



The Hague, 23 February 2023  
EDOC #1268633v8A

**Binding document defining the general scope for the research and innovation projects (application of Article 8 of the Management Board Decision further specifying procedures for the processing of information for the purposes listed in Article 18(2) (e) of the Europol Regulation)**

**Table of Contents**

1.	Introduction.....	2
2.	Europol Overall Strategic Objectives in relation to R&I projects .....	2
3.	Strategic Goals.....	4
4.	Safeguards to be implemented and decision-making process.....	4
5.	Subject Matters to be addressed by the research and innovation projects .	6
A.	Image and Video Analytics .....	6
B.	Text Analytics.....	9
C.	Audio Analytics .....	9
D.	Data Representation and Processing.....	10
6.	Conclusion .....	11

**PUBLIC**  
Document made public on:  
**31 JUL 2023**

## 1. Introduction

In accordance with Article 18(2)(e) of the Europol Regulation<sup>1</sup> (“the Regulation”), Europol may process personal data for the purpose of research and innovation projects regarding matters covered by the Regulation for the development, training, testing and validation of algorithms for the development of specific tools, and other specific research and innovation projects relevant for the achievement of Europol’s objectives set out in Article 3 of the Regulation.

It is also defined, that if necessary to achieve the objectives of Europol’s research and innovation projects, the processing of personal data for the purpose defined in Article 18(2)(e) of the Regulation<sup>2</sup>, shall be carried out only in the context of research and innovation projects with clearly defined purposes and objectives. These projects shall be subject to the additional safeguards provided for in Article 33a of the Regulation.

Article 33a(3) of the Regulation provides that “The Management Board shall establish in a binding document the general scope for the research and innovation projects. Such document shall be updated where appropriate and made available to the EDPS for the purpose of its supervision.”

This provision is further implemented in Article 8(3) of the MB Decision which states that “the Management Board shall define the general scope for the research and innovation projects that Europol can establish in accordance with Article 33a(3) of the Regulation in a binding document by defining the overall strategic objectives, goals and subject matters to be addressed by the research and innovation projects. The document shall be updated where appropriate and made available to the EDPS for the purpose of its supervision.”

Finally, in accordance with the procedure outlined in Article 8(4) of the MB Decision, the Management Board shall be consulted prior to the launch of those research and innovation projects which are not yet covered by the binding general scope for the research and innovation projects and for research and innovation projects within that scope, the Management Board shall only be informed.

In accordance with the legal provisions mentioned above, this binding document defines the general scope for the Europol research and innovation projects that Europol can establish in accordance with Article 33a(3) of the Regulation by defining the overall strategic objectives and goals to be addressed by the research and innovation projects. The binding document also lists specific subject matters to be addressed by the research and innovation projects.

Operational needs for research and development in the fields mentioned below were gathered by the Europol Innovation Lab, from the EU Member States representatives for innovation (EuCB – European Clearing Board for Innovation) and from Europol staff from the Operations and Capability Directorates.

## 2. Europol Overall Strategic Objectives in relation to R&I projects

Based on the October 2019 JHA Council Conclusions<sup>3</sup> and the Europol Strategy 2020+<sup>4</sup>, Europol aims to “be at the forefront of law enforcement innovation and research” for the benefit of the Member States.

---

<sup>1</sup> Regulation (EU) 2016/794 of the European Parliament and the Council of 11 May 2016 as amended.

<sup>2</sup> Please note that the processing purpose described in Article 18(2)(e) (for research and innovation purposes) entails different processing methods (research and innovation projects) and set of safeguards (Article 33a) than the processing purpose described in Article 18(2)(c) (for operational analysis).

<sup>3</sup> *Outcome of Council Meeting*, Justice and Home Affairs (JHA) Council, Brussels, 7-8 October 2019, 12837/19, EDOC#1074519v1.

<sup>4</sup> *Europol Strategy 2020+*, Vienna, 13 December 2018, EDOC#1000001.

## Europol Unclassified – Basic Protection Level

To that end, the Europol Innovation Lab was created in January 2020 by the Executive Director (ED) with the overall strategic objective to foster innovation in the EU law enforcement community and to support Europol and its Member States in identifying and adopting innovative solutions in close collaboration with all relevant Europol units and departments. The Europol Innovation Lab should act as a central driver for innovation for Member States in the field of internal security.

Furthermore, in order to address the emerging and future challenges in the field of EU Security, Europol's objectives on innovation are directly linked to the strategic priorities as laid out in Europol Strategy 2020. The updated Regulation strengthens the mandate of Europol to support the European law enforcement community in fostering innovation and enabling them to make the most of new technologies and innovative solutions. Europol and specifically the Europol Innovation Lab aims to monitor new technological developments and identify, promote and develop practical solutions supporting the EU Member States' operational work.

The Europol Work Programme 2023<sup>5</sup> foresees that: "As part of the implementation of the Europol Regulation Recast, the Lab will provide the environment for Europol and MS to develop, train and validate AI models and tools using operational data." The draft Europol Work Programme 2024<sup>6</sup> reiterates the same objective and further foresees that "Europol Research and Innovation projects will develop AI tools, trained with data provided by MS for that purpose, to facilitate investigations."

With the R&I projects, Europol and its Member States aim at improving and enriching law enforcement agencies' (LEAs) knowledge and awareness of how new technologies can reinforce their operational capabilities, at both EU and national level. They intend to improve the preparedness of LEAs for investigating, combating, preventing the malicious use of new technologies for criminal activities and terrorism. Europol aims at increasing the capacities of LEAs across the EU, by facilitating the uptake of the results of research and development projects, and bridging the technological skills gap, leveraging the work of relevant past and current projects and the European Union Commission's funding aimed at the fight against crime and terrorism. Europol employs a co-design and co-creation approach to reduce duplication of work and create synergies and collaboration between EU LEA partners.

Europol and its Member States focus on adopting a human-centric approach to new technologies in order to design and develop advanced tools, but carefully addressing potential ethical and legal implications. Europol ensures that the development, use and deployment of new technologies are guided by the principles of transparency, explainability, fairness and accountability<sup>7</sup>, do not undermine fundamental rights and freedoms and are in full compliance with Union law.

Before considering processing personal data within research and innovation projects, Europol and its Member States will ensure that the processing of personal data is strictly required and justified and that the objective of the relevant project cannot be achieved through the use of non-personal data, such as synthetic or anonymous data, and that full respect for fundamental rights, in particular non-discrimination, is ensured.<sup>8</sup>

Similarly, the processing of special categories of personal data for research and innovation purposes will only be considered by Europol where it is strictly necessary. Given the sensitivity of such processing, appropriate additional safeguards, including

---

<sup>5</sup> *Draft Europol Programming Document 2023-2025*, The Hague, 21 October 2022, EDOC#1192222v16.

<sup>6</sup> *Draft Europol Programming Document 2024-2026*, The Hague, 21 October 2022, EDOC#1258316v2.

<sup>7</sup> Europol and 3 other JHA agencies are working together to operationalise these principles in the framework of the project : *Accountability Principles for Artificial Intelligence in the field of Internal Security (AP4AI)* [www.ap4ai.eu](http://www.ap4ai.eu).

<sup>8</sup> In line with the safeguards and requirements listed in Article 33a of the amended Europol Regulation.

pseudonymisation, will be applied. To prevent bias in algorithmic decision-making, Europol should be allowed to process personal data that do not relate to the categories of data subjects listed in Annex II of the Regulation. Europol plays a key role in assisting Member States in developing new technological solutions based on artificial intelligence that are relevant to the achievement of Europol's objectives and that benefit competent authorities of the Member States throughout the Union. That assistance is provided while fully respecting fundamental rights and freedoms, including non-discrimination.

### **3. Strategic Goals**

The strategic goals that Europol aims to achieve with the research and innovation activities over the next years are as follow:

Europol aims:

- to strengthen the technological capabilities of the competent authorities of the Member States to identify, secure and analyse the data needed to investigate criminal offences, by obtaining and if needed developing and disseminating technologies that can help the operational work of Europol and the Member States.
- to provide software solutions that reduce the manual work of operational analysts, in order to enhance efficiency, reduce human errors, quickly identify leads, and improve the workforce productivity.
- to prepare EU LEAs to deal with the increasing volume, heterogeneity and variability of the operational data.
- to support Member States in the use of emerging technologies, in exploring new approaches and in developing common technological solutions for Member States to better prevent and counter crimes.
- to assist the Commission in drawing up and implementing the Union framework programmes for research and innovation activities that are relevant to achieve Europol's objectives.
- to drive innovation and reinforce synergies in research and innovation projects by stepping up cooperation with relevant networks of Member States' practitioners and other Union agencies within their respective competences in that area, and support interagency cooperation such as the EU Innovation Hub for Internal Security, a collaborative network of innovation labs.
- to play a key role in promoting the development and deployment of ethical, trustworthy and human-centric artificial intelligence that is subject to robust safeguards in terms of security, safety, transparency, explainability and fundamental rights.

### **4. Safeguards to be implemented and decision-making process**

Processing of personal data within research and innovation projects falling in the scope of the subject matters mentioned in the list below will only take place when strictly necessary and when the objective of the relevant project cannot be achieved through the use of non-personal data.

Europol shall anonymise data for the development, training, testing and validation of algorithms for the development of specific tools, and other specific research and innovation

projects whenever this is possible for the purpose of the research and innovation project in question. Implementing these essential safeguards also requires the use of AI-tools.

### ***Anonymisation of images (face removal)***

Images that need to be analysed for example to identify an object (Automated object detection) may contain persons, either suspects, victims, or by-standers, who are not of immediate interest for the analysis. As an example, to train a firearms detection algorithm, images in the training dataset will often portray a person holding a gun, but for the training and assessment of the algorithm, the face of the person is irrelevant. A tool can be developed to detect and automatically remove (block) the faces of the persons, so that the anonymised image can be used for further processing, including training or testing of AI-based algorithms. In addition to that, in certain cases it is necessary to automatically remove (block) the faces of specific age groups only, like minors.



*Figure 1 Examples of image anonymisation results.*

### ***Removal of persons from scenes***

Similarly, images of locations or scenes relevant for an investigation, may contain persons, either suspects, victims, or by-standers, who are not of interest for the investigation. As an example, one can use a scene localisation tool from scene textual descriptions. A tool can be developed to detect and automatically remove the persons, while maintaining the integrity of the scene so that the anonymised image can be used for further processing, including training or testing of AI-based algorithms, or sharing it with the public in a request for information. For example, it could help in the Europol Trace an Object campaign, where it would be easier to share the pictures and people could see the whole scene and not just one object.



*Figure 2 Example of automated people removal from scenes.*

In application of Article 33a(5) of the Regulation, Europol will only process data for research and innovation projects with the consent of the provider of the data obtained explicitly via a SIENA message.

Whenever an anonymisation of data for the development, training, testing and validation of algorithms for the development of specific tools, and other specific research and innovation projects, is not possible for the purpose of the research and innovation project in question, prior to launching any processing operations upon personal data in the

## Europol Unclassified – Basic Protection Level

framework of the research and innovation project, Europol will prepare the documentation describing the safeguards to be applied as mentioned in Article 33a(2) of the Regulation.

This documentation will include:

- a description of the objectives of the project with explanation of how the project assists Europol or competent authorities of the Member States in their tasks,
- envisaged processing activity,
- objectives, scope and duration of the processing,
- the necessity and proportionality to process the personal data,
- description of categories of personal data to be processed,
- assessment of the compliance with the data protection principles,
- time limit for storage,
- conditions for access to personal data,
- result of a data protection impact assessment (DPIA), including risks to rights and freedoms of data subjects, risks of any bias in the personal data to be used for the training of algorithms and in the outcome of the processing, measures envisaged to address those risks and violation of human rights.

This documentation will be submitted to:

- the Europol Executive Director for prior authorisation, in consultation with the Data Protection Officer and the Fundamental Rights Officer,
- both the MB and the EDPS will be informed prior to the launch of the research and innovation project.

Europol will keep logs of the processing of personal data in the context of the research and innovation project until two years after the conclusion of it, solely for the purpose of and only as long as necessary for verifying the accuracy of the outcome of the data processing as stated in Article 33a(2)(e) of the Regulation.

### **5. Subject Matters to be addressed by the research and innovation projects**

In order to achieve the overall strategic objectives and goals mentioned in this binding document, Europol intends to establish research and innovation projects on subject matters mentioned in the list below in application of Article 33a(3) of the Regulation and of Article 8 of the MB Decision.

The subject matters are clustered around four functionalities: A) Image and Video Analytics, B) Text Analytics, C) Audio Analytics, D) Data Representation and Processing.

#### **A. Image and Video Analytics**

Video and image analytics aim to use computer vision algorithms and software tools to extract information and knowledge from video and images files, and reduce the need for manual processing of these large volumes of files. The section below provides an overview of the research and innovation projects Europol and its Member States aim to launch in the field of automated video and image analysis in order to develop and train tools which could help Europol and its Member States to derive actionable information, and collect useful operational insights, optimising, therefore, the operational value derived from the video and images.

## **1. Pre-processing and preparation**

Initiatives under this section aim to pre-process and prepare the data for further analysis and information extraction. This for example can aim to improving the quality of datasets, or removing sensitive information from the dataset, if it is not specifically needed for the further processing.

### ***Video and Image Super Resolution***

Video Super-Resolution is incorporating methods to generate high-resolution video frames from the original ones with lower-resolution, enhancing the fine details in the video while preserving the motion characteristics. Super-resolution imaging similarly concerns a set of methods that enhances the resolution of images. Law enforcement is often in possession of images and videos of low quality, and the planned work aims at developing methods and tools to improve them.

### ***Video/image similarity detection***

Detection of duplicate images and videos, or partial usage of known videos, is critical for Europol and MS. Europol and MS will explore technologies (both commercial and open source) to put in place a reliable system for de-confliction and detection.

## **2. Object Detection**

Object detection aims at locating instances of predefined objects in images or video frames. Europol and MS intend to work on automatically detecting objects such as guns, rifles, knives, drugs, money, criminal organisation and terrorist group logos, flag detection (i.e. ISIS flags in jihadi related videos), container labels, cocaine logos, identity documents or license plates. In the future, Europol and MS also plan to investigate the possibility to use satellite images for segmentation and object detection (i.e. in order to locate illegal garbage dumps and marijuana plantations).

## **3. Image similarity search**

This task aims at facilitating advanced search scenarios, in which analysts can search for images similar to a given image. Models trained for specific contexts can facilitate more detailed scenarios, like searching for images depicting similar cocaine logos or similar weapons. This type of technology can also facilitate and inform other techniques, like clustering.

## **4. Image Classification**

The purpose of image classification is to identify what information is represented in an image. An image classification model can be trained to recognise various classes of images. For example, Europol and MS intend to train models to recognise and classify CSEM (child sexual exploitation material) images, and detect and classify memes used in CSE and estimate the severity of the content.

There is also the need for image classification for classes such as drugs and drug types (cocaine, heroin, hashish, cannabis, cannabis plantations, synthetic drugs, etc.), money stashes, bank notes, weapons, ammunition, locations (satellite images, maps, street signs, etc.), smartphone screenshots (chat conversations, contact details, etc.), identity documents, etc.

## **5. Face Recognition**

Face recognition technology is part of biometrics, and is based on the identification of facial features of a person. Faces are detected in operational images and matched against databases of known suspects. Europol and its Member States plan to launch research and

innovation projects in this field, in full compliance with the safeguards to be applied as mentioned in Article 33a(2) of the Regulation.

#### **6. Personal features extraction**

The aim of such initiatives would be to detect personal features on images, such as the age, the gender, or origin of a victim in order for example to detect CSEM material and provide an indication of the potential location of the victim and the crime. This information helps the investigator detect minors at risk, prioritise cases and search cases with similar victim characteristics. Human corroboration of the information extracted is a pre-requisite for the further use of such extracted information in the investigation.

#### **7. Fingerprint detection from fingertips in images**

Europol and its Member States are interested in developing new methods to extract and match fingerprints from images that include fingertips. This process involves multiple steps, such as detecting and segmenting the fingertips, aligning them, segmenting the fingerprints, extracting and storing minutiae and texture representations and matching them.

#### **8. Scene recognition**

Europol and its Member States are interested in developing tools to automatically describe what happens in an image or a video scene. The results would help investigators to relate the image with other facts and situations and would participate in the automatic extraction of information from media. Recognising scenes would enable a pre-classification of media that typically takes a long time to analyse (for instance long videos) to make the selection of the most relevant media, much faster.

#### **9. Scene classification**

Europol and its Member States are planning to work on automated scene classification, for example the automated violence detection, explosion detection or detection of torture scenes in terrorist propaganda videos.

#### **10. Indoor and outdoor image geo-localisation**

The aim of such work will be to help identify the exact location the image or video is taken from the information portrayed on it. For an indoor scene items such as the electric plugs, objects and products in the scene can be extracted to identify the country. In outdoor scenes, the aim is to identify landscapes and landmarks in image or video so that possible geo-localisation is achieved.

#### **11. Optical character recognition (OCR)**

Optical character recognition even if it is considered an established technology, still require further work for specific challenging cases. For example automatically extracting handwritten text, or extracting text from outdoor scene-photo (for example the text on signs and billboards) to be used for geo-localisation of the image. Another challenging example is extracting text from low quality images of identity documents, when reflections, flash glares impact the performance of the commercial text extractors.

#### **12. Deep fake and image/video tampering detection**

The aim of this work will be to detect manipulated images and videos using the current state-of-the-art methods.

## **B. Text Analytics**

Text analytics is the field of automated processing of large volumes of unstructured text and its transformation into structured information in order to store it, and analyse it further. It is often combined with data visualisation tools, in order to enable investigators to understand the events and entities involved and make sense of large volumes of such data and identify leads.

### **13.Named entity extraction**

Named-entity recognition and extraction entails the entity identification, localisation and classification of named entities mentioned in unstructured text into pre-defined categories such as person names, organisations, locations, events, means of communication, financial information, monetary values, etc. As the volume and variability of the data Europol receives and processes continues to grow, the entity extraction techniques will need to be continuously updated and improved.

### **14.Link extraction**

Link extraction is related and following the entity extraction, where the links between extracted entities are automatically extracted and presented to the analyst for corroboration.

### **15.Robust translation of informal text**

Commercial text translators exist for most common languages but often rely on online tools, while LEAs require software that are securely installed on premises, and are unable to use online translation services. Furthermore, certain languages are still difficult to translate, especially when different dialects or regional peculiarities exist on the text. Furthermore, law enforcement is often in possession of informal conversations, either from chats or forums. The understanding and translation of such text remains a major challenge, especially due to the use of slang, abbreviations, and coded language.

### **16.Text Classification**

The purpose of text classification is to identify specific content and intentions in the text. Text classification models can be trained to recognise various types of text, for example detect grooming, threat to life from chats and forum discussions. For example, Europol and its MS intend to investigate the feasibility of such models to recognise and classify suspected grooming from CSEM forums, and detect and threat to life and cases of priority.

### **17.Stylometric analysis**

Stylometric analysis is studying the linguistic style, in written language in order to detect texts that are potentially written by the same user. Europol and its MS intend to investigate the feasibility of such models to recognise and group grooming activities from CSEM forums. Combined with other techniques, like facial recognition, voiceprint analysis or graph analytics, stylometric analysis can help identify accounts managed by the same user.

## **C. Audio Analytics**

Audio Analytics is used to extract information from audio files. As a first step, it is often required to improve the quality of the audio and the signal-to-noise ratio, before further analysis.

### **18.Speech to text in low quality audio**

Europol and its Member States are aiming at further developing technical capabilities for speech-to-text transcription in low-quality audio and translation from various languages and dialects so that the content can be further processed. For example, detecting and extracting the narrative in video and audio of terrorism propaganda in order to identify if the video has been used before, by whom, and link it to a specific terrorist group.

### **19.Language Detection**

Detecting the language from low-quality audio, including very short segments of speech from video (less than 5 seconds), or from screams or whispers is needed in order to help identify the location and the language spoken by perpetrators and victims in cases of CSEM.

### **20.Age and gender detection from voice**

The aim of this work would be to detect personal features in voices, such as the age, the gender of a victim or the perpetrator. This information would help the investigator detect the location of the crime (for example in relation to CSE), link it to other cases, identify minors at risk. Human corroboration of the information extracted is a pre-requisite for the further use of such extracted information in the investigation.

### **21.Voiceprint analysis**

Voiceprint analysis is the processing of the spectrum of a voice sample in order to compare it with other known or unknown voice spectrums. The methods are based on the analysis of frequency, and intensity of the sounds caught on the spectrograph and Europol and MS wish to investigate the feasibility of such methods to link perpetrators. Human corroboration of the information extracted and any possible links would be a pre-requisite for the further use of such extracted links in any operational work.

### **22.Accent recognition in voice**

This aims to detect personal features on the voice, the accent and the origin of a victim or the perpetrator. This information is to help the investigator detect the location of the crime (for example in relation to CSE), link it to other cases, and help locate minors at risk. Human corroboration of the information extracted would be a pre-requisite for any further use of such extracted information in an investigation.

## **D. Data Representation and Processing**

This work is also linked to the Data Refinery Area concept including, ways to process large volumes of data both structured and unstructured, in a forensically sound environment, data enrichment with OSINT, commercial databases and internal resources, creation and enrichment of ETL pipelines<sup>9</sup> including graphical, SQL (Structured Query Language) and relational databases, and data visualization tools including Virtual Reality/Augmented Reality (VR/AR).

### **23.Graph Analytics**

Graph analytics is used to analyse very large volumes of data and is seen as an alternative to the traditional data representation models. It offers a framework to absorbing both structured and unstructured data from various sources and enables investigators to view

---

<sup>9</sup> ETL stands for "Extract Transform and Load". It is the pre-processing, clearing and normalisation process necessary before data can be loaded to databases.

and interrogate the data in a non-prescribed manner. Europol and its Member States wish to further explore the use of the technology, to enhance EU LEA's analytical capabilities.

#### **24. Clustering in network of nodes**

Clustering nodes (i.e. elements) in a network consists in the automatic identification of communities based on the network topology and/or properties of the nodes. This clustering would enable, for instance, the discovery of potential associates in an organisations for which a certain numbers of members are already known or a better understanding of the "subgroups" constituting a certain social network.

#### **25. AR/VR data analysis**

Europol and its Member States aim to investigate the use of Augmented reality (AR) and Virtual Reality (VR) data analysis, in order to improve the analysis process and offer the investigator a new and interactive experience to visualise and interrogate the data.

## **6. Conclusion**

In accordance with Article 8(3) of the MB Decision, this binding document shall be updated where appropriate and made available to the EDPS for the purpose of its supervision.

In accordance with Article 8(4) of the MB Decision, the Management Board shall be informed prior to the launch of those research and innovation projects covered by the scope of this binding document.

The Management Board shall be consulted prior to the launch of those research and innovation projects, which are not covered by the scope of this binding document.

Done at The Hague on 22 March 2023

Jérôme Bonet  
Chairperson of the Management Board



