EUROPEAN DATA PROTECTION SUPERVISOR

Redacted version for public release

WOJCIECH RAFAŁ WIEWIÓROWSKI
SUPERVISOR

Mrs Catherine DE BOLLE
Executive Director
Europol
P.O. Box 908 50
2509 LW The Hague
The Netherlands

Brussels, 18th September 2020
WW/xx/vm/ D(xxx) xxx **C 2019-0370**
Please use edps@edps.europa.eu for all correspondence

Subject: **EDPS Decision on the own initiative inquiry on Europol's big data challenge**

Dear Mrs De Bolle,

Please find attached a decision of 17 September 2020 relating to EDPS own inquiry on Europol's big data challenge.

This Decision, addressed to Europol, will also be transmitted to the Europol Cooperation Board because this case is highly relevant for national data protection authorities as a large part of the information it refers to is shared by national law enforcement authorities in the first place.

Yours sincerely,

[**E-signed**]

Wojciech Rafał WIEWIÓROWSKI

CC:

- Mr. Oliver Rüß, Chair of the Europol Management Board
- Mr. Daniel DREWER, Data Protection Officer, Europol
- Mr. François PELLEGRINI, Chair of the Europol Cooperation Board

EDPS DECISION
of 17 September 2020
relating to EDPS own inquiry on Europol's big data challenge

1. INTRODUCTION

1.1. This decision concerns the processing by Europol of "large datasets" received as contributions from Member States (MS), from other operational partners or collected in the context of open source intelligence activities. "Large datasets" are defined for the purpose of this decision as datasets, which because of the volume, the nature or the format of the data they contain, cannot be processed in the Europol Operational Network (OPS NET) [1] with regular tools, but require the use of specific tools and/or storage facilities.

1.2. This decision is addressed to Europol. Under Article 43 of Regulation (EU) 2016/794 of 11 May 2016 ("the Europol Regulation")[2], the EDPS is responsible for monitoring and ensuring the application of the provisions of the Europol Regulation relating to the protection of fundamental rights and freedoms of natural persons with regard to the processing of personal data by Europol.

2. BACKGROUND

2.1. On 1 April 2019, Europol's Executive Director informed the EDPS of major compliance issues with the Europol Regulation in relation to the processing of personal data taking place in the [...] also referred to as "Europol's big data challenge". She informed that Europol had taken two actions to tackle the issues encountered:

- a Security Assessment Review was commissioned […]; and
- a taskforce was created, […] to deal with four areas of concern: capabilities, policies, […] data review and access rights.

2.2. On 11 April 2019, Europol and EDPS staff met at EDPS' premises. Europol informed that the [...] is hosting more than [...] of operational data. Europol also shared the findings of the first report of the taskforce issued on 28 March 2019. […] The subsequent reports were shared with the EDPS on 9 April 2019 […], 3 May 2005 […], 28 June 2019 […] and 31 July 2019 […]. Three issues are still ongoing: the update of the Forensic IT Environment (FITE) policy, the deletion of datasets older than three years by the Internet Referral Unit and the accreditation of the [...]. The last report presents a proposal for changes to the FITE environment in order to integrate operational needs, data protection and compliance requirements, as well as to provide a sufficient information security level.

2.3. On 30 April 2019, the EDPS decided to open an own initiative inquiry on the use of Big Data Analytics by Europol for purposes of strategic and operational analysis. The evolution of Europol's personal data processing activities towards Big Data Analytics

[1] OPS NET is the IT environment where Europol performed operational analysis, excluding data stored and processed on the [...].

[2] Regulation (EU) 2016/794 of the European Parliament and of the Council of 11 May 2016 on the European Union Agency for Law Enforcement Cooperation (Europol) and replacing and repealing Council Decisions 2009/371/JHA, 200/934/JHA, 2009/935/JHA, 200/936/JHA and 2009/968/JHA.

raised concerns linked to the compliance with the Europol's data protection framework, in particular with the principles of purpose limitation, data minimisation, data accuracy, storage limitation, with the impact of potential data breaches, location of storage, general management and information security.

2.4. [...].

2.5. On 4-6 June 2019, the EDPS conducted its annual inspection at Europol, including the use of the [...]from a legal and technical perspectives.

2.6. On 17 September 2019, the EDPS informed Europol of the preliminary findings of the annual inspection with regard to the use of the [...]. In this letter, the EDPS also requested additional information, to be provided by 8 November 2019, namely:

- a mapping of all the scenarios where Europol is processing large datasets.
- a mapping of all the scenarios where Europol is processing data about individuals not linked to any criminal activity, specifying the legal basis, assessing the risks for those data subjects and indicating the mitigation measures in place.
- a list of any current or planned development, including research activities, to process big data sets in an automated (or semi-automated i.e. an automated part followed by a manual one) manner.

2.7. On 18 November 2019, Europol replied to the letter:

- providing an indicative and non-exhaustive list of data repositories for operational information [...][3], whose creation is justified based on the volume, the nature and the format of the data.
- specifying that the collection of datasets is always linked to a criminal activity and their legality has been assessed at national level according to the applicable national criminal procedure. At Europol, their use fall under the established policies, such as e.g. the data review process.
- informing that most of research activities relating to the processing of large datasets are taking place in the context of [...].[4] They aim at organising non-structured data, indexing data to make it searchable, facilitating the extraction of entities, translating text, or using *ad hoc* dictionaries to flag relevant content through machine learning techniques.

2.8. On 4 December 2019, the reply of Europol was discussed during a bi-monthly meeting at staff level.

2.9. On 19 December 2019, the EDPS issued a series of recommendations on technical aspects of the [...] in the 2019 Annual Inspection Report.

2.10. On 31 January 2020, the EDPS and Europol met again at staff level in order to discuss a series of additional questions put forward by the EDPS.

3. FINDING OF FACTS

[3] [...] are OPSNET main databases used for the processing of operational data

[4] The [...] is aimed at reviewing the current Europol's IT infrastructure.

[...]

3.1. In 2002, Europol started assisting MS with the analysis of computer data by forensic means. The High Tech Crime Centre (HTCC) and later the European Cybercrime Centre (EC3) played central roles in this process.

3.2. In 2008, Europol created the Forensic IT Environment (FITE). FITE is composed of an operational environment, the [...], and a Research & Development environment. [...]

3.3. [...]

3.4. According to the FITE policy adopted in 2012 (upon recommendation of the Joint Supervisory Body) […] [...], (2) to process operational information including personal data contained on digital devices where this is not possible in the existing OPS NET infrastructure, in particular due to the volume and/or format of the data, [...]. [...]

3.5. The Research & Development environment of FITE is used to acquire best practices and knowledge in digital forensic activities and share it with other peers at Europol, in the MS and Third Parties. […]

[…]strategic and operational and analysis purposes

3.6. The FITE policy does not provide a definition of what digital forensic services entail. Digital forensics are however usually defined as the collection and analysis of data from computer systems, networks, wireless communications, and storage devices in a way that is admissible as evidence in a court of law.[5]

3.7. This inquiry has shown that the [...], but also for broader strategic and operational analysis purposes, in the same way as the OPS NET. [...]

3.8. This phenomenon concerns all APs […] [...]

3.9. The interviews of Europol analysts during the 2019 Annual Inspection confirmed the findings of the DPF. They explained that this evolution in the use of the [...] is due to the fact that **MS now send larger volumes of data to Europol**. The nature of the data collected at national level in the context of criminal investigations and criminal intelligence operations is not limited anymore to targeted data collection but also increasingly includes the collection of large datasets. More digital content is generated and thus available for law enforcement in the context of criminal investigations, which, in turn, impacts the methods used to produce criminal intelligence.

3.10. **Techniques of digital forensics and big data are used in order to exploit these larger volumes of information**. This requires specialised expertise and entails processing information in a way that is codified by best practices in digital forensics. Forensic experts' objective in this context is to process **all the data received** so as to provide a subset of data to the operational analysts as indicated in the corresponding SIENA message (which specifies what data should be found in these big datasets).

[5] See e.g. Suneeta Satpathy, Sachi Nandan Mohanty, Big Data Analytics and Computing for Digital Forensic Investigations, CRC Press, 7 March 2020.

This in turn entails that multiple copies of datasets are created in a specific order, each one refining more and more the data so as to meet the objectives. These copies usually decrease in size as the refinement process is applied. Furthermore, as creating these refined copies is resource intensive, and their storage is required to establish the chain of evidence to ensure that the data is admissible as evidence in a court of law, the copies are retained so that forensic experts may go back to one of the copies as needed (for example, as new information is provided by MS and new analysis is possible based on this new information).

3.11. This creates risks of loss of technical[6] and factual[7] context and of increased bias in the analysis. Criminal analysts are thus currently faced with three challenges: to ensure (1) data veracity, (2) the reliability of the analysis and (3) the traceability of the decision-making process by the analyst. According to the explanations provided by Europol analysts, addressing these challenges require the continuous storage of the datasets until the investigation is concluded, and in particular beyond the process of entity extraction.

3.12. **Expectations of MS towards Europol have not changed in substance**. MS do not share large datasets with Europol because they do not have the tools to process these data at national level but because these data are seized/collected in the context of criminal investigations/criminal intelligence operations, which have a cross-border element. MS thus expect Europol to provide them back with intelligence products.

3.13. **In the course of this inquiry, the EDPS has identified three scenarios, where Europol processes large datasets in the** [...]**.**

3.14. **Scenario 1: Collection of evidence linked to an identified or identifiable suspect.** The first case relates to the collection by Europol of large volume of data linked to a suspect, such as: [...]. In that context, even if Europol or MS cannot ensure that all the information processed relate to individuals linked to the criminal activity investigated, all of them have at least a link with the suspect of a crime.

3.15. **Scenario 2: Collection of "extensive lists of indiscriminate data"**. The second case relates to the collection of large amounts of data [...] which mainly refer to individuals not linked in any capacity to any criminal activity. For example, Europol collects […] data linked to a series of supposedly connected […] attacks and crosschecks this information in order to identify individuals potentially involved in these attacks. […].

3.16. **Scenario 3: Full involvement of Europol in a criminal investigation from beginning to end**. [...] This includes a mix of information described in scenario 1 and 2, [...].

4. LEGAL ANALYSIS

4.1. The personal data processing activities taking place in the [...] are performed for strategic and operational analysis purposes, in accordance with Article 18(2)(b) and (c)

[6] Steps taken by the forensics processes filter the data with a certain level of accuracy. The work done by the forensics experts requires a significant effort. Furthermore, each step is performed with a view to keep chain of custody. All this establishes a technical context for the data processed.

[7] Factual context refers to the source of the information and the knowledge of which kind of information can be retrieved from this type of source.

of the Europol Regulation. […] Such data processing activities should thus comply with the provisions of Article 18(3), 18(5) and Annex II B of the Europol Regulation.

4.2. The Europol Regulation, and in particular Article 18(2)(b) and (c), does not prescribe any requirement as regards the structure of Europol's information systems. Recital 24 specifies that Europol databases should be structured in such a way as to allow Europol to choose the most efficient IT infrastructure. The only criteria retained by the Europol Regulation to define which data protection rules are applicable is the purpose of the processing. It is thus irrelevant in which database the processing of personal data takes place. Data protection rules related to the processing of personal data for purposes of strategic and operational analysis apply both the data processed in OPS NET and in the [...].

4.3. Article 18(3) of the Europol Regulation states that the processing of personal data for the purpose of operational analysis should be performed by means of AP in compliance with specific safeguards defined by the article. One of these safeguards is the requirement to define, for each AP, its specific purpose, the categories of personal data and categories of data subjects, the participants, the duration of storage and conditions of access, transfer and use of the data concerned. This is implemented by Europol through Opening Decisions (OD). Europol must not process any information beyond what is established in each OD to which the dataset contributed by the operational partner is assigned.

4.4. Article 18(5) of the Europol Regulation limits the categories of personal data and categories of data subjects whose data may be collected and processed for purposes of strategic and operational analysis by Europol as listed in Annex II B. Annex II B (1) limits the categories of data subjects about whom Europol can process data to suspects, potential future criminals, contacts and associates, victims, witnesses and informants. Annex II B (2), (3), (4), (5), (6) define which categories of personal data Europol can process in relation to each of the categories of data subjects mentioned above. Europol must not process personal data beyond these categories of data subjects and of personal data.

4.5. These provisions apply and specify the principle of data minimisation for the processing of personal data for operational analysis purposes, as defined under Article 28(1)(c) of the Europol Regulation. They implement the necessary safeguards to limit the processing of personal data to data that are adequate, relevant and limited to what is strictly necessary for the purposes for which they are processed.

4.6. According to Article 28(4) of the Europol Regulation, Europol is responsible for compliance with the principle of data minimisation for all personal data processed by it.

4.7. This inquiry has shown that it is not possible for Europol, from the outset, when receiving large data sets to ascertain that all the information contained in these large datasets comply with these limitations. The volume of information is so big that its content is often unknown until the moment when the analyst extracts relevant entities for their input into the relevant database in OPS NET.

4.8. These large datasets are further stored in the [...] even after the analysts has completed the extraction process in order to ensure that they, potentially with the support of a

forensic expert, can come back to the contribution in case of a new lead and to ensure the veracity, reliability and traceability of the criminal intelligence process. The retention period is defined in accordance with the provisions of Article 31 of the Europol Regulation. Compliance with Article 31 is already subject to separate recommendations from the EDPS.

4.9. This leads to a situation where large amounts of personal data for which it is uncertain that they comply with the requirements set up by Articles 18(3), 18(5) and Annex II B of the Europol Regulation, are stored on Europol systems for several years. As such, the continued storage of personal data that might go beyond the limits contained in these articles undermines the principle of data minimisation, as defined by Article 28(1)(c) of the Europol Regulation. Indeed, there is a high likelihood that Europol continually processes personal data on individuals for whom it is not allowed to do so and retain categories of personal data that go beyond the restrictive list provided in Annex II B of the Europol Regulation. While the exact amount cannot be quantified, the increase in the use of the [...] observed for the last years clearly shows that the amount of large datasets shared by MS with Europol is rapidly growing. […]

4.10. The processing of data about individuals in an EU law enforcement database can have deep consequences on those involved. Without a proper implementation of the data minimisation principle and the specific safeguards contained in the Europol Regulation, data subjects run the risk of wrongfully being linked to a criminal activity across the EU, with all of the potential damage for their personal and family life, freedom of movement and occupation that this entails.

4.11. In light of the above, the EDPS considers that **the processing of large datasets, as defined in §1.1 of this Decision, by Europol, does not comply with Articles 18(3), 18(5) and Annex II B of the Europol Regulation, as well as the principle of data minimisation (Article 28(1)(c) of the Europol Regulation)**.

5. ADMONISHMENT

5.1. The EDPS, pursuant to Article 43(3)(d) of the Europol Regulation, **admonishes Europol.**

5.2. The EDPS exercises its powers under Article 43(3) of the Europol Regulation, taking due account of its obligation to act in the public's interest as laid out in Article 3 of the EDPS Rules of Procedure.

5.3. The personal data processing activities at stake are linked to the evolution of the datasets contributed by operational partners to Europol or collected in the context of OSINT. The legal concerns identified above are thus structural as they relate to Europol's core working methods.

5.4. Europol, as data controller, is in a better position to devise mitigation measures that can both reduce the risks for data subjects and ensure that Europol does not lose its operational capabilities.

5.5. The EDPS therefore considers that, at that stage, making proposals for remedying that issue, imposing an order of erasure of personal data or a ban, pursuant to Article 43 (3) (b), (e) and (f) of the Europol Regulation, is not proportionate.

5.6. However, as described above, the risks for data subjects are high and the impact on their fundamental's rights and freedoms is severe.

5.7. The EDPS thus urges Europol to implement all necessary and appropriate measures to mitigate the risks created by such personal data processing activities to data subjects.

5.8. The EDPS invites Europol to inform of the **action plan to address this admonishment within two months** and of the **measures taken within six months** since the date of this Decision.

6. JUDICIAL REMEDY

6.1. Pursuant to Article 48 of the Europol Regulation, any action against a decision of the EDPS shall be brought before the Court of Justice of the European Union within two months from the adoption of the present Decision and according to the conditions laid down in Article 263 TFEU.