



A European Positive Sum Approach towards AI tools in support of Law Enforcement and safeguarding privacy and fundamental rights

Practical Ethics Toolbox for the Use of AI by LEAs: Technology Ethics Briefs



This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 101022001.

Executive Summary of the Ethics Toolbox

Practical Ethics Toolbox for the Use of AI by LEAs includes a taxonomy of ethical frameworks and critical analysis of their gaps and implementability and guidelines for the work of T3.2. T2.4 from ethical frameworks to ethics in practice [M3-M16] is developed by ERI, TRI, TU/e and NCSR partners, where the current ethical frameworks in the LEA and AI space will be documented in a systematic, extensible taxonomy, drawing on published materials and internal reports used by security actors and technology providers. This toolbox highlights the main common points between frameworks, identifies differences and challenges of the AI implementation through LEAs contexts. The task therefore will include a novel taxonomy of ethics principles in the LEA/AI space, but also identify gaps and challenges that will feed into WP3, where stakeholder attitudes will be explored and tested. The final outcome responds to the specific needs of LEAs currently and in the near future. The practical ethics toolbox will explore, for instance, how to determine the acceptable amount of FPR and FNR in AI security solutions, or establish the baseline needs of algorithmic auditing to fulfil the legal requirement for algorithmic explainability, providing specific examples and cases.

List of Terms & Abbreviations

Abbreviation	Definition
AI	Artificial Intelligence
AI HLEG	High Level Expert Group on Artificial Intelligence
D	Deliverable
DPA	Data Protection Authority
DPIA	Data Protection Impact Assessment
FRT	Facial Recognition Technologies
HRIA	Human Right Impact Assessment
LEAs	Law Enforcement Agencies
UN	United Nations
VR	Virtual Reality
WP	Work Package
GDPR	General Data Protection Regulation

1. Introduction

PopAI is an EU-funded project within the H2020 framework, addressing the topic SU-AI03-2020: Human factors, and ethical, societal, legal and organisational aspects of using Artificial Intelligence in support of Law Enforcement. It began on the 1st of October 2021 and will finish on the 1st of October 2023.

Pop AI aims to foster trust in AI for the security domain via increased awareness, ongoing social engagement, consolidating distinct spheres of knowledge (including theoretical & empirical knowledge by academics & non-academics) and offering a unified European view across Law Enforcement Agencies (LEAs), and specialized knowledge outputs (recommendations, roadmaps, etc.), while creating an ecosystem that will form the structural basis for a sustainable and inclusive European AI hub for Law Enforcement.

PopAI is divided into six Work Packages (WPs). This document includes part of the fifth deliverable within Work Package 2, titled “Security AI in the next 20 years: trends, practices and risks”. It is based on work carried by the following partners: Eticas Research and Innovation (ERI), National Centre for Scientific Research Demokritos (NCSR), Trilateral Research Ltd (TRI) and Technische Universiteit Eindhoven (TU/e). The aim of Ethics Toolbox is to document in a systematic and extensible taxonomy all the possible published materials and internal reports used by security actors and technology providers regarding LEAs and AI. In this sense, D2.5 Practical ethics toolbox for the use of AI by LEAs (M16), integrates the work done in multiple deliverables with a more applicable and practical perspective for the use of LEAs.

The Practical Ethics Toolbox respond to the crucial necessity of AI introduced in Law Enforcement with the potential to generate benefits in efficiency and accuracy, bringing positive change to European LEAs operational capacity. However, at the same time, AI generates also great risks for fundamental rights and democracy. To mitigate risks and create a just, sustainable and inclusive European AI culture for Law Enforcement, it is important to understand what AI ethics is and how it applies to Law Enforcement.

The Toolbox becomes the first resource that integrates the multiple perspectives of Ethics and AI through LEAs space. In this sense, the ethics toolbox is composed by three main products: eight educational videos, taxonomy collection and technological ethics briefs. This technological briefs are composed by a series of three practical cases.

The briefs explore how artificial intelligence is used in law enforcement, specifically focusing on predictive analytics, natural language processing (NLP), and image recognition. The aim of the briefs is to provide law enforcement personnel with better awareness and information about the various uses and applications of artificial intelligence, including the potential benefits and risks of these technologies in policing, and to consider the ethical implications of their use.



D2.5: Practical Ethics Toolbox for the Use of AI by LEAs

The ethical briefs would cover the definition, examples of usage, and ethical considerations related to each AI application. For instance, the predictive analytics brief would include details on its use in policing, ethical implications, and examples of its usage in Europe. The aim of these briefs is to provide a comprehensive overview of technology ethics related to specific AI applications, enabling better-informed decision-making.

Finally, **Section 3** would serve as a basis for creating an **interactive taxonomy visualization** on the PopAI web page of the four AI taxonomies: ethics, legal, functionalities, and social controversies. Each taxonomy would provide a categorization framework for better understanding the field of AI and LEAs ethics. For example, the ethics taxonomy would include a collection of relevant documents about AI in policing from an ethical perspective. Users can search the taxonomies around the visualization in the four different categories. This tool pretends to be a hub of multiple documents or cases valuable for researchers, practitioners, and policymakers. The main objective of this visualization is to help the multiple final users to get a quick, accessible and useful tool to get the main discussions, applications and cases of AI and LEAs.

2. Technology Ethics Briefs

The rapid advancement of technology brings with it countless opportunities to improve various aspects of society, such as predicting crime hotspots and digitizing old documents, to enhance public safety and efficiency. However, the use of technology in these areas also raises ethical dilemmas and considerations that must be carefully examined. This is particularly true in the field of law enforcement, where the application of technologies like predictive analytics, image recognition, and natural language processing can significantly impact privacy, civil liberties, and the potential for biased outcomes. In this series of three ethics briefs, we will explore how artificial intelligence is used in law enforcement, specifically focusing on predictive analytics, natural language processing (NLP), and image recognition. Our aim is to provide law enforcement personnel with better awareness and information about the various uses and applications of artificial intelligence, including the potential benefits and risks of these technologies in policing, and to consider the ethical implications of their use.

2.1 Predictive Analysis

What is Predictive Analytics (in Policing)?

Predictive analytics in policing is the use of advanced data analysis techniques to identify patterns and predict future events related to crime and public safety. These techniques involve the application of machine learning algorithms to historical crime data, along with other relevant data to identify potential crime hotspots or individuals who may be at risk of committing a crime.

Some examples of how predictive analytics can be used in policing include:

- **Predicting crime hotspots:** By analysing historical crime data and other relevant data sources, predictive analytics can identify areas that are at higher risk of crime. This information can be used to guide the deployment of police resources to those areas, allowing law enforcement agencies to be more proactive in preventing crime.
- **Predicting repeat offenders:** Predictive analytics can be used to identify individuals who are at higher risk of committing a crime based on their past behaviour. By focusing resources on these individuals, law enforcement agencies can intervene before a crime is committed.
- **Identifying emerging crime trends:** By analysing data on crime patterns and trends, predictive analytics can help law enforcement agencies identify new or emerging crime trends that may require a proactive response.

Overall, the use of predictive analytics in policing holds the promise to improve public safety and enhance the effectiveness of law enforcement agencies, while also raising important questions about bias, privacy, fairness, and accountability.

How is predictive analytics being used in Europe?

- **ProKid 12_SI** is used by the Dutch police and intends to assess the criminality of children under 12 years old and is used. ProKid employs a combination of historical data and statistical analysis to assign a risk to a child. It utilizes available police data on children, including information on their address, police contacts, and "living environments" which includes information about parents and cohabitants, their police contacts, and whether they have been victims of violence, to categorize them into one of four colour-coded risk categories for future criminal activity. The system assesses children based on their relationships with others and their supposed risk levels, allowing individuals to be classified as higher risk if they are associated with others who have a high-risk assessment, such as siblings or friends. Additionally, a parent's assessed risk can also affect a child's risk level (La Fors, 2020).
- **Key Crime Delia** has been used by the Italian State Police in Milan. The software analyses data on past criminal events to predict future events and help allocate police resources. Based on previous crimes, targets and repeating patterns Delia purports to help to plan their routes Radar -ite The German Federal Criminal Police Office and the University of Konstanz have developed a risk assessment tool to assess the risk of individuals committing serious acts of violence related to terrorism.
- **RADAR-iTE** compares an individual's behaviour with data on the behaviour of people who committed crime. The evaluation procedure consists of two stages. The initial stage involves the police officer in charge collecting all available information about the extremist individual, while the second phase comprises a questionnaire consisting of 73 questions. The questionnaire includes queries about personal and social life events, social network, proof of terrorism-inspired travel, history of violence, with response options of "Yes," "No," or "Unknown."

Ethical dilemmas and considerations

The use of predictive analytics in policing raises important ethical considerations that need to be carefully addressed. One major concern is the potential for **biased outcomes**, particularly if the algorithms used to analyze the data are trained on biased historical data. This could lead to unfair profiling of certain groups, exacerbating existing social inequalities. Another concern is the impact on **individual privacy**, as predictive analytics relies on collecting and analyzing large amounts of personal data. There is also a risk that individuals may be **stigmatized or falsely accused** based on predictions that are not accurate. Additionally, there are questions around **accountability and transparency**, particularly regarding the use of proprietary algorithms and the lack of oversight or regulation. **Finally**, the use of predictive analytics can lead to a shift in the role of police officers from being reactive to proactive, potentially leading to a loss of community trust if not implemented in a fair and transparent manner.

2.2 Natural Language Processing (NLP)

What is Natural Language Processing (in Policing)?

Natural language processing (NLP) is the field of AI covering knowledge and techniques involved in the processing of linguistic data by a computer (ISO, 2019), used to determine and identify key words and phrases within processing audio data (e.g. call centres) and free-form text (e.g. the body of an email) (ISO, 2018; ISO, 2021). The purpose of this technology is to make machines capable of reading and reasoning with human language and therefore, automatically process it. Some common tasks involving NLP include information extraction, document categorization and semantic text matching (European Commission, 2022), which can help police agencies in crime investigations, intelligence gathering, and public safety operations. The ultimate goal of NLP in policing is to improve the efficiency and accuracy of police operations by automating the analysis of linguistic data and facilitating the decision-making process.

Some examples of how NLP can be used in policing include:

- **Crime analysis:** Police agencies can use NLP algorithms to analyze crime data and extract useful insights, such as identifying patterns and trends in criminal activity, analyzing the language used in social media posts to detect threats or identify potential suspects, and identifying high-risk areas for criminal activity based on text data from various sources.
- **Investigative support:** NLP can be used to automatically extract and categorize data from police reports, interviews, and other sources, making it easier for investigators to quickly identify key information and potential suspects. NLP can also help in identifying the language patterns used in crime reports, such as determining the level of risk and severity of the incident, which can help in prioritizing investigations.
- **Community engagement:** Law Enforcement Authorities can use NLP algorithms to analyze social media and other online platforms to gauge public sentiment and detect potential threats or unrest. This can help authorities to anticipate and respond to potential public safety issues, as well as to identify areas where community outreach and engagement may be needed.

Overall, NLP has a range of potential applications in Law Enforcement, from improving crime analysis and investigative support to enhancing community engagement and predicting the risk of recidivism. As NLP technology continues to evolve, it is likely that we will see even more innovative and effective uses of this technology in the field of Law Enforcement.

How is NLP being used in Europe?

- **VeriPol**, is a model developed for the detection of false robbery reports based solely on their text, which combines Natural Language Processing and Machine Learning methods in a

decision support system¹. The tool was developed in collaboration with the Spanish National Police and provides police officers the probability that a given report is false. The tool has been tested on more than 1000 reports from 2015 provided by the Spanish National Police, and empirical results show that it is extremely effective in discriminating between false and true reports with a success rate of more than 91% (Quijano-Sánchez et al., 2018).

- **Automatic speech analysis software** is being used by Germany's Federal Office for Migrants and Refugees to verify the dialects of refugees seeking asylum, with the aim of reducing the number of language analysis experts needed for the process. The software works by analyzing the refugees' speech patterns and providing a report on the probability of the dialect they speak².
- **Kamu**³, a chatbot developed by the Finnish Immigration Service that helps citizens with general advice related to immigration processes and application procedures. The chatbot uses NLP and machine learning algorithms to understand user queries and provide appropriate responses, enabling the Immigration Service to provide more efficient and timely support to citizens.
- The **Swedish Land Registry (SLR)** developed by border authorities to increase efficiency in handling land registry requests⁴. Many old handwritten documents cannot be digitized easily as the text is hard to read. The SLR uses NLP and AI to analyze and interpret these documents, thereby making the process of digitization faster and more accurate. This has improved the quality of service provided to citizens by reducing the processing time for requests and ensuring the accuracy of records.

Ethical dilemmas and considerations

When implementing NLP technology in policing or other fields, it is important to consider the ethical implications and potential risks associated with its use. For instance, there may be concerns regarding **privacy** and the use of **personal data**, as well as the **potential for bias and discrimination** in the algorithms used. Additionally, there is a **risk of over-reliance** on NLP technology, which could lead to a decrease in human decision-making and critical thinking skills. Therefore, it is crucial to ensure that NLP systems are designed and implemented with transparency, fairness, and accountability in mind. Policymakers, law enforcement agencies, and developers should work together to establish ethical frameworks and guidelines for the use of NLP technology in order to ensure its responsible and effective use in society.

1 <https://eticasfoundation.org/veripol-spotting-false-complaints-made-to-the-police/>

2 <https://www.dw.com/en/automatic-speech-analysis-software-used-to-verify-refugees-dialects/a-37980819>

3 <https://migri.fi/en/chat1>

4 <https://www.gov.uk/government/case-studies/natural-language-processing-for-land-registry-documentation-in-sweden>

2.3 Image Recognition

What is Image Recognition (in Policing)?

Image recognition or image classification is a computer vision process that involves identifying, detecting, and categorizing objects, patterns, or concepts in digital images or photographs. This process typically involves the use of machine learning algorithms and deep neural networks to train a model on a large dataset of labeled images. Once trained, the model can be used to analyze and classify new images based on the patterns and features it has learned. Google's definition of image recognition or image classification is "*a process that classifies object(s), pattern(s), or concept(s) in an image.*"⁵ This definition emphasizes the key role of classification in image recognition, which involves identifying and categorizing objects or patterns within an image. The ISO/IEC 22989:2022 standard provides a similar definition of image recognition, describing it as "*a process that classifies object(s), pattern(s), or concept(s) in an image.*" However, the ISO standard goes into more detail, defining image classification as "*a process that assigns a label or multiple labels to an image based on its content or a portion of its content.*" This definition emphasizes the role of labels in image classification, which are used to identify and categorize different objects, patterns, or concepts within an image.

Thus, image recognition in policing is a technology that utilizes computer vision algorithms and machine learning to analyze and classify images or video footage obtained from cameras or other visual sources. This technology can help law enforcement agencies in a variety of ways, such as identifying suspects, tracking criminal activity, and monitoring public safety. Image recognition systems can be used to automatically detect and recognize faces, license plates, and other objects of interest, allowing police to quickly and accurately identify individuals and track their movements. In addition, these systems can be trained to identify specific types of criminal behavior, such as gun violence or drug trafficking, enabling police to respond more quickly and effectively to these incidents. However, image recognition in policing also raises concerns around privacy and potential biases in the algorithms, which must be carefully monitored and regulated to ensure that they are used in a fair and ethical manner.

Some examples of how Image Recognition can be used in policing include:

- **Crime Scene Analysis:** Image recognition technology can be used to analyze crime scenes and help investigators identify potential evidence, such as weapons or other objects. This can help to solve crimes more quickly and efficiently.
- **Crowd Surveillance:** Image recognition can be used to monitor crowds and identify individuals who may be acting suspiciously. This can be useful in detecting potential threats or criminal activity.

⁵ See, <https://developers.google.com/machine-learning/glossary/image>

- **Predictive Policing:** Image recognition can also be used in conjunction with machine learning algorithms to predict areas where crime is likely to occur based on past data. This can help police departments allocate resources more effectively.

Overall, image recognition technology is increasingly being used in policing to aid in investigations and improve public safety. The technology can be used for various purposes such as facial recognition, license plate recognition, crime scene analysis, crowd surveillance, and predictive policing. However, it's important to ensure that the technology is used ethically and with appropriate safeguards in place to protect individual privacy and civil liberties.

How Image Recognition is being used in Europe?

- **ANPR system**, which stands for Automatic Number Plate Recognition, is a software used by the Dutch police that is designed to recognize and read number plates of vehicles⁶. It has a wide range of applications including paid parking, traffic control, tax avoidance, environmental protection and many more. This software is currently being used in over 2000 administrations and is considered to be a powerful and versatile AI technology.
- **Henri Coandă International Airport** in Bucharest has awarded a contract worth EUR 4.675.000 (excluding VAT) for the modernization and expansion of its video surveillance system⁷. As part of the project, a face recognition system will be implemented to alert security personnel of the airport when a wanted person or terrorism suspect listed by intelligence services is detected. The system aims to help detect and prevent terrorism, and it will be integrated with the existing video surveillance cameras at the airport. In addition, an intelligent image analysis system will be implemented to detect suspicious behavior and alert security staff. The security systems deployed will use state-of-the-art technology, while maintaining data protection and confidentiality mechanisms in compliance with GDPR regulations.
- **S.A.R.I.** (Sistema Automatico di Riconoscimento Immagini), or Automated Image Recognition System, is an algorithmic facial recognition tool used by Italy's national police force. The software has the capacity to process live footage and identify recorded subjects through a process of facial matching (AlgorithmWatch 2019). In 2018, SARI made headlines when it correctly identified two burglars in Brescia as a result of its algorithmic matching process. Despite being successful in that instance, questions have been raised regarding the accuracy of the software (the risks it poses to justice and safety due to a susceptibility to create false-positives and false-negatives), as well as cybersecurity and privacy guidelines it follows (AlgorithmWatch, 2019).

⁶ <https://www.politie.nl/onderwerpen/anpr.html>

⁷ <https://www.uti.eu.com/press-releases/uti-will-implement-face-recognition-and-intelligent-image-analysis-systems-with-otopeni-airport/>

Ethical dilemmas and considerations

When it comes to image recognition in policing, there are ethical dilemmas and considerations that need to be addressed. In Europe, image recognition technology is being used in various applications, such as ANPR systems, face recognition systems at airports, and algorithmic facial recognition tools used by national police forces. However, there are concerns around **privacy** and **potential biases** in the algorithms, which must be carefully monitored and regulated to ensure that they are used in a fair and ethical manner. Moreover, the **accuracy of the software** and its susceptibility to create false-positives and false-negatives pose risks to justice and safety, which need to be addressed through proper cybersecurity and privacy guidelines. Therefore, it's essential to ensure that the technology is used ethically and with appropriate safeguards in place to protect individual privacy and civil liberties.

3. References

- AlgorithmWatch. (2019). Automating Society Report 2019: Taking Stock of Automated Decision-Making in the EU. Retrieved from https://algorithmwatch.org/de/wp-content/uploads/2019/02/Automating_Society_Report_2019.pdf
- European Commission (DIGIT, Directorate-General for Informatics). (2022). Natural Language Processing for Public Services. Retrieved from https://joinup.ec.europa.eu/sites/default/files/inline-files/D02.01_Natural%20Language%20Processing%20for%20Public%20Services_4.pdf
- International Organization for Standardization. (2018). ISO/TR 22957:2018(en), Artificial intelligence -- Categorizing and tagging [Technical Report]. <https://www.iso.org/standard/64728.html>
- International Organization for Standardization. (2018). ISO 31000:2018 Risk management – Guidelines. <https://www.iso.org/standard/74296.html>
- International Organization for Standardization. (2019). ISO 24613-1:2019(en), Language resource management -- Semantic annotation framework (SemAF) -- Part 1: Time and events [Standard]. <https://www.iso.org/standard/72582.html>
- International Organization for Standardization. (2021). ISO/TR 24291:2021(en), Artificial intelligence -- Learning, education and training -- Guidance [Technical Report]. <https://www.iso.org/standard/78787.html>
- La Fors, K. (2020). Legal Remedies For a Forgiving Society: Children's rights, data protection rights and the value of forgiveness in AI-mediated risk profiling of children by Dutch authorities. *Computer Law & Security Review*, 38, 105430. <https://doi.org/10.1016/j.clsr.2020.105430>
- Quijano-Sánchez, L., Liberatore, F., Camacho-Collados, J., & Camacho-Collados, M. (2018). Applying automatic text-based detection of deceptive language to police reports: Extracting behavioral patterns from a multi-step classification model to understand how we lie to the police. *Knowledge-Based Systems*, 150, 102-116. <https://doi.org/10.1016/j.knosys.2018.03.010>