; the context of data collection; the nature of the personal data; possible consequences of the proposed repurposing; and the existence of safeguards.[43]

Depending upon the circumstances of the research, it could be difficult to demonstrate compatibility. For example, the repurposing for research of covert surveillance data where the data subject not only does not consent but does not know of the existence of the data would likely be difficult to justify. Further, because the provision in Recital 50 that compatible repurposing does not require a separate legal basis (see text to footnote 40 above) is in a recital, rather than an article in the GDPR, and the tests in Articles 5(1) and 6(4) are about the compatibility of purposes and not

[40] GDPR (n 3), Recital 50.
[41] Emphasis added.
[42] European Data Protection Supervisor, 'A Preliminary Opinion on Data Protection and Scientific Research' (2020), 23, <https://edps.europa.eu/sites/edp/files/publication/20-01-06_opinion_research_en.pdf> accessed on 14 August 2024.
[43] GDPR (n 3), Art. 6(4).

the compatibility of a legal basis, this potentially poses issues in demonstrating that LEA case data can be repurposed on the same legal basis.[44]

Consequently, it could be preferable, and more legally secure, for data controllers carrying out research on LEA-controlled personal data to determine a legal basis separate from the original legal basis used for the original purpose.[45] Assuming that consent is not feasible, an LEA could rely on the public interest legal basis if research is within its public function.[46] Or, it could consider legitimate interest as a legal basis where research is not part of their public function.[47]

If the personal data is special category data, the scientific research condition would likely apply under Article 9(2)(j) ('research' exemption) assuming data minimisation techniques required under Article 89(1) are in place and use of the exemption is based on Union or Member State law. It is arguable that research could constitute a 'substantial public interest' under Article 9(2)(g); however, the EDPS views this as 'difficult, if not impossible' unless it is specifically allowed in national laws of Member States.[48] If the data relate to criminal convictions/offences, then processing of these data can take place 'only under the control of official authority or when the processing is authorised by Union or Member State law'.[49] Thus, an LEA researcher could potentially process such data on the condition of the official authority of the LEA; or, if this is not possible, researchers would need to have an applicable condition in EU or Member State law.

In LED, the regulation of repurposing is structured differently. The purpose limitation principle (Article 4(1)(b)) does not mention repurposing for scientific reasons, and the article on lawfulness of processing (Article 8), does not set out a general test for repurposing. Consequently, the most relevant provision here is the previously mentioned Article 4(3), which opens the door for scientific research to be seen as falling within the law enforcement purposes of the Directive's Article 1(1). Then again, it would be advisable to distinguish between different types of research and types of personal data, by assessing whether the research activity using data collected for an earlier Article 1(1) purpose (such as data gathered during an investigation) falls within Article 1(1). Automatic assumption of purpose compatibility should be avoided, as it prevents reflection on whether processing actually has to fall within the LED.

---

[44] Jessica Bell et al, 'Balancing data subjects' rights and public interest research: Examining the interplay between UK law, EU human rights law and the GDPR' (2019) 5 *European Data Protection Law Review* 43.
[45] Ibid, pp. 47–48.
[46] GDPR (n 3), Art. 6(1)(e).
[47] GDPR (n 3), Art. 6(1)(f); Article 29 Working Party, 'Opinion 06/2014 – Overview of results of public consultation on Opinion on legitimate interests of the data controller' (2014) <https://ec.europa.eu/justice/article-29/press-material/public-consultation/notion-legitimate-interests/files/20141126_overview_relating_to_consultation_on_opinion_legitimate_interest_.pdf> accessed 14 August 2024.
[48] EDPS (n 42), 23.
[49] GDPR (n 3), Art. 10.

4.2 Sharing LEA-Controlled Personal Data with Research Partners

A further question is whether LEAs can share the personal data they control with other non-LEA and LEA researchers during research projects. For a hypothetical example, let us consider passing a set of closed case files to a technology developer, to train a machine learning algorithm to identify crime patterns. The Court of Justice of the European Union (CJEU) case law has dedicated attention to situations where data is flowing *to* the LEAs *from* other controllers (such as electronic communications providers),[50] but the reverse situation is increasingly relevant to crime prevention research and development. This section focuses on the legality of LEAs sharing their data with non-LEA researchers, while noting the possibility that LEA-to-LEA sharing could occur. Researchers and technology developers often want direct access to data, as it is easier (or at least more familiar) to work with it directly. While mock, synthetic or pseudonymised data should be used where possible, this article's interest is on situations where this is not the case.

Assuming the LEA has a legal basis to process the data for research purposes, two main issues arise. The first issue is the appropriate data-sharing arrangement between the various research partners, i.e., a joint controllership or a controller−processor relationship; these concepts − existing in both GDPR and LED − are of fundamental importance when it comes to allocation (and execution) of obligations to entities processing personal data. The second is whether the data recipient requires a separate legal basis. While the order of these issues might at first appear counterintuitive, it is the controller−processor arrangement which determines the need for the second. The appropriate data-sharing arrangement will be determined by the specific context of the research in question. There are, however, certain 'red lines' in relation to personal data which restrict the partners' discretion.

The definition of a controller is essentially the same in the GDPR and LED, differing only in the legal personhood (data controllers are required to be a 'competent authority' under the LED),[51] whilst the definition of a processor is identical.[52] The important consideration for controllership is 'who' should rightfully define the purpose of the data processing. In the context of this article, this is formally the LEA. However, in many security research projects, the goal of the project is typically formulated as one aimed at aiding the LEA end user to combat crime. Such projects can be LEA-led or comprise mixed consortia. Due to the difference in expertise and

[50] Joined Cases C-293/12 and C-594/12, *Digital Rights Ireland Ltd v Minister for Communications, Marine and Natural Resources* (2014) ECLI:EU:C:2014:238; Joined Cases C-203/15 *Tele2 Sverige AB v Post-och telestyrelsen* and C-698/15 *Secretary of State for the Home Department v Watson* (2016) ECLI:EU:C:2016:970; Case C-207/16 *Ministerio Fiscal* (2018) [2018] ECLI:EU:C:2018:788; Case C-311/18 *Data Protection Commissioner v Facebook Ireland Ltd and Maximilian Schrems* (2020) ECLI:EU:C:2020:559.

[51] GDPR (n 3), Art. 4(7) and LED (n 2), Art. 3(8).

[52] GDPR (n 3), Art. 4(8) and LED (n 2), Art. 3(9).

resource availability, it may be the case that the project is often not conceptualised by the LEAs, but instead drafted by researchers, even if guided by LEAs' needs.

The required constellation is most clearly defined when the LEA sharing the data is processing it under the LED (and the respective national implementing law). As mentioned, under the LED, only 'competent authorities' can fulfil the role of controller or joint controller. A competent authority under the LED can also include 'any other body or entity entrusted by Member State law to exercise public authority and public powers' for the purposes set out in Article 1(1) LED.[53] According to Purtova, this second concept was intended to provide for potential future privatisation of LEA functions, including through public–private partnerships (PPPs).[54] While this can in theory include non-LEA entities such as universities and private companies, the emphasis is that such bodies would be entrusted to exercise a public authority function (such as prevention of crime or ensuring public security) via Member State law.[55] Absent this, non-LEA partners accessing LEA personal data processed under the LED should do so as processors, since the LED does not envisage joint controllership except with other competent authorities. In this context, a contract or other legal act is required which defines the processing, necessary safeguards, and obligations and rights of the controller as per Article 22(3) LED.[56] This limitation would, of course, not apply to the context where LEAs share data with other LEA partners in the research.

Where processing by the LEA occurs under the GDPR, there is slightly more scope for LEA case data to be shared via a joint controllership or controller-processor relationship. Yet even in this context, this paper argues that the more legally secure (and less demanding for data-sharing partners) approach is still the controller–processor constellation, for two reasons. The first reason has parallels with the LED focus on competent authorities. Under the GDPR, data-sharing must comply with Article 10, which states that criminal convictions and offences data (broadly defined) can be carried out '*only* under the control of official authority' or when otherwise authorised by Union or Member State law.[57] The 'only' here suggests that a joint controllership with other entities, such as private enterprises, may not be possible without a separate statutory basis. The second GDPR-based argument in favour of a controller–processor constellation pertains to whether the recipient research partner needs a separate legal basis. The issue is that, assuming a clear legal

[53] LED (n 2), Art. 3(7)(b).
[54] Nadezhda Purtova, 'Between the GDPR and the Police Directive: Navigating Through the Maze of Information Sharing in Public–Private Partnerships' (2018) 8(1) *International Data Protection Law*, 62.
[55] Purtova (n 54), 62 and 65.
[56] LED (n 2), Recital 11 and Art. 22(3). The LED provides that the controller-processor relationship could also be set out in another legal act. See also in the context of the GDPR, European Data Protection Board, 'Guidelines 07/2020 on the concepts of controller and processor in the GDPR Version 1.0' (2020) <https://edpb.europa.eu/sites/edpb/files/consultation/edpb_guidelines_202007_controllerprocessor_en.pdf> accessed 14 August 2024, paras 98–112.
[57] Emphasis added.

basis for LEA repurposing of personal data for research, can this legal basis be extended to a recipient partner? The short answer is that such an extension can only apply in the context of a controller−processor constellation. According to the European Data Protection Board (EDPB) Guidelines 07/2020 on the concepts of controller and processor in the GDPR, 'each joint controller has the duty to ensure that they have a legal basis for the processing'.[58] Whereas in the context of a controller−processor relationship:

> The lawfulness of the processing according to [the GDPR] will be derived from *the controller's activity* and the processor must not process the data otherwise than according to the controller's instructions.[59]

For a valid joint controllership, the recipient research partners must themselves have a legal basis for processing the LEA-held data. This is not impossible, as Union or Member State law could authorise the use of criminal convictions and offences data by other entities, and research partners could rely either on the public interest, in case of universities, or legitimate interests, in case of private researchers. Arguably, however, this would be a demanding exercise, as legitimate interests especially require a thorough balancing of commercial interests with data subject rights under Article 6(1)(f) GDPR and the outcome of the balancing would be heavily context-specific.[60] A controller−processor constellation is therefore preferable given that non-LEA research partners would struggle to identify an independent legal basis, within the GDPR or within Member State law, on which they could process real LEA case files. This is taking into account Articles 10 and 9(2)(j) GDPR, which suggest that the non-LEA entity would need a clear domestic legal basis for processing this form of special category data; as well as the Article 6(1)(f) GDPR difficulty in justifying the legitimate interests of the researcher above that of the data subject. This does not apply in a controller−processor context where the further processing by the (non-LEA/technical) processor would instead be covered by a contract (Article 28(3) GDPR). The consortium agreements commonly used in research projects do not typically go into sufficient detail about data processing to be a controller−processor contract themselves.[61]

Some projects will be tempted by joint controllerships. While the nature of the research will vary, where there is a group of public and private researchers working together as equal partners on a funded project (such as often occurs in e.g., the EU Horizon programmes context) the legal agreement between partners (i.e., the

[58] EDPB Guidelines 07/2020 (n 56), 4 (see also para. 164).
[59] EDPB Guidelines 07/2020 (n 56), para. 78, emphasis added.
[60] Stergios Aidinlis et al, 'Lawful Grounds to Share Justice Data for LawTech Innovation in the UK' (2024) 140 *Law Quarterly Review* 544−569.
[61] See European Commission, 'DESCA Model Consortium Agreement' (2024) <https://www.desca-agreement.eu/desca-model-consortium-agreement/> accessed 14 August 2024.

consortium agreement) is suggestive of a joint controllership constellation.[62] However, whether or not partners jointly determine the purposes and means of processing is to be established through a factual test, resulting in a focus on the substance of the partnership in practice. Where a funding body, such as the EC, funds partners jointly, the arrangement in practice does not at this level often appear to be one of vertical control. Instead, the emphasis is on the determination of the purpose and means of processing in practice. It is often the case (but not exclusively so) that LEA partners join research projects as 'end users' – essentially stakeholders and potential future customers/recipients of a technology, whose needs guide the development of that technology. However, they can ultimately say yes or no to the data processing proposed, even if the idea, or the design of the processing, comes from elsewhere. The risk here is that LEAs might not have sufficient technical expertise and resources to assess what is being proposed to them, including credibility, safeguards and its eventual link to prevention of crime. Such situations, where there is a potential information asymmetry between controller and processor, might be increasingly common as technologies such as machine learning are widely deployed.

5. The Importance of the Distinction between the GDPR and the LED

Having discussed the immediately visible questions of applicability, national interpretations, legal basis and controllership, we can now move towards the perspectives of data subjects and the affected research environment – both of which public bodies, regulators and hopefully data controllers in general should consider. The subsequent sections describe five key differences between GDPR and LED which make the distinction matter for data subjects[63] and identify four key categories of risks stemming from the current lack of clarity on the legal nature of processing personal data for research carried out by LEAs.

5.1 Key Differences for Data Subjects

The question in this section is whether the data subjects' position is different when covered by one of the two contesting data protection frameworks. As a preliminary point, it is worth reiterating that the GDPR is a directly applicable EU regulation, while the LED requires national implementation. In practice, a regulation generally leads to a greater degree of harmonisation because it accords less discretion to Member States.[64] Hence, Member States' interpretations of the LED may vary to a greater

[62] See European Commission, 'DESCA Model Consortium Agreement for Horizon Europe' (2024) <https://www.desca-agreement.eu/assets/helmholtz_gemeinschaft/user_upload/Brussels_Office/DESCA/20240206_DESCA_HorizonEurope_v.2.0_with_elucidations.pdf> accessed 14 August 2024, Section 4.4, 16.

[63] This paper uses 'research data processing' here as a shorthand for 'data processing carried out by a LEA as part of a research activity', with prejudice to which law applies.

[64] Christopher Kuner, Lee Bygrave and Christopher Docksey, 'Background and Evolution of the EU General Data Protection Regulation (GDPR)', in Christopher Kuner et al (eds), *The EU General Data Protection Regulation (GDPR): A Commentary* (OUP, 2020) 11.

extent than in the context of GDPR (as evidenced in section 3.2 above), and this may affect the position of data subjects.

Beyond this observation, there are significant, visible differences in the degree and nature of protection granted to data subjects by the two legislative instruments, particularly regarding transparency, consent, data minimisation, special category data and automated decision-making. In creating a legal framework for data protection comprised of the GDPR and the LED, EU lawmakers recognised the distinct social role of law enforcement authorities and the different requirements they have around the processing of personal data, as well as national sensitivities around law enforcement. The purpose of the LED was to 'protect the right of individuals to the protection of their personal data while guaranteeing a high level of public security'.[65] However, the legal protections for the data subject are lesser under this legal instrument than under the GDPR.

Firstly, processing of personal data under the LED is not supported by the principle of transparency; while Article 5(1)(a) GDPR states that personal data should be 'processed lawfully, fairly and *in a transparent manner in relation to the data subject*',[66] Article 4(1)(a) LED limits itself to the statement that such data should be 'processed lawfully and fairly'. Secondly, the data subject's consent as a legal basis for processing is a powerful tool in the GDPR (Article 6(1)(a)) for exercising control over data. In contrast, the LED does not recognise 'consent' as a legal basis for processing personal data; the processing here is allowed to the extent it is necessary for competent authorities to fulfil the purposes set out in Article 1(1) LED and under law (Article 8(1) LED). Thirdly, the principle of data minimisation is softened in the LED. While Article 5(1)(c) GDPR provides that personal data shall be 'adequate, relevant and limited to *what is necessary* in relation to the purposes for which they are processed' (emphasis added),[67] its sister Article 4(1)(c) LED relies on the less restrictive legal expression 'adequate, relevant and *not excessive* in relation to the purposes for which they are processed'.[68]

Fourthly, in Article 9 GDPR, the processing of special category data is forbidden, unless one of ten purpose-driven exceptions applies. In seven of those, the processing must be 'necessary'. However, in Article 10 LED, processing of special categories of personal data is *a priori* allowed, but only where 'strictly necessary', and only in three further circumstances: 'where authorised by Union or Member State law' (Article 10(a)); 'to protect the vital interests of the data subject or of another natural person' (Article 10(b)); or 'where such processing relates to data which are manifestly made public by the data subject' (Article 10(c)). Article 10(a) LED is construed very openly; on a literal reading of this provision, it seems that the EU or Member State law

[65] See European Council, 'Data protection in law enforcement' (2024) <https://www.consilium.europa.eu/en/policies/data-protection/data-protection-law-enforcement/> accessed 14 August 2024.
[66] Emphasis added.
[67] Emphasis added.
[68] Emphasis added.

warranting the processing of special categories data can have any purpose whatsoever, as long as it is strictly necessary and accompanied by appropriate safeguards. That being said, the test of strict necessity is a rigid and well-established one, supported by multiple decisions of the CJEU. As the EDPB described it, referring to the CJEU decision in Case C-623/17 *Privacy International*,[69] strict necessity is 'closely linked to the requirement of objective criteria in order to define the circumstances and conditions under which processing can be undertaken, thus excluding any processing of a general or systematic nature'.[70]

Finally, for automated decision-making purposes, Article 22(3) GDPR states that the key *de minimis* safeguard to data subject's rights, freedoms and interests should be 'the right to obtain human intervention on the part of the controller, *to express his or her point of view and to contest the decision*'.[71] Conversely, while Article 11(1) LED also supports 'the right to obtain human intervention on the part of the controller', it does not expand on this notion like Article 22(3) GDPR, nor does it explain that the right should enable the expression of opinion and a challenge to the automated decision. In consequence, the right to obtain human intervention within LED is considerably weaker. This is extremely pertinent to the police development and adoption of AI tools, the drive towards which is evidenced by initiatives such as Interpol's Artificial Intelligence Toolkit, which provides 'guidance on the development, procurement and use of responsible AI in law enforcement agencies'.[72]

5.2 Key Risks for Research Quality and Integrity

Legal uncertainty concerning whether the GDPR or the LED is the applicable European data protection instrument may halt the research before it even starts, or indeed during the research project itself. There is a risk that LEAs might simply avoid engaging in research with personal data (such as case files) altogether due to the perception that sharing personal data for research will increase their exposure to legal liabilities and reputational risks. The more unclear the organisation's lead researcher (or worse, data protection officer (DPO) feels regarding the applicable legal framework, the greater is this reluctance to engage – or, if engaged, to share personal data. Indeed, a common requirement of publicly funded research is that research organisations provide sufficient clarity about the legal basis for processing personal data, in compliance with EU data protection law.[73]

A lack of legal certainty can present a clear barrier to the EC's 'fighting crime and terrorism' policy objectives and to publicly sponsored research and development

[69] Case C-623/17, *Privacy International v Secretary of State for Foreign and Commonwealth Affairs and Others* (2020) ECLI:EU:C:2020:791, para. 78.
[70] European Data Protection Board (n 27) 21.
[71] Emphasis added.
[72] See INTERPOL, 'Artificial Intelligence Toolkit' (2023) <https://www.interpol.int/en/How-we-work/Innovation/Artificial-Intelligence-Toolkit> accessed 14 August 2024.
[73] For example, see n 28.

agendas generally.[74] In the health context, some authors are making estimates of financial burdens arising from the non-use of datasets.[75] Such costs should not be underestimated; the inability to use real public authority data to test and evaluate new research and technologies may constitute a loss of opportunity to improve the functioning of law enforcement bodies and their response to crime and terrorism. This could be due to the data quality and accuracy difficulties presented through superficially recreating this data, as well as not participating in the research at all, due to preliminary, unexplored legitimacy concerns. Unfortunately, obtaining information on LEAs' decision-making processes in this regard is exceedingly difficult.

Secondly, inconsistent treatment of research activities as 'research' also creates doubts as to the scientific accuracy of the analytical outputs produced from LEA data. Policing data-science risks attracting the same questions around rigour, quality and reliability that have been acknowledged in forensic science.[76] The 'research ecosystem' at its best includes rigorous scrutiny of the scientific standards and methodological robustness that a specific research activity must meet. In essence, the public could be concerned with LEAs deploying technology tools not based on reliable and sufficiently scrutinised research.

Thirdly, legal uncertainties risk blurring the lines between research activities and LEA operations. Researchers are encouraged to work in partnership with 'end-users' (here LEAs) to have impact.[77] However, a lack of clear delineation between policing and research is problematic in both directions. LEAs could conduct research activities without defining them as such. Moreover, research (whilst regulated by Member State or EU law) is also governed by its own 'ecosystem', i.e., guidelines and frameworks, including professional training, often created and monitored by research ethics committees (RECs) and other organisations performing a stewardship role in the research environment.[78] If LEAs conduct what are essentially research activities without characterising them as such, any potential guidelines and frameworks that could cover or at least inform LEA activities are unlikely to be applied in practice. Furthermore, governance within research organisations often involves RECs, which scrutinise the conformity of research activities with fundamental ethical principles. The extent to which LEAs have a similar governance structure, or a culture of adhering to research ethics, is not as clear, and LEA researchers may not necessarily be part of

[74] The European Commission's Horizon Europe programme allocates c. €1.5 billion to the Pillar 2 cluster on 'civil security for society'. See, European Union, 'Horizon Europe, budget' (2021) <https://op.europa.eu/en/publication-detail/-/publication/1f107d76-acbe-11eb-9767-01aa75ed71a1> accessed 14 August 2024.

[75] Kerina H Jones et al, 'The Other Side of the Coin: Harm Due to the Non-use of Health-related Data' (2017) 97 *International Journal of Medical Informatics* 43.

[76] National Academy of Sciences, *Strengthening Forensic Sciences in the United States: A Path Forward* (National Academies Press, 2009).

[77] 'Impact' being one-third of the scoring criteria for a H2020 or Horizon Europe research proposal.

[78] Graeme T. Laurie et al, 'Charting Regulatory Stewardship in Health Research: Making the Invisible Visible' (2018) 27(2) *Cambridge Quarterly of Healthcare Ethics* 333.

the same tradition. Whilst LEAs may have their own codes of ethics, as shown in section 3.1 above, these rarely include details on the ethical conduct of research activities. Hence, the ethical safeguards that conventionally apply to research activities may not be available, raising risks for data subjects' rights.

Finally, yet importantly, without a clear demarcation of 'research' from 'prevention' activities, the fear of LEA 'mission creep' and abuse of personal data for surveillance purposes under the rubric of 'research' will remain present, threatening overall public trust in LEAs. A good example is Sensing, a 2019 project conducted by the Dutch police. It was located in the city of Roermond and focused on the prevention and detection of 'mobile banditry' (property crime, such as pickpocketing, shoplifting, doorstep distraction crimes and home burglaries).[79] The predictive policing system made use of police records and data collected through new and existing sensors installed in public spaces to gather information about vehicles and movement patterns.[80] The Sensing project had an operational objective and a learning objective. To fulfil them, high-risk scores ('hits') were identified during data processing, based on which the police would decide if and how to intervene. According to Amnesty International, for the learning objective, the data was used to learn more about data-driven policing, thus data collected by the sensors was processed and analysed to train crime-predicting algorithms.[81] Consequently, the line between the two objectives of the Sensing project could be seen as blurred, illustrating well the concerns about mission creep and clarity of legal provisions seeking to prevent it.

6. Conclusion

This article has demonstrated that research involving LEAs and personal datasets such as LEA-held personal data poses a variety of real, legal challenges demanding answers. It has explored the nature of such activities and illuminated the blurry line between research and operational activities, noting how some technologies may be too nascent to be realistically conceived as for the purpose of crime prevention.

It suggested that the GDPR should be understood as the default regime for the discussed research activities. Where they are conducted for Article 1(1) LED purposes – most realistically the purpose of crime 'prevention' – they can fall within the scope of the LED, where: 1) prevention is understood broadly; 2) the national implementation of the LED supports this approach; and 3) such implementation lays out suitable safeguards for data subjects. This is most convincingly the case where

[79] Politie, *Plan van Aanpak Operationele Proeftuin Sensing Roermond* (Dutch, 12 October 2017), available at <https://www.politie.nl/binaries/content/assets/politie/wet-open-overheid/11-landelijke-eenheid/overige-documenten/2021/sensing-roermond/20210803---8457---besluit.pdf> accessed 14 August 2024.

[80] Dutch Police (n 79).

[81] Amnesty International, 'We Sense Trouble: Automated Discrimination and Mass Surveillance in Predictive Policing in the Netherlands' (2020) <https://www.amnesty.org/download/Documents/EUR3529712020ENGLISH.PDF> accessed 14 August 2024.

research entails linking different criminal offences. There are examples of national legislation that explicitly include research as part of an LEA function that offer admirable clarity (and curious breadth), but this is far from common.

Regarding the legal basis for processing, this article has argued that for both GDPR- and LED-covered activities a basis separate from the original should be found. As for the identification of controllers and processors in collaborations between LEAs and researchers, the controller–processor arrangement was found be the most reliable basis for research using personal data collected by LEAs. This requires the LEAs engaging with this research to be skilled, aware controllers. Consequently, it is vital that LEAs understand the predicted technology readiness of the tools they develop or test, to accurately define, as 'controllers', the purpose of the data processing. Joint controllership arrangements suffer from multiple pitfalls in this area.

This article then proved that the choice between GDPR and LED has a material impact on the position of affected data subjects, noting five key differences between the instruments in question. Where a choice arises for LEAs with respect to which instrument should be invoked, research activities should be analysed in a granular manner, so that – where feasible – GDPR might be chosen to apply, for the benefit of data subjects and public perception of LEAs. A good example of this would be research which may rely on consent as a legal basis. The article also evidenced how a lack of clarity with respect to data protection laws, and failure to recognise research as research, have a tangible impact on this important area and the opportunities therein.

In conclusion, harmonisation work (taking into account the nature and scope of national divergences) in this area is still needed, and LEA DPOs' attention should be drawn to issues such as the definitions of research and crime prevention, as well as highlighting the risk that LEA research with personal data escapes the coverage of appropriate legal and ethical systems. There are also further challenges and open questions remaining. It is still an open question of how much use we, as a society, want to make of historical LEA data, given what is known about its potential biases and discriminatory impacts. We should also be asking how we can make usable tools for LEA data analysis without using such data in a direct manner (e.g., making properly anonymised data sets more available or sharing the cost burden of creating realistic synthetic data). Practically, there is education work to be done regarding the limitations of LEA data analytics.

In the absence of further legal reform at either the European level or the national level there are potential ways forward. An EDPB Opinion on the topic, although politically challenging, could help to bring greater clarity. Notably the lack of definition around processing personal data for 'research' is not unique to the LEA context – a good example of this being the request from the EC for EDPB clarification

on the GDPR's application in the health research context.[82] Guidance similar to that provided in response to the EC's request would be welcomed here.[83] Taking EC-funded programmes specifically (such as Horizon Europe[84]), there is an opportunity for the European Commission to produce clear guidance on how it would expect LEA-controlled data – and similar – research to proceed, including models for how data-sharing agreements should be structured, and the safeguards they must include.

Moreover, Article 40 GDPR allows for associations of categories of data controllers to draw up codes of conduct and have these reviewed and approved by supervisory authorities (and the EDPB if the code would apply to processing in multiple Member States). It could be possible for EU LEAs to develop a code of conduct regarding their processing of personal data in the context of research activities beyond that of facial recognition. Adherence to such codes of conduct would give LEAs and researchers greater clarity about what is expected. Developing such a code would be a substantial project, given the possibility of divergent LEA perspectives. The LED does not envisage such codes, meaning such an effort could be based upon the GDPR, but at the same time extend to how LEAs intend to process data for research purposes under the directive, including matters such as relationships with processors, and particularly, the safeguards they are expected to deploy.

Finally, LEAs engaged in research should be encouraged to focus upon research ethics, developing deeper professional commitment to ethical research practices, and robust research governance processes. This could happen through establishing independent ethics committees ensuring that ethics and people's rights are protected in the research process ,or engaging people with expertise similar to that often present in RECs in the conversation regarding research, development and adoption of new interventions. In the absence of specific research ethics codes, law enforcement researchers should be bound by general principles of research ethics, such as those laid down in the European Charter for Researchers.[85] LEAs as researchers are obliged to follow principles and rules of scientific research, including respect for human rights, research integrity, research ethics, responsibility towards communities, individuals and society, as well as data protection and confidentiality protection requirements. LEA DPOs and legal teams should be closely engaged with LEA research activities.

---

[82] See European Data Protection Board, 'EDPB Document on response to the request from the European Commission for clarifications on the consistent application of the GDPR, focusing on health research' (2021) <https://edpb.europa.eu/our-work-tools/our-documents/other-guidance/edpb-document-response-request-european-commission_en> accessed 14 August 2024.

[83] See EDPB (n 82).

[84] See European Commission, 'Horizon Europe' (2024) <https://ec.europa.eu/info/research-and-innovation/funding/funding-opportunities/funding-programmes-and-open-calls/horizon-europe_en> accessed 14 August 2024.

[85] See <https://euraxess.ec.europa.eu/jobs/charter/european-charter> accessed 14 August 2024.