



EDPS follow-up of the own-initiative inquiry on Europol's Big Data Challenge

(Cases 2019-0370 & 2021-0699)

References:

EDPS Decision D(2020)2036 on the own initiative inquiry on Europol's big data challenge of 17 September 2020;
Europol Reply MBS 146.2020 of 16 October 2020;
Europol 'Action plan' EDOC# 1131384 of 17 November 2020;
EDPS letter D(2020) 2821 of 4 December 2020;
Europol 'Progress Report', EDOC# 1156601 of 17 March 2021;
EDPS letter D(2021)882 of 19 April 2021;
Europol Reply 116904v5B of 2 June 2021.

1. Introduction

This document sets out the EDPS' provisional analysis resulting from the follow-up to the EDPS Decision¹ and admonishment on Europol's processing of large datasets (referred to as 'Europol's Big Data Challenge.')

The EDPS issues this document based on Article 43(2)(b) of Regulation (EU) 2016/794² ('the Europol Regulation', or 'ER' abbreviated).

The focus of this follow-up to the inquiry is whether Europol's retention policy for the storage and deletion of large datasets lacking a Data Subject Classification ('DSC'), as set out in the Europol Action Plan addressing the risks raised in the EDPS Decision on Europol's big data challenge³ and further detailed in the Progress Report of 17 March 2021⁴, complies with Article 18(3), 18(5) and Annex II.B of the Europol Regulation.

Large datasets are defined for the purpose of this assessment as datasets which, due to the volume, nature or format of the data they contain, cannot be processed in the Europol

Large datasets "lacking a DSC" refers to datasets which, because of their characteristics, could not undergo the regular data classification process as provided for in the Europol

¹ European Data Protection Supervisor (EDPS) Decision of 17 September 2020 on the EDPS own initiative inquiry on Europol's big data challenge (case 2019-0370).

² Regulation 2016/794 of the European Parliament and 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, 2009/934/JHA, 2009/935/JHA, 2009/936/JHA and 2009/968/JHA, OJ L 135, 24.05.2016.

³ EDOC#1131384v14A.

⁴ EDOC#1156601v13A.

Regulation (the so-called ‘data subjects categorisation’ or ‘DSC’) and extraction of data categories according to Annex II ER and the Opening Decision Orders.

2. Investigative actions

2.1. EDPS own-initiative inquiry on Europol’s Big Data Challenge

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 (case 2019-0370). The evolution of Europol’s personal data processing activities towards forms of ‘Big Data Analytics’ raised concerns linked to the compliance with the applicable 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.

On 17 September 2020, the EDPS concluded his investigation and issued an admonishment to Europol, on the basis of his findings that the processing of large datasets did 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).

The EDPS considered that Europol, as data controller, was in a better position to devise mitigation measures that would both reduce the risks for data subjects and ensure that Europol does not lose its operational capabilities. The EDPS therefore considered that imposing an order of erasure of personal data or a ban, pursuant to Article 43 (3) (b), (e) and (f) of the Europol Regulation would, at that time, not have been proportionate.

However, in view of the high risks for data subjects and potentially severe impact on their fundamental’s rights and freedoms, the EDPS urged Europol to implement all necessary and appropriate measures to mitigate the risks created by such personal data processing activities to data subjects. He invited Europol to provide an action plan to address the admonishment within two months and to report on the measures taken within six months.

2.2. Europol Action Plan and follow-up to the EDPS Decision and admonishment on Europol’s big data challenge

On 17 November 2020, Europol sent an Action Plan⁵ detailing Europol’s measures to address the risks raised in the EDPS Decision. The Action Plan set out five categories of measures intended to strengthen Europol’s data review activities and build enhanced controls into the [REDACTED], i.e. the new operational environment foreseen to replace the [REDACTED] as the technical platform for handling large datasets:

- (1) the flagging of large datasets lacking [REDACTED] in [REDACTED] by the data provider;

⁵ EDOC#1131384v14A.

- (2) the labelling of these datasets in Europol's environment before the start of the extraction process;
- (3) the definition of strict access rights in order to limit access to these datasets only to the persons in charge of performing the extraction process;
- (4) the implementation of more regular reviews in order to assess whether the large datasets which have not been through the whole extraction process should be retained or deleted;
- (5) the appointment of a Data Quality Control Coordinator in order to closely monitor the data review process.

On 4 December 2020, the EDPS provided comments on the Action Plan, and requested further clarifications on a number of elements. Among these, the EDPS expressed reservations regarding the efficacy of the review process of large datasets lacking a DSC and requested more information regarding the time limit to be put on the fulfilment of the extraction task. The EDPS also asked for clarifications regarding the exact content of the review performed by the analysts/specialist of the Analysis Project ('AP') at the level of extraction, and the policy applied by Europol regarding deletion, i.e. whether the raw data are deleted once the relevant data have been extracted.

The EDPS stated that he considered that the review process under Action 4 should be stricter and contain clear criteria, for instance by imposing a maximum time limit for the storage of these large datasets, as there is no guarantee that these datasets are processed in line with the restrictions of Annex II.B of the Europol Regulation.

On 17 March 2021, Europol sent a Progress Report⁶ detailing the state-of-play concerning measures being put in place under the Action Plan. In the Progress Report, as regards the question of whether Europol will define a maximum retention period for datasets pending a DSC, Europol states that it will "*keep the data for as long as is necessary and proportionate for the support to the investigation concerned.*"⁷

On 19 April 2021, the EDPS sent a letter to Europol commenting on the progress report and expressing serious concerns relating to data minimisation and retention of datasets lacking a DSC. In particular, the EDPS noted Europol's failure to specify a time limit for the extraction process or a maximum retention period for datasets lacking a DSC. The letter instructs Europol to define and put in place a maximum retention period for datasets lacking a DSC, beyond which datasets must be deleted, and to do so with immediate effect⁸.

On 2 June 2021, Europol replied to the EDPS. As regards the request to implement a maximum retention period for datasets without a defined DSC, Europol re-stated that the datasets will be subject to increased reviews to determine the necessity and proportionality of retaining the data for the purpose of an ongoing investigation and to ascertain the time required to determine the DSC when this request comes from a Member State. Europol stated that a default retention period would not be appropriate to its core business and would hamper the success of live investigations.

⁶ EDOC#1156601v13A.

⁷ EDOC#1156601v13A, p. 12.

⁸ "*The EDPS requests Europol to implement a maximum retention period for datasets lacking a DSC. Once that retention period is exceeded, datasets lacking a DSC must be deleted. The EDPS requests that the new retention policy concerning large datasets pending a DSC be implemented as of receipt of this letter, and procedures and guidance to staff be updated accordingly*" Letter of 19 April 2021, p. 3.

3. Preliminary findings

A detailed explanation of the EDPS findings related to use of the ██████ for strategic and operational analysis purposes are described in the EDPS Decision of 17 September 2020 on the own initiative inquiry on Europol's big data challenge.

3.1. Mitigation actions for the processing of large datasets pending a DSC in Europol's forensic environment

Under the Action Plan responding to the EDPS admonishment on Europol's big data challenge, a number of technical controls are being implemented in the ██████, the ██████ and integrated into the new data environment, to mitigate the risk that data without a DSC is processed or integrated into Europol's analysis work.

An update of Europol's ██████ will include the requirement for those contributions pending a DSC to be flagged by the contributor to indicate whether the DSC is 'Completed', 'Not Completed' or 'Not applicable.'

Europol will perform a second assessment upon acceptance of the contributions. After the second assessment, the following steps will be performed:

- a) All messages which are accepted with a second assessment of 'DSC Completed', will be accepted into the data environment for further processing.
- b) All messages which are accepted with a second assessment of 'DSC Not Completed', will be accepted into the data environment but with limited access rights given.
- c) All messages which are accepted with a second assessment of 'DSC Not Applicable', will be accepted into the data environment if relevant for case management purposes (these messages do not contain personal data).

In the new data environment, scheduled to replace ██████ in Q3 of 2021 and replace ██████ in 2022, all accepted ██████ contributions where the DSC is not completed will be automatically labelled as such and these labels will be visible on the contributions as a file property. All data in the data environment will be accessed via the new ██████ and the labelling will be clearly visible to those accessing the data. Pending the launch of the new data environment, Europol will put in place a temporary, interim solution for the automatic labelling of new Serious and Organised Crime (SOC) data.

Specific folders have been created in the ██████ for the storage of data lacking a DSC. Access rights are currently limited in the CFN and will be further limited in the ██████ upon its deployment. Europol states that a limited number of designated Europol staff with a dedicated forensic training will be given access to the ██████. It will be their responsibility to liaise with the relevant Analysis Projects in order to identify the DSC and extract the data where the DSC has been identified.

Europol has confirmed that with the progressive implementation of the above controls, data without a DSC will not be further processed for analysis, included in a Europol analysis product, be subject to a general search against new information received by Europol, or be shared with a Member State or third party.

3.2. Extraction, retention and deletion of large datasets lacking a DSC

As stated in the EDPS Decision of 17 September 2020, it is not possible for Europol, when receiving large datasets, to ascertain that all information contained in these large datasets comply with the limitations in the ER. Due to the volumes of information concerned, their content is often unknown until the moment when the analyst extracts the relevant entities for their input into the relevant database in [REDACTED].

During the extraction, a filtering takes place (data minimisation) based on the restrictions in the respective AP [REDACTED] (categories of data subjects, crime area, operational relevance, and in agreement with the data provider according to what is required). Where the DSC is not provided, designated staff work to identify the DSC during the extraction process. The extracted data undergoes another review by the analysts/specialists of the AP, in order to further reduce the amount of data and ensure compliance. Once the data has been properly reviewed, it becomes available and accessible to a larger group of users, via the [REDACTED] and the [REDACTED].

There is no time limit imposed by Europol for the fulfilment of the extraction task. Europol states that once extraction is completed, data where no DSC can be identified is deleted. However, the lack of a deadline for this task means that in principle the extraction process may last years, allowing analysts to return to datasets where extraction has not been officially completed in case of a new lead or new development in the criminal investigation.

Under the new approach to data review (Action 4 of the Action Plan), Europol will keep a record of the data without a DSC in [REDACTED] and APs will review these contributions on a quarterly basis to ensure the necessity and proportionality of their continued retention. This process is overseen by the Data Quality Control Coordinator (appointed under Action 5 of the Action Plan).

Once it is established that the data is no longer needed for the purpose of an ongoing investigation or for the purpose of determining the DSC, the data is deleted, except where it is stored to preserve the chain of evidence (required to ensure that the data is admissible as evidence in a court of law). According to Europol, datasets retained to preserve the chain of evidence will remain stored in the same secure location as datasets lacking a DSC, be subject to similar access controls (available to the analysts and forensic specialists that have worked on the investigation), and will continue to be periodically reviewed.

Europol's stated policy for the retention of large datasets lacking a DSC is to store the datasets for "as long as is necessary and proportionate for the support to the investigation concerned"⁹. Europol states that it will take all necessary measures to ensure that this is done in the shortest time possible and in line with the rules on data retention as defined under the Europol Regulation.

In response to written requests by the EDPS of 4 December 2020 and 19 April 2021 to specify a more restrictive retention period for large datasets lacking a DSC, Europol states that gauging the necessity and proportionality of continued storage is subject to an assessment on the merits of the individual contribution. Europol does not consider that a default retention period imposing a more restrictive deadline than that provided by the ER in Article

⁹ EDOC#1156601v13A, p. 12.

31 is appropriate to fulfil Europol's core business and contends that it would hamper the success of live investigations.

4. Legal assessment

Article 18(3) of the Europol Regulation states that the processing of personal data for the purpose of operational analysis should be performed in compliance with specific safeguards. In particular, the Executive Director must define "the categories of personal data and categories of data subjects ... [the] duration of storage..."

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.

Article 28(1)(e) provides that personal data shall be kept for no longer than necessary for the purposes for which the personal data are processed.

All the provisions mentioned above give effect to 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.

Article 18(6) of the Europol Regulation provides for a maximum period of six months for Europol to temporarily process data for the purpose of determining whether such data are relevant to its tasks and, if so, for which of the purposes referred to in paragraph 2 (cross checking, strategic or operational analysis, facilitating exchange of information between Member States, Europol, Union bodies, third countries and international organisations).

Article 31 provides for the standard time-limits for the storage and erasure of personal data processed in compliance with the Europol Regulation.

As established in the EDPS decision of 17 September 2021, the retention of large datasets lacking a DSC for an undefined period and potentially throughout the duration of a criminal investigation is in breach of Article 18(5) and Annex II.B ER. This practice generates a high risk that Europol processes data of persons who do not fall under any of the categories of data subjects listed in Annex II of the Europol Regulation and whose link with a criminal activity has not been established.

The EDPS acknowledges that, for technical reasons, Europol needs some time to process these datasets for the sole purpose of extracting relevant information according to the limits laid down in the Europol Regulation.

However, the Europol Regulation does not provide for a situation where Europol would be authorised to temporarily process personal data not falling in the categories foreseen in Annex II.B for purposes of operational analysis (Article 18(3)ER) for the sole purpose of

verifying if personal data received fall into the specific categories of data subject listed in Annex II. B ER and extracting relevant data for operational analysis.

In light of the above, Europol is of the opinion that the standard retention period foreseen in Article 31 ER applies to these datasets. This article requires Europol to review every three years the need for continued storage, taking into account the specificities of the criminal analysis process, which is continuous and requires analysts to go back to the information at their disposal whenever they obtain a new lead.

On the contrary, the EDPS considers that Article 31 ER only applies to datasets whose processing conform to the provisions of the Europol Regulation, i.e. datasets that have been lawfully shared with Europol and comply both with Article 4 and Article 18 ER, including the restrictions contained in Article 18(5) ER.

The Europol Regulation does not contain specific provisions concerning the retention period of personal data that appear to be relevant to Europol's tasks in accordance with Article 18(6) but where there is no certainty as to whether their processing is compliant with other fundamental requirements for processing of the Europol Regulation.

While the Europol Regulation does not provide explicitly for the transmission to Europol of large datasets, including datasets whose compliance with the Regulation has not been assessed, nor envisage their processing for the purpose of extraction, the EDPS does not consider that there is a legislative intention to prohibit such a practice. However, under the combined reading of Article 18(3) and (5) read in the light of Article 28(1)(c) and (e), the processing of such large datasets of persons not having undergone the categorisation should be limited to the shortest time necessary to materially proceed to such categorisation. This is important to ensure that data of persons not related to the crimes for which the operational analysis is being performed ceases as soon as possible.

Under EU law, in the absence of a clear, directly applicable legal provision, and in order to avoid legal lacunae, it is possible to rely on the contextual or teleological interpretation of the EU legal provisions in question.

Article 18(6) ER allows Europol to process datasets only in view of determining whether said datasets are relevant to its task (as defined in Article 4 ER) and if so, for which of the purposes listed in Article 18(2) ER. The large datasets which are the subject of this preliminary analysis are shared with Europol for purposes of operational analysis (Article 18(3) ER). They thus do not fall, in principle, under the scope of Article 18(6) ER.

However, in the absence of an explicit legal provision laying down the retention period of personal data processed for this purpose, the EDPS considers it appropriate to apply Article 18(6) by analogy. This article regulates the case where there is no certainty that the data received are relevant for Europol's tasks and thus in compliance with the ER, a situation which is comparable to the one at stake: the data received are relevant for Europol's tasks in line with Article 4 ER and have been accepted under Article 18(3) ER, i.e. for operational analysis purposes, but there is no certainty that they comply with the rest of the requirements of the Europol Regulation, in particular with Article 18(5) and Annex II.B ER.

Extending the application of Article 18(6) by analogy would allow for an initial processing of data in a pre-analysis phase, while imposing the maximum time limitation necessary to allow for this task to be fulfilled in compliance with the data minimisation and the storage limitation principles. The EDPS notes that this approach is in accordance with the

conclusions reached by the European Commission in its Impact Assessment accompanying the legislative proposal for an amendment to the Europol Regulation.¹⁰

Furthermore, such initial processing, subject to a 6 month retention period, should take place prior to, and be entirely separate from, Europol's processing for purposes of cross-checking, strategic analysis, operational analysis or exchange of information, as it would be ensured by the application of the set of mitigation measures by Europol under its Action Plan of 17 November 2021.

In light of the above, the EDPS considers that:

- By providing for the continuous storage of large datasets lacking a DSC in accordance with Article 31 ER, the Action Plan provided by Europol does not address the legal breach identified in the admonishment decision of 17 September 2020 and represented by Europol's continued processing of large amounts of personal data on individuals and data categories, beyond the restrictive list provided in Annex II.B of the Europol Regulation. Nor does the enhanced review process, foreseen under the Action Plan, prevent the continuous storage of this data, for prolonged time periods.
- The processing by Europol of large datasets lacking a DSC, as defined above, for the sole purpose of extracting relevant information in compliance with Articles 18(3), (5), and Annex II.B of the Europol Regulation, as well as the principles of data minimisation and storage limitation (Article 28(1)(c) and (e) of the Europol Regulation), cannot exceed a maximum period of six months.

5. Proposed corrective measures

As a consequence, on the basis of the preliminary findings and legal analysis as described above, the EDPS is considering **to order Europol, on the basis of Article 43(3)(e) ER, to carry out the destruction of any datasets which do not comply with Article 18(3) and (5) and Annex II.B ER read in the light of Article 28(1)(c) and (e) after six months, as from the date of reception of the contribution.**

This is without prejudice to any follow-up or other actions the EDPS might undertake in the future with regard to the supervision of Europol.

Done at Brussels, 26 July 2021

[e-signed]

Wojciech Rafał WIEWIÓROWSKI

¹⁰ See the Commission Staff Working Document Impact Assessment Report accompanying the 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, SWD(2020) 543, 09.12.2020, in particular Annex I p. 19.