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

D4.1: White paper for LEAs

Grant Agreement ID	101022001	**Acronym**	popAI
Project Title	A European Positive Sum Approach towards AI tools in support of Law Enforcement and safeguarding privacy and fundamental rights		
Start Date	01/10/2021	**Duration**	24 Months
Project URL	https://www.pop-ai.eu/		
Document date	30/09/2023		
Nature	R: Document, report	**Dissemination Level**	PU: Public
Authors	Georgia Melenikou (KEMEA), Dimitra Papadaki (KEMEA)		
Contributors	Vasiliki Zomenou (KEMEA), Panagiotis Douris (KEMEA), Luis Gonzalez (ETICAS), Claire Morot-Sir (ECAS), Xenia Ziouvelou, Dimitris Kyriazanos (NCSRD), Pinelopi Troullinou (TRI), Eleonora Fiori (PLTO), Max Hausner (BayHfoeD), Paola Fratantoni (Z&P)		
Reviewers	Anthi Bania (HPOL), Lilian Mitrou (External Ethics Advisor), Andreas Ikonomopoulos (NCSRD)		

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

Executive Summary

Nowadays, the advancements in the field of Artificial Intelligence (AI) as well as the number and variety of its applications keep increasing at a rapid pace. The dynamics and the impact of AI on the society and the environment necessitate the existence of harmonised legal rules that will accompany the technological developments in order to mitigate the potential risks and smoothly incorporate the AI-based technologies into our lives. Therefore, the role of policymakers is crucial and decisive. At the same time, apart from the risks that may derive from new technological solutions, the benefits are also indisputable. One of the examples is the high benefits for Law Enforcement Agencies (LEAs) that will use AI as part of their everyday tasks and operational activities. LEAs will start deploying more and more AI applications to assist them in carrying out law enforcement activities and meeting the demanding needs of their job more efficiently. Due to the position of LEAs in the society and the significant degree of power imbalance that characterises their actions, it is critical that AI technologies are deployed in a way that will not only facilitate the work of LEAs but will also -most importantly- prioritise fundamental rights and foster the trust of citizens.

This deliverable presents and analyses recommendations for LEAs and policymakers related to the ethical use of AI in law enforcement. The recommendations are created following a specific methodology. The opinions of citizens, technology developers, ethics and legal experts, other interested stakeholders and of the LEAs as such have been taken into consideration for the production of emerging best practices in line with the current ethical and legal framework. As a result, a White Paper in the form of a public report is delivered to the European Commission.

Table of Contents

Table of Figures

List of Terms & Abbreviations

Abbreviation	Definition
AI	Artificial Intelligence
AFRIA	ALIGNER Fundamental Rights Impact Assessment
AIA	Artificial Intelligence Act
AI HLEG	High-Level Expert Group on Artificial Intelligence
ALTAI	Assessment List for Trustworthy AI
CAHAI	Ad hoc Committee on Artificial Intelligence
CCTV	Closed-Circuit Television
Charter	Charter of Fundamental Rights of the European Union
CO	Confidential (amongst the Consortium and the EC)
CoE	Council of Europe
Council	Council of the European Union
CSAM	Child Sexual Abuse Material
DPIA	Data Protection Impact Assessment
DPO	Data Protection Officer
EAB	Ethics Advisory Board
EC	European Commission
EP	European Parliament
ESIA	Ethical and Social Impact Assessment
EU	European Union
GA	Grant Agreement
GDPR	General Data Protection Regulation
HRESIA	Human Rights, Ethical and Social Impact Assessment
HUDERIA	Human Rights, Democracy, and the Rule of Law Impact Assessment
LEA	Law Enforcement Agency
LED	Law Enforcement Directive
PU	Public
SAB	Stakeholder Advisory Board
SMEs	Small and Medium-sized Enterprises

Terminology

"artificial intelligence system" (AI system) means:

software that is developed with one or more of the techniques and approaches listed in Annex I (of the AI Act Proposal)[1] and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with (definition proposed by the European Commission)[2];

a system that is designed to operate with elements of autonomy and that, based on machine and/or human-provided data and inputs, infers how to achieve a given set of objectives using machine learning and/or logic- and knowledge-based approaches, and produces system-generated outputs such as content, predictions, recommendations, or decisions influencing the environments the system interacts with (definition proposed by the Council of the European Union)[3];

a machine-based system designed to operate with varying levels of autonomy and that can, for explicit or implicit objectives, generate outputs such as predictions, recommendations, or decisions, that influence physical or virtual environments (definition proposed by the European Parliament)[4].

"affected person" or **"AI subject"** means any natural person or group of persons who are subject to or otherwise affected by an AI system[5].

"Draft AI Act" or **"AI Act Proposal"** is the Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, 21.4.2021 COM (2021) 206 final 2021/0106 (COD) as issued by the European Commission and as it has been amended so far by the Council of Europe and the European Parliament.

"instructions for use" means the information provided by the provider to inform the deployer of an AI system's intended purpose and proper use, as well as information on any precautions to be taken; inclusive of the specific geographical, behavioural or functional setting within which the high-risk AI system is intended to be used[6].

"law enforcement" means activities carried out by law enforcement authorities or on their behalf for the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, including the safeguarding against and the prevention of threats to public security[7].

[1] The techniques and approaches listed in Annex I are the following:
(1) Machine learning – including supervised, unsupervised and reinforcement learning;
(2) Logic- and knowledge-based approaches;
(3) Statistical approaches, Bayesian estimation, search and optimisation methods.

[2] European Commission, Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, 21.4.2021 COM (2021) 206 final 2021/0106 (COD)

[3] Council of Europe, Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts - General approach, 25.11.2022

[4] European Parliament, Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts

[5] Ibid.

[6] Ibid.

[7] Ibid.

"Law Enforcement Authority or Agency (LEA)" means:

(a) any public authority competent for the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, including the safeguarding against and the prevention of threats to public security; or

(b) any other body or entity entrusted by Member State law to exercise public authority and public powers for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, including the safeguarding against and the prevention of threats to public security[8].

"lifecycle" means all phases of existence of an artificial intelligence system ranging from its design and development (including sub-phases such as requirement analysis, data collection, training, testing, integration), installation, deployment, operation, maintenance, to its decommissioning[9].

"personal data" means any information relating to an identified or identifiable natural person **("data subject"**); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person[10].

"policymaker" means a member of a government department, legislature, or other organisation who is responsible for making new rules, laws, policies etc. Examples include EU and national legislative bodies, mayors and municipalities.

"provider" or **"technology developer"** means a natural or legal person, public authority, agency or other body that develops an AI system or that has an AI system developed with a view to placing it on the market or putting it into service under its own name or trademark, whether for payment or free of charge[11].

"user" or **"deployer"** means any natural or legal person, public authority, agency or other body using an AI system under its authority, except where the AI system is used in the course of a personal non-professional activity[12].

[8] European Commission, Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, 21.4.2021 COM (2021) 206 final 2021/0106 (COD)

[9] Committee on Artificial Intelligence (CAI) (2023), Revised Zero Draft Framework (Convention) on Artificial Intelligence, Human Rights, Democracy and the Rule of Law, available at https://rm.coe.int/cai-2023-01-revised-zero-draft-framework-convention-public/1680aa193f and AI HLEG (2020), Assessment List for Trustworthy Artificial Intelligence (ALTAI), available at https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment

[10] Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation)

[11] European Commission, Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, 21.4.2021 COM (2021) 206 final 2021/0106 (COD)

[12] Ibid. and European Parliament, Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts

1. Introduction

The core vision of the popAI project is to enhance and promote trust in the application of Artificial Intelligence (AI) in the security domain. To meet its vision, the project follows a cross-disciplinary approach and aims at increased awareness, ongoing social engagement and the consolidation of diverse spheres of knowledge offering a unified European view across Law Enforcement Agencies (LEAs) and specialised knowledge outputs, including recommendations, best practices and roadmaps.[13]

In that context, WP4 "The pandect of recommendations for the ethical use of AI by LEAs" is a Work Package that, having considered the main actors involved in the AI lifecycle and the potential impact of law enforcement AI on fundamental rights and freedoms, consists of:

- T4.1 "Recommendations for and from policymakers and LEAs",
- T4.2 "Recommendations for and from the Civil Society",
- T4.3 "Recommendations for and from Technology Developers",
- T4.4 "Cross-disciplinary interchange of best practices".

1.1 Scope and objectives of the deliverable

D4.1 presents the results of T4.1 and is mainly addressed to:

- **LEAs (and entities acting on their behalf)** that are planning to deploy or have already started deploying AI systems in order to carry out activities for the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, including the safeguarding against and the prevention of threats to public security and

- **policymakers** since the dynamics and the impact of AI on the society and the environment necessitate the existence of harmonised legal rules that will accompany the technological developments and will establish appropriate procedures and mechanisms to incorporate smoothly AI in the society by efficiently mitigating relevant risks.

Its main purpose is to present and analyse recommendations for LEAs and policymakers that will constitute emerging best practices for the ethically and legally compliant use of AI in law enforcement, i.e., for the use of AI systems by LEAs in a way that will fully respect and prioritise fundamental rights and freedoms and will foster the trust of citizens, also including vulnerable individuals and groups.

The recommendations are of multi-stakeholder and multidisciplinary origin seeking to cover the needs and expectations of all groups of interest. In particular, the present deliverable takes advantage of the knowledge obtained from the results of WP1, the literature review of WP2, the empirical research of WP3 with the broader ecosystem, the results of the other WP4 tasks, the opinions of ethics experts forming the popAI Ethics Advisory Board (EAB), the opinions of the members of the popAI Stakeholder Advisory Board (SAB) and the results of the sibling projects and,

[13] A European Positive Sum Approach towards AI tools in support of Law Enforcement and safeguarding privacy and fundamental rights (pop AI) Grant Agreement 101022001

based on these, creates a library of target group-specific recommendations focused on LEAs and policymakers that will be in line with the applicable ethical and legal framework. Furthermore, we captured the views of LEAs and policymakers, both from within the popAI Consortium and outside of it, to collect further input and to also evaluate our findings.

1.2 Relation to other tasks and deliverables

The dependence of T4.1 and, consequently, of this deliverable on other WPs, Tasks and Deliverables can be briefly outlined as follows:

D4.1 receives input from:

WP1: D1.6 “Policy briefs - 1st year”

WP2:

- D2.1 “Functionality taxonomy and emerging practices and trends”
- D2.2 “Legal casework taxonomy: emerging trends and scenarios”
- D2.4 “Ethical frameworks for the use of AI by LEAs”
- D2.5 “Practical ethics toolbox for the use of AI by LEAs”
- D2.6 “AI meets organisational cultures: Human-machine interaction at the police station”

WP3:

- D3.1 “Map of AI in policing innovation ecosystem and stakeholders”
- D3.3 “Citizen produced priorities and recommendations for addressing AI in the security domain”
- D3.4 “Stakeholder attitudes, priorities, and recommendations for addressing AI in the security domain in practice”
- D3.5 “Foresight Scenarios for AI in Policing”

WP4:

- D4.2 “White Paper for Civil Society”
- D4.3 “White Paper for Technology Developers”

D4.1 provides output to:

WP1: D1.7 “Policy briefs - 2nd year”

WP4:

- D4.3 “White Paper for Technology Developers”
- D4.4 “Synthesis: a collection of the best multidisciplinary practices”

WP5:

- D5.2 “Final community building and ecosystem engagement activities plan”
- D5.6 “Communication & Dissemination plan – final”
- D5.8 “popAI roadmaps”

1.3 Structure of the deliverable

The introductory part, the scope and objectives of this deliverable and the relation of this deliverable to other popAI tasks and deliverables were presented above. The rest of the deliverable is structured as follows:

Section 1.4 contains important terms and definitions.

Section 1.2 presents the methodological approach adopted for the production and delivery of recommendations for LEAs and policymakers, including the overall procedure followed, the guidelines and criteria set as well as the sources of information.

In accordance with the methodology, *Section 3* details the sources one by one starting with the applicable ethical and legal framework. Based on these, recommendations are produced for LEAs and policymakers on the ethical use of AI in law enforcement.

Section 4 lists and analyses the produced recommendations for the LEAs that constitute emerging best practices for the ethical use of AI in law enforcement.

Section 5 lists and analyses the produced recommendations for the policymakers to ensure through legislation the ethical use of AI in law enforcement.

Section 6 concerns the evaluation of the presented emerging best practices through the involvement of LEAs and policymakers, the popAI EAB and the popAI SAB. As the SAB’s feedback is pending, this deliverable is submitted as is, and the SAB’s feedback will be incorporated into D4.4.

Section 7 concludes the deliverable.

2 Methodology

2.1 Objectives

The overall procedure defined, adopted, and followed towards the production and delivery of the recommendations for LEAs and policymakers is based on a methodology which seeks to achieve the following goals:

From a methodological standpoint:

- To remain compliant with the Grant Agreement (GA) and the milestones set therein.
- To follow a multi-stakeholder and multidisciplinary approach.
- To consider and include the outputs of all associated popAI tasks, deliverables, workshops, as well as the outputs from other sibling projects, and thus satisfy all the appropriate interdependencies related to the content of the present deliverable.
- To (cross-)validate the results produced within the WP4 ecosystem (D4.1 for LEAs and policymakers with D4.2 for citizens and D4.3 for technology developers), to ensure their validity from different perspectives and according to the needs of LEAs and the target groups of interest.
- To effectively capture and elicit the target groups’ input and resolve potential contradictory answers / feedback received.

- To feed the recommendations and useful output into D4.4 and the dissemination activities appointed to the stakeholders under WP5.
- To consider the feedback provided by the SAB which consists also of members from the sibling projects ALIGNER and STARLIGHT.
- To consider the feedback provided by the EAB.

From a substantial standpoint:

- To ensure that the produced results represent the needs and expectations of LEAs.
- To ensure that the produced results represent the needs and expectations of the persons affected by law enforcement AI.
- To ensure that the produced results have considered the respective ethics principles and the applicable legal framework, such as the principles set by EU for a trustworthy AI, the fundamental human rights and freedoms framework, the data protection legislation, the societal as well as environmental values and needs.
- To ensure that as many fields and cases of interest as possible have been covered by taking into consideration the perspectives of all groups of interest.

From a dissemination standpoint:

- To ensure that the results of this deliverable are valid, useful and of practical use to LEAs and policymakers, but also to the whole ecosystem of stakeholders (including LEAs, policymakers, civil society, technology developers, research and academia),
- To present, deliver and disseminate the produced recommendations appropriately so that they are informative and clear and also communicated to the targeted audiences in collaboration with D1.7 “Policy briefs” and WP5 “Dissemination, Communications and Sustainable Community Engagement”.

2.2 Methodological approach

WP4 applies a combination of doctrinal and empirical research in order to answer the question of what the emerging best practices or recommendations for the ethical use of AI by LEAs would be. D4.1, in particular, examines the recommendations for the ethical use of AI in law enforcement by focusing on two main categories of stakeholders: LEAs and policymakers.

The main source where the present deliverable draws the theoretical framework from is WP2 ‘Security AI in the next 20 years: trends, practices and risks’ as revised based on the latest legislative developments at European level (most importantly the AI Act Proposal, the proposed AI Liability Directive and the CoE Draft Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law) to ensure that it stays up-to-date due to the numerous and essential latest changes in the forthcoming legal landscape. Furthermore, it makes use of the WP3 ‘Empirical Knowledge Collection and Management Framework’ results, with a focus on the Policy Labs of Task 3.4 in which stakeholders including LEAs and policymakers, civil society representatives, ethics experts and technology developers participated. In addition to the above sources, auxiliary sources were utilised to enhance the recommendations, such as D1.6.

The aforementioned sources were used in order to create the initial "entries" of recommendations. According to these entries, the main trends were identified, and a set of recommendations, categorised according to their thematic areas, was produced. The existing and applicable ethical and legal framework (especially the ALTAI principles, GDPR, and LED) was the benchmark to filter out popAI findings, as compliance with the ethics requirements and the applicable legal provisions was the criterion according to which the entries would be considered as emerging best practices/recommendations or not. As for the draft EU legislation (e.g., the Draft AI Act, the CoE Draft Convention on AI, the AI Liability Directive Proposal), it was studied for us to keep up with the latest developments in the legal landscape on AI.

Furthermore, T4.1 questionnaires based on the WP2 and WP3 taxonomies, functionalities, and controversies were developed under WP4. The questionnaires were addressed to LEAs and policymakers of the Consortium and limited externals, and were used in WP4 as an assistant tool, to support, update, crosscheck and evaluate the emerging best practices for the ethical use of AI on law enforcement. Lastly, D4.1 took into consideration the feedback of the SAB and the EAB to enhance the recommendations as well as the sibling projects' (ALIGNER and STARLIGHT) proposals on the issue of ethical AI for law enforcement purposes.

The justification behind the combination of theoretical and empirical research is based on the GA, according to which WP2 provides the theoretical framework, WP3 conducts the empirical research, while WP4 draws conclusions, based on the findings of the previous WPs, in the quest of ethical AI in law enforcement.

Taking into account the nature of the terms "recommendations" and "emerging best practices", their purpose in the present deliverable is to illuminate the existing concerns around law enforcement AI, specify the obligations of its deployers and complement the legal framework by serving as a practical guide that will help (a) LEAs make ethical use of AI systems and (b) policymakers smoothly incorporate AI in the society in a way that is valued by both LEAs and affected persons.

2.3 Sources

The methodology can be divided into three main phases: (1) elicit/collect existing data, (2) analyse new data/input, (3) produce new recommendations and update or discard existing recommendations.

Towards the direction of those steps, outcomes from the following sources have been utilised to produce recommendations to and from LEAs and policymakers:

- popAI deliverables
- literature, bibliography
- ethics guidelines, applicable legal framework
- draft EU legislation on AI
- popAI workshops
- popAI Policy Labs
- popAI crowdsourcing platform

- popAI Consortium meetings (e.g., plenary meetings)
- popAI, STARLIGHT, ALIGNER workshops and deliverables
- WP4 questionnaires
- EAB and SAB feedback

Following-up, chapter 3 of the present deliverable elaborates on the sources used to extract the recommendations addressed to LEAs and policymakers.

3 Sources for the production of recommendations

This chapter presents, analyses and comments on the main sources that we have used for the extraction of conclusions and the provision of recommendations that constitute emerging best practices on the ethical use of AI by LEAs.

3.1 Ethical and Legal Framework

This section presents the ethical and legal framework that surrounds (and has to surround) the design, development and use of AI systems until their decommissioning. All technological developments bring changes that have the potential to influence society and the environment. In particular, AI-enabled technologies, characterised also as "innovative and disruptive technologies" have the power to bring significant social and economic opportunities but also a plethora of risks and insecurity if not developed and used properly. The role of legislation is critical and decisive as it sets out core ethical rules that put humans in the centre, prioritise and protect fundamental rights during the entire lifecycle of AI. Policymakers are responsible for smoothly incorporating AI into society through the establishment of mechanisms, institutions and procedures that aim at efficiently mitigating the risks.

It needs to be clarified that all recommendations described in chapters 4 and 5 have been filtered in order to be compliant with the existing applicable ethical and legal framework and to complement the draft EU legislation on AI (mainly the Draft AI Act) wherever vague points or gaps have been identified.

3.1.1 Ethics Guidelines for trustworthy AI

In 2019 the High-Level Expert Group on Artificial Intelligence (AI HLEG)[14] established a framework for the creation and implementation of trustworthy AI. A trustworthy AI system must be designed and operate in compliance with applicable laws (lawful), it must respect ethical standards and ethical values (ethical) and it must ensure technical robustness, safety, accuracy and resilience by design (robust). The seven key requirements are the following:

- **Human agency and oversight**
    - Respect for human autonomy by allowing humans to make informed decisions.
    - Proper oversight mechanisms (human-in-the loop, human-on-the loop, human-in-command approaches) to ensure that AI systems act as enablers for a democratic society and foster fundamental rights.

[14] The AI HLEG is an independent expert group that was established by the European Commission in June 2018.

- **Technical robustness and safety**
  - Resilience to attack and security.
  - General safety by following a preventative approach to risks.
  - Accuracy to ensure that training data are up-to-date, of high quality, complete and representative of the environment (by also monitoring false positives, false negatives) and communication of the accuracy metrics.
  - Reliability of the system to operate based on its intended goals, fall-back plans and reproducibility and relevant verification methods e.g., through logging.

- **Privacy and data governance**
  - Respect to the fundamental rights of privacy and data protection.
  - Data governance through appropriate mechanisms (DPIA, DPO consultation, data minimisation, privacy by design and by default through anonymisation, pseudonymisation or encryption of personal data, security standards etc.).

- **Transparency**
  - Traceability mechanisms for documenting and monitoring the complete trajectory of the AI system, from design and development to deployment and usage.
  - Explainability, i.e., the ability to explain both the technical processes and the reasoning behind the predictions, recommendations or decisions made (opposite example: black boxes).
  - Communication to the users that they are interacting with an AI system and implementation of mechanisms to inform users about the purpose, criteria and limitations of the predictions, recommendations or decisions made.

- **Diversity, non-discrimination and fairness**
  - Inclusion and diversity. Avoidance of unfair bias through a set of procedures to avoid creating or reinforcing unfair bias in the AI system, both regarding the use of input data as well as for the algorithm design, also including mechanisms that allow for the flagging of issues related to bias, discrimination or poor performance of the AI system. The continuation of unfair biases could lead to unintended (in)direct prejudice and discrimination against certain groups or people, potentially leading to prejudice and marginalisation.
  - Accessibility and universal design in a way that allows a user-centric approach and people to use AI products or services, regardless of their age, gender, abilities or characteristics. AI systems should consider universal design principles addressing the widest possible range of users to enable equal access and active involvement of all people.
  - Active participation of stakeholders affected by the AI system from its design and development and even after its deployment.

- **Societal and environmental well-being**
  - Environmental well-being. Operation of the AI system in the most environmentally friendly way possible during its lifecycle.
  - Impact on work and skills. AI systems to support humans in the working environment and aim for the creation of meaningful work. Provision of information to the workers about the AI system's operation and impact.
  - Impact on society and democracy. AI systems to benefit all human beings, including future generations, to maintain and foster democratic processes and to respect the plurality of values and life choices of individuals.

- **Accountability and auditability**
  - Auditability through accessible mechanisms for accountability that ensure an adequate possibility of redress by design in case unjust or adverse impacts occur.
  - Risk management that identifies and mitigates risks in a transparent way that can be explained to and audited by third parties, i.e., ability to report on actions or decisions that contribute to the AI system's outcome, and to respond to the consequences of such an outcome[15].

3.1.2 Artificial Intelligence Act

3.1.2.1 Objectives and status update

A Proposal for a Regulation laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) was issued by the European Commission on 21 April 2021. The AI Act constitutes the world's first comprehensive AI regulatory framework, puts the human in the centre (human-centred approach) and is a horizontal regulation, applicable to AI systems of all sectors which are classified according to the risk they pose to the affected individuals (risk-based approach). Following the example of the GDPR, the AI Act's territorial scope extends beyond the EU also to providers placing on the market or putting into service AI systems in the EU irrespective of whether those providers are established within the Union or in a third country as well as to providers and users of AI systems that are located in a third country, where the outputs (i.e., predictions, recommendations or decisions) produced by the AI system are used in the Union.

It is important to mention that the AI Act is not applicable in case of AI systems specifically developed for the sole purpose of scientific research and development. However, under all circumstances, any research and development activity should be carried out in accordance with the Charter, Union law

[15] AI HLEG (2019), Ethics Guidelines for Trustworthy AI, available at https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai and AI HLEG (2020), Assessment List for Trustworthy Artificial Intelligence (ALTAI), available at https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment

as well as the national law[16] and based on the AI HLEG Ethics Guidelines presented in the previous section.

The Commission's specific objectives for the AI Act are to:

1. ensure that AI systems used in the EU are safe and respect existing law on fundamental rights and EU values,
2. ensure legal certainty to facilitate investment and innovation in AI,
3. enhance governance and enforcement of the law on fundamental rights and applicable safety requirements and
4. facilitate the development of a single market for lawful, safe and trustworthy AI applications and prevent market fragmentation[17].

Further to the text proposed by the European Commission, the Council of the European Union issued its General Approach on 25 November 2022[18] with the amendments being formally adopted by the Council on 6 December 2022.

On 9 May 2023, the European Parliament issued its Draft Compromise Amendments[19] with the amendments being formally adopted by the Parliament on 14 June 2023.

For the finalisation and adoption of the AI Act, a trilogue among the Council, the European Parliament and the European Commission is required which is expected to start soon and constitutes the last phase of the negotiations before the law is passed.

At this stage of the pop-AI project, we can only be based on the current developments around the AI Act, i.e., on three different texts as they have been presented above, and we are committed to closely monitoring the progress. Updated information, if any, will be provided in D4.4.

3.1.2.2 Prohibited AI practices

As a consequence of the humancentric approach on which the Draft AI Act has been based, the AI systems are categorised according to the risks that may derive from their use and the impact on society and the environment. According to this risk-based approach, there are AI systems that create:

- Unacceptable risk
- High risk
- Limited risk
- Minimal risk

[16] European Parliament, Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, Recital 2 f

[17] Explanatory Memorandum of the Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, 21.4.2021 COM (2021) 206 final 2021/0106 (COD)

[18] Council of Europe, Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts - General approach, 25.11.2022

[19] European Parliament, Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts

AI systems that create unacceptable risk are these that contravene the EU values of respect for human dignity, freedom, equality, democracy and the rule of law and Union fundamental rights, including the right to non-discrimination, data protection and privacy and the rights of the child. Therefore, such systems are prohibited.

The list of prohibited AI, as it has been extended by the European Parliament through its proposed amendments to the Draft AI Act[20], is the following:

1. Manipulative AI systems:

(a) AI systems that deploy subliminal techniques beyond a person's consciousness or purposefully manipulative or deceptive techniques, with the objective to or the effect of materially distorting a person's or a group of persons behaviour by appreciably impairing the person's ability to make an informed decision, thereby causing the person to take a decision they would not have taken otherwise in a manner that causes or is likely to cause that person, another person or group of persons significant harm[21].

(b) AI systems that exploit any of the vulnerabilities of a person or a specific group of persons, including characteristics of such individual's or group of persons' known or predicted personality traits or social or economic situation, age, physical or mental ability, with the objective or to the effect of materially distorting the behaviour of that person or a person pertaining to that group in a manner that causes or is likely to cause that person or another person significant harm.

2. Social scoring:

AI systems used for the evaluation/classification of the trustworthiness of natural persons or groups thereof based on their social behavior or known, inferred or predicted personal or personality characteristics, with this social score leading to either/both:

(a) unfavorable treatment in social contexts which are unrelated to the contexts in which the data was originally generated or collected;
(b) unfavorable treatment that is unjustified or disproportionate to social behaviour or its gravity.

3. Real-time[22] remote biometric identification in public spaces:

Complete prohibition without the exemptions stipulated in the previous versions of the Commission and the Council[23]

[20] We are waiting for the conclusions that will be made after the trilogue amongst the European Commission, the Council and the European Parliament on the final list of prohibited AI practices.

[21] Only relevant exemption: AI systems intended to be used for approved therapeutical purposes on the basis of specific informed consent of the individuals that are exposed to them or, where applicable, of their legal guardian.

[22] Real-time means that the identification process, from the moment of the collection of data to the identification, has to occur in real time, or without any significant delay.

[23] According to the initial text proposed by the EC, real-time remote biometric systems can be used by LEAs in the following cases: (i) the targeted search for specific potential victims of crime, including missing children; (ii) the prevention of a specific, substantial and imminent threat to the life or physical safety of natural persons or of a terrorist attack; (iii) the detection, localisation, identification or prosecution of a perpetrator or suspect of a criminal offence referred to in Article 2(2) of Council Framework Decision 2002/584/JHA62 and punishable in the Member State concerned

4. **Biometric categorisation systems**:

AI systems that categorise natural persons according to sensitive/protected attributes or characteristics or based on the inference of those attributes or characteristics

5. **Predictive policing:**

AI systems that perform risk assessments of natural persons or groups to assess the risk of offending or reoffending or for predicting the occurrence or reoccurrence of an actual or potential criminal or administrative offence based on profiling of a natural person or on assessing personality traits and characteristics (location, past criminal behaviour)

6. **Facial recognition databases:**

AI systems that create or expand facial recognition databases through untargeted scraping of facial images from internet/CCTVs

7. **Emotion recognition:**

AI systems to infer emotions of natural persons in the areas of law enforcement, border management, in workplace and education[24]

3.1.2.3 Obligations for high-risk AI systems

Although it is important to know which types of AI systems are prohibited by law due to the unacceptable risk that they are likely to pose to the AI subjects, it is even more critical to have knowledge about the development and deployment of high-risk AI systems which can pose significant risks to the health and safety or fundamental rights of persons. These AI systems are the ones that are permitted to be placed and used on the internal market but only if specific obligations are met during their lifecycle. Therefore, it is of high significance that providers, importers, distributors and users of such AI systems fulfil the relevant requirements stipulated by law in order to fully prevent or mitigate the risks that may emerge and protect the affected persons to the biggest extent.

As explained previously, the final AI Act will be issued after the trilogue amongst the EC, the Council and the EP. According to each EU institutional body involved in the legislative process the types of ‘high-risk AI systems” are listed in Annex III of the AI Act Proposal. The list is as follows:

EC’s high-risk AI systems (8 categories in total, 4 categories most relevant to LEAs):

1. Biometric identification and categorisation of natural persons, including “real-time” and “post” remote biometric identification;
2. Law enforcement, including predictive policing tools, polygraphs or similar instruments, tools to detect deep fakes, systems used to evaluate the reliability of criminal evidence and, in general, profiling tools and systems used for crime analytics using large datasets to identify unknown patterns;

by a custodial sentence or a detention order for a maximum period of at least three years, as determined by the law of that Member State.

[24] The final four categories (biometric categorisation systems, predictive policing, facial recognition databases and emotion recognition) were added by the EP.

3. Migration, asylum and border control management;
4. Administration of justice and democratic processes.

Council's high-risk AI systems (8 categories in total, 4 categories most relevant to LEAs):

1. Remote biometric identification system;
2. Law enforcement, including predictive policing tools, polygraphs or similar instruments, systems used to evaluate the reliability of criminal evidence and, in general, profiling tools;
3. Migration, asylum and border control management;
4. Administration of justice and democratic processes.

EP's high-risk AI systems (8 categories in total, 4 categories most relevant to LEAs):

1. Systems used for biometric identification of natural persons and systems used to make inferences about personal characteristics on the basis of biometric data, including emotion recognition[25] (also non-remote biometric identification systems);
2. Law enforcement (incl. also other entities acting on behalf of LEAs or EU agencies), including polygraphs or similar instruments, systems used to evaluate the reliability of criminal evidence and, in general, profiling tools and systems used for crime analytics using large datasets to identify unknown patterns;
3. Migration, asylum and border control management;
4. Administration of justice and democratic processes including systems used to influence elections and recommender systems used by social media.

Specific obligations must be fulfilled for these categories of AI systems as explained in Chapter 2 of the AI Act Proposal. The obligations are the following:

- **Risk management system** (establishment, implementation, documentation and maintenance).
- **Data governance** for the training, validation and testing data.
- Preparation of **technical documentation** prior to placing the AI system on the market and modifications to keep it updated.
- **Record-keeping** through automatic recording of events (logs) during the design and development phases in conformity with recognised standards or common specifications.
- **Transparency and provision of information to the users** through instructions for use and by making available the provider's contact details, the system's characteristics, capabilities and limitations of performance, pre-determined changes to the system described in the initial conformity assessment, the human oversight measures and the expected lifetime of the system along with any necessary maintenance and care measures to ensure its proper functioning.
- **Human oversight measures** which are either implemented by design (if technically possible) or can be implemented by the user of the AI system to prevent or minimise the risks to health,

[25] It needs to be reminded that (i) real-time remote biometric identification in public spaces, (ii) predictive policing, as well as (iii) emotion recognition by LEAs are banned in the EP's Draft AI Act (see the prohibited AI practices in the previous section).

safety or fundamental rights that may emerge. The persons to whom human oversight is assigned must fully understand the capabilities and limitations of the system, remain aware of automation bias, be able to interpret the outputs of the system, be able to decide when to use or not use the outputs of the system, as well as be able to intervene or stop the system. Verification and confirmation of the outputs is required by at least two natural persons on behalf of the AI system's user.

- **Accuracy, robustness and cybersecurity** by design and during the system's development in a way that clarifies to the user the levels of accuracy and the relevant accuracy metrics, ensures resilience of the system with respect to errors, faults or inconsistencies that may occur within the system or the environment in which the system operates and ensures resilience of the system as regards malicious attempts by unauthorised third parties to modify their use or performance by exploiting the system vulnerabilities.
- Conducting of a **conformity assessment** by the provider[26] before specific types of high-risk AI systems are placed on the market or put into service where compliance with all aforementioned obligations is demonstrated and drafting of a relevant EU declaration of conformity.

The LEAs or entities acting on their behalf as deployers of high-risk AI systems have specific obligations that are stipulated in Chapter 3[27] of the Draft AI Act such as the carrying out of a fundamental rights impact assessment, the implementation of appropriate technical and organisational measures to ensure that they use such systems in accordance with the instructions for use, the implementation of human oversight, the involvement of skilled and trained staff in the human oversight process, the regular monitoring of the robustness and cybersecurity measures, confirmation that input data is relevant and sufficiently representative in view of the intended purpose of the high-risk AI system, communication with the providers, the distributors and the relevant national supervisory authorities, conducting of a data protection impact assessment etc[28].

3.1.3 AI Liability Directive

On 28 September 2022, the Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence ("AI Liability Directive") was published[29]. The AI Liability Directive complements the AI Act.

[26] "With the only exception of AI systems intended to be used for the remote biometric identification of persons, or AI systems intended to be used to make inferences about personal characteristics of natural persons on the basis of biometric or biometrics-based data, including emotion recognition systems for which *the involvement of a notified body in the conformity assessment* should be foreseen, to the extent they are not prohibited", European Parliament, Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, Recital 64

[27] Chapter 3 specifies separately the obligations of the providers, importers, distributors, users of high-risk AI systems and of other third parties.

[28] European Parliament, Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, Article 29

[29] Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive), COM (2022) 496 final. See also European Commission, Impact assessment

Its purpose is to harmonise non-contractual fault-based liability rules, in order to ensure that persons claiming compensation for damage caused to them by an AI system enjoy a level of protection equivalent to that enjoyed by persons claiming compensation for damage caused without the involvement of an AI system[30].

The AI Liability Directive does not apply to criminal liability. The scope of the AI Liability Directive includes common rules on (a) the disclosure of evidence on high-risk AI systems to enable a claimant to substantiate a non-contractual fault-based civil law claim for damages and (b) the burden of proof in the case of non-contractual fault-based civil law claims brought before national courts for damages caused by an AI system[31]. The AI Liability Directive follows a minimum harmonisation approach[32]. Also, the AI Liability Directive uses the same definitions as provided in the AI Act.

3.1.4 CoE Framework Convention on AI, Human Rights, Democracy and the Rule of Law

Following the mandate of the Council of Europe Committee of Ministers for the period 2022-2024 to set up an international negotiation process for drawing up a Convention on the development, design and application of AI, the Council of Europe Framework Convention on AI, Human Rights, Democracy and the Rule of Law is prepared by the Committee on Artificial Intelligence (CAI) and is in progress[33].

The purpose of the CoE Convention is to ensure that during their lifecycle, AI systems are fully consistent with respect for human dignity and individual autonomy, human rights and fundamental freedoms, the functioning of democracy and the observance of the rule of law[34].

It is worth mentioning that according to the current Consolidated Working Draft, the Convention shall not apply to research and development activities regarding AI systems (same as the Draft AI Act) unless the systems are tested or otherwise used in ways that have the potential to interfere with human rights and fundamental freedoms, democracy and the rule of law (difference to the Draft AI Act)[35].

Chapters II and III of the Consolidated Working Draft Convention provide for the general obligations and principles. The general obligations are the following two: (a) respect for human rights and

report accompanying the document: Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence, SWD (2022) 319 final and European Parliament, EPRS, Artificial intelligence liability directive, Briefing, February 2023.

[30] Proposal for an AI Liability Directive, Recital 7. See also Recital 9: “Such harmonisation should increase legal certainty and create a level playing field for AI systems, thereby improving the functioning of the internal market as regards the production and dissemination of AI-enabled products and services”.

[31] Proposal for an AI Liability Directive, Article 1 par. 1

[32] Ibid., Recital 14

[33] During its 6th plenary meeting of 31 May to 2 June 2023, the Council of Europe’s Committee on Artificial Intelligence (CAI) reached a significant milestone in the negotiation of the (framework) convention on artificial intelligence (AI) by completing the first reading of the draft text. The upcoming phase of the CAI negotiations will require the negotiating parties to reach a compromise on the convention’s text. Some key questions include the scope of the convention, the level of granularity and content of the obligations, the principles, as well as the approach for risk identification and mitigation. The next plenary meeting is scheduled for 23 to 26 October 2023 and aims to bridge the positions of the negotiating parties on the aforementioned issues.

[34] Committee on Artificial Intelligence (CAI), Consolidated Working Draft of the Framework Convention on Artificial Intelligence, Human Rights, Democracy, and the Rule of Law (7 July 2023), Article 1

[35] Ibid., Article 4 par.2; See also Article 12 on the principle of safe innovation.

fundamental freedoms and (b) integrity of democratic processes and respect for rule of law. The principles to be respected during the design, development, use and decommissioning of AI systems are those of transparency and oversight, accountability and responsibility, equality and non-discrimination, privacy and personal data protection, safety, security and robustness, and safe innovation[36].

Chapter V is about the assessment and mitigation of risks and adverse impacts through the establishment of a risk and impact management framework and through the training of the actors responsible for the design, development, use and decommissioning of AI systems.

Chapter VI refers to the implementation of the Convention highlighting the necessity of non-discrimination, respect of the rights of persons with disabilities and of children, public consultation and digital literacy and skills for all segments of the population and for those responsible for the design, development, use and decommissioning of AI systems.

In order to ensure effective implementation of its provisions, the Convention establishes in Chapter VII a follow-up mechanism that consists of conference of the Parties, international cooperation and effective oversight mechanisms.

3.1.5 EU Data Protection Law

The fundamental right to the protection of personal data is enshrined in Article 8 of the Charter of Fundamental Rights of the European Union (the Charter) and safeguarded in particular and as regards the content of the present deliverable by:

(a) Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation or GDPR) and

(b) Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA (Law Enforcement Directive or LED) as transposed in the national legal frameworks of the EU Member States.

In cases where personal data are processed through an AI system, data protection law applies and binds the controllers and processors[37]. Some of the core provisions are presented below.

[36] Ibid., Articles 5-12

[37] 'Controller' means the natural or legal person, public authority, agency or other body which, alone or jointly with others, *determines the purposes and means* of the processing of personal data; where the purposes and means of such processing are determined by Union or Member State law, the controller or the specific criteria for its nomination may be provided for by Union or Member State law;
'Processor' means a natural or legal person, public authority, agency or other body which processes personal data *on behalf of the controller* (GDPR, Article 4).

3.1.5.1 General Data Protection Regulation

The GDPR governs how the personal data of individuals in the EU may be processed and transferred by defining individuals' rights related to the fundamental right to personal data protection, the obligations of those processing data, procedures and measures for ensuring compliance and sanctions for those in breach of the rules.

The main principles to ensure personal data protection are stipulated in Article 5 of the GDPR according to which personal data shall be:

a. processed lawfully, fairly and in a transparent manner in relation to the data subject (principle of "lawfulness, fairness and transparency");
b. collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes; further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall, in accordance with Article 89(1) GDPR, not be considered to be incompatible with the initial purposes (principle of "purpose limitation");
c. adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed (principle of "data minimisation");
d. accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that personal data that are inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay (principle of "accuracy");
e. kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed; personal data may be stored for longer periods insofar as the personal data will be processed solely for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) GDPR subject to implementation of the appropriate technical and organisational measures in order to safeguard the rights and freedoms of the data subject (principle of "storage limitation");
f. processed in a manner that ensures appropriate security of the personal data, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures (principle of "integrity and confidentiality").

Simultaneously, the controller shall be responsible for, and be able to demonstrate compliance with the aforementioned principles (principle of "accountability").

The lawful bases for the processing of personal data are listed in Article 6 GDPR according to which processing shall be lawful only if and to the extent that at least one of the following applies:

a. the data subject has given consent to the processing of his or her personal data for one or more specific purposes;
b. processing is necessary for the performance of a contract to which the data subject is party or in order to take steps at the request of the data subject prior to entering into a contract;
c. processing is necessary for compliance with a legal obligation to which the controller is subject;

d. processing is necessary in order to protect the vital interests of the data subject or of another natural person;
e. processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller;
f. processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data, in particular where the data subject is a child. Point (f) shall not apply to processing carried out by public authorities in the performance of their tasks.

As regards LEAs, the GDPR is applicable in cases where the Law Enforcement Directive is not (see below).

3.1.5.2 Law Enforcement Directive

The Law Enforcement Directive applies to the processing of personal data by competent authorities[38] for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, including the safeguarding against and the prevention of threats to public security. The main principles relating to the processing of personal data are also stipulated in LED[39]. In contrast to the GDPR, in order to be lawful, the processing of personal data under the LED should be necessary for the performance of a task carried out in the public interest by a competent authority based on Union or Member State law for one or more of the said purposes.

Therefore, in cases where personal data are processed by LEAs through AI systems deployed for the aforementioned purposes, applicable is solely the LED as transposed in the national legal framework. In other cases, such as when LEAs are using AI systems in the context of their participation in research projects or to be assisted in carrying out their administrative tasks, applicable is the GDPR.

3.2 Interdependent popAI Work Packages and Tasks

3.2.1 Work Packages 1, 2 and 3

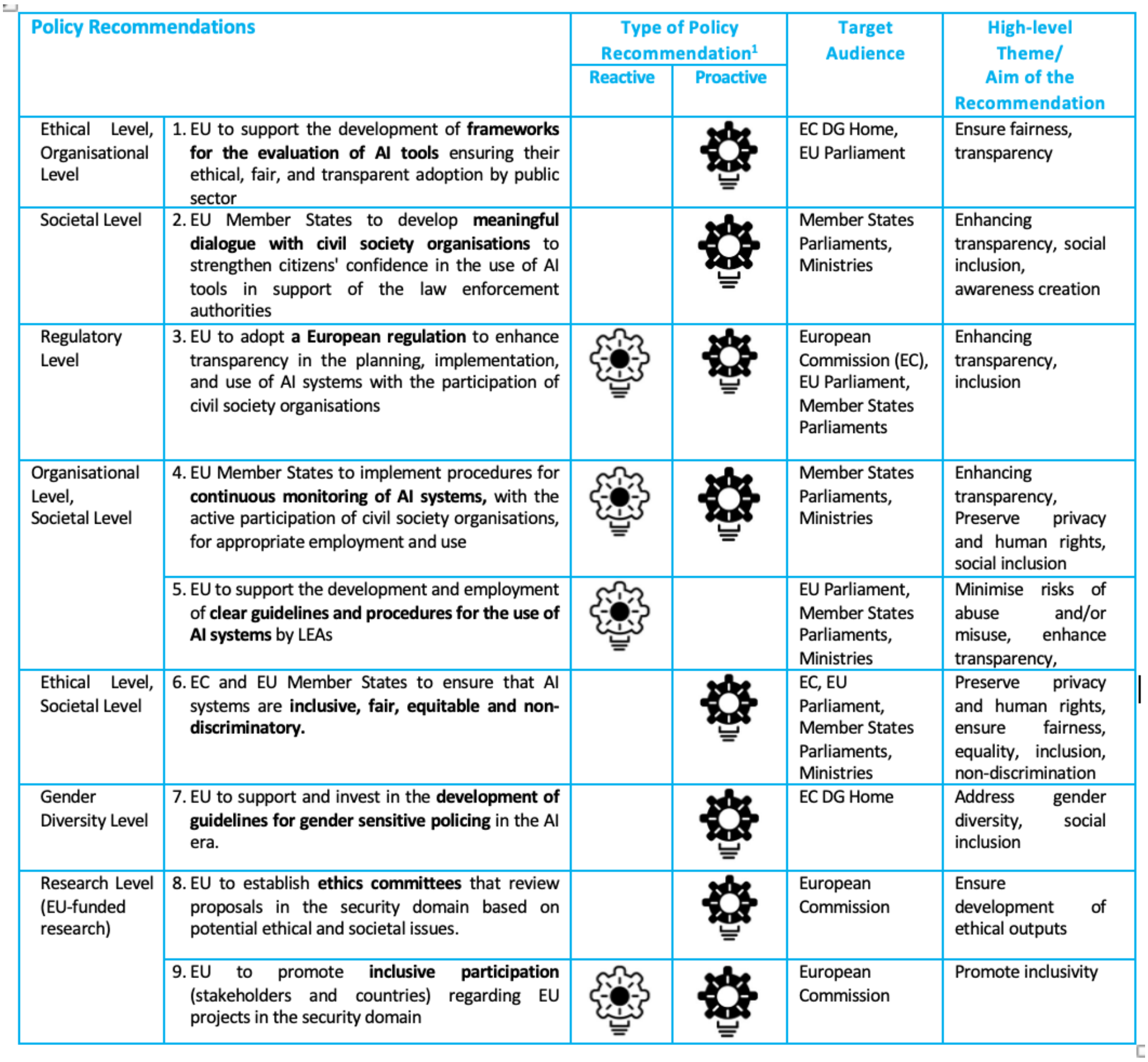
WP4, due to its nature, needs to receive input from previous work done as part of the project. In particular, the present deliverable is based, amongst others, on other project's tasks as they are shortly presented below.

[38] 'Competent authority' means: (a) any public authority competent for the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, including the safeguarding against and the prevention of threats to public security; or (b) any other body or entity entrusted by Member State law to exercise public authority and public powers for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, including the safeguarding against and the prevention of threats to public security (LED, Article 3).

[39] Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, Article 4, Recital 26, Recital 61

3.2.1.1 T1.5 - Policy briefs

The first deliverable of this task (D1.6)[40] provided a set of early evidence-based policy recommendations that emerged from the popAI work in the course of WP2, WP3 during the first year of the project, aligned with popAI's aim towards the creation of a European AI hub for LEAs; a hub that will foster a dynamic, ongoing dialogue in Europe between civil society and public authorities, fostering trust in AI systems for Law Enforcement. These initial policy recommendations addressed a wide range of policy levels, including ethical, organisational, societal, regulatory, gender-diversity as well as research level. Furthermore, they addressed distinct high-level themes (core aim) targeting diverse audiences as can be seen in Figure 2 below.



Policy Recommendations		Type of Policy Recommendation[1]		Target Audience	High-level Theme/ Aim of the Recommendation
		Reactive	Proactive		
Ethical Level, Organisational Level	1. EU to support the development of **frameworks for the evaluation of AI tools** ensuring their ethical, fair, and transparent adoption by public sector		✓	EC DG Home, EU Parliament	Ensure fairness, transparency
Societal Level	2. EU Member States to develop **meaningful dialogue with civil society organisations** to strengthen citizens' confidence in the use of AI tools in support of the law enforcement authorities		✓	Member States Parliaments, Ministries	Enhancing transparency, social inclusion, awareness creation
Regulatory Level	3. EU to adopt **a European regulation** to enhance transparency in the planning, implementation, and use of AI systems with the participation of civil society organisations	✓	✓	European Commission (EC), EU Parliament, Member States Parliaments	Enhancing transparency, inclusion
Organisational Level, Societal Level	4. EU Member States to implement procedures for **continuous monitoring of AI systems,** with the active participation of civil society organisations, for appropriate employment and use	✓	✓	Member States Parliaments, Ministries	Enhancing transparency, Preserve privacy and human rights, social inclusion
	5. EU to support the development and employment of **clear guidelines and procedures for the use of AI systems** by LEAs	✓		EU Parliament, Member States Parliaments, Ministries	Minimise risks of abuse and/or misuse, enhance transparency,
Ethical Level, Societal Level	6. EC and EU Member States to ensure that AI systems are **inclusive, fair, equitable and non-discriminatory.**		✓	EC, EU Parliament, Member States Parliaments, Ministries	Preserve privacy and human rights, ensure fairness, equality, inclusion, non-discrimination
Gender Diversity Level	7. EU to support and invest in the **development of guidelines for gender sensitive policing** in the AI era.		✓	EC DG Home	Address gender diversity, social inclusion
Research Level (EU-funded research)	8. EU to establish **ethics committees** that review proposals in the security domain based on potential ethical and societal issues.		✓	European Commission	Ensure development of ethical outputs
	9. EU to promote **inclusive participation** (stakeholders and countries) regarding EU projects in the security domain	✓	✓	European Commission	Promote inclusivity

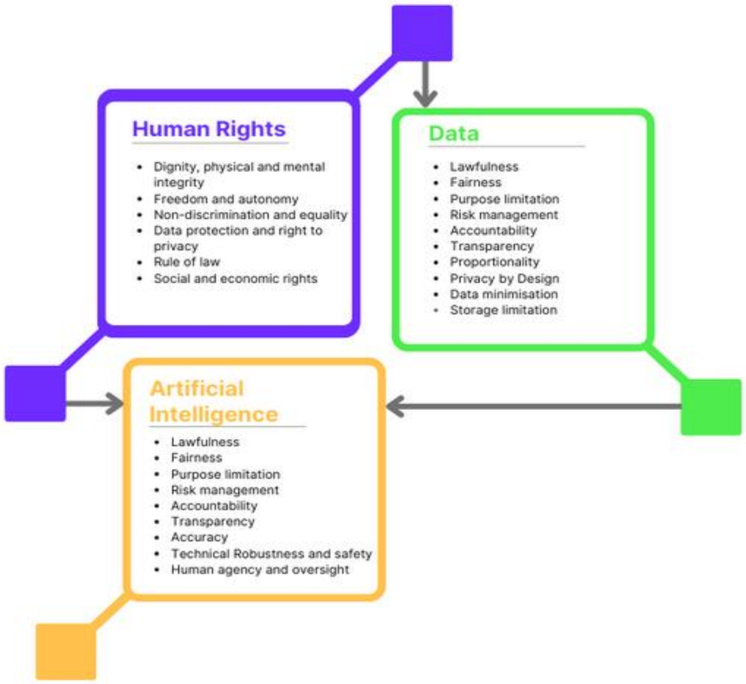
Figure 1: popAI Policy Recommendations Overview (D1.6)

3.2.1.2 T2.2 - Legal framework and casework taxonomy: emerging trends and scenarios

As part of this task, the legal taxonomy was presented which classified EU legal framework into three high-level classes:

[40] popAI D1.6, "Policy Briefs – 1st Year"

- Human rights: This class refers to regulations that define fundamental human rights that must not be violated by AI applications. When they are infringed, people are entitled to legal remedies. AI applications in the security domain can negatively impact a wide range of fundamental rights such as the right to privacy, freedom of assembly, non-discrimination, presumption of innocence, right to a fair trial, right to an effective remedy among others. Making an ethical and trustworthy AI in the security domain involves primarily assessing risks and setting standards ensuring civil, political, economic, social and cultural rights.
- Personal data: This class refers to data protection regulations that apply to AI and includes an in-depth analysis of European instruments. The processing of data is governed by data protection laws, which started to emerge in the last decades of the 20th century. Many principles of those laws (e.g., purpose limitation, data minimisation) are also relevant to AI applications and have been used by data protection authorities to sanction and ban AI applications. AI applications can process personal data to profile individuals or make predictions of their behaviours. Therefore, to develop ethical and trustworthy AI, it is also vital to ensure that data protection laws and their principles are fulfilled.
- Artificial Intelligence: This is a domain that poses its own challenges when it comes to regulations. These challenges include, for example, ensuring accountability in the AI-human interaction, appropriate levels of understandability and transparency regarding AI systems' purpose, how they work and how they use data to produce an output. Therefore, this class refers to regulations that are specific to Artificial Intelligence. The regulatory effort on AI is very recent and ongoing. This class includes 1) EU binding instruments, 2) EU non-binding instruments showing how to apply existing non-specific AI regulations (e.g., data protection laws) to AI, 3) US laws destined to govern AI and AI applications.



Human Rights

- Dignity, physical and mental integrity
- Freedom and autonomy
- Non-discrimination and equality
- Data protection and right to privacy
- Rule of law
- Social and economic rights

Data

- Lawfulness
- Fairness
- Purpose limitation
- Risk management
- Accountability
- Transparency
- Proportionality
- Privacy by Design
- Data minimisation
- Storage limitation

Artificial Intelligence

- Lawfulness
- Fairness
- Purpose limitation
- Risk management
- Accountability
- Transparency
- Accuracy
- Technical Robustness and safety
- Human agency and oversight

Figure 2: Visual representation of the taxonomy and the key principles within each class

For each class, the taxonomy indicated several principles related to the ethical and lawful use of AI in the security domain aiming also to the increase of public trust. All relevant information is included in D2.2[41].

3.2.1.3 T2.4 - From ethical frameworks to ethics in practice

As part of this task, ethics, AI ethics and the main ethics mechanisms were defined and the key ethical concerns regarding the use of AI for crime prevention and investigation, migration management, administration of justice, cyber operations, and LEAs' training were analysed. In addition, the methodology that was used to develop the popAI taxonomy and guidelines were outlined and the gaps between the current LEA practices and the AI ethical and legal framework were identified.

The ethical frameworks in the LEA and AI space were documented in a systematic, extensible taxonomy, drawing on published materials and internal reports used by security actors and technology providers. This taxonomy highlighted the main common points between frameworks, identify differences and challenges to their implementation. The outcome of this task therefore included a novel taxonomy of ethics principles in the LEA/AI space, but also identified gaps and challenges that fed into WP3, where stakeholder attitudes were explored and tested. All relevant information is included in D2.4 “Ethical frameworks for the use of AI by LEAs”, a public pop-AI deliverable that is publicly accessible through the official project’s website[42].

In addition, a practical Ethics Toolbox for the use of AI by LEAs is created as part of this task[43]. This is the final stage of a joint effort between the popAI functionality taxonomy and emerging practices and trends (D2.1), legal taxonomy (D2.2), the controversies report (D2.3) and the ethical frameworks for the use of AI by LEAs (D2.4)[44]. Extended reference to the Ethics Toolbox is made in section 4.2.2.

3.2.1.4 T2.5 - AI meets organisational cultures: Human-machine interaction at the police station

This task provides a comprehensive mapping of the current and future challenges along three axes:

1. How artificial intelligence (AI) is incorporated in the organisation.
2. Perception of AI by LEAs.
3. Expectations around the introduction of AI.

These axes are integrated by three main information sources composed by a literature review on the main debates about AI and LEAs, results of multiple Policy Labs to discuss the main debates on the area, and interviews to the personnel of LEAs. The theoretical discussion of the main literature debates integrating AI into LEAs are mapped from the current and future challenges of human-AI interaction in LEAs. Likewise, this discussion emphasises the actual situation on predictive policing, automation bias and the role of “human-in-the-loop”.

[41] popAI D2.2 “Legal framework and casework taxonomy: emerging trends and scenarios”
[42] popAI D2.4 “Ethical frameworks for the use of AI by LEAs”
[43] PopAI D2.5 “Practical ethics toolbox for the use of AI by LEAs”
[44] popAI D2.1 “Functionality taxonomy and emerging practices and trends”, D2.2 “Legal framework and casework taxonomy: emerging trends and scenarios”, D2.3 “The controversies and risks that have shaped innovation and will shape AI in the next 20 years”, and D2.4 “Ethical frameworks for the use of AI by LEAs”

As a result of the academic research, Policy Labs, interviews and interaction with LEAs important conclusions were drawn including: the decisive role of humans and the supportive role of AI systems, the requirement of holistic education to ensure that they are used responsibly, social benefit as the ultimate goal, the necessity of transparency and, consequently, acceptance of AI systems by society. A debate was generated about the potential misuse of AI systems, high costs for proper implementation and adaptation difficulties. Extended information can be found in D2.6[45].

3.2.1.5 T3.1 - Map the controversy ecosystems of AI tools in the security domain

To map the ecosystem of AI tools in security domain and explore the involved stakeholders, the key expressed benefits, concerns, and risks, this task identified and analysed relevant controversial cases. The mapping of controversy ecosystems was evolved around six broad civil security domains: crime prevention, crime investigation, migration, asylum, and border control, administration of justice, cyber operations for law enforcement, and LEAs' training. The main controversial issues identified referred to privacy, gender and racial discrimination, lack of transparency and accountability both in the design of the technologies as well as in the policies and procedures of their employment, and violation of fundamental rights. The stakeholders' categories emerged were namely, LEAs, police academies, social and humanities research, policy makers, government and public bodies, technologists and data scientists, civil society organisations, national and local authorities, and ICT and software companies. It is important to note that research findings indicated asymmetries in the level of involvement in technology development and policy making of different stakeholder categories and the need to include more silent voices. Extended information is provided in D3.1[46].

3.2.1.6 T3.3 - Crowdsourcing stakeholder attitudes and pro-active solutions ideation

This task actively engaged EU citizens in identifying key concerns and best practices for avoiding harm and ensuring benefit from AI in policing. This happened in two phases. The first phase was based on the controversies identified in WP2, from which targeted questions were prepared to assess and compare how citizens from different countries and backgrounds understand and experience the controversies. The second phase engaged the crowd to suggest solutions for the priority concerns. Citizens suggested through forums and open questions concrete solutions on how they would want their local police and the EU to deal with the issues. Finally, e-voting was conducted to find the most appropriate solution for each challenge and controversy identified. Extended information is provided in D3.3[47].

3.2.1.7 T3.4 - Engaging LEAs and relevant experts through policy labs

As part of this task, five Policy Labs in five different countries (Greece, Germany, Slovakia, Spain, Italy) were organised that involved LEAs and other experts such as technology developers, policymakers and NGO representatives. In September 2023 the sixth Policy Lab will take place in Brussels where all previous results will be presented and discussed as part of a final conference. During the five Policy Labs, fruitful discussions took place in the mother language of the participants and information was exchanged in order for best practices to be identified that can be shared with other actors throughout

[45] popAI D2.6 "AI meets organisational cultures: Human-machine interaction at the police station".
[46] popAI D3.1 "Map of AI in policing innovation ecosystem and stakeholders"
[47] popAI D3.3 "Citizen produced priorities and recommendations for addressing AI in the security domain"

the EU, ideas to be developed that will overcome controversies, the outcome of such development processes to be tested in an experimental setting and, finally, in order to assess whether or not public policy change is needed to ensure smart innovation. Each Policy Lab was focused on policy needs in relation to fundamental human rights, liability, proportionality, gender and diversity and it also covered organisational challenges. The outputs included both region-specific and EU-wide recommendations which the present deliverable takes into consideration. Extended information is provided in D3.4[48].

Due to the multidisciplinary structure of the Policy Labs and the variety of the items discussed, special emphasis has been placed on this source and detailed information per Policy Lab as well as presentation of their key outcomes can be found in Annex A of the present deliverable.

3.2.1.8 T3.5 - Multidisciplinary foresight scenarios

This task adjusted the foresight scenario methodology to the purposes and objectives of popAI to develop scenarios that will support the development of a "popAI Roadmap of AI in Law Enforcement 2040". Five foresight scenarios were developed depicting plausible futures in the next 5 years responding to the AI employment in the five main domains of civil security, namely, crime prevention, crime investigation, cyber operations, migration, asylum, and border control, and administration of justice. The scenarios demonstrated the capabilities of different technologies that are currently in the development or in pilot phase to support LEAs' operation. The scenarios also considered the AI Act under development and expected to impact the use of AI by LEAs. The main technological advancement addressed in all five scenarios is the capacity of algorithms to provide predictions, rankings, and recommendations based on identified patterns or specific design and databases. The risk of mass surveillance and the uncertainties regarding both the technological feasibility and an appropriate regulatory framework emerged as key points of discussion.

3.2.2 Work Package 4

3.2.2.1 Recommendations from citizens for the ethical use of AI by LEAs (T4.2)

Within this task, the challenges for citizens towards the deployment of AI by LEAs, as identified under WP2 and WP3, were carefully examined by following an anthropocentric approach, i.e., in the sense of using AI at the service of the citizens and aiming to social benefit. The rationale behind this is that AI should be developed and deployed in an inclusive and trustworthy way that reflects citizens' values and fosters their trust. Therefore, specific recommendations were provided by citizens and other interested stakeholders utilising the content of D3.3 and D3.4 and were addressed to LEAs, policymakers, technology developers and to the citizens as such. Extended information is provided in D4.2[49].

3.2.2.2 Recommendations from technology developers for the ethical use of AI by LEAs (T4.3)

As part of this task, a set of recommendations were delivered to the technology developers when designing and developing AI-based technologies and processing data. The principles for a trustworthy

[48] popAI D3.4 "Stakeholder attitudes, priorities, and recommendations for addressing AI in the security domain in practice"

[49] popAI D4.2 "White Paper for Civil Society"

AI as well as the outcomes of WP2 that were LEA-oriented were taken into consideration, to lead to the design and development of AI tools which are accepted and valued by both LEAs and citizens. An interaction among legal and technical actors to translate the legal rules and the ethical principles to technical specifications and vice versa was of high priority. Extended information is provided in D4.3[50].

3.3 Communication with LEAs and policymakers

3.3.1 Questionnaire

Since T4.1 focuses on the provision of recommendations to and also *from* LEAs and policymakers, further to the involvement of these categories in the aforementioned Policy Labs, a specific anonymous questionnaire was prepared and addressed to them in order to collect valuable insights on LEAs' and policymakers' perception of AI (see Annex B of the present deliverable).

The questionnaire was sent to LEAs and policymakers within and outside of the popAI Consortium. Ultimately, it was completed by police officers of the Hellenic Police (4), the University of Applied Science - Police Affairs, in short BayHfoeD, (3) and Madrid Municipal Police (1) that belong to the popAI Consortium and police officers of the Vilnius Municipal Police (2), Krakow Municipal Police (2) and Valencia Local Police (3) that are external to the popAI Consortium, as well as it was completed by the city of Turin, in short PLTO, (1) that is a member of the popAI Consortium.

The questionnaire consisted of three main sections. Through the first set of questions, we tried to understand the level of awareness and training of the participants on AI matters and to learn whether AI-enabled tools are currently used by them. Through the second set of questions, we aimed at finding out on which fields (recognition, communication, prediction & analytics, surveillance or other fields proposed) and for what purposes (crime prevention, crime investigation, cyber operations, migration, asylum & border control, administration of justice, LEA's training or other purposes proposed) the participants would find helpful the use of AI, what kind of risks they can identify and what types of mitigating measures they recommend. Through the third and final set of questions, we aimed at assessing the level of organisational readiness and compliance of the participants with the current ethical and legal framework as well as we wanted to learn what kind of support or assistance they need to fulfil the ethical and legal requirements.

Based on the types of the questions (some of them provided to the participants predefined answers to choose and evaluate) and the responses received, we can claim that the questionnaire has a twofold function, mainly as a source for the production of recommendations/emerging best practices for the ethical use of AI by LEAs and policymakers but also as a way that enables us to evaluate the recommendations presented in the following chapters 4 and 5 (see below section 7.1).

3.3.2 Results

3.3.2.1 Introductory questions on the current use of AI by LEAs and the level of LEA's awareness

With respect to the first set of questions:

[50] popAI D4.3 "White Paper for Technology Developers"

Only in Greece there is currently in force a national legislation on AI[51]. The law establishes obligations for public bodies and private entities that produce, distribute, and use these technologies as well as guarantees for the protection of natural and legal persons against the risks involved in the use and operation of these systems in the public and private sectors. It is explicitly not applicable to LEAs[52].

Two participants replied that their departments have started deploying AI-based tools in the field of crime prevention (with one of these tools aiming to support victims of gender-based violence) and that the whole process is still at an early stage. Another participant replied that their department has started using an AI chatbot aimed at tackling sexual harassment in transportation, with the Committee on the Ethical Use of AI[53] having already approved the tool and the whole process being at a test phase. Furthermore, the same participants responded that a specific department exists that makes use of such tools. In one of the cases, it was explained that a multidisciplinary team of experts is responsible for the evaluation of the AI outcomes, while in the other two cases a more general reference to the IT and technological-investigational departments was made. Citizens have not yet been informed of the use of these AI tools given that their deployment is still at an early/test phase, but recommendations for effective citizen-awareness tactics were provided. The rest of the participants are not using any AI technologies or tools.

Finally, most responses related to the level of awareness on the use of AI-enabled technologies and tools were negative. Only a few participants stated that they have received theoretical knowledge (e.g., through workshops, conferences, seminars) and only one of the participants has had practical knowledge through training. All participants, with no exemption, highlighted the necessity of educational and training courses on AI offered by experts with different backgrounds.

3.3.2.2 Questions on the usefulness of AI, identified risks and proposed mitigating measures

Regarding the second set of questions:

Most participants found the use of AI helpful in all fields included in the relevant question of the questionnaire (i.e., recognition, communication, prediction & analytics, surveillance). One of the participants added that AI would be useful to generally support decision-making. In addition, most participants found the use of AI helpful for all the purposes included in the relevant question (i.e., crime prevention, crime investigation, cyber operations, migration, asylum & border control, administration of justice, LEA's training).

Furthermore, most participants replied that the ethical and legal compliance of the AI technologies would make them appropriate for operational use and put special emphasis on human oversight, transparency, technical robustness, accuracy and reliability (this in conjunction with the capability of the system to provide quick outcomes), data protection, high quality of datasets and mitigation of biases.

[51] Greek Law 4961/2022 "Emerging information and communication technologies, strengthening digital governance and other provisions", Government Gazette 146/A/27-07-2022

[52] Law 4961/2022, Article 4 par.2

[53] The Committee on the Ethical Use of AI is a group of experts that will decide whether to use or not a certain AI technology considering its ethics. It will be formally presented within the city of Turin in September 2023.

Similarly, most participants replied that lack of ethical and legal compliance and, most importantly, overreliance on the AI outcomes without proper user validation, lack of transparency (black box), inaccuracy (false positive alerts), creation of unfair or discriminatory outcomes, security and privacy concerns would make AI inappropriate or unpractical for operational use. Further to this, very high technical and operational cost, lack of appropriate training of the LEAs on the use of AI, not user-friendly interfaces and inability of the AI system to handle real-world complexities were considered as negative factors increasing inappropriateness.

With respect to the risks identified by the participants on the use of AI by LEAs, almost all of them were related to the potential infringement of fundamental human rights and, mainly, the right to privacy and data protection, the right to equity and non-discrimination (e.g., by stigmatising specific individuals or groups, or by leading to over-policing in certain communities and exacerbating existing social inequalities) and presumption of innocence. Additionally, further concerns were expressed about overreliance on the AI outcomes without proper human judgment, black box AI, LEAs' AI-supported decisions through the prism of admissibility and reliability of evidence in court and about malicious attacks against AI systems that are not technically robust and resilient.

As regards the mitigating actions proposed by the participants, almost all of them placed emphasis on the necessity of a strong harmonised legislation, humancentric AI, transparent and explainable AI, regular tests, monitoring and evaluation of the AI systems in order to verify that they function safely and properly, to prevent or early detect biases and to reduce false positives, implementation of security measures by design to confront cyberattacks and strict data governance, clear policies, protocols and training on the ethical use of AI. Also, the implementation of tailored measures based on the nature of each identified risk was recommended.

Finally, all participating LEAs supported the use of "real-time" remote biometric identification systems in publicly accessible spaces. Based on the responses, they prefer their classification as high-risk AI systems that are subject to strict obligations including prior acquisition of judicial authorisation (as proposed by the EC and the Council) and their use in specific or urgent cases[54] instead of their absolute prohibition (as proposed by the EP). Open dialogue with the policymakers was recommended. Only the participating policymaker could not come up with any cases where the use of such systems would be useful.

3.3.2.3 Questions on the level of organisational readiness and ethical & legal compliance

As for the final set of questions:

Most participants claimed that they do not consider their department ready to fulfil the key ethics requirements for trustworthy AI (AI HLEG Ethics Guidelines) and to meet the obligations stipulated by the Draft AI Act and that they also find their implementation somewhat difficult, apart from the

[54] E.g., in emergency situations where individuals are missing, such as natural disasters or terrorist attacks, to assist in locating and identifying missing persons, enabling authorities to reunite them with their families or provide necessary assistance; in large public gatherings (stadiums, festivities, demonstrations) to spot missing individuals or known criminals and intervene timely and effectively to enhance public safety; at border crossings, airports, or ports to assist in the identification of known offenders and to improve border control and immigration processes.

requirements for human oversight and data protection which were considered easier to be implemented by some participants. Interesting was the opinion of one of the participants stating that, to be able to determine whether a department is ready to comply with the obligations, several questions need to be answered starting with this of whether the department is aware of the ethical and legal requirements and their significance. Only a few participants, including those that are currently using AI, felt more prepared to meet the requirements but, still, found the implementation difficult in some cases. The procedures that are followed by these participants include human oversight, GDPR compliance, impact assessments and collaboration with experts.

Further to this, the participants were asked to choose from a list of predefined recommended practices what kind of support or assistance they need to fulfil the ethical and legal requirements (see more about the validating character of this question in section 7.1 below). In addition, the participants expressed their opinion on how they are planning to involve citizens and increase their trust towards the use of AI tools. The recommendations, which are also considered to be emerging best practices to ensure the ethical use of AI-enabled technologies by LEAs, are listed and analysed in chapters 4 and 5.

Finally, as a concluding remark, most participants replied that they need time to establish the necessary procedures and implement the appropriate measures and that they will start using AI tools only after they have ensured compliance. Only two of them responded that they have already started or will immediately start to take the necessary actions.

3.4 Involvement of Ethics Advisory Board and Stakeholders Advisory Board

The EAB has been established as part of the popAI project with the intent to cover ethical aspects of the research and development activities and to help ensuring that the relevant standards for responsible research and development will be met throughout the project. It consists of three internal members with expertise on ethics and law and the Project Coordinator.

The SAB has been established as part of the popAI project to provide feedback to the projects' results. It consists of 6 members that are external to the popAI Consortium with 2 of them being members of the sibling projects ALIGNER and STARLIGHT.

As regards the recommendations that are provided through the WP4 deliverables for the ethical use of AI by LEAs, there is a twofold function of the EAB and SAB involvement both:

(a) as a source for the production of best practices through:

- the active participation of the EAB in project meetings where the AI-related legislative developments and relevant ethical concerns were communicated to the attendees and a dedicated online meeting between KEMEA researchers and two members of the EAB (EAB's chair and EAB's member from KEMEA) where a discussion on the content of the present deliverable took place and guidance was provided on the proposed best practices,
- questions addressed to the SAB members (along with externals present) during the popAI project meeting in Rome where the SAB members had the opportunity to express their opinion on the use of AI by LEAs. Some of the input collected includes highlighting the importance of the human oversight principle, that cases of fight against terrorism and criminal

investigation could be exceptions where the use of more intrusive AI tools may be justified, that the adoption of AI technologies may require effort, costs and time, especially for SMEs, and that LEAs could be more outward looking when it comes to the adoption of AI tools. The SAB mentioned that it would be important for LEAs, before using a new AI tool, to publish an open paper, sharing quantitative and qualitative data about the challenges that they may face.

(b) as evaluating bodies (see below chapter 6).

3.5 Sibling projects (ALIGNER, STARLIGHT)

The developments of ALIGNER and STARLIGHT have been monitored from the starting date of popAI, while the SAB consists of -among others- three representatives from ALIGNER and STARLIGHT. The popAI researchers have participated in the sibling projects' workshops. While previous workshops served as an initiative of collaboration expressing the need for exchange of best practices, the following workshops contributed to the present deliverable: ALIGNER, popAI, STARLIGHT, AP4AI projects Joint Workshop: "Ethical and Legal Aspects of AI for Law Enforcement"[55] which included an exchange of opinions on the ethical use of AI and suggested best practices and the 5th ALIGNER Workshop[56] where popAI, STARLIGHT and ALIGNER presented their Policy Recommendations. In addition, communication with the Coordinator of ALIGNER was initiated to gain access to two essential deliverables regarding Policy Recommendations.[57] In sum, the recommendations produced from the workshops which relate to the suggestions herein are:

- the conduct of an impact assessment even when not obliged by law,
- the ALIGNER FRIA as described below in section 4.3.1.2,
- the need for harmonisation of the legal framework at EU level,
- the standardisation of the AI procurement procedure,
- and the collaboration among disciplines for the ethical use of AI by LEAs.

4 Recommendations for LEAs (emerging best practices)

This chapter aims at helping LEAs benefit from AI while using AI-enabled technologies in conformity with the ethical and legal framework so that the fundamental rights are prioritised and respected and trust of the affected persons is fostered. Therefore, specific recommendations have been produced for the ethical use of AI by LEAs. The recommendations emerge from the popAI stakeholder community, involving LEAs, policymakers, civil society representatives, legal and ethics experts, technology developers (bottom-up recommendations), but also by the popAI researchers based on their background (top-down recommendations).

[55] ALIGNER, popAI, STARLIGHT, AP4AI projects, Joint Workshop: "Ethical and Legal Aspects of AI for Law Enforcement", January 25th and 26th 2023, CEA premises in Brussels, Press release: https://www.pop-ai.eu/wp-content/uploads/2023/02/Ethical-and-legal-aspects-of-AI-for-law-enforcement-Conclusive-Statement.pdf

[56] 5th ALIGNER Public Workshop, June 2023

[57] ALIGNER D2.3 "Policy recommendations" and "D5.5 First Update of the Research Roadmap for AI in Support of Law Enforcement and Policing"

In particular, the recommendations derive from the following groups:

- the LEAs themselves (Policy Labs discussions, questionnaires completed by police officers within and outside of the popAI Consortium),
- technology developers (Policy Labs discussions, questionnaires completed by industry, content of D4.3),
- citizens (crowdsourcing platform of D3.3, content of D4.2) as well as
- researchers working for the popAI WP1, WP2, WP3 and WP4 deliverables and
- ethical and legal experts and stakeholder advisors involved in the popAI project and its sister projects (dedicated discussions with the project's EAB and SAB, Policy Labs discussions, participation of popAI in sibling projects' workshops).

The current ethical and legal framework on AI constitutes the basis on which all the aforementioned sources have been filtered to ensure ethical and legal compliance.

4.1 Ethics, privacy, and security by design

To be able to use high-risk AI in an ethically and legally compliant manner, it must be first ensured and demonstrated by the providers of such systems that they have been designed and developed in compliance with the applicable ethical and legal framework (see above section 3.1.2.3). Furthermore, in order to ensure a high level of trustworthiness of high-risk AI systems, some high-risk AI systems should be subject to a conformity assessment prior to their placing on the market or putting into service[58]. High-risk AI systems that have already been subject to a conformity assessment procedure shall undergo a new conformity assessment procedure whenever they are substantially modified, regardless of whether the modified system is intended to be further distributed or continues to be used by the current deployer[59].

Extended information about the design and development of high-risk AI can be found in D4.3 where relevant recommendations for providers of high-risk AI systems are presented and analysed.

Albeit the ethically and legally compliant design and development of AI technologies is not a recommendation, yet an obligation addressed by the EU legislator to the providers, it is necessary to start with this reference as it constitutes the first step and was also highlighted in several popAI deliverables as well as by the Policy Labs and questionnaire participants.

Therefore, it is strongly recommended that the deployers of high-risk AI systems have the necessary knowledge on what to ask for and what to expect by the providers as well as the capability to monitor the operation of such systems efficiently and regularly on the basis of the instructions of use. The deployers must without undue delay get in contact with the providers (as well as the distributors and the competent national supervisory authorities) when they have reasons to consider that the use in

[58] See Article 43 par.1 and Annex III point 1 of the Draft AI Act.

[59] European Parliament, Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, Article 43 par.4

accordance with the instructions of use may result in the AI system presenting a risk or in cases of a serious incident or malfunctioning and interrupt or suspend the use of the system[60].

4.2 AI literacy

4.2.1 Education and training

The use of AI systems by LEAs should be accompanied by prior and regular comprehensive education and training to ensure that the users and, most importantly, the natural persons to whom the operation of an AI system is assigned, understand how the system functions as well as what their role and their ethical and legal obligations are and to guarantee that they are skilled enough to use it correctly, safely and responsibly in conformity with the regulatory framework.

The necessity of education and training was the most repeated recommendation provided by the participants in the Policy Labs and in the questionnaire of section 3.3. This reveals not only the admitted lack of relevant knowledge but also the strong interest of the LEAs in the deployment of AI in an ethically and legally compliant manner.

The necessity of AI literacy, awareness raising and information communication was also added to the latest version of the Draft AI Act proposed by the EP where it is stipulated that literacy measures through education, training, skilling and re-skilling programmes should include the teaching of basic notions and skills about AI systems and their functioning, including the different types of products and uses, their risks and benefits[61]. Additionally, the EP proposed that to the extent deployers exercise control over the high-risk AI system, they shall ensure that the natural persons assigned to ensure human oversight of the high-risk AI systems are competent, properly qualified and trained, and have the necessary resources in order to ensure the effective supervision of the AI system[62]. The emphasis put on AI education and training by the LEAs and policymakers participating in the Policy Labs and the questionnaire heavily supports the amendments made by the EP and the inclusion of the said provisions in the final version of the AI Act.

Types:

- **Ethical and legal education and training:** educational programmes (webinars, seminars and workshops, written guidelines) and training/skilling courses, either organised by or indicated for LEAs, to provide theoretical and practical knowledge on the ethical and legal considerations and risks that may derive from the use of AI by LEAs (e.g., on fundamental rights, on admissibility of evidence at court), the ethical rules, the applicable legal framework and the obligations, limitations and prohibitions stipulated by law and, soon, on issues clarified, revised and regulated through case law. At this stage, considering that the AI Act is in progress but will be finalised soon, it is of high importance that the users obtain sufficient information about the basic notions, the legislative developments so far and, once issued, about the content of the AI Act in order to understand to what extent and how it affects them.

[60] Ibid., Article 29 par.4
[61] Ibid., Article 4b
[62] Ibid., Article 29 par.1a(ii)

- **Technical education and training:** educational programmes (webinars, seminars and workshops, written guidelines) and training/skilling courses, either organised by or indicated for LEAs, to provide theoretical and practical technical knowledge enabling users to understand AI algorithms and best practices for data collection, data preparation and model training in general and also with respect to the deployment of a specific AI system by LEAs, as well as to be informed about that AI system's purposes and technicalities including its capabilities and limitations of performance, pre-determined changes, human oversight measures, potential bias, accuracy metrics, expected lifetime along with any necessary maintenance and care measures, etc (see section 3.1.2.3 above) and, ultimately, to use it correctly and to be able to evaluate properly its outputs.

Frequency:

- **Before** the deployment of an AI technology or tool for the users to be well informed and prepared and for all required actions and measures to be taken timely.
- **Regularly**[63] **(before and after the deployment of an AI system)**. The frequency of educational programmes and trainings can vary depending on the gaps and needs identified in each deployer's police department, the potential modification of an AI system in use or of its purpose, the introduction of new AI systems in a police department, or the pace of the AI advancements and of the legislative and case-law developments in general. The organisation of courses on a regular basis and of follow-up sessions are considered necessary to keep the users constantly informed and updated and to ensure that operators remain properly and adequately trained and supervised.

Educators/trainers:

Considering the recommended types of educational and training programmes, the educators and trainers should be persons with relevant knowledge, experience and expertise (see also below the recommendation of section 4.5).

- **Ethics and legal advisors along with policymakers** will provide information on the applicable ethical and legal framework and discuss with LEAs about the AI-related ethical and legal considerations and risks as well as they can provide guidance and consultation to LEAs for the minimisation of the risks and the trustworthy deployment of AI.
- Next to them, the **involvement of society representatives** in educational and training programmes would be highly beneficial for the identification of all potential risks from the AI subjects' perspective and for the proposal of best practices and measures that will minimise them and, consequently, will increase the trust of the affected persons.
- **Technology developers** will share their knowledge and expertise on the design, development and implementation of AI systems. Providers of AI systems planned to be used by LEAs are

[63] Every six months or at least once a year were some of the recommendations provided by the questionnaire participants, while others preferred to mention that the frequency may vary depending on the circumstances.

obliged to provide information about the functioning of the systems and the reasoning behind the outcomes and to give instructions for use and, therefore, they are ideal to train LEAs.

- **Representatives of other law enforcement agencies** that are using AI could also provide their valuable insights and share their experience on the procedures that they are following in order to implement and be in line with the AI Act[64].

Implementation:

- **At national level:** Educational and training programmes on AI can be an initiative of LEAs in collaboration with the national competent authorities (e.g., Ministries). AI education and AI training can be provided through the organisation of relevant seminars and courses by LEAs to LEAs to ensure that the staff will be educated, well-trained and skilled, and, consequently, capable of using AI systems and dealing with AI-related issues. Furthermore, the addition of relevant courses to the police academies' curriculum for the preparation of future police officers would be highly beneficial.
- **At European level:** The participation of LEAs in EU-funded research and innovation projects that aim to the design and development of AI systems will help LEAs familiarise themselves with AI-enabled technologies and tools, define user requirements and system requirements of such technologies, test them through training sessions and pilot demonstrations and, finally, evaluate them prior to their deployment[65].

4.2.2 The example of the popAI Ethics Toolbox

The Ethics Toolbox for the use of AI by LEAs is part of the objective of the popAI project for the creation of an EU AI innovation hub for LEAs and the broader community[66] and is composed by three main sections:

1. **Educational Videos on AI and Ethics**:

   Eight educational videos provide a collection of multimedia material that explores the intricate relationship between AI and ethics within the context of policing. These videos are created to offer valuable insights and facilitate a deeper understanding of the ethical implications surrounding the use of AI technologies in law enforcement. The objective is to equip viewers (students, educators, and the general public) with basic knowledge to engage in informed discussions and make sound decisions regarding the responsible integration of AI in policing. Indeed, each video explores different aspects of AI in policing, addressing key questions and shedding light on its implications.

   The first video introduces the concept of AI in law enforcement and explores its practical applications. In the second video the focus shifts to ethics. It delves into the meaning and

[64] "Case studies of how other entities apply the AI Act" was a way of support recommended by many questionnaire participants.

[65] For the implementation of educational and training programmes at both national and European levels, 'additional funding to cope with the additional efforts' was a way of support recommended by many questionnaire participants (see also below section 5.2.6).

[66] More information about the AI Hub will be provided in popAI deliverable D5.7 "Sustainability and exploitation plan".

complexities of ethics, particularly within the realm of policing. The third video dives deeper into the complex relationship between AI and ethics, exploring the ethical considerations of AI used in security issues. The fourth video addresses the risk of bias and its mitigation in AI. In the fifth video, the spotlight is on ensuring traceability and accountability in AI systems. The sixth video tackles the question of responsibility in advancing and monitoring AI development. The seventh video examines how citizens view AI usage by the police. Lastly, the eighth video focuses on the impact of AI on police organisational needs.

2. **Technology Ethics Briefs**

   The technology ethics briefs engage the discussion on the use of technology in public security areas and the 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. The ethics briefs contain three cases that explore how artificial intelligence is used in law enforcement, specifically focusing on the aforementioned predictive analytics, natural language processing, and image recognition. The aim of these 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. Each of the briefs contains real cases or applications of AI, for example Swedish Land Registry model, Kamu chatbot, Key Crime Delia software, SARI image recognition software, and more.

3. **Interactive Visualisation on AI and LEAs Ethics Taxonomies**

   These taxonomies are integrated into an online webpage which serves as an open reference that can be consulted through an interactive visualisation on the popAI webpage. The visualisation will serve as a translation tool for the work accomplished under popAI deliverables D2.1, D2.2, D2.3, and D2.4[67]. The website shows to the user the multiple types of cases, applications, concepts, and terms in which the taxonomies can be displayed. This living tool for analysis and discussion is an important resource for anyone interested in the ethical implications of AI in law enforcement. Finally, the visualisation offers an interactive platform for searching and accessing relevant documents, supporting researchers, practitioners, and policymakers in navigating through AI and LEA ethics. The image below represents the online version proposed for the taxonomy visualisation.

[67] popAI D2.1 “Functionality taxonomy and emerging practices and trends”, D2.2 “Legal framework and casework taxonomy: emerging trends and scenarios”, D2.3 “The controversies and risks that have shaped innovation and will shape AI in the next 20 years”, and D2.4 “Ethical frameworks for the use of AI by LEAs”

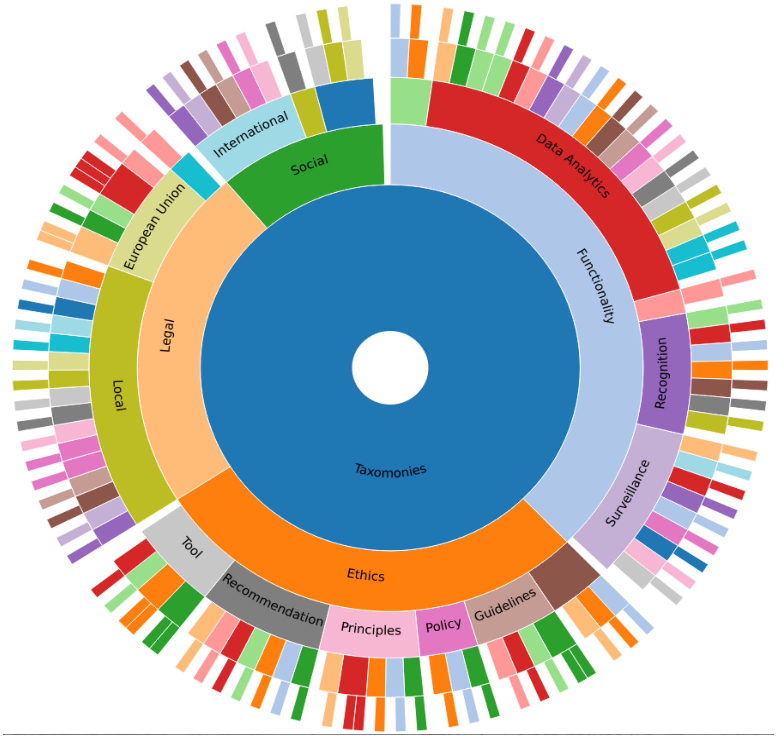


Figure 3: Taxonomy visualisation

4.3 Impact Assessments

Risks related to AI systems can result not only from the way such systems are designed, but also from the way such AI systems are used. Therefore, deployers of high-risk AI systems play a critical role in ensuring that fundamental rights are protected, complementing the obligations of the providers during the AI development phase. Deployers know for what purposes a high-risk AI system will be used, hence, they can identify potential significant risks that were not foreseen in the development phase, by taking into account the context of use, the people likely to be affected, including marginalised and vulnerable groups. To this end, in order to effectively ensure that fundamental rights are protected, the deployers of high-risk AI systems should conduct a fundamental rights impact assessment prior to putting such systems into use. The impact assessment should include a detailed plan describing the measures or tools that will help minimising the relevant risks identified at the latest from the time of putting the high-risk AI systems into use, otherwise the deployers should refrain from their utilisation. When performing this impact assessment, the deployers should notify the national supervisory authority and, to the best extent possible relevant stakeholders as well as representatives of groups of persons likely to be affected by the AI systems in order to collect relevant necessary information. They are also encouraged to make the summary of their fundamental rights impact assessment publicly available on their website[68].

Further to this, for compliance with the applicable data protection legislation, where applicable, a data protection impact assessment must be conducted and, in order to ensure that the affected

[68] European Parliament, Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, Recital 58a and Article 29a

persons are informed and heard, the fundamental rights impact assessment should be accompanied by an ethical and social impact assessment.

Finally, it is worth pointing out that all types of impact assessments must be reviewed on a regular basis and updated whenever needed as well as that as regards new, disruptive technologies the carrying out of impact assessments is highly recommended even in cases where this is not obligatory by law.

4.3.1 Recommended templates of fundamental rights impact assessments

Due to the emphasis put by the Policy Lab and the questionnaire participants on carrying out impact assessments prior to the deployment of a high-risk AI system and considering their wish to have relevant templates, this section aims at facilitating LEAs by presenting recommended templates of fundamental right impact assessments.

4.3.1.1 Human Rights, Democracy and the Rule of Law Impact Assessment (HUDERIA)

HUDERIA forms part of the proposal for a Human Rights, Democracy, and the Rule of Law Assurance Framework for AI systems prepared by the Alan Turing Institute[69] based on the Model for a Human Rights, Democracy, and the Rule of Law Impact Assessment (HUDERIA) presented in CAHAI-PDG(2021)05rev[70] and following the relevant request of the Secretariat of the Council of Europe's Ad hoc Committee on Artificial Intelligence (CAHAI)[71]. The purpose of HUDERIA according to CAHAI is to "define a methodology to carry out impact assessments of AI applications from the perspective of human rights, democracy, and the rule of law, based on relevant CoE standards and the work already undertaken in this field at the international and national level, and to develop an impact assessment model"[72].

The HUDERIA template can be found in the aforementioned proposal of the Alan Turing Institute[73] and provides an opportunity for project teams and engaged stakeholders to come together in order to produce detailed evaluations of the potential and actual impacts that the design, development, and use of an AI system could have on human rights, fundamental freedoms and elements of democracy and the rule of law.

[69] The Alan Turing Institute, Human Rights, Democracy, and the Rule of Law Assurance Framework for AI Systems: A proposal prepared for the Council of Europe's Ad hoc Committee on Artificial Intelligence, available at https://rm.coe.int/huderaf-coe-final-1-2752-6741-5300-v-1/1680a3f688

[70] Ad Hoc Committee on Artificial Intelligence (CAHAI), Policy and Development Group (CAHAI-PDG), Human Rights, Democracy and Rule of Law Impact Assessment of AI systems, 21 May 2021, available at https://rm.coe.int/cahai-pdg-2021-05-2768-0229-3507-v-1/1680a291a3

[71] Ad Hoc Committee on Artificial Intelligence (CAHAI), Policy and Development Group (CAHAI-PDG), Human Rights, Democracy, and the Rule of Law Assurance Framework (HUDERAF) for AI systems, Executive Summary, 8 October 2021, available at https://rm.coe.int/cahai-pdg-2021-09-huderaf-executive-summary/1680a416de

[72] Ad Hoc Committee on Artificial Intelligence (CAHAI), Policy and Development Group (CAHAI-PDG), Human Rights, Democracy and Rule of Law Impact Assessment of AI systems, 21 May 2021, available at https://rm.coe.int/cahai-pdg-2021-05-2768-0229-3507-v-1/1680a291a3

[73] The Alan Turing Institute, Human Rights, Democracy, and the Rule of Law Assurance Framework for AI Systems: A proposal prepared for the Council of Europe's Ad hoc Committee on Artificial Intelligence, available at https://rm.coe.int/huderaf-coe-final-1-2752-6741-5300-v-1/1680a3f688, p.247-271

4.3.1.2 The example of the ALIGNER Fundamental Rights Impact Assessment (AFRIA)

As part of the sibling H2020 project ALIGNER, a template of a fundamental rights impact assessment was released which is exclusively addressed to LEAs who are planning to deploy AI systems for law enforcement purposes within the EU. The main objective of the AFRIA is to help LEAs make sure and prove that their use of AI is ethically and legally compliant.

The AFRIA consists of two connected and complementary templates:

(a) The Fundamental Rights Impact Assessment template which aims at helping LEAs identify the risks and mitigate the impact of law enforcement AI on those fundamental rights of the affected persons that are most likely to be infringed when LEAs deploy AI systems (presumption of innocence, right to an effective remedy, right to a fair trial, right to equality and non-discrimination, freedom of expression and information, right to respect for private and family life, right to protection of personal data), and

(b) The AI System Governance template which aims at helping LEAs identify the key requirements for trustworthy AI and mitigate the impact of law enforcement AI on fundamental rights.

The first template is divided in four parts. Each part includes a group of fundamental rights which is used as the basis for the following assessment. First, related challenges of the AI system are listed in the template, i.e., characteristics that may have negative impact on fundamental rights (challenge column). Based on these, LEAs need to describe whether and how the challenges relate to the AI system (evaluation column) and to assess the level of negative impact of the AI system on fundamental rights (estimated-impact-level column).

The second template is divided in seven parts. Each part includes a key requirement for trustworthy AI which is used as the basis for the following assessment. First, for each subcategory of a key requirement (component column) some necessary characteristics of the AI system are listed in the template (minimum-standards-to-be-achieved column) and when the minimum standard is suitable to minimise the negative impact of the AI system on fundamental rights the template connects it to the relevant previously estimated challenge (initial-impact-estimate column). Based on these, LEAs need to describe how the minimum standard is planned to be implemented (additional-mitigation-measures-implemented column) and to assess the final level of negative impact following the implementation of the mitigation measures (final-assessment column)[74].

The AFRIA template can be found online on the official ALIGNER website[75].

4.3.2 Data Protection Impact Assessment (DPIA)

Where applicable, as explained below, deployers of high-risk AI systems shall use the information provided under Article 13 of the AI Act[76] to comply with their obligation to carry out a data protection

[74] Donatella Casaburo, KU Leuven, 'The ALIGNER Fundamental Rights Impact Assessment: Mitigating the impact of law enforcement AI', accessible at https://www.law.kuleuven.be/citip/blog/the-aligner-fundamental-rights-impact-assessment-mitigating-the-impact-of-law-enforcement-ai/ 20 June 2023

[75] https://aligner-h2020.eu/fundamental-rights-impact-assessment-fria/

[76] Article 13 is about transparency and provision of information by the providers to the users of AI systems.

impact assessment under Article 35 of Regulation (EU) 2016/679 or under Article 27 of Directive (EU) 2016/680[77] (see above in section 3.1.5 when GDPR or LED apply).

According to Article 35 GDPR and Article 27 LED, where a type of personal data processing in particular using new technologies, and taking into account the nature, scope, context and purposes of the processing, is likely to result in a high risk to the rights and freedoms of natural persons, the data controller (i.e., the user of the AI system) shall, prior to the processing (i.e., prior to the deployment of the AI system), carry out an assessment of the impact of the envisaged processing operations on the protection of personal data. The DPIA is a living document that must be reviewed on an ongoing basis and updated based on any changes related to the scope, nature, context or purposes of the processing in question by following the opinion and advice of the DPO designated in the law enforcement agency.

A DPIA of Article 35 GDPR and, by analogy, of Article 27 LED shall in particular be required in case of:

(a) systematic and extensive evaluation of personal aspects relating to natural persons which is based on automated processing, including profiling, and on which decisions are based that produce legal effects concerning the natural person or similarly significantly affect the natural person;

(b) processing on a large scale of special categories of data referred to in Article 9(1), or of personal data relating to criminal convictions and offences referred to in Article 10; or

(c) systematic monitoring of a publicly accessible area on a large scale[78].

Therefore, with respect to the fundamental right of the affected persons to personal data protection and as regards data processing operations that are likely to result in high risks to the rights and freedoms of the AI and data subjects, a DPIA must be carried out by the user of the AI system prior to its deployment[79].

4.3.3 Ethical and Social Impact Assessment (ESIA)

Further to carrying out a fundamental rights impact assessment and a data protection impact assessment and to ensure that the affected persons' opinions are taken into account by the users of high-risk AI systems prior to the deployment of AI in law enforcement, the process of impact assessment and evaluation needs to include an ethical and social impact assessment (ESIA) in order to be complete. Certainly, this process is equally important prior to the development of a high-risk AI system.

The significance of ethical and social perspectives gives the opportunity to the providers and deployers of AI systems to have a broader perception which may be partially restricted by the

[77] European Commission, Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, 21.4.2021, Article 29 par.6; The EP also added the following to the said Article "[…] a summary of which (i.e., DPIA) shall be published, having regard to the specific use and the specific context in which the AI system is intended to operate. Deployers may revert in part to those data protection impact assessments for fulfilling some of the obligations set out in this article, insofar as the data protection impact assessment fulfils those obligations".

[78] See also the Article 29 Data Protection Working Party Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679, wp248rev.01

[79] See also the newly added by the EP Article 29a par.6 of the Draft AI Act.

individual dimension of fundamental rights and freedoms. Particularly in the Big Data and AI era when further ethical and societal concerns emerge, for example in terms of unforeseen bias or social acceptability, public participation is imperative.

The example of the Human Rights, Ethical and Social Impact Assessment (HRESIA) model which includes a self-assessment questionnaire in line with the traditional impact assessment approach, and an ad hoc committee[80] takes also into consideration the social and collective dimension of data use and seems to constitute a comprehensive, participatory and transparent tool which, albeit oriented to cover the relevant gaps of data protection laws, can help AI providers and deployers foster the trust of the AI subjects towards them.

4.4 Inclusion of the civil society

4.4.1 Awareness raising and transparency towards the AI subjects

Trustworthy AI means that the AI systems are designed and developed in a way that makes them understood, accepted and valued by the users and the affected persons as well as they are used in a way that makes them understood, accepted and valued by the affected persons.

As explained in section 4.2, a sufficient level of AI literacy is necessary for providers and users of AI systems in order to equip them with the notions and skills required to ensure compliance with and enforcement of the AI Act.

At the same time, it is equally necessary that the affected persons obtain information about the purposes, risks and impact of AI-based technologies and tools on the society and the environment and have the right to give their feedback. Especially, in the case of law enforcement, since the LEA actions that involve certain uses of AI systems are characterised by a significant degree of power imbalance and may lead to surveillance, arrest or deprivation of a person's liberty as well as other adverse impacts on fundamental rights guaranteed in the Charter, it is critical that:

(a) sufficient information about the AI systems planned to be used and about AI systems already being used in law enforcement is provided to the affected persons to enhance public understanding and trust and
(b) transparent information about the datasets used and the way in which the AI outcomes are produced is provided to the affected persons to ensure that no discrimination or stigmatisation is made against them and that their rights to presumption of innocence, defence, fair trial, effective remedy and personal data protection are not violated.
(c) a feedback mechanism is established in order to collect input on how to improve the system directly from those potentially affected thereby[81].

As added by the EP on this matter, it is necessary that, in view of allowing a democratic control of AI systems, the Commission, the Member States as well as providers and users of AI systems, in

[80] A. Mantelero, 'AI and Big Data: A blueprint for a human rights, social and ethical impact assessment', available at https://www.sciencedirect.com/science/article/pii/S0267364918302012

[81] Ad Hoc Committee on Artificial Intelligence (CAHAI), Possible elements of a legal framework on artificial intelligence, based on the Council of Europe's standards on human rights, democracy and the rule of law, 3 December 2021, p.12

cooperation with all relevant stakeholders, promote the development of a sufficient level of AI literacy, in all sectors of society, for people of all ages, including women and girls[82].

Further to this, the EP also added that given the potential impact and the need for democratic oversight and scrutiny, users of high-risk AI systems that are public authorities (such as LEAs) or EU institutions, bodies, offices and agencies should be required to register the use of any high-risk AI system in a public database[83].

What constitutes "sufficient" level of AI literacy is a notion that needs to be defined by policymakers or by case law along with the establishment of awareness raising and transparency tactics that are considered necessary and adequate to reach that goal and the establishment of feedback mechanisms. In addition, more information about the required public registry of high-risk AI systems used by public authorities needs to be provided by the policymakers (see in sections 5.2.1 and 5.2.5 relevant recommendations to the policymakers).

4.4.2 Recommended tactics

To start with what needs to be highlighted is that equality, diversity and non-discrimination must be prioritised by including all members of the civil society and, most importantly, vulnerable individuals or groups affected by AI by also ensuring proper gender and age balance.

Raising awareness and transparency tactics may vary depending on who needs to provide the information (e.g., provider, deployer), what are the most common communication means which citizens are familiar with and what the target audience is.

Some of the recommended practices as indicated in the work done as part of other popAI tasks and deliverables listed and presented above[84], through the Policy Labs and by the questionnaire participants are:

- The organisation of events (physical, online or hybrid) to inform citizens about AI technologies, their benefits, limitations, purposes, potential risks and impact and engage them in an open and transparent dialogue.
- Visits and talks at schools or universities to also involve younger people and foster their critical thinking.

[82] European Parliament, Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, Article 4d par.1

[83] Ibid., Recital 58a. See also a relevant reference in the Ad Hoc Committee on Artificial Intelligence (CAHAI), Possible elements of a legal framework on artificial intelligence, based on the Council of Europe's standards on human rights, democracy and the rule of law, 3 December 2021, p.12: "The CAHAI considers that the establishment of public registers listing AI systems used in the public sector, containing essential information about the system such as, its purpose, actors involved in its development and deployment, basic information about the model, and performance metrics, where appropriate, and the result of a HUDERIA, should be addressed in the context of a legally binding or non-legally binding instrument on AI in the public sector".

[84] Special attention was given to D4.2 'White Paper for Civil Society' as we value the opinion of citizens on this matter.

- Drafting of protocols and code of conduct governing the use of AI tools in law enforcement and making them publicly available (e.g., via the LEA official websites and other governmental websites).
- The creation of educational videos and campaigns (e.g., see above the Ethics Toolbox).
- The establishment of communication channels, public registers and feedback mechanisms informing citizens about the use of AI in law enforcement, the exact purposes of each AI system in use, the way in which the decisions are taken and the rights of the AI subjects as well as enabling citizens to interact with the users, provide feedback, ask questions, express concerns, raise objections about the use of AI tools and to exercise their rights.
- Conducting of ethical and social impact assessments prior to the development (for providers) and deployment (for users) of an AI system and prior to any changes made to that system or to its purposes, collecting valuable feedback about the citizens' expectations, concerns, fears and objections and actively involving the affected persons in the evaluation and validation of the AI systems used in law enforcement (see also above section 4.3.3).
- The establishment of multidisciplinary and diverse teams in the user's entity that also communicate with civil society representatives through communication channels and feedback mechanisms (see next section 4.5 for more information).

It is a task of policymakers to decide what the most suitable means are to ensure proper communication and interaction with the AI subjects and, consequently, foster citizens' trust.

4.5 Establishment of a multidisciplinary team

4.5.1 Structure and objectives

Collaboration is the key. To understand new disruptive technologies, their functioning, benefits, risks, and impact on society and the environment and to achieve the ethical and secure design, development and deployment of trustworthy AI, the active participation of persons with different backgrounds is imperative.

Therefore, for the realisation of all aforementioned emerging best practices it is highly recommended to adopt a multidisciplinary approach during the entire lifecycle of AI through the establishment of a multidisciplinary and diverse team of people that have knowledge and expertise on AI-enabled technologies, ethics and law and care for inclusion, diversity and social benefit. In case of AI planned to be used by LEAs, or on their behalf, for law enforcement purposes, such a team should include:

- Industry representatives,
- ethicists and lawyers,
- law enforcement agents,

that will give voice through communication channels and feedback mechanisms to

- civil society representatives.

Technology developers have the knowledge, skills and experience to design, develop and provide AI-based technological solutions based on the needs of their clients. They are aware of what is possible or not from a technical perspective and are able to implement it by making sure that the technologies and tools work efficiently and securely.

Ethicists and lawyers with knowledge and expertise on new technologies need to stand next to technology developers, guide and advise them on the obligations and limitations set out by law when designing AI systems. In addition, they need to stand next to deployers of such systems, guide and advise them on the ethical and lawful use of AI.

LEAs are deployers, using AI systems to support their daily operations, while specific police officers will be assigned to implement proper human oversight. They need to obtain knowledge from technology developers on the characteristics, capabilities, limitations and levels of accuracy of the system, remain aware of automation bias and be able to interpret the outputs of the system. At the same time, based on the specific purposes of the AI system, they need to identify the potential risks and minimise them by taking into consideration the ethics and legal experts' opinions on the "dos and don'ts" stipulated by law in order to be able to interpret also from an ethical and legal perspective the outputs of the system and decide how and when to use or not use them.

Representatives of the civil society need to be involved in the process due to the impact of AI on society and the environment and the requirement to create trustworthy AI systems accepted and valued by the affected persons. In this way, their expectations and objections will be heard and taken into account. Diverse and vulnerable groups must be represented, particularly, as regards the risks posed by AI systems on the fundamental rights of equity and non-discrimination (e.g., the Fundacion Secretariado Gitano participated in the Spanish Policy Lab representing the vulnerable group of Romanies).

Therefore, multidisciplinary and diverse teams need to be established both at the AI development and deployment stages with the rationale that in this way the AI lifecycle will be efficiently monitored from different perspectives by people with different backgrounds.

4.5.2 The example of the popAI Policy Labs

In the spirit of ensuring proper human oversight and fostering equity, non-discrimination, transparency and, consequently, citizens' trust, the popAI project proposes the establishment of a multidisciplinary team that consists of LEAs, legal and ethics experts, policymakers and technology developers which will be in communication with the civil society at each critical stage of the AI system engagement, from its design and development to its use and validation by the deployers.

The organisation of the popAI Policy Labs (see above section 3.2.3) which involved LEAs, industry, ethics experts, policymakers and civil organisations sets a successful example of pluralism with added value in the discussion on AI. As explained in the relevant section above, during the Policy Labs the participants had the opportunity to express their opinions on AI in the context of specific case studies. Prior to and during the deployment of AI systems for law enforcement purposes, the participants of a structure similar to that of the Policy Labs will have the opportunity to examine and constantly monitor the functioning of AI, its value and impact in the context of real cases.

For all aforementioned reasons, it is highly recommended that, for the ethical use of AI by law enforcement and compliance with the ethical and regulatory framework, LEAs establish an AI department/body/committee composed of a multidisciplinary team of LEAs, ethics and legal experts and technology developers that will exclusively deal with the use of AI-based technologies, will provide guidance and consultation to LEAs on AI matters, organise educational and training courses for the LEAs, will conduct impact assessments, inform and interact with the citizens via established communication channels and feedback mechanisms and schedule regular meetings to discuss relevant concerns, address potential risks early and recommend the appropriate mitigating actions. Close collaboration or association of such a body with the Data Protection Officer designated in the law enforcement agency will ensure higher levels of conformity of the AI systems used by the LEAs with the key requirement of data protection and governance. Finally, following the example of the GDPR and the role of the DPO[85], this body should be in cooperation with the national supervisory authority stipulated in Article 59 of the Draft AI Act[86].

5 Recommendations for policymakers (emerging best practices)

As described in the popAI Policy Briefs of the first year[87] and in the terminology of the present deliverable, the policymakers to which the recommendations of this chapter are addressed, are envisioned as key actors in shaping policies and creating new rules both at European level, such as the EC, EC DG Home, EU Parliament and EU Council and at national level, such as the Member States along with the Member State Parliaments, Ministries and Municipalities. In addition, the policymakers also include the civil society in the sense that the engagement of citizens in decision-making is one of the building blocks of democracy.

The recommendations emerge from the popAI stakeholder community, involving LEAs, policymakers, civil society representatives, legal and ethics experts, technology developers (bottom-up recommendations), but also by the popAI participants based on their background (top-down recommendations).

The main sources of the following recommendations are:

- The content of D1.6 regarding the first version of the Policy Briefs of the popAI project and the recommendations extracted by the popAI "Inclusion and diversity workshop"[88],
- The Policy Labs and the draft Reports of each Policy Lab as part of T3.4 with the inclusion of various stakeholders from the popAI community,
- The answers to the T4.1 questionnaires by LEAs and policymakers as mentioned above,
- The recommendations from a citizen-centric point of view as presented in D4.2,

[85] Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation), Article 39

[86] Article 59 stipulates the requirement of each Member State to designate a national supervisory authority responsible to ensure the application and implementation of the AI Act.

[87] popAI D1.6 "Policy Briefs-1st year"

[88] popAI "Inclusion and diversity workshop", 3rd popAI Plenary Meeting in Dublin (October 2022)

- Ethical and legal experts and stakeholder advisors involved in the popAI project and its sibling projects (dedicated discussions with the project's EAB and SAB, Policy Labs discussions),
- The ALIGNER and STARLIGHT policy recommendations as communicated during the 5th ALIGNER Public Workshop (June 2023) and ALIGNER "D2.3 Policy recommendations" and "D5.5 First Update of the Research Roadmap for AI in Support of Law Enforcement and Policing",[89]
- The ALIGNER, popAI, STARLIGHT, AP4AI projects Joint Workshop: "Ethical and Legal Aspects of AI for Law Enforcement"[90].

The current ethical and legal framework on AI is the threshold according to which the findings are selected as recommendations and emerging best practices.

The recommendations addressed to policymakers are thematically categorised under two main parts: the harmonisation of the regulatory framework along with the adoption of recommendations at EU level and the institutional safeguarding and procedure building.

5.1 Harmonisation of the regulatory framework at EU level

The harmonisation of laws regulating AI, primarily at EU level, was noted as a recurring recommendation, to avoid fragmentation and different levels of minimum protection to citizens at national level.[91] Indeed, the Proposal for a Regulation at the EU level (AI Act), especially considering its direct effect, is a necessary initiative towards this goal.

It is further on suggested that the EU legislator leaves as little space as possible to the national legislator to deviate from the maximum level of protection, while only limited and clearly specified derogations should be envisaged. On the other hand, attention should be paid to the preservation of and respect to Member States' national identity and cultural heritage as well as to the different levels of AI technological development and use among Member States. In other words, AI law needs to strike a balance between the aspiration for uniformity on the one hand, and national particularities on the other, especially regarding the technological development stages of each Member State.

5.1.1 The requirement of compliance to the applicable data protection framework

To elaborate more on the identified requirements for the harmonisation of AI at EU level, it was emphasised that the AI legal framework shall also be complementary to the data protection legislation, with due respect to obligations stemming from GDPR and LED when applicable and does not undermine the level of privacy and personal data protection as guaranteed by the EU Charter of Fundamental Rights of the European Union. One of the main concerns regarding personal data protection was the processing of special categories of personal data, and especially biometrics, by

[89] ALIGNER "D2.3 Policy recommendations", "D5.5 First Update of the Research Roadmap for AI in Support of Law Enforcement and Policing"

[90] ALIGNER, popAI, STARLIGHT, AP4AI projects, Joint Workshop: "Ethical and Legal Aspects of AI for Law Enforcement", January 25th and 26th 2023, CEA premises in Brussels.
Press release: https://www.pop-ai.eu/wp-content/uploads/2023/02/Ethical-and-legal-aspects-of-AI-for-law-enforcement-Conclusive-Statement.pdf

[91] popAI D3.4 "Stakeholder attitudes, priorities, and recommendations for addressing AI in the security domain in practice"

LEAs[92]. Furthermore, the need for conducting data protection impact assessments and implementing technical and organisational measures in the scenario of citizens' personal data processing in public spaces was raised.[93] To this regard, it is mentioned that, according to the latest amendments to the AI Act the latter does not seek to affect the applicable personal data protection framework[94].

5.1.2 The requirement of legislation adaptation to the technological developments

The new challenges of the rapid technological developments and globalisation have always been EU's primary concern highlighting the need for adopting a strong and more coherent legal framework in the Union[95].

The need for law to keep up with the technological advancements was expressed quite literally during the popAI Policy Labs, as even the development of an AI system itself is described as a dynamic rather than a static procedure. It was pointed out that AI systems need to be constantly improved and updated, seeking optimality[96]. To this regard, the conditions on which the systems will be constantly improved and/or updated shall be defined by the EU legislator.

In the quest for law being in context, the legislator should have great knowledge of the available state-of-the-art technology, while constantly monitoring the technological changes to proceed to the necessary amendments in legislation. Further on, the legislator could use terminology or wording which allows for a level of flexibility and adaptability to the technological changes, in order to cover existing and future cases to the extent that they do not compromise the notion of legal certainty.

5.1.3 The adoption of recommendations in complementarity to the AI Act

The need for a common EU legislative approach towards AI was expressed within the ALIGNER and STARLIGHT projects accompanied, in addition, by the recommendation to adopt guidelines that are complementary to the legislation. ALIGNER suggests "*to encourage the EU to provide a tailored legislative framework, designed specifically to guide and support Member States' Police and LEAs in their adoption of AI technology*", while STARLIGHT refers to the establishment of "*clear and consistent ethical approaches, principles and guidance based on a common European understanding"*[97]*.* The introduction of such guidelines could be in the form of "Recommendations"

[92] popAI D4.2 "White Paper for Civil Society"

[93] popAI D3.4 "Stakeholder attitudes, priorities, and recommendations for addressing AI in the security domain in practice"

[94] European Parliament, Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, Recital 2 b

[95] Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation) Recital 6, 7; Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, Recital 3

[96] popAI D3.4 "Stakeholder attitudes, priorities, and recommendations for addressing AI in the security domain in practice"

[97] 5th ALIGNER Public Workshop (June 2023)

at EU level[98]; however, they shall be consistent with the EU legal order, including the final text of the AI Act, LED and GDPR to avoid creating a patchwork of rules conflicting with and not complementing each other.

5.1.4 Three additional special legal regimes for harmonisation

5.1.4.1 Legal regime for children

During the popAI Policy Labs, it was mentioned that, when AI systems are deployed for law enforcement purposes, in the field of criminal prevention, investigation, detection and prosecution, special attention shall be paid to whether the suspect or accused person is an adult or a child[99]. In case of them being a child, Juvenile Law applies as *lex specialis* in the national legal orders of the Member States.

In a similar fashion, the latest draft Amendments to the Draft AI Act, state that when implementing the described in Article 9 risk management system, the providers of high-risk systems shall give specific consideration to whether the latter are likely to adversely impact vulnerable groups of people or children[100]. It is evident that emphasis is placed on the protection of children from potential adverse consequences caused by an AI system used for law enforcement purposes. However, there is no harmonisation for the minimum age of criminal responsibility for children in the EU, leading to a non-uniform treatment of children across the Member States' jurisdictions[101].

5.1.4.2 A framework for the procurement of AI systems and their social acceptance

The procurement of an AI system by LEAs was mentioned during the popAI Policy Labs as one of the most critical phases for assessing its appropriateness and suitability to achieve its designated purpose. Further on, the issue of ensuring the social acceptance of the technical specifications during the AI system procurement was pointed out[102]. The ALIGNER project has similarly proposed that the procurement, utilisation and in-service development of all AI-enabled technologies by LEAs are carried out with full awareness of their positive and negative effects on society[103].

Towards the same direction, an important initiative has been taken by the City of Amsterdam developing a set of contractual clauses for the procurement of AI to create a framework for the

[98] EUR-Lex, Official Website of the EU, Glossary of summaries, Recommendations, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=LEGISSUM:recommendations

[99] popAI D3.4 "Stakeholder attitudes, priorities, and recommendations for addressing AI in the security domain in practice"

[100] European Parliament, Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, Article 9 – paragraph 8

[101] ES Think Tank, Leah Rea, 'The EU Strategy on the Rights of the Child: A missed opportunity to introduce harmonisation for the age of criminal responsibility in Europe?' available at: https://esthinktank.com/2023/04/04/the-eu-strategy-on-the-rights-of-the-child-a-missed-opportunity-to-introduce-harmonisation-for-the-age-of-criminal-responsibility-in-europe/

[102] popAI D3.4 "Stakeholder attitudes, priorities, and recommendations for addressing AI in the security domain in practice"

[103] ALIGNER "D5.5 First Update of the Research Roadmap for AI in Support of Law Enforcement and Policing"

information that suppliers need to provide on algorithms used to ensure citizens' trust in these services and for the city to provide transparent information on how AI is used[104].

In that sense, specific legal provisions should be formulated to devise a potential standardised EU AI procurement procedure for LEAs. This procedure should also include a step for conducting a thorough and comprehensive impact assessment or revising the impact assessment already conducted by the provider prior to the design and development phases (see above section 4.3) of the envisaged AI system as part of the procurement process[105].

5.1.4.3 Minimum limits of administrative penalties for LEAs

The EC in its AI Act Proposal stipulates that *"Member States should take all necessary measures to ensure that the provisions of this Regulation are implemented, including by laying down effective, proportionate and dissuasive penalties for their infringement"[106].* During the sibling projects' joint workshop, the opinion that the law should have the strictest standards possible for public authorities when they deploy AI technologies, because they are authorised to exercise force, was expressed[107].

The EP in the latest Amendments to the AI Act Proposal states that it lays down the upper limits for the administrative fines for certain specific infringements while it leaves to the national competent authorities to decide based on certain criteria[108].

Member States' discretion to regulate the amount of the penalties for AI Regulation's infringement raises some concerns, as it enables national legislators to be less strict towards the public sector's infringements, such as infringements by LEAs. For that reason, it is recommended that the AI Act sets the respective minimum limits for administrative penalties on public authorities.

5.2 Institutional safeguarding and procedure building

The following recommendations reflect the identified needs for institutional safeguarding and procedure building, either at EU level or at national level, depending on the level of discretion provided by the EU to national policymakers (including legislators).

[104] AI Procurement, Develop EU standard contractual clauses for the procurement of ethical AI https://living-in.eu/groups/solutions/ai-procurement

[105] Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, Article.27

[106] European Parliament, Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, Recital 84

[107] ALIGNER, popAI, STARLIGHT, AP4AI projects, Joint Workshop: "Ethical and Legal Aspects of AI for Law Enforcement", January 25th and 26th at the CEA premises in Brussels.
Press release: https://www.pop-ai.eu/wp-content/uploads/2023/02/Ethical-and-legal-aspects-of-AI-for-law-enforcement-Conclusive-Statement.pdf

[108] European Parliament, Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, Recital 84

5.2.1 Establishment of training and awareness raising educational programmes under the AI literacy notion

The demand for appropriate and continuous training on AI tools and their legal and ethical use, both to users and technology developers was repeatedly mentioned in the popAI Policy Labs[109]. Specifically, with respect to the operators, the need for acquisition of respective certifications was highlighted. Moreover, a general request is that of raising the citizens' awareness on AI technologies.

In accordance with the notion of AI literacy, as prescribed in the AI Act Proposal[110], it is recommended that the Member States add courses on "ethical and lawful AI" to the educational curriculums of technological institutions (such as Universities or equivalents). Furthermore, the above courses should also constitute part of the national educational programmes of LEAs during their studies at the police academies and part of their life-long learning and regular training in the context of their duties (see also section 4.2.1).

Furthermore, and in accordance with the LEAs Rules of Procedure, the Member States are advised to establish an AI department/body/committee in the LEA forces, as explained in section 4.5.2, dedicated to train and regularly evaluate the level of training of the LEA staff to which the use of AI and human oversight is assigned.

In addition, citizens' awareness and knowledge could be raised and developed through the establishment of relevant courses as part of the national general education provided by the Member States at public and private schools, but also through governmental awareness-raising campaigns (see sections 4.4.1 and 4.4.2).

Finally, what constitutes "sufficient" level of AI literacy is a notion that needs to be defined by policymakers or by case law along with the establishment of awareness raising and training tactics that are considered necessary and adequate to reach that goal.

5.2.2 Building bridges for EU cooperation among stakeholders (AI Hub)

Exchanging lessons learnt and knowledge generated from one case to another across LEAs and the broader community is of crucial importance. The establishment of a platform for interchanging best practices, encouraging the usage of ethical and secure-by-design AI tools has been provided in popAI through the AI Hub[111]. The AI Hub gathers in one place the findings of the project and could be theoretically updated to include future results under a continuous learning model. Such a platform could be established at the policy level to support cooperation among Member States' LEAs and the broader community.

[109] popAI D3.4 "Stakeholder attitudes, priorities, and recommendations for addressing AI in the security domain in practice"

[110] European Parliament, Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, Recital 9b, Article 4d

[111] PopAI "D5.7: Sustainability and exploitation plan"

5.2.3 Institutionalisation of multidisciplinary collaboration

The need for active participation and collaboration of experts from different backgrounds throughout the entire lifetime of an AI system was highlighted during the popAI activities.

As mentioned under 4.5.2, the popAI Policy Labs engaging LEAs, ethics and legal experts, policymakers, technology developers and civil society representatives, set a successful example of pluralism and give added value to the discussion on AI. Further on, in line with a diversity and inclusion-by-design approach, diverse teams (indicatively in terms of gender, race, nationality) should be encouraged[112]. The stakeholders involved in the ecosystem of AI in the civil security domain are also described in D3.1.

The STARLIGHT project recommends the dynamic collaboration among technical, legal, ethical experts, while ALIGNER proposes the involvement of a competent and knowledgeable "human-in-the loop", in order to assist the decision-making processes for law enforcement purposes[113].

Elaborating upon the principle of human oversight, close to a "human-over-the-loop" approach, the active collaboration among LEAs, ethics and legal experts, policymakers, technology developers and civil society representatives throughout the lifetime of the AI system, from its design and development to its testing, validation, implementation and improvement is proposed.

Therefore, the institutionalisation of multidisciplinary collaboration at EU and national levels among legal and ethics experts, developers of AI tools, end-users, and the affected persons (incl. vulnerable groups) is highly recommended.

5.2.4 Standardisation of the Impact Assessment procedure

As already described in section 4.3, it is necessary that an impact assessment is conducted by the providers prior to the design and development stages of high-risk AI systems (a revision of the impact assessment is recommended during the procurement stage) as well as by LEAs prior to putting such systems into use. Conducting an impact assessment is highly recommended even when there is not an obligation set by law. Hence, it is suggested that the EU legislator via "Recommendations", guidelines or other appropriate means provides a regulatory framework accompanied by a standard template, specifically for the use of AI by LEAs for conducting a thorough impact assessment. It is highly recommended that such a template draws inspiration from the methodologies presented in section 4.3 (i.e., HUDERIA, AFRIA, DPIA and ESIA). It is essential to report that the impact assessments (or at least a summary) should be made publicly available, where appropriate, as well as they need to be multidisciplinary and inclusive in order to engage citizens in policymaking[114].

[112] popAI "Inclusion and diversity workshop", 3rd popAI Plenary Meeting in Dublin (October 2022)

[113] ALIGNER "D2.3 Policy Recommendations"

[114] UNESCO, Recommendation on the Ethics of Artificial Intelligence, 23 November 2021, Paris, France available at: https://www.unesco.org/en/legal-affairs/recommendation-ethics-artificial-intelligence

5.2.5 Establishment of transparency and accountability protocols for LEAs

Among other concerns regarding citizens, the needs for transparency about them being subjects to the use of an AI-system, the provision of information to them for their data processing by LEAs and the establishment of a procedure to object to unjust decisions were raised during the popAI activities.

In accordance with the transparency and accountability principles, as expressed in ALTAI, it is imperative that, the EU legislator along with the Member States regulate and ensure that appropriate mechanisms are put in place so that sufficient information about the AI systems already used and planned to be used by LEAs, and about the data processed and the decision-making process by LEAs is provided by the providers and/or the LEAs to the affected persons[115].

A broader field of application of Article 52 of the latest Amendment to the AI Act Proposal is suggested, in the sense that persons affected by law enforcement AI shall be informed that they are subject to an AI system, its purpose, the humans responsible for making the decision, the decision-making process, the adherence to the ALTAI principles and about their rights (including their right to object, redress and the right to seek explanation). In addition, the exercise of the data subjects' right to be informed about and request further information about (indicatively) the types of data processed, sources of data (or how they are collected), data controller, purposes and types of data processing, legal basis for processing, retention period as well as the exercise of the other data protection rights stipulated in the GDPR could be accomplished through this mechanism/procedure. The mechanism or procedure to inform the exposed to an AI system person and data subject, about (indicatively and as a minimum) the above, is proposed to be conducted timely and automatically, if possible.

Moreover, and in accordance with the accountability principle, the development and establishment of an easy-to-follow, yet well-defined procedure towards enabling the review of information, provision of feedback and the objection against unjust decisions by citizens, for example through a centralised online platform, is considered of paramount importance, as discussed in the popAI Policy Labs. Such a platform could be the "feedback mechanism" mentioned in section 4.4.1.

In support of the latest draft Amendments to the AI Act Proposal, the users of high-risk AI systems that are public authorities (such as LEAs) or Union institutions, bodies, offices and agencies should be required to register the use of any high-risk AI system in a public database[116]. More information about the required public registry of high-risk AI systems used by public authorities needs to be provided by the policymakers.

[115] European Parliament, Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, Amendment 484, Article 52

[116] Ibid., Recital 58a. See also a relevant reference in the Ad Hoc Committee on Artificial Intelligence (CAHAI), Possible elements of a legal framework on artificial intelligence, based on the Council of Europe's standards on human rights, democracy and the rule of law, 3 December 2021, p.12: "The CAHAI considers that the establishment of public registers listing AI systems used in the public sector, containing essential information about the system such as, its purpose, actors involved in its development and deployment, basic information about the model, and performance metrics, where appropriate, and the result of a HUDERIA, should be addressed in the context of a legally binding or non-legally binding instrument on AI in the public sector".

It is a demand that the citizens are informed timely, appropriately and automatically about their exposure to AI-systems used by LEAs and the processing of their data by LEAs and that they can exercise their respective rights through established and accessible to them procedures, considering their specific conditions and particularities.

5.2.6 EU funding

Additional funding is requested, in order to cope with the extra efforts required for the use of AI (e.g., research funding).

It is recommended that the EU and consequently the Member States provide dedicated financial resources:

- to the national educational institutions and the competent national authorities responsible for their supervision in order to offer a sufficient level of AI literacy,
- to the LEAs and the competent authorities responsible for their supervision in order to acquire a sufficient level of AI literacy and to upgrade their technological progress, especially considering that the Member States do not share the same level of technology development in the field of AI[117],
- for research and development taking into account the importance of regulatory sandboxes for the development of reliable and mature AI systems.

6 Evaluation of recommendations

The recommendations provided in the present deliverable are addressed to LEAs (chapter 4) and policymakers (chapter 5) and have been based on the sources listed and analysed in chapter 3.

Although we examined and filtered them through the current ethical and legal framework, we consider important their further evaluation by experts and interested stakeholders.

6.1 Evaluation by LEAs and policymakers

As described in section 3.3, a questionnaire was drafted and addressed to LEAs and policymakers within and outside of the popAI Consortium. The questionnaire was distributed after we had collected valuable input from other sources (ethical and legal framework, popAI tasks, Policy Labs, discussions with the EAB and the SAB, sibling projects). Amongst others, the questionnaire participants were asked to choose from a list of predefined recommended practices what kind of support or assistance they need to fulfil the ethical and legal requirements and be in line with the AI HLEG Ethics Guidelines and the AI Act (see the questionnaire in Annex B). Based on this, the recommendations in chapters 4 and 5 were evaluated and enhanced by the participants.

In addition, the deliverable was reviewed by the Hellenic Police and following their review, modifications and additions were made to its content in order to be in line with their comments.

[117] ALIGNER "D2.3 Policy Recommendations"

6.2 Evaluation by experts

The EAB chair reviewed the present deliverable and following the review, modifications and additions were made to its content in order to be in line with her remarks.

Also, the EAB member from KEMEA contributed to the finalisation of the recommendations.

Finally, a dedicated report will be drafted by the External Ethics Advisor which will include her opinion on the recommendations provided in the present deliverable as well as in D4.2 and D4.3.

6.3 Evaluation through the last deliverable of WP4

The last deliverable of WP4 is D4.4. It will collect and examine the recommendations included in D4.1, D4.2 and D4.3 and evaluate them in order to present them as the best multidisciplinary practices emerging from interdependent and collaborative work of people with different specialties and experiences. Within D4.4, the feedback of the EAB and SAB will be incorporated.

7 Conclusion

The advancements in the field of AI as well as the number and variety of its applications keep increasing at a rapid pace. It is foreseen that LEAs will start deploying more AI applications to assist them in carrying out law enforcement activities and meeting the demanding needs of their job more efficiently. Despite the indisputable benefits, due to the role and position of LEAs in the society, the use of AI in law enforcement is likely to result in arrest or deprivation of a person's liberty as well as other adverse impacts on fundamental rights guaranteed in the Charter. Therefore, it is critical that AI technologies are used in law enforcement in a way that will not only facilitate the work of LEAs but will also prioritise fundamental rights and foster the trust of citizens.

In the same spirit, considering the dynamics and impact of AI on society and the environment, the existence of harmonised legal rules that will accompany the technological developments is a necessity for the potential risks to be mitigated and the AI-based technologies to be smoothly incorporated into our lives. Therefore, the role of policymakers is crucial and decisive.

This deliverable presents and analyses recommendations to both LEAs and policymakers related to the ethically and legally compliant use of AI in law enforcement. The recommendations are of multi-stakeholder and multidisciplinary origin, created following a specific methodology seeking to cover the needs and expectations of all groups of interest. The opinions of citizens, technology developers, ethics and legal experts, LEAs and other interested stakeholders were taken into consideration for the production of comprehensive best practices that are in line with the current applicable ethical and legal framework. At this point, it is worth mentioning that, at this stage of the popAI project and as explained above in section 3.1, the AI Act, the AI Liability Directive and the CoE Convention are still in progress. Hence, based on the popAI findings, some of the recommendations aim at improving the current versions of the aforementioned drafts with a view to minimise the risks to fundamental rights and foster the trust of citizens towards law enforcement AI.

The sources that have been used were listed and analysed in chapter 3 of the present deliverable.

The recommendations for the LEAs were presented and analysed in chapter 4 of the present deliverable. This chapter aims to serve as a useful guide for LEAs that are planning to deploy or have already started deploying AI systems.

The recommendations for the policymakers were presented and analysed in Chapter 5 of the present deliverable. This chapter aims to serve as a useful guide for policymakers to be informed of the emerging recommendations for the ethically and legally compliant use of AI in law enforcement.

The present recommendations along with the related outputs of Task 4.2 (Recommendations for and from the Civil Society as presented in D4.2) and Task 4.3 (Recommendations for and from Technology Developers as presented in D4.3) will form a set of multidisciplinary best practices that will be presented in D4.4 “Synthesis: a collection of best multidisciplinary practices” and will be evaluated by the EAB and SAB members.

8 References

A.Mantelero, "AI and Big Data: A blueprint for a human rights, social and ethical impact assessment", available at https://www.sciencedirect.com/science/article/pii/S0267364918302012

Ad Hoc Committee on Artificial Intelligence (CAHAI), Policy and Development Group (CAHAI-PDG), Human Rights, Democracy, and the Rule of Law Assurance Framework (HUDERAF) for AI systems, Executive Summary, 8 October 2021, available at https://rm.coe.int/cahai-pdg-2021-09-huderaf-executive-summary/1680a416de

Ad Hoc Committee on Artificial Intelligence (CAHAI), Policy and Development Group (CAHAI-PDG), Human Rights, Democracy and Rule of Law Impact Assessment of AI systems, 21 May 2021, available at https://rm.coe.int/cahai-pdg-2021-05-2768-0229-3507-v-1/1680a291a3

Ad Hoc Committee on Artificial Intelligence (CAHAI), Possible elements of a legal framework on artificial intelligence, based on the Council of Europe's standards on human rights, democracy and the rule of law, 3 December 2021

AI HLEG (2019), Ethics Guidelines for Trustworthy AI, available at https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai and AI HLEG (2020), Assessment List for Trustworthy Artificial Intelligence (ALTAI), available at https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment

AI HLEG (2020), Assessment List for Trustworthy Artificial Intelligence (ALTAI), available at https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment

AI Procurement, Develop EU standard contractual clauses for the procurement of ethical AI https://living-in.eu/groups/solutions/ai-procurement

ALIGNER "D2.3 Policy recommendations"

ALIGNER "D5.5 First Update of the Research Roadmap for AI in Support of Law Enforcement and Policing"

ALIGNER "D5.5 First Update of the Research Roadmap for AI in Support of Law Enforcement and Policing"

ALIGNER, popAI, STARLIGHT, AP4AI projects, Joint Workshop: "Ethical and Legal Aspects of AI for Law Enforcement", January 25th and 26th 2023, CEA premises in Brussels, Press release: https://www.pop-ai.eu/wp-content/uploads/2023/02/Ethical-and-legal-aspects-of-AI-for-law-enforcement-Conclusive-Statement.pdf

Article 29 Data Protection Working Party Guidelines on Data Protection Impact Assessment (DPIA) , https://ec.europa.eu/newsroom/article29/items/611236

Committee on Artificial Intelligence (CAI) (2023), Revised Zero Draft Framework (Convention) on Artificial Intelligence, Human Rights, Democracy and the Rule of Law, available at https://rm.coe.int/cai-2023-01-revised-zero-draft-framework-convention-public/1680aa193f

Committee on Artificial Intelligence (CAI), Consolidated Working Draft of the Framework Convention on Artificial Intelligence, Human Rights, Democracy, and the Rule of Law (7 July 2023)

Council of Europe, Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts - General approach, 25.11.2022

Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data

Donatella Casaburo, KU Leuven, 'The ALIGNER Fundamental Rights Impact Assessment: Mitigating the impact of law enforcement AI', accessible at https://www.law.kuleuven.be/citip/blog/the-aligner-fundamental-rights-impact-assessment-mitigating-the-impact-of-law-enforcement-ai/ 20 June 2023 https://aligner-h2020.eu/fundamental-rights-impact-assessment-fria/

Draft Compromise Amendments on the Draft Report Proposal for a regulation of the European Parliament and of the Council on harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/CJ40/DV/2023/05-11/ConsolidatedCA_IMCOLIBE_AI_ACT_EN.pdf

ES Think Tank, Leah Rea, 'The EU Strategy on the Rights of the Child: A missed opportunity to introduce harmonisation for the age of criminal responsibility in Europe?' available at: https://esthinktank.com/2023/04/04/the-eu-strategy-on-the-rights-of-the-child-a-missed-opportunity-to-introduce-harmonisation-for-the-age-of-criminal-responsibility-in-europe/

Ethics Guidelines for Trustworthy AI issued by the European Commission's High-Level Expert Group on AI (https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai)

EUR-Lex, Official Website of the EU, Glossary of summaries, Recommendations, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=LEGISSUM:recommendations

European Commission, Impact assessment report accompanying the document: Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence, SWD (2022) 319 final and European Parliament, EPRS, Artificial intelligence liability directive, Briefing, February 2023.

European Commission, Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, 21.4.2021 COM (2021) 206 final 2021/0106 (COD)

European Commission, Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, 21.4.2021.

European Parliament, Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts.

Explanatory Memorandum of the Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, 21.4.2021 COM (2021) 206 final 2021/0106 (COD).
Fifth (5th) ALIGNER Public Workshop, June 2023

Greek Law 4961/2022 "Emerging information and communication technologies, strengthening digital governance and other provisions", Government Gazette 146/A/27-07-2022

popAI D1.6 “Policy Briefs-1st year”

popAI D2.1 “Functionality taxonomy and emerging practices and trends”

popAI D2.2 “Legal framework and casework taxonomy: emerging trends and scenarios”

popAI D2.3 “The controversies and risks that have shaped innovation and will shape AI in the next 20 years”

popAI D2.4 “Ethical frameworks for the use of AI by LEAs”

popAI D2.5 “Practical ethics toolbox for the use of AI by LEAs”

popAI D2.6 “AI meets organisational cultures: Human-machine interaction at the police station”

popAI D3.1 “Map of AI in policing innovation ecosystem and stakeholders”

popAI D3.3 “Citizen produced priorities and recommendations for addressing AI in the security domain”

popAI D3.4 “Stakeholder attitudes, priorities, and recommendations for addressing AI in the security domain in practice”

popAI D4.2 “White Paper for Civil Society”

popAI D5.7 “Sustainability and exploitation plan”

popAI-A European Positive Sum Approach towards AI tools in support of Law Enforcement and safeguarding privacy and fundamental rights (pop AI) Grant Agreement 101022001

Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive), COM (2022)

Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation).

The Alan Turing Institute, Human Rights, Democracy, and the Rule of Law Assurance Framework for AI Systems: A proposal prepared for the Council of Europe's Ad hoc Committee on Artificial Intelligence, available at https://rm.coe.int/huderaf-coe-final-1-2752-6741-5300-v-1/1680a3f688

UNESCO, Recommendation on the Ethics of Artificial Intelligence, 23 November 2021, Paris, France available at: https://www.unesco.org/en/legal-affairs/recommendation-ethics-artificial-intelligence

ANNEX A: Policy Labs

The Policy Labs are an initiative of the popAI project organised under T3.4. The objective of this task was to gather experts with different backgrounds and involve them in discussions related to the ethical development and application of AI-enabled technologies by LEAs. Five Policy Labs have taken place so far in five countries and each one of those was focused on different case studies. More information about the Policy Labs and the key outcomes can be found in the relevant deliverable D3.4[118]. The recommendations that emerged from the Policy Labs will be incorporated into the map of best practices (WP4), including also T4.1 and the present deliverable (see below chapters 4 and 5), and will form the basis of creating an ecosystem for a sustainable and inclusive social hub on the sound and ethical use of artificial intelligence by LEAs.

Policy Lab (Greece)

Date: 25 May 2022

Location: Online

Organiser: Hellenic Police

Language: Greek

Participants: LEAs (14), KEMEA (4), ECAS (1), CERTH (1), TRI (1), European Union Agency for Asylum (1), Municipality (1), National Commission for bioethics & techno ethics (1), National Technical University of Athens (1), Special Secretary for Long-Term Planning (1), My Data Greece (1), Ubitech (2), BYTE computer (1)

Case studies:

1. The system should use crime data (what, where, when) from an existing crime recording system, on the one hand to predict the commission of offences and therefore use it for the appropriate deployment of police forces, and on the other hand to investigate and solve offences, since the methodology followed by offenders in specific periods of time and geographical areas may constitute serious evidence.
2. The AI system will detect dangerous driving using video footage from traffic management cameras or other real-time footage.

Key outcomes:

1. First, it is worth pointing out that AI used in the context of predictive policing is considered as a high-risk AI system by the EC and the CoE and, hence is heavily regulated, while it constitutes a prohibited AI practice by the EP (see above sections 3.1.2.2 and 3.1.2.3). Therefore, the key outcomes of the Policy Lab related to this category of AI system will be useful only in the first case, i.e., if the final version of the AI Act does not ban yet it includes predictive policing AI in the list of high-risk AI systems.

[118] popAI D3.4 “Stakeholder attitudes, priorities, and recommendations for addressing AI in the security domain in practice”

The participants discussed about the benefits and then, they identified the potential challenges of such system by putting special emphasis on the potential biases and negative feedback loop as well as on overreliance of the LEAs on the AI outcomes. Recommendations were provided by the participants from an organisational and technical viewpoint in order to minimise the identified risks and meet the key requirements for a trustworthy AI. Extended reference is made in D3.4.

2. All participants agreed that the prevention of traffic accidents is of high importance and that the focus should be on the social benefit. Then, they identified the potential challenges of such system that were mostly related to the creation of mass surveillance. Recommendations were provided by the participants from an organisational and technical viewpoint in order to minimise the identified risks and meet the key requirements for a trustworthy AI. Extended reference is made in D3.4.

Policy Lab (Germany)

Date: 15 September 2022

Location: Munich, Germany

Organiser: University of Applied Science - Police Affairs (HfoeD)

Language: German

Participants: LEAs (8), Logobject (2), Munich Innovation (1), Adesso SE (2)

Case studies:

1. AI to support decision making in patrolling: An emergency call is received at the operations centre. Apparently, there was a dispute between two neighbours. One person was injured by a knife.
2. AI to process child sexual abuse material:
    - Various hard disks and data carriers are seized from one suspect.
    - Within the framework of international police reporting systems, hundreds of suspicious online contents are reported to the German police every day.
    - All suspicious and seized material is individually visually inspected manually by the officers.

Key outcomes:

1. Command and control centres face significant challenges in the deployment of units for operation. The location of the units, the available equipment, the skills and experience of the officers involved are critical for the success of the operations. This is why an AI system to support mission control would be helpful. The views of the participants on the potential challenges were expressed from an operational, technical and ethical perspective and respective recommendations were provided to minimise the risks and meet the key requirements for a trustworthy AI. Extended reference is made in D3.4.

2. Processing a big number of suspected CSAM cases poses both psychological and organisational challenges to LEAs. This is why an AI system could help identify perpetrators more easily. The views of the participants on the potential challenges were expressed from an operational, technical and ethical perspective and respective recommendations were provided to minimise the risks and meet the key requirements for a trustworthy AI. Extended reference is made in D3.4.

Policy Lab (Slovakia)

Date: 22 December 2022

Location: Hybrid (Bratislava and online)

Organiser: Police Academy of Bratislava

Language: Slovak

Participants: LEAs (27), Institute of Administrative and Security Analysis of the Ministry of the Interior of the Slovak Republic (1), National Security Office (1), National Crime Agency (2), Department of Computer Crime Presidium of the Police presidium (2), Kempelen institute of intelligent technologies (2), Comenius University Bratislava, Faculty of Law (1)

Case studies:

1. AI in support of monitoring the social networks based on a real situation of two people from the LGBTQ+ community being murdered by the perpetrator who had posted relevant tweets on the social network (hate speech against minorities): All suspicious and seized material is individually visually inspected manually by the officers.
2. Use of the Ethics Toolbox

Key outcomes:

1. The participants agreed on the need for AI support when it comes to monitoring of social networks for crime prediction and crime prevention. Without the help of modern technologies, it is objectively impossible to search and analyse huge amount of data on social networks and on the Internet in general. At the same time, it was considered necessary to set clear ethical and legal limits on the use of AI and create the necessary counterbalance so that fundamental rights including the right to personal data protection are not endangered. Extended reference is made in D3.4.
2. The Ethics Toolbox for the use of AI by LEAs has been developed as part of T2.4 of the popAI project and was presented during the Policy Lab. Extended information about it can be found in D2.5 “Practical ethics toolbox for the use of AI by LEAs” which is available to the public through the official popAI website.

Policy Lab (Italy)

Date: 20 April 2023

Location: Online

Organiser: City of Turin (PLTO)

Language: Italian

Participants: LEAs (8), Think Legal (1), Ethic Solution (1), Privacy Network (1), Studio Legale Ciccia (1), Member of Expert.ai (1), AI Tech Vision (1), Studio Legale Iafolla (1), Associazione Italiana per l'IA (1), University of Freiburg (1)

Case studies:

1. Following a brutal murder where the murderer struck a random victim among passers-by, an AI system has been set up in your city in the video surveillance network, with the adoption of algorithms for data recognition, extraction and analysis, in real time from video streams, which allows the production of massive amounts of value-added information (metadata) in the domain of security, monitoring, analysis and planning. This will allow police, starting from information derived from witness accounts, which is fragmentary and qualitative, and to the exclusion of using biometric data, to extract frames of interest that need to be validated. By way of example only, we mention in relation to vehicles: vehicle type; colour, lettering, markings; license plate and country of registration; direction and speed etc.; and to pedestrians: distinction between adult/child; colour of clothing and shoes; presence of objects such as bags, backpacks, hats, glasses etc. The system will be able to process the video streams acquired from the city's cameras and from unconnected private cameras and - once appropriately uploaded to the platform - will be able to “metadatabase” the information by comparing and integrating it with that present in the video streams generated by the connected camera system.
2. Use of the Ethics Toolbox

Key outcomes:

1. The participants acknowledged the undeniable advantages of using AI for surveillance purposes, the most important being that an AI system can analyse videos in a very short time while it would take several days and several police officers to do the same. Then, they expressed their concerns and focused on the risk of misuse of such systems, e.g., for social scoring (which is a prohibited AI practice) or due to long-term data retention, the risk of creating prejudices and biases and risks related to the right to personal data protection and other civil liberties. Recommendations were provided in order to minimise the identified risks and meet the key requirements for a trustworthy AI. Extended reference is made in D3.4.
2. The Ethics Toolbox for the use of AI by LEAs has been developed as part of T2.4 of the popAI project and was presented during the Policy Lab. The participants highlighted the need for basic training to the user who will interact with this toolbox. Extended information about it can be found in D2.5 “Practical ethics toolbox for the use of AI by LEAs” which is available to the public through the official popAI website.

Policy Lab (Spain)

Date: 27 April 2023

Location: Online

Organiser: Municipal Police of Madrid

Language: Spanish

Participants: LEAs (24), CIDALIA (1), University of Alcala (1), City Council Department (3), Ministry of Interior (1), Fundacion Secretariado Gitano (1), OBERAXE (government organisation) (1), University Complutense of Madrid (1)

Case studies:

1. In the field of security, CCTV systems are part of the tools used by the police in their daily work, both as crime prevention and as a tool for locating suspects. There is a wide range of CCTV technology on the market and the implementation of AI in these systems, exponentially increases their effectiveness in the scope of the public safety. We have a European legal framework that guarantees the rights and freedoms in these matters, in addition to the internal regulations of each country, which must be in line with the common framework of the European Union. However, the ethical questions about its use and limitations are on the table of debate, both for its ethical implications and its impact on citizenship in the field of privacy.
2. A 75-year-old male is reported missing, suffering from episodes of memory loss. It is believed that he may have had access to his vehicle and could be driving it. The biometric data of this person are requested: e.g., age, skin colour, eye colour, as well as the data concerning the clothes he was wearing at the time of his disappearance, such as the colour of his clothes, if he was wearing a hat, shoes, sneakers, etc. And the vehicle's license plate, model, colour, etc. Once the drone unit has this data, it proceeds to use the drones in different areas of the city in their search, so that, using the artificial intelligence software, they match the data entered to search for this person, while the data they have of the license plate of the vehicle.

Key outcomes:

1. The participants placed special emphasis on the right to data protection and expressed their relevant concerns. Recommendations were provided to ensure that personal data are properly used for the purpose for which they were initially collected and in accordance with the principle of proportionality as well as to establish a relationship of trust amongst LEAs and society. Extended reference is made in D3.4.
3. First, it is worth pointing out that AI used in the context of real-time remote biometric identification is considered as a high-risk AI system by the EC and the Council and, hence is heavily regulated, while it constitutes a prohibited AI practice by the EP (see above sections 3.1.2.2 and 3.1.2.3). Therefore, the key outcomes of the Policy Lab related to this category of AI system will be useful only in the first case, i.e., if the final version of the AI Act does not ban yet it includes real-time remote biometric identification in the list of high-risk AI systems.

   The participants commented that drones are similar to CCTV cameras in terms of privacy concerns, but their mobility and versatility increase the related challenges. Recommendations were provided to ensure that personal data are properly used for the purpose for which they were initially collected and in accordance with the principle of proportionality as well as to establish a relationship of trust amongst LEAs and society. Extended reference is made in D3.4.

Conclusions

This section summarises the main benefits and the main risks identified by the participants of the Policy Labs in relation to the use of AI by LEAs. Based on these, recommendations were provided during the Policy Labs and the most repeated ones, which are also considered to be emerging best practices to ensure the ethical use of AI-enabled technologies by LEAs, are listed and analysed below in chapters 4 and 5.

Main benefits:

- Improving the existing system of crime recording that provides statistics based on collected data, including type of crime, location, gender and age of the offender
- Identifying leads and detecting patterns not discernible by humans
- Analysing huge amount of data much faster
- Analysing social media (e.g., monitoring to detect hate speech and prevent hate crimes)
- Reducing investigation time
- Assisting organisations not only to predict but also to act pre-emptively and even guiding policymaking through an evidence-based approach

Main risks:

- Data misuse either due to extensive data retention or use for discriminatory purposes
- Bias and discrimination due to the risk of impartial control and biases of the AI system
- Overreliance on AI technologies without sufficient human involvement or supervision
- Threat to fundamental human rights and freedoms, such as to the right to equality and non-discrimination for specific groups targeted based on their gender/ethnical origin/political beliefs/sexual orientation, freedom of expression, presumption of innocence etc.
- Transparency issues
- Lack of citizens’ trust in AI technologies and processing of their personal data by LEAs

ANNEX B: Questionnaire for LEAs and policymakers

WP4 – T4.1 Questionnaire

The present questionnaire is addressed to **LEAs and policymakers**. It aims at collecting valuable input from them in order to:

- Produce recommendations for LEAs & policymakers that will be presented in D4.1 'White Paper for LEAs' (PU) of the popAI project.
- Lead to the design and use of AI tools, accepted & valued by LEAs as well as citizens.

All the following questions are always associated with and focused on the use of Artificial Intelligence (AI) or Machine Learning (ML) as a subcategory of AI.

You can answer the questionnaire individually or in groups, i.e., in collaboration with your colleagues, if it is easier or more convenient for you.

For questions that are answered with a *Yes* or *No* or with pre-defined answers, simply underline or highlight your response(s).

Please try to justify your answer wherever requested.

Definitions:

*"**Artificial intelligence (AI) systems** are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions. As a scientific discipline, AI includes several approaches and techniques, such as machine learning (of which deep learning and reinforcement learning are specific examples), machine reasoning (which includes planning, scheduling, knowledge representation and reasoning, search, and optimisation), and robotics (which includes control, perception, sensors and actuators, as well as the integration of all other techniques into cyber-physical systems)".*[119]

*"**Artificial intelligence system (AI system)** means a machine-based system designed to operate with varying levels of autonomy and that can, for explicit or implicit objectives, generate outputs such as predictions, recommendations, or decisions, that influence physical or virtual environments".*[120]

[119] Ethics Guidelines for Trustworthy AI issued by the European Commission's High-Level Expert Group on AI (https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai)

[120] Draft Compromise Amendments on the Draft Report Proposal for a regulation of the European Parliament and of the Council on harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts (https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/CJ40/DV/2023/05-11/ConsolidatedCA_IMCOLIBE_AI_ACT_EN.pdf)

Introductory questions:

1. Please state (a) your professional background, i.e., LEA / policymaker / other (in this case please specify) and (b) the city and country where your entity is established.

   Answer:

2. At national level, is there an **AI law** currently in force in your country (to your knowledge)?

   Answer:

3. Is/Are there any **specific police department(s)** (for LEAs) or any **specific department(s) in your entity** (for policymakers) which have the expertise to deal with AI-related issues (by e.g., providing relevant consultation, establishing necessary procedures, implementing appropriate measures, using AI tools)?

   Answer:

4. Are you **currently using** any AI-enabled technologies or tools in your department? Or are you aware of such technologies being already **used by other police departments/other departments like yours in your country?**

   Yes
   No

   <u>If yes, please answer the questions from (a) to (f) below:</u>

(a) What type(s), on which operational field(s) and for which purposes?

Answer:

(b) Do you find these AI-enabled tools useful and why?

Answer:

(c) Do you have any concerns (e.g., do you believe that the users are attached to the recommendations provided by the AI tools or that biases are inevitable or that the results are not accurate and may lead to incorrect decisions and consequently harm fundamental rights…)?

Answer:

(d) How can these concerns be mitigated? What would you propose?

Answer:

(e) Are the citizens aware of the use of such tools?

Yes
No

<u>If yes</u>, how have they been informed?

Answer:

If not, what would you propose as adequate ways to inform the citizens?
Answer:

(f) To your knowledge, are **citizens more satisfied or concerned** with the use of the AI system by LEAs/municipalities? Have you received any relevant queries?
Answer:

5. Have you **attended any seminars (theoretical knowledge)** so far, organised by your department, on the use of AI tools?

Yes
No

If yes, could you describe in more detail what was the seminar about and whether it also provided guidance about the **ethical use of AI**?
Answer:

If not, have you attended any other relevant seminars/workshops/educational sessions (not organised by your department) and were they useful?
Answer:

6. Have you had any **training (practical knowledge)** so far, organised by your department, on the use of AI tools?

Yes
No

If yes, could you describe in more detail what was the training about and whether it also included training on the **ethical use of AI**?
Answer:

If not, have you had any other relevant training (not organised by your department) and was it useful?
Answer:

7. Do you believe that educational seminars and trainings on the use of AI systems are necessary? Please describe what the **ideal seminars/trainings** should include, **how frequently** they should take place and **who** the ideal speaker(s)/trainer(s) would be, e.g., technology providers, policymakers, ethics and legal advisors, Law Enforcement Agents, all of them collaboratively.
Answer:

Assume that you are provided with an AI system that can produce recommendations to assist the user.

1. On **which particular fields would you prefer to be assisted by an AI system**? You may choose (underline or highlight) more than one area and/or indicate a new area**.**

- **Recognition:** This category concerns functionalities related to recognition / identification / verification / validation tasks either real-time or offline. Examples are voice recognition, suspects identification, etc.
- **Communication:** This category comprises of interaction with humans such as communication robots, translation bots, chatbots, etc.
- **Prediction & Analytics:** This category comprises all the data processing and information analysis and knowledge extraction operations, real-time or offline, such as: digital forensics, agent-based simulations, suspicious behaviour detection, pattern recognition, etc.
- **Surveillance:** This category includes all the surveillance patrolling monitoring functionalities, such as: surveillance drones, patrol robots, AI-generated Patrol Live Stream, etc.
- **Other (please specify):**

2. For **which particular purposes would you prefer to be assisted by an AI system**? You may choose (underline or highlight) more than one area and/or indicate a new area**.**

- **Crime Prevention:** Functionalities that contribute to the prevention of a potential criminal offence.
- **Crime Investigation:** Functionalities that contribute to the support of the investigation procedures after a criminal offence takes place.
- **Cyber Operations:** Functionalities concerning the network cloud and digital communications infrastructure.
- **Migration, Asylum, Border Control:** Functionalities that contribute to the facilitation of the asylum and migration procedures and/or the improvement of border surveillance and border control operations.
- **Administration of Justice:** Functionalities that support jural and/or judicial procedures.
- **LEAs Training:** e.g., AI-assisted training applications for LEAs skill improvement.
- **Other (please specify):**

3. Based on your experience or the needs that you may have identified, what are some characteristics or procedures that would make the AI system **appropriate** for operational use?

Answer:

4. Based on your experience or the needs that you may have identified, what are the characteristics or procedures that would make the AI system **inappropriate or unpractical** for operational use?

Answer:

5. Based on the aforementioned fields and purposes, can you think of potential **risks per field/purpose** that may derive from the use of an AI system by your department?

- **Recognition:** <add risk(s)>
- **Communication:** <add risk(s)>
- **Prediction & Analytics:** <add risk(s)>
- **Surveillance:** <add risk(s)>
- **Crime Prevention:** <add risk(s)>
- **Crime Investigation:** <add risk(s)>
- **Cyber Operations:** <add risk(s)>
- **Migration, Asylum, Border Control:** <add risk(s)>
- **Administration of Justice:** <add risk(s)>
- **LEAs Training:** <add risk(s)>
- **Other (please specify)**: <add risk(s)>

6. Could you come up with **any measures or procedures** that will help mitigate the risks identified above?

- **Recognition:** <add recommendation(s)>
- **Communication:** <add recommendation(s)>
- **Prediction & Analytics:** <add recommendation(s)>
- **Surveillance:** <add recommendation(s)>
- **Crime Prevention:** <add recommendation(s)>
- **Crime Investigation:** <add recommendation(s)>
- **Cyber Operations:** <add recommendation(s)>
- **Migration, Asylum, Border Control:** <add recommendation(s)>
- **Administration of Justice:** <add recommendation(s)>
- **LEAs Training:** <add recommendation(s)>
- **Other (please specify)**: <add recommendation(s)>

7. Based on the current version of the AI Act proposal, the use of "**real-time" remote biometric identification systems in publicly accessible spaces** is **prohibited**. Can you think of **cases** where the use of such systems on the fields and for the purposes mentioned above would be useful and helpful for your department and the society?

   Answer:

Questions to assess the level of organisational readiness and compliance with the applicable legal framework and the ethical standards:

1. Based on the **Ethics Guidelines for Trustworthy Artificial Intelligence**, specific principles must be respected for AI technologies and tools to be trustworthy (**see the first reference for more details about each principle)**:
    - human agency and oversight,
    - technical robustness and safety,
    - privacy and data governance,
    - transparency,
    - diversity, non-discrimination and fairness,
    - societal and environmental wellbeing,
    - accountability and auditability.

    (a) Do you consider your department being in a good position to operate in conformity with the aforementioned principles?

    Yes
    No

    (b) What procedures are you currently following and what measures are you implementing to this end (e.g., human as the final decision maker, impact assessments, close collaboration with legal advisors and technology providers, transparency tactics, training, other)? Please describe **per principle**.
    Answer:

    (c) Please use the list of question 1 above and rate **per principle** how difficult / easy you consider the implementation of these requirements.
    Scale: Very difficult - Somewhat difficult - Indifferent - Somewhat easy - Very easy

    Answer:

2. Based on the **Proposal for an Artificial Intelligence Act**, **AI systems that will be used by LEAs** are considered **high-risk** and are subject to **strict obligations** as follows:
    - conducting of a conformity assessment,
    - establishment of a risk management system,
    - appropriate testing procedures,
    - high quality of the datasets feeding the system to mitigate risks and discriminatory outcomes, activity logging to ensure traceability of results,
    - technical documentation,
    - record-keeping (“logs”),

- transparency and provision of clear and adequate information to the user,
- appropriate human oversight,
- high level of robustness, security and accuracy.

(a) Do you consider your department being in a good position to meet the aforementioned obligations?

Yes
No

(b) What procedures are you following and what measures are you implementing to this end (e.g., human as the final decision maker, impact assessments, close collaboration with legal advisors and technology providers, transparency tactics, training, other)? Please describe **per obligation**.

(c) Please use the list of question 2 above and rate **per obligation** how difficult / easy you consider the implementation of these requirements.
Scale: Very difficult - Somewhat difficult - Indifferent - Somewhat easy - Very easy

Answer:

3. What kind of support or assistance would you like to have for meeting the legal obligations? You may choose (underline or highlight) more than one type of assistance and/or indicate a new one.

- Training & education through training courses, seminars, guidelines with best practices
- Regular consultation and close collaboration with experts
- Supervision by a competent independent authority
- Direct communication with policymakers
- Tools that have been developed by following an ethics-, security- and privacy-by-design approach
- Case studies of how other entities apply the AI Act
- Provision of templates of impact assessments (data protection impact assessment, human rights impact assessment, democracy impact assessment, societal impact assessment)
- Additional funding to cope with the additional efforts
- Other (please elaborate)

4. How are you planning to involve citizens and increase their trust towards the use of AI tools (e.g., through dialogue, societal impact assessments, informational events, school visits, other)?

Answer:

5. Underline or highlight what your opinion and reaction to the aforementioned obligations are.

- Positive: We embrace the new obligations as we believe they add value for us and for the society and we have already started/will immediately start to take the necessary actions.
- Slow down: We need time to establish the necessary procedures and implement the appropriate measures and we will start using AI tools only after we have ensured compliance.
- Negative: The time and cost for compliance outweigh the benefits.
- Shutdown: We will not use AI-enabled technologies/tools.