Douwe Korff
Emeritus Professor of International Law, London Metropolitan University
Associate, Oxford Martin School, University of Oxford

THE EU'S OWN 'SNOWDEN SCANDAL': ILLEGAL MASS SURVEILLANCE AND BULK DATA DATA MINING BY EUROPOL AND THE EU MEMBER STATES

"This is about the desire of Europol and EU Member States to collect, 'in a generalised manner', vast stores of personal data on overwhelmingly innocent people in order to 'mine' the bulk data to single out by AI-based algorithms individuals who the algorithms say 'may' be involved in crime – in other words, it is about mass surveillance of entire populations 'without any differentiation, limitation or exception being made in the light of the objective of fighting against serious crime', without regard for the inherent serious dangers and deficiencies in the data mining technologies, and in clear breach of EU law. It is a scandal on a par with the indiscriminate surveillance by the US agencies, exposed nearly a decade ago by Edward Snowden." Douwe Korff

1. Introduction

On 3 January 2022, the European Data Protection Supervisor (EDPS), which supervises *inter alia* the processing of personal data by the EU's law enforcement agency, Europol, notified Europol of **an order to delete data it held in its databases on individuals with no established link to a criminal activity**.[1]

The order does not spell out how big those databases are. However:[2]

> According to internal documents seen by the Guardian, Europol's cache contains at least 4 petabytes – equivalent to 3m CD-Roms or a fifth of the entire contents of the US Library of Congress. ...
>
> Among the quadrillions of bytes held are sensitive data on at least a quarter of a million current or former terror and serious crime suspects and a multitude of other people with whom they came into contact. It has been accumulated from national police authorities over the last six years, in a series of data dumps from an unknown number of criminal investigations.

The EDPS order was a follow-up from an EDPS inquiry into "**the use of Big Data Analytics by Europol for purposes of strategic and operational analysis**", launched by the EDPS of its own motion on 30 April 2019.[3] In the context of that inquiry, in September 2020, the EDPS had already issued an **admonishment** to Europol, in which it considered that the massive data sets in question posed "**high risks for data subjects**" and could have "**potentially severe impact on their fundamental's rights and freedoms**" – but in which he still left it to Europol itself "to devise mitigation measures that would both reduce the risks for data subjects and ensure that Europol does not lose its operational capabilities", merely asking that Europol "provide an action plan to address the admonishment within two months and to report on the measures taken within six months".[4]

This was followed by exchanges between Europol and the EDPS over a period of more than a year – September 2020 to October 2021,[5] when Europol asked the Supervisor in a letter to:[6]

> Defer a decision on whether to issue an order to delete until the proposed [Revised Europol Regulation, Regulation 2016/79413] enters into force; and

> In the meantime, agree to "allow Europol to store large datasets where the DSC is not completed [i.e., which still include extensive data on innocent people] for a period of twelve months, extendable by up to a further six months in duly justified cases, where further processing is necessary and proportionate."

The January 2022 order makes clear that the EDPS does not find this acceptable. Rather, the EDPS ordered Europol to carry out "data subject categorisation" of all new data received within six months of the receipt of the data; data already held must be "categorised" within 12 months.[7] As the press release on the order explains:

> This means that Europol will no longer be permitted to retain data about people who have not been linked to a crime or a criminal activity for long periods with no set deadline. The EDPS has granted a 12-month period for Europol to comply with the Decision for the datasets already received before this decision was notified to Europol.

2. What this is all about: "mining" bulk data sets to "identify" (serious) criminals

This is all about the trend – discernable for many years now – towards "pre-emptive", "predictive" law enforcement: the increasing approach by national law enforcement (and national security) agencies towards trying to "identify" individuals who "may be" involved in crime, especially serious crime (as rather variably defined) or terrorism (as also increasingly expansively defined).

I have recently again written about this trend, in my opinion on the PNR case that is pending before the Court of Justice, where I noted that:[8]

> One more recent element of the EU strategy is **the facilitating of bulk access by law enforcement and national intelligence agencies to large-scale databases and data sets held by the private sector on whole – of course, overwhelmingly innocent – populations** (what I will call "mass surveillance").[9] The data sets to which law enforcement and intelligence agencies (working increasingly closely together: see above) want access in bulk include data on electronic communications (both on the actual communications and "metadata"), financial data – and data on people travelling to and in the EU, in particular air passengers.

Closely linked to this is the also increasing trend at the level of the EU Member States in various "law enforcement depositories" and at EU level in the Europol SIS II database, of data on so-called "**persons of interest**" (or "**subjects of interest**" in the UK terminology) – which includes **individuals on whom there is a so-called "Article 36 alert" in the SIS II database**. An alert entered into the SIS database requires the EU Member States' authorities to carry out "**discreet checks, inquiry checks or specific checks**" on those persons although they are not (yet) – and cannot (yet) – formally be declared a suspect under criminal or criminal procedure law, i.e., **on persons against whom there is as yet no real evidence that they will commit a relevant crime**.[10]

The EDPS reports that:[11]

> Europol considers that processing [of] non-DSC data [i.e., of not-yet-categorised data that includes information on persons with no established links to criminal activity] for the purpose of **extracting relevant data** in compliance with the [Europol Regulation] is included under the objectives, tasks and competences [of Europol] as set out in the [Europol

> Regulation]. According to Europol, the act of **reducing, filtering and extracting data for the purposes of criminal intelligence** is integral to the activity of law enforcement analysis –

In this, the phrases "extracting relevant data" and "extracting data for the purposes of criminal intelligence" refer especially and specifically to the filtering of the massive data sets on mostly entirely innocent individuals in order to not only identify (in the proper sense of the word) known criminals and other "known" persons clearly linked to criminal activity, but also to "identify" – for which one must read here: single out and label – "persons of interests" that were not previously linked to crime (and against whom the only indication of such a link is the fact that they are filtered out).[12] Crucially, this is done by means of what are coyly called "pre-determined criteria", which increasingly take the form of AI-based self-learning algorithms that are supposed to be good at such "identifying" of "suspicious" people (or activities).

Most pertinently, as I also noted in my PNR opinion**, Europol has become increasingly involved in such algorithm/AI-based data analysis (or at least in the research underpinning those technologies)**:[13]

> Europol is [already – DK] described as the EU's 'criminal information hub'[14] and the main 'information broker',[15] as it facilitates information exchange between EU Member States, Europol, other EU bodies, international organisations and third countries, and produces criminal intelligence on the basis of information acquired from various sources, including Member States and its partners. **Amongst its many tasks, Europol** also supports and coordinates cooperation on cross-border police work and **produces** regular assessments that offer comprehensive, **forward-looking analyses of crime and terrorism in the EU**.

Last year the Commission proposed to further expand this role by *inter alia*:[16]

- enabling Europol to receive personal data (including bulk data) directly from private parties on a more regular basis, inform such private parties of missing information, and ask Member States to request private parties to share further information;
- enabling Europol to process large and complex datasets and to carry out "pre-analyses" of those large and complex datasets; and
- strengthening Europol's role on research and innovation, aimed at the development of tools, including the use of AI for law enforcement, and involving the development of new technologies for which extensive processing of large quantities of personal data may be required, e.g., to create and test algorithms or for encryption;

As EDRi put it:[17]

> the proposal suggests that Europol should play a bigger role in developing and shaping future policing technologies that will be deployed in Europe in the coming years. For that, the Agency would help set the priorities in terms of EU funding for research projects in the field of security and train and test itself algorithms to develop tools for EU law enforcement authorities.

I discuss algorithm/AI-based data mining and profiling in some detail in my PNR opinion.[18] Here, it must suffice to reiterate that **this kind of processing suffers from fundamental, inescapable flaws that pose great risks to the rights and freedoms of individuals**. Specifically:[19]

- because the data sets are so large and the number of actual serious criminals in them s relatively low, the algorithm/AI-based profiling is likely to result in tens of thousands of "**false positives**": individuals who are wrongly labelled to a be person who "may be" involved in serious crime or terrorism – and in such contexts, this cannot be avoided by tweaking the algorithm;
- although various provisions in relevant rules and laws may limit or prohibit the collection or use of sensitive data, or stipulate that individual decisions and matches may not be "solely based on" sensitive aspects of the individuals concerned, this do not protect those individuals from **discriminatory outcomes** of the profiling;
- the algorithm/AI-based outcomes of the processing are almost **impossible to challenge** because those algorithms are constantly dynamically changed ("improved" through self-learning) and therefore in effect impossible to fully comprehend even by those carrying out the analyses/risk assessments; and
- the outputs and outcomes of the algorithm/AI-based profiling and data mining and matching are **not subject to proper scientific testing or auditing**, and extremely unlikely to made subject to such testing and auditing.

In regard of the latter point, it is notable that **the authorities – the EU Member States, the European Commission and Europol – consistently refuse to provide serious evidence of the efficacy of their data mining and profiling technologies**, and even refuse to establish a suitable system for the capturing of reviewable data. This was noted in respect of mandatory data retention a decade ago,[20] was repeated in relation to PNR data (as noted in my opinion),[21] and is again clear in the present context: Europol merely asserts that it needs to "extract relevant data" from non-classified massive data sets and that "the act of reducing, filtering and extracting data for the purposes of criminal intelligence is integral to the activity of law enforcement analysis"[22] – but it provides no serious, peer-reviewable evidence of those assertions.

3. Europe's own "Snowden scandal"?

Edward Snowden's revelations were mainly about the scope of the US global surveillance operations; he provided less detail about what was actually done with the massive data troves the agencies accumulated. However, it has become clear through subsequent investigations and exposures that the hoovering up of data (especially e-communications data) in bulk, by the US National Security Agency, NSA, working closely with its UK counterpart, the UK Government Communications Headquarters, GCHQ, was precisely for the above purposes: to analyse and filter the massive data sets in order to try and seek out – "identify" – individuals who "*might be*" involved in nefarious activities (or who just "*might be*" "*of interest*" for political or other purposes).[23]

The UK-USA activities and the associated technologies were described in a 2015 report by the UK House of Commons Intelligence and Security Committee (ISC), albeit with many details blocked out in the published version.[24] Many of those further details were set out in an Open Rights report in the same year.[25] They are summarised in a 2020 submission by Ian Brown and myself to the EU authorities on the inadequacy of UK law, and in particular in Part Two of that submission that

analysed the UK surveillance activities in the light of the *Schrems II* judgment of the Court of Justice of the EU.[26]

The Snowden revelations caused widespread, furious denunciations of the USA by European politicians including Angela Merkel ("friends do not spy on each other"), MEPs, the European Commission, etc..

However, the Americans, with some justification, accused the Europeans of hypocrisy and "double standards".[27] As Ian Brown and I pointed out:

> [The most pertinent claim of hypocrisy] can be laid against the EU and the EU Member States in relation to [their own] actual compliance with either the ECtHR or the CJEU standards. … [T]here have been serious questions about the general effectiveness of enforcement by the supervisory authorities in the EU Member States. More specifically, [in relation to] CJEU requirements relating to personal data retention obligations imposed on entities that are subject to EU data protection law, for the purpose of allowing access to the retained personal data by EU Member State intelligence agencies for national security purposes, the EU and the Member States have been reluctant to come up with laws that actually meet the CJEU standards laid down in its *DRI* and *Tele-2/Watson* judgments, in spite of strong criticism by the European Parliament and civil society. The current laws in the EU Member States have also not yet been brought into line with CJEU judgments in *PI* and *LQDN*. And the surveillance laws and practices in many EU Member States would clearly fail the tests applied to the laws and practices of the USA in *Schrems II*.

We therefore suggested that not only should the USA bring its surveillance laws into line with international human rights standards, but the EU and the EU Member States should ensure this in the EU and Member States' legal orders as well:[28]

> The EU institutions and in particular the European Parliament should stand up for the rule of law and demand that both the Member States and third countries bring their intelligence practices and domestic law frameworks fully in line with international human rights law.

Until this is done, the human rights violating activities of the EU Member States' intelligence agencies, and the now clear active support for these by Europol and the European Commission – and the now explicit aim to extend these nefarious practices into the area of law enforcement, which is of course squarely within the scope of EU law – is as much as, and in many details identical to the scandalous behaviour of the US and the UK intelligence agencies.

4. Conclusion

As the EDPS rightly noted, the indiscriminate bulk data collection and algorithmic/AI-based analyses of the "uncategorised" data by Europol violates EU data protection law, the Charter of Fundamental Rights and the Treaties. However, these activities not only continue in spite of the EDPS strong criticism – Europol and the Commission (and at least some Member States) are trying to legitimise those activities under the proposed revised Europol Regulation, even though the compatibility of any such new rules with the Charter and the Treaties is beyond doubtful.

It is clear that Europol has been deliberately procrastinating in its lengthy exchanges with the European Data Protection Supervisor, stretching over two years. Europol's demand that the EDPS

defer its order even further, until after the revised Europol Regulation is in place, is simply an attempt to "save" the dangerous technologies it is wedded to (without evidence of the efficacy of those technologies or of the negative effects on innocent people and on the rule of law).[29]

For the maintenance of the rule of law and for the protection of the fundamental rights and freedoms of European citizens it is essential that the mass surveillance activities of the EU Member States' law enforcement agencies and of Europol – and of the EU Member States' intelligence agencies that are increasingly involved in this too![30] – be brought under control.

The judgment of the Court of Justice in the PNR case will hopefully send a clear signal that mass surveillance by means of data mining and profiling of bulk datasets on mostly innocent people is incompatible with the EU Charter of Fundamental Rights and the Treaties. If it does send that signal, that should also kill of the dangerous activities of Europol discussed in this article. The opinion of the Advocate-General on the PNR case is due later this month.

- o – O – o -

Douwe Korff (Prof.)
Cambridge (UK), 17 January 2022

NOTE:

[1] EDPS press release: *EDPS orders Europol to erase data concerning individuals with no established link to a criminal activity*, 10 January 202, available at:
https://edps.europa.eu/press-publications/press-news/press-releases/2022/edps-orders-europol-erase-data-concerning_en

Full text of the EDPS order ("**the EDPS Order**" or "**the Order**"):
https://edps.europa.eu/system/files/2022-01/22-01-10-edps-decision-europol_en.pdf

FAQs (Frequently Asked Questions) information sheet:
https://edps.europa.eu/system/files/2022-01/22-01-10-europol-order_faqs_en.pdf

[2] *A data 'black hole': Europol ordered to delete vast store of personal data*, Guardian, 10 January 2022, available at:
https://www.theguardian.com/world/2022/jan/10/a-data-black-hole-europol-ordered-to-delete-vast-store-of-personal-data

[3] Case 2019-0370.

[4] EDPS Order, paras. 2.1 – 2.4.

[5] *Idem*, paras. 2.5 – 2.11.

[6] *Idem*, para. 2.14.

[7] *Idem*, order points 1 and 2 (at the end of section 5, on p. 13).

[8] Douwe Korff, Opinion on Core Issues in the PNR CJEU Case, prepared at the request of the Fundamental Rights European Experts Group (FREE Group), November 2021, p. 8 (original emphasis), available at:
https://www.ianbrown.tech/wp-content/uploads/2021/12/KORFF-FREE-Paper-on-Core-Issues-in-the-PNR-Case.pdf
The trend is described in some detail in section 2 of the opinion, under the heading *PNR in context*.

[9] I explain why I call the analyses of bulk data sets "mass surveillance" in a box on p. 15 of the opinion.

[10] For further details, see section 2.3 of the opinion.

[11] EDPS Order, para. 2.13, quoting a Europol letter of 20 October 2021, EDOC#1191646v7B. Emphases added.

[12] On the ambiguous meanings of the word "identify" in this context, see the introduction to section 4.9 of the PNR opinion.

[13] Niovi Vavoula and Valsamis Mitsilegas, Strengthening Europol's mandate: A legal assessment of the Commission's proposal to amend the Europol Regulation, study requested by the European Parliament's Civil Liberties (LIBE) Committee, May 2021, p. 12, emphasis added, available at:
https://www.europarl.europa.eu/RegData/etudes/STUD/2021/694200/IPOL_STU(2021)694200_EN.pdf

[14] 'Europol Strategy 2020+' (*Europol*, 5 February 2019) < https://www.europol.europa.eu/publications-documents/europolstrategy-2020 > accessed 3 May 2021, 4 [original footnote]

[15] Thomas Wahl, 'The European Union as an Actor in the Fight Against Terrorism' in Marianne Wade and Almir Maljevic (eds), A War on Terror? (Springer 2010) 144. [original footnote]

[16] European Commission, Proposal for a Regulation of the European Parliament and of the Council amending Regulation (EU 2016/794, as regards Europol's cooperation with private parties, the processing of personal data by Europol in support of criminal investigations, and Europol's role on research and innovation, COM(2020) 796 final (Europol proposal of 2020), available at:
https://beta.op.europa.eu/en/publication-detail/-/publication/0d3eb9fd-3b02-11eb-b27b-01aa75ed71a1/language-en/format-PDF/source-search
In my brief indents, I draw on the LIBE study, referenced in note 13, above.

[17] EDRi, Recommendations on the revision of Europol's mandate, position paper on the European Commission's proposal amending Regulation (EU) 2016/794, as regards Europol's cooperation with private parties, the processing of personal data by Europol in support of criminal investigations, and Europol's role on research and innovation, 10 June 2021, pp. 3 – 4, available at:
https://edri.org/wp-content/uploads/2021/06/Recommendations-on-the-revision-of-Europols-mandate.pdf
For a discussion of the broad criticisms from civil society, see section 2.4 of the PNR opinion.

[18] See section 3, sub-section 4.9(f) of the opinion.

[19] Cf. the box on p. 16 of the PNR opinion.

[20] See the 2011 Max Planck Institute study referenced in footnote 250 of the PNR opinion.

[21] See the whole of section 5.1 in the PNR opinion.

[22] See the quote from the Europol letter of 20 October 2021, on p. 3, above.

[23] Under the US Foreign Intelligence Surveillance Act (FISA), the US agencies can look for any "foreign intelligence information", defined in US 50 U.S. Code § 1801 "information with respect to a foreign power or foreign territory that relates to the national defense, national security, or conduct of foreign affairs of the United States." The UK definition, while seemingly more limited, in fact comes close to this: see Ian Brown & Douwe Korff, The inadequacy of UK data protection law in general and in view of UK surveillance laws, Part Two (note 26, below), section 2.2.2, *Indiscriminate collection in bulk*, sub-section I, What is collected and how?, on p. 8.

[24] UK House of Commons' Intelligence and Security Committee, Privacy and Security: A modern and transparent legal framework (HC 1075), 12 March 2015,available at:
http://isc.independent.gov.uk/news-archive/12march2015

[25] Open Rights Group, Collect It All: GCHQ and mass surveillance, 2015, available at:
https://www.openrightsgroup.org/publications/collect-it-all/
See in particular Part One, Chapter One, *Passive Collection*, available separately at:
https://www.openrightsgroup.org/app/uploads/2020/03/01-Part_One_Chapter_One-Passive_Collection.pdf

[26] Ian Brown & Douwe Korff, The inadequacy of UK data protection law in general and in view of UK surveillance laws, submission to the European Union bodies involved in assessing whether under the EU General Data Protection Regulation (GDPR) the United Kingdom should be held to provide "adequate" protection to personal data:
Part One on general inadequacy of UK data protection law, submitted on 9 October 2020, available at:
https://www.ianbrown.tech/2020/10/09/the-uks-inadequate-data-protection-framework/
Part Two on UK surveillance law, submitted on 30 November 2020, available at:
https://www.ianbrown.tech/wp-content/uploads/2020/11/Korff-Brown-Submission-to-EU-re-UK-adequacy-Part-Two-DK-IB201130.pdf
Executive Summary and discussion of the implications, also submitted on 30 November 2020:
https://www.ianbrown.tech/wp-content/uploads/2020/11/Korff-Brown-Submission-to-EU-re-UK-adequacy-ExecSumm-DK-IB201130.pdf

[27] See Ian Brown & Douwe Korff, Exchanges of personal data after the Schrems II judgment, study carried out at the request of the European Parliament's Civil Liberties (LIBE) Committee into the future of EU – US flows of personal data, July 2021, available at: https://www.europarl.europa.eu/RegData/etudes/STUD/2021/694678/IPOL_STU(2021)694678_EN.pdf, section 2.3.1.3, *Requirements relating to access to personal data by state authorities*, under the heading "Hypocrisy and 'Double standards'? US criticism of the EU and the EU Member States", on pp. 58 – 60.

[28] *Idem*, section 4.2.4, *Recommendations*, on p, 128.

[29] On the lack of evidence, see the final paragraph in section 2, above. On the possibly very serious consequences of being "identified" as a person who "may be" involved in serious crime or terrorism as a result of a "match" against "pre-determined" algorithm/AI-based criteria or profiles, see the PNR opinion (note 8, above), section 4.12, *The consequences of a 'match'*.

[30] See section 4.5 of the PNR opinion, where it is noted that in some Member States the PNR Information Units are actually established within the intelligence agencies, and section 4.10, where it is noted that in at least one Member State (the Netherlands), the intelligence agencies have actual direct access to the PNR data.

- o – O – o -