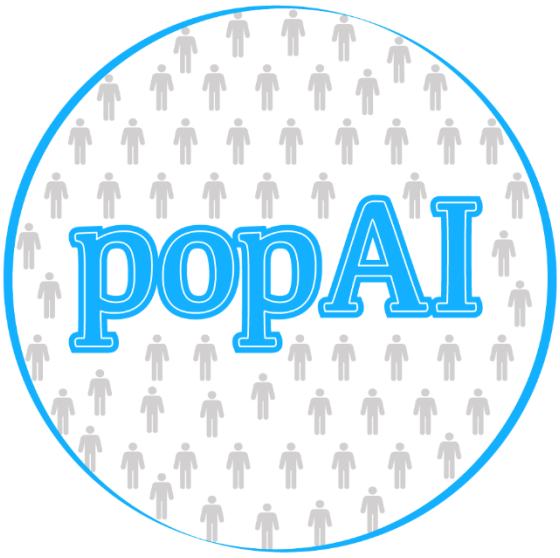


A European Positive Sum Approach towards AI tools in support of Law Enforcement and safeguarding privacy and fundamental rights

D4.2: White Paper for Civil Society

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	12/08/2023		
Nature	R: Document, report	**Dissemination Level**	PU: Public
Author	Claire Morot-Sir (ECAS)		
Contributors	Assya Kavrakova (ECAS), Dimitra Papadaki (KEMEA)		
Reviewers	Judith Jorda Frias (ETICAS), Pinelopi Troullinou (TRI), Lilian Mitrou (EAB), Eleftherios Chelioudakis (SAB), 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

The analysis presented in this deliverable makes use of the findings of popAI Work Package 2: "Security AI in the next 20 years: trends, practices and risks", and Work Package 3: "Empirical Knowledge Collection and Management Framework", and mainly the outcomes of the crowdsourcing exercise carried out under Task 3.3 and the five Stakeholder Policy Lab sessions that took place under Task 3.4. Results of other tasks under WP3 such as the photo competition (Task 3.6) or the social listening exercise (Task 3.2) have also been taken into consideration. The purpose of this deliverable is to contribute to the general goal of the popAI project of enhancing trust of citizens in the use of AI by LEAs, through a citizen-centric approach, aiming at empowering citizens in the exercise of their rights. Based on the attitudes observed both on the crowdsourcing platform and amongst participants who took part in the Stakeholder Policy Lab sessions, the main objective of this deliverable is to produce a set of recommendations which includes citizens' thoughts and values. These recommendations aim at voicing citizens' and civil society's identified concerns and may enable civil society to address key challenges.

Table of Contents

List of Figures

List of tables

List of Terms & Abbreviations

Abbreviation	Definition
AI	Artificial Intelligence
ALTAI	Assessment List for Trustworthy AI
CSA	Coordination and Support Action
EAB	Ethics Advisory Board
EC	European Commission
EU	European Union
GDPR	General Data Protection Regulation
IPR	Intellectual Property Rights
LEA	Law Enforcement Agency
LED	Law Enforcement Directive
LEO	Law Enforcement Officer
NGO	Non-governmental Organisation
SAB	Stakeholder Advisory Board
WP	Work Package

1 Introduction

Understanding citizens' concerns constitutes a cornerstone on how the future development of AI technologies in policing should be approached, to build up tools that are in line with the needs of citizens affected by AI assisted decision making, both from a security and an ethical standpoint. The findings analysed in this White Paper compile citizens' and stakeholders'- including civil society representatives- perspectives on the choice of technologies used by Law Enforcement Agencies (hereafter referred to as LEAs) and what their scopes and limitations should be. For the purposes of the present deliverable, recommendations for and from the civil society are presented from the perspective of citizens and civil society representatives and aim at empowering them, thanks to a comprehensive mapping of their concerns and expectations that should be considered in contemporary policy developments.

This deliverable presents the results stemming from different sources on understanding citizens' discourse and extracting their recommendations to foster public trust in the use of AI by law enforcement authorities.

The purpose of this deliverable, on one hand, is to take into account the opinions of the public which is a step towards the use of AI in compliance with democratic values with respect for fundamental rights and the rule of law (within the limitations of the popAI project). Building upon the recommendations described below, the opinions of the affected community must be taken into consideration before implementing the respective technologies. On the other hand, the input gathered from citizens and stakeholders -including civil society representatives- is going to enrich the recommendations appointed to the other stakeholders in D4.1 "White paper for LEAs" and D4.3 "White paper for Technology Developers" from the citizens' and civil society representative point of view.

Lastly, the recommendations for the citizens can be a step towards a guide for their empowerment and an exercise of their rights (e.g. to exercise their right to lodge a complaint). Being able to dispute a decision affecting them and the establishment of obligations (including transparency) for the authorities could at least provide for a minimum standard for citizens when affected by decisions assisted by AI technologies.

The findings of those empirical exercises allowed us to analyse and examine the main challenges identified by the citizens who participated and extract recommendations from our activities that reflect their values, in an attempt to build up AI tools that ensure the security of democratic societies while at the same time remaining fair, ethical, transparent and compliant with data protection laws and international human rights law.

This report provides an overview of citizens' recommendations classified by thematic categories to ensure that all types of challenges highlighted by citizens and stakeholders including civil society representatives -which is the focal point of the present deliverable- are equally represented.

1.1 Aim and scope of the recommendations for and from citizens and stakeholders - including civil society representatives- for the ethical use of AI for LEAs

Work Package 4 "The pandect of recommendations for the ethical use of AI for LEAs" aims at taking advantage of the knowledge obtained from the literature review performed in Work Package 2 "Security AI in the next 20 years: trends, practices and risks" and mainly the empirical research of Work Package 3 "Empirical Knowledge Collection and Management Framework" (especially deliverables D3.3 and D3.4 as described in the popAI Grant Agreement) with the broader ecosystem to create a library of target group-specific recommendations appointed to:

1. policymakers and LEAs as in D4.1
2. the civil society as in D4.2
3. technology developers as in D4.3

Such proposals have drawn inspiration from each community's insights on the use of AI for LEAs, their potential drawbacks and opportunities.

In this attempt, the popAI consortium has consulted the European Commission's Assessment List for Trustworthy AI (ALTAI).[1] The innovation of such a library lies upon the conjunction of the existing legal, ethical, and technical AI frameworks, the future trends, the positive and negative feedback provided by the users, developers, and communities affected by AI for LEAs but also on the extensive focus on target-specific recommendations.

Finally, it will create a synthesis of the emerging best practices by the refinement of the work conducted under T4.1, T4.2, and T4.3. This composition will also be facilitated by the exchange of best practices among popAI, ALIGNER and STARLIGHT which has been initiated through the three sibling projects' joint and separate Workshops.

WP4 recommendations will be communicated via the popAI extended network of stakeholders as well as the full eco-system in Work Package 5 to create bridges with related projects and their outcomes to publicise the popAI research output depicted in WP4. Based on the findings of Work Package 2 and Work Package 3, this Work Package aims to build trust in the use of AI technology in the security sector via:

1. Improving citizens' perception of security and specifically the use of AI technology for LEA purposes, while considering their concerns to improve, limit or enhance such use
2. Safeguarding the legitimacy and trust of LEAs in using AI tools by improving their accuracy
3. Drawing the attention of all involved parties to the principles and requirements ensuring a lawful, trustworthy and legitimate use of AI for Law Enforcement purposes in a democratic society
4. Issuing recommendations appointed to each target group and a combination of best practices

The main objective is to draw attention to the conditions under which the use of AI for LEAs is acceptable while exchanging knowledge and opinions, considering the risks and benefits and

[1] High-Level Expert Group on Artificial Intelligence (AI HLEG), Ethics Guidelines for Trustworthy Artificial Intelligence (AI) available online at: https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment

mitigating distrust towards AI used for security purposes by providing advice to the identified ecosystem, from citizens and stakeholders (including civil society representatives) to LEAs and AI technology developers.

The results are based on the challenges identified under Work Package 2: "Security AI in the next 20 years: trends, practices and risks", and mainly two empirical activities performed under Work Package 3: "Empirical Knowledge Collection and Management Framework" : a crowdsourcing online platform developed by ECAS to understand citizens' attitudes and collect pro-active solutions; and a second activity where ECAS and popAI partners engaged LEAs and relevant experts in online workshops referred to below as Stakeholder Policy Labs (including LEAs and policy-makers, technology developers and civil society representatives). The above-mentioned activities took place in the context of Task 3.3 "Crowdsourcing stakeholder attitudes and pro-active solutions ideation" and 3.4 "Engaging LEAs and relevant experts through Stakeholder Policy Labs", respectively.

Results of Tasks 3.2 "Understanding citizen discourses around AI and security controversies" (the social listening exercise) and 3.6 "Engaging New Citizens through student photo and caption competition" (photo competition) have also been incorporated into the findings of the present deliverable. Outcomes of an additional social media listening activity (social sensing) carried out by CERTH have also been incorporated. Interaction with sibling projects have been sought throughout the projects' workshops, also via ALIGNER and STARLIGHT representatives, as members of popAI Stakeholder Advisory Board.

Through the two above-mentioned activities (Task 3.3; crowdsourcing platform and Task 3.4; Stakeholder Policy Labs), citizens and stakeholders including civil society representatives have been able to express their main concerns, define their limits and identify the support that should be provided to ensure a fair, ethical and transparent use of AI tools. Recommendations defined by citizens and for citizens, under the present deliverable, have fully integrated these concerns, leading to solutions that would enable an efficient use of AI tools while safeguarding human rights and fundamental freedoms. The support of Task 3.2 "Understanding citizen discourses around AI and security controversies" (the social listening exercise) and 3.6 "Engaging New Citizens through student photo and caption competition" (photo competition) as well as the interaction with sibling projects have been a valuable added value to collect opinions from citizens and stakeholders, including civil society representatives. The additional social sensing exercise's results have also provided additional support to complete the recommendations.

The recommendations gathered under this deliverable constitute a solid basis to compile suggestions from citizens and stakeholders including civil society representatives, that can benefit LEAs, policymakers, technology developers and citizens. It seems crucial, as part of the holistic approach of Work Package 4, that we put the accent on recommendations that aim at involving citizens in all steps of the decision-making process, starting from the design phase of the AI tools.

1.2 Structure of the Deliverable

The deliverable introduced, under Section 1, the aim and scope of Task 4.2, which is followed by the deliverable structure and a description of the methodology chosen to draw up this report, explaining more in detail which guidelines we followed, what sources we relied upon, the relation with other tasks and WPs and whose input was provided throughout the different tasks that have allowed us to gather and compile those recommendations.

Section 2 will focus on the recommendations **from** citizens and stakeholders including civil society representatives. It will start by extracting recommendations appointed to LEAs, explaining what type of technical and psychological support should be provided to AI users, at the same time highlighting the need for training that has been broadly suggested by the different participants involved in Tasks 3.3 and 3.4. The second part of the 2nd section will describe recommendations suggested for policymakers, notably the need for an improved legal framework, the necessity to create intermediary bodies to avoid discrimination and bias, and recommendations that aim at ensuring the respect of fundamental human rights and freedoms. Finally, we will show recommendations that have been suggested for technical developers, with the goal to ensure a more ethical development of AI tools.

The 3rd Section will demonstrate recommendations **empowering citizens**, in particular how we could raise awareness on the scope and purpose of AI tools in policing, and better inform citizens on the circumstances under which such tools are used. We will report on solutions put forward by participants to ensure that citizens are more involved in decision-making processes. Additionally, we will delve into the reporting on the processing of personal data, outlining recommendations that suggest informing citizens on how, when and for how long their data can be processed and stored.

The conclusions drawn from these works are summarized in the last section.

1.3 Methodological Approach and relation to other Work Packages and Deliverables

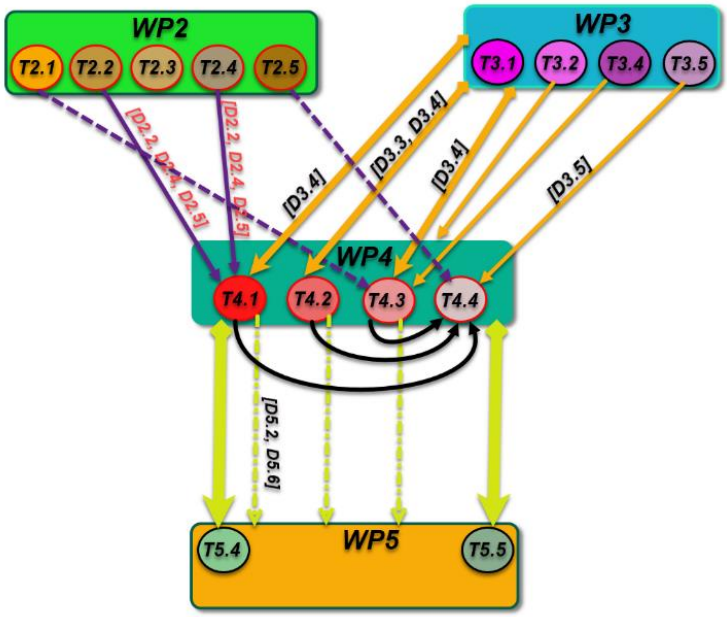


Figure 1 - Interdependence of WP4 with other WPs and tasks

Deliverable 4.2 mainly incorporates the outcomes of two tasks that have been performed under WP3: crowdsourcing stakeholder attitudes and pro-active solutions ideation (Task 3.3) and engaging LEAs and relevant experts through Stakeholder Policy Labs (Task 3.4). The main results of the **social listening** exercise have been incorporated in this deliverable. These activities have been used to feed the recommendations in support of the rest of the findings.

Additionally, we have been following up on sibling projects developments (ALIGNER and STARLIGHT) in order to inform popAI recommendations. Questions about citizens have also been raised to our Stakeholder Advisory Board during the popAI plenary meeting presentation in Rome. Their feedback was received in verbal form, and incorporated into this deliverable. An example of the questions posed is given in Annex 6.7.

The **crowdsourcing exercise** was used as a primary tool for actively engaging citizens within popAI in order to understand their attitudes towards AI in the security domain. Through ECAS' crowdsourcing platform,[2] citizens have been invited to share their thoughts in three different phases:

1. Share input on main controversies and experiences
2. Put forward solutions
3. Vote on the best ideas proposed

Controversies suggested in the first phase of the crowdsourcing exercise were based on the legal framework and casework taxonomy performed under WP2; emerging trends and scenarios have been defined in Task 2.2. Controversies have also been based on Task 2.3 which aimed at mapping controversies and risks that have shaped innovation in AI and will shape AI in the next 20 years.

[2] popAI (ecas.org)

The added-value of using tools such as the crowdsourcing platform is to give the opportunity for people of different backgrounds to take part in a common, cross-border effort, which goes beyond engaging the usual stakeholders. Information indicating the type of participants in our activities, listed in Annex 6.3 (Crowdsourcing platform participants' statistical data) and 6.4 (Countries of residence of crowdsourcing platform participants) show that a large variety of profiles have been involved. To that end, an invitation to join the crowdsourcing platform has been broadly promoted as part of Work Package 5 (Dissemination, Communications and Sustainable Community Engagement). More information and data regarding the participation of citizens can be found below in the ANNEX (6.1-6.4) and in D3.3 "Citizen produced priorities and recommendations for addressing AI in the security domain" in detail.[3] Moreover, in a joint effort to make the crowdsourcing platform more accessible, all popAI partners contributed to the translation of the questionnaires in the languages covered by the project, namely Italian, Slovak, Greek, Spanish, German, and Dutch.

Case studies and scenarios presented as a basis for discussion at the beginning of each session are the result of a reflection between the LEA partners and the other organisations involved in Task 3.4. Discussions between participants covered human rights, liability, proportionality, gender and diversity topics as developed in Tasks 2.2 "Legal framework and casework taxonomy: emerging trends and scenarios" and 2.4 "From ethical frameworks to ethics in practice".

The goal of the Stakeholder Policy Labs was to gather NGOs, policymakers, technology developers, LEAs and academics to reflect together on a specific scenario that helps exchange and identify best practices and solutions in an experimental format. More information and data regarding the participants to the Policy Labs are found below in the ANNEX (6.5-6.6) and in D3.4 "Stakeholder attitudes, priorities, and recommendations for addressing AI in the security domain in practice".[4] To foster active participation during the Stakeholder Policy Lab sessions, it was decided that all workshops would take place in the local language. The decision to conduct the Policy Labs in the local languages of the country where the Policy Labs were organised, promoted the participation and inclusion of stakeholders and facilitated the free exchange of opinions without the limitations or barriers to communication between non-native English speakers. To support the moderation of the Stakeholder Policy Labs, partners from different organisations have stepped in to help with the language requirements.

Controversies mapped under Task 3.1 have supported both activities, as well as **the social listening** and social sensing exercise performed under Task 3.2. Finally, case studies and recommendations extracted from the three first Stakeholder Policy Labs have served as a basis to elaborate foresight scenarios under Task 3.5. Outcomes of other tasks such as 3.6 have been also taken into account in this deliverable.

This deliverable will help completing other tasks under WP4 as it includes similar challenges and controversies as other tasks, however focusing on the perspective of citizens. By approaching problems from a different angle, Task 4.2 puts the emphasis on proposals that have emerged from

[3] popAI D3.3 "Citizen produced priorities and recommendations for addressing AI in the security domain"

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

citizens and gathers recommendations that can support both citizens and civil society in taking an active role in defining the future evolution of AI applied to policing activities.

1.4 Procedure, Guidelines, Criteria

The tasks mentioned are the primary however not exclusive sources of this deliverable, as explained above. In addition to the results of the empirical exercises , the findings were informed by the theoretical framework of WP2 and our following up on recent developments regarding the AI Act Proposal, the Workshops of the sibling projects and especially the ALIGNER, STARLIGHT, pop AI (SU-AI cluster meeting: "Ethical and legal aspects of AI for Law Enforcement") which took place in Brussels on 25 - 26 January 2023, the ALIGNER and popAI Workshops, and especially the ALIGNER's 5th workshop which took place in June 2023 where the identified best practices/recommendations of STARLIGHT, ALIGNER and popAI were presented, the popAI Plenary meetings and especially the popAI Plenary meeting in Rome with the participation of the popAI SAB. In addition, the current deliverable was reviewed by popAI **SAB** and **EAB** as in the document revision history table and their valuable feedback and comments were incorporated into the drafting of recommendations.

1.4.1 Crowdsourcing platform procedure

The results of the crowdsourcing platform (Task 3.3 Crowdsourcing stakeholder attitudes and pro-active solutions ideation) have allowed us to analyse the outcomes of three different phases:

1. Phase 1: citizens were able to rank the controversies presented on the platform from a very negative to a very positive sentiment; an example is given in Annex 6.1
2. Phase 2: idea generation – citizens were asked to provide their ideas and solutions to use AI tools in a fair, ethical, lawful, transparent and yet efficient manner. An example is given in Annex 6.2
3. Phase 3: citizens were asked to vote for the best solutions proposed in phase 2

Five topics were discussed on the platform, each one introduced by an explanatory paragraph to ensure that participants with a limited knowledge on the topic can easily understand:[5]

1. Biometric identification
2. AI systems used to prevent crime (predictive policing)
3. AI systems used in cyberoperations[6]
4. Police hacking
5. Justice decision-making tools

[5] See also popAI D2.1 "Functionality taxonomy and emerging practices and trends" : Areas of Application

[6] As in popAI D2.1 "Functionality taxonomy and emerging practices and trends", **Cyber Operations** are defined as: Functionalities that regard network cloud and digital communication infrastructure and popAI D3.1 "Map of AI in policing innovation ecosystem and stakeholders", where cyberoperations are defined as : " fighting crime not only in a physical world, but also on a digital environment.".

For each of the five topics citizens were asked to rate their level of agreement on several aspects of their implication and management:

1. Respect to human rights
2. Human oversight
3. Accuracy
4. Reliability
5. Respect (for the right) to privacy
6. Legitimate access to people's data (respect to the right to personal data protection)
7. Transparency
8. Prejudice and (respect to the principle of non-) discrimination
9. Benefit to society
10. Sustainability
11. Accountability

Respondents had to place a numerical value to a statement on each of the eleven aspects, ranging from 1 (totally disagree, highest concern) to 7 (totally agree, least concern) where 4 is neither agreement nor disagreement, for example:

1. Biometric identification tools are reliable
2. Biometric identification tools have enough human oversight
3. Biometric identification tools are accountable

More information can be found in Deliverable 3.3 "Citizen produced priorities and recommendations for addressing AI in the security domain".

1.4.2 Stakeholder Policy Labs procedure

The following methodology applied to Stakeholder Policy Labs:

1. Case studies (listed in ANNEX: 6.5,6.6) have been elaborated by LEAs with support from other partners involved in the task to launch discussions around the use of AI in policing (e.g. the use of systems for predicting dangerous driving using video footage from traffic management cameras or other real-time footage to prevent traffic accidents)
2. Participants were divided into separate break-out rooms to discuss the potential risks and challenges
3. Results were shared in a plenary session
4. Participants were divided again to discuss main solutions to overcome the potential risks in using the presented AI tool
5. Recommendations and conclusions were shared in a plenary session

Further information can be found in Deliverable 3.4 "Stakeholder attitudes, priorities, and recommendations for addressing AI in the security domain in practice".

In view of the high number of solutions to the use of AI in policing provided throughout those two activities, and the rest of popAI sources, reporting and comparison of the recommendations has been performed. To ensure a continuous and joint effort in sharing recommendations on a regular

basis under WP4, we have created an assistant tool, i.e. an "indicative list of emerging recommendations /best practices". The list (Table 1) consists of columns under which we have added entries to support us in identifying the emerging recommendations, especially from the Stakeholder Policy Labs in addition to the rest of the identified sources. The columns refer to the number of entry, the target group for which the recommendation is appointed (a. Recommendations for LEAs and policymakers, b. recommendations for citizens and c. recommendations for technology developers), the source group from which the recommendations were extracted (if applicable/specified), the source or reference of the recommendation, and an indication of compliance or not with the ALTAI principles as a minimum threshold.

Table 1 Indicative list of recommendations/emerging best practices columns

No	**Recommendation/ Best Practice Title and Summary**	**Target Group (for) (a. LEAs & policymakers b. citizens c. technology developers d. other (please specify)**	**Source Group (from) -if applicable**	**Source (Reference)**	**ALTAI principle as a minimum threshold** 1 Human agency and oversight 2 Technical robustness and safety 3 Privacy and data governance 4 Transparency 5 Diversity, non-discrimination and fairness 6 Societal and environmental wellbeing 7 Accountability	**Comments**

We noticed a pattern in the indicative list of the recommendations/emerging best practices, in the sense that certain entries were similar or repetitive, so they fit under the same thematic category, providing the recommendations described in the context of this deliverable. According to this procedure, recommendations from and for citizens and stakeholders (including civil society representatives), have been thematically categorised, to represent the main identified trends. Finally, we have examined each thematic category to understand citizens' needs, and draw conclusions that can benefit citizens, policymakers, technology developers and end-users. The ALTAI principles were referenced as a minimum threshold for the emerging recommendations, along with the rest of the applicable framework, as it will be elaborated in D4.1 for policymakers and LEAs and D4.3 for technology developers. However, as this deliverable refers to the civil society, it is characterised by particularities compared to T4.1 and 4.3, in the sense that it aims to voice citizens' concerns which may *de lege ferenda* influence the legal developments in the field of AI. Lastly, the extensive feedback of the **SAB** (mentioning that they believe that this work will be "*very beneficial for enhancing the public dialogue on the use of AI tech by LEAs*") and **EAB** regarding the current deliverable, as it was reviewed by them as in the document history table, and the

developments of the sibling projects were taken into consideration to inform the recommendations identified.

1.4.3 Social listening procedure

Although the main sources used to complete this deliverable are the above-mentioned crowdsourcing and Stakeholder Policy Lab activities, results of the social listening exercise have served as a valuable complement to recommendations for and from citizens.

The methodology applied as part of the popAI project has been following the same logic as the crowdsourcing exercise when determining main keywords and controversies.

To complete the exercise, a database called CommonCrawl[7]- an open web repository for the last 7 years, containing 3.1 billion pages, where each month's worth of data totals more than 300 terabytes- has been searched to extract the main tendencies observed on the web on the use of AI in policing. The CommonCrawl database was used to crawl and record the internet: blogs, publications, research papers and news articles. An algorithm was used to search through the internet for strings of keywords that identify topics of interest to the popAI project, previously determined with the support of contributors to Task 3.2.

It is important to note that while CommonCrawl provides access to a vast amount of data, it does not cover the whole internet. For example, the Google search is about hundreds of times bigger than CommonCrawl. The social listening exercise has therefore been completed by a **social media sensing** conducted exclusively on Twitter.

More information on the methodology applied to collect and analyse the results of the scanning can be found in D3.3 while additional information on the results can be found on the digital dashboard.[8]

1.5 Sources, Input and Data

Through the different activities which have allowed us to gather recommendations aiming at understanding the discourse of both citizens and civil society, policymakers, LEAs and academics, we tried reaching out to a diverse audience in order to collect the points of view and perception of stakeholders from different backgrounds.

The main sources we used to feed this report are the crowdsourcing questionnaires submitted to citizens as part of Task 3.3, and the Stakeholder Policy Labs as described above, which correspond to Task 3.4. On top of these two activities, we used the results of the ethical social listening activity carried out by ECAS (Task 3.2).

[7] popAI D3.3: Citizen produced priorities and recommendations for addressing AI in the security domain
[8] PopAI Dashboard (citizens.is)

Concerning the crowdsourcing platform, we have collected statistical data from participants at each phase that is available in Annex 6.3. It appears that the three versions of the questionnaires were answered in their majority by citizens identifying themselves as men, followed by a high percentage of female participants and a rather small number of non-binary participants. The participants' education level varied from phase to phase, with the majority of them holding a Master's degree, followed by those holding a Bachelor's degree and those who have completed high school education, among others. The participants' ages were mainly between 20 and 49 years old and the majority of the participants had a yearly income which amounted to less than 20.000 euros. Whether participants consider themselves as religious or not varied from one phase to another. There are of course limitations to the research study dependent upon the participants to it, which are recognised by the researchers; however, the results are informative to the extent they express the opinions of the sample of participants and shall be taken into consideration along with additional sources. Further on, Annex 6.4 shows the countries of residence of participants to the crowdsourcing platform: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Egypt, Estonia, Ethiopia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Lithuania, Netherlands, Philippines, Poland, Portugal, Romania, Slovenia, Spain, Sweden, Tanzania, UAE, United Kingdom and USA. In that sense, when mentioning the term "citizens" under the present deliverable, especially as part of the crowdsourcing platform, we refer to a broader term than that of an "EU citizen" which is important in terms of representation.

Attendees of the different Stakeholder Policy Lab sessions are summarised in Annex 6.6. Stakeholders involved in the exercise had various levels of technological literacy, depending on their professional backgrounds (LEAs, NGOs, civil society representatives, policymakers, technology developers, academics etc.).

More information on the data regarding the empirical exercises of WP3 which provided input to the present deliverable may be found in deliverables: popAI D3.2 "Report on citizen discourses and attitudes towards controversies", D3.4 "Stakeholder attitudes, priorities, and recommendations for addressing AI in the security domain in practice", D3.3 "Citizen produced priorities and recommendations for addressing AI in the security domain", D3.6 "Photo Competition Results".[9]

2 Recommendations from citizens and stakeholders (including Civil Society representatives)

In this section, we are going to summarise the recommendations provided by citizens expressing their opinions in the Crowdsourcing platform and different stakeholders, including civil society representatives participating to Stakeholder Policy Labs activities, according to their field of application. We will also present recommendations provided as part of the photo competition

[9] popAI D3.2 "Report on citizen discourses and attitudes towards controversies", D3.4 "Stakeholder attitudes, priorities, and recommendations for addressing AI in the security domain in practice", D3.3 "Citizen produced priorities and recommendations for addressing AI in the security domain", D3.6 "Photo Competition Results".

organised under Task 3.6, as well as the social listening exercise conducted under Task 3.2. and the additional social sensing exercise completed by CERTH.

For the scope of the present deliverable, we are presenting the recommendations from a citizen-centric point of view. Throughout the above mentioned exercises, solutions have been proposed to support LEAs in the use of AI, to guide policymakers in structuring future AI regulations, and to provide advice to technical developers so AI tools are developed in an ethical manner.

2.1 Recommendations for LEAs

It was identified that the use of technology to predict crimes, prevent crimes, identify perpetrators or support decision-making processes are generally perceived by the participants (especially in the Policy Labs), as an asset in policing activities. AI tools increase human capabilities, can support big data analysis and identify patterns that might be difficult for a person to discern. Yet, it was highlighted that, in order to ensure a smooth adaptation, proper training is required so that technologies are used with a clear purpose, respecting limitations that could potentially hinder individual liberties if they were to be violated.

As underlined by stakeholders during discussions over the use of AI by police officers, "it is impossible to halt the progress of AI technologies, and it is unwise to rely solely on technology due to time-saving advantages".[10] In light of this consideration, stakeholders have provided recommendations from the citizens' perspective, aiming at enabling the use of AI tools by LEAs, while securing that AI users receive adequate support to avoid any misuse.

One of the main trends which has been brought up by stakeholders is the need for LEOs to receive **appropriate and regular training** on the use of AI tools. According to Stakeholder Policy Lab participants: "it should not be assumed that users possess a basic understanding of technology, and particularly of artificial intelligence.[11]"

It has been recommended that both technical and psychological support is provided to LEA staff, as to avoid a misuse of technology that could lead to errors, to the so-called "feedback loop" and consequently to discrimination. In that respect, LEOs should receive adequate training, supporting them in the **identification of potential biases**, and instructing them on how to correct those errors. Ethics and legal (data protection, fundamental human rights) oriented training should be provided to all end-users to increase their knowledge and ensure that tools are being used to the maximum of their capacity while limiting the risks of discriminatory application (e.g. using AI to target specific group(s) of people based on the protected grounds of the anti-discrimination principle).

To make appropriate training happen, it is therefore crucial that sufficient resources are dedicated to support LEAs in the evaluation of their tools, notably predictive policing systems. It is noted that although in the Stakeholder Policy Labs, the case of predictive policing was studied (e.g. Case Study 1-Greece, Annex 6.5), we would like to acknowledge that according to the latest Draft Compromise Amendments on the AI Act Proposal, predictive policing AI systems that make "risk assessments of

[10] Italian Stakeholder Policy Lab, April 20th 2023.
[11] Italian Stakeholder Policy Lab, April 20th 2023.

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), are proposed to be among the prohibited AI practices.[12]

Nonetheless, the quest for training of the LEOs in the technological tools, ethical conduct and how to exercise oversight towards AI systems, remains valid.[13] It will not only help ensuring a safe and ethical use of the tools but will also allow mapping irregularities and gaps to improve the policies and procedures they need to follow.

With regards to the trainers, their expertise, professionalism and ethical conduct should be given very careful consideration, under the "train the trainers" notion.

To mitigate the risks of misapplication, police officers should gain a sound knowledge of the **applicable laws** surrounding the use of the AI technologies they have to use to perform their work. In order to minimize the risk of potential human biases, it is suggested that certain safeguards or guarantees should be put in place. For instance, a procedure for LEOs to complete and obtain certifications for the ethical use of AI should be established and comprehensively described in a legal framework. Legislation should also play a key role in identifying **who can access the system data**, how this access should be limited, and under which circumstances and conditions access should be authorised.

As part of an updated legal framework, accent should be put on reinforcing privacy and data protection, on how to protect human rights and how to comply with the principles of non-discrimination. **Data protection** and privacy laws should be fully integrated in the training, therefore increasing the knowledge of LEOs, and enabling them to have a proper understanding of the legal limitations and its consequences when using these technologies. Also, understanding the notion of non-discrimination, the protected grounds of discrimination, along with bias identification should be part of the training for LEOs, so that the input data which they provide to the algorithm are fair and do not lead to discriminatory decisions.

In addition, training should not be limited to the first use of AI tools; continuous training programs should be established to keep users up-to-date with evolving AI technologies.

In that sense, a periodic psychological and/or psychiatric evaluation for LEA officers who use AI systems which can potentially affect their behaviour and ability to make decisions is recommended. Consequently, considering the effects that the use of AI tools, as well as their quick evolution can have, not only technical, but also psychological and psychiatric training of LEO's is indispensable in a

[12] European Parliament, 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, 9.5.2023 available at : https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/CJ40/DV/2023/05-11/ConsolidatedCA_IMCOLIBE_AI_ACT_EN.pdf ; 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 available at : EUR-Lex - 32016L0680 - EN - EUR-Lex (europa.eu)

[13] High-Level Expert Group on Artificial Intelligence (AI HLEG), Ethics Guidelines for Trustworthy Artificial Intelligence (AI) available online at: https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment, page 8

complementary way. The aim is to ensure that misuse or discrimination, whether consciously or unconsciously or due to the speed of their progress, is avoided. The training would further contribute to the mental health of the LEAs staff and avoid that the use of these tools negatively affect it. There are, however, certain reservations regarding such evaluations, related to the processing of the health data of the employees affecting their professional development.

Moreover, our crowdsourcing exercise has shown that not solely LEAs at the operational level must be trained. Indeed, LEOs but also judges and magistrates shall also receive training, notably on the use of "**hacking" as an investigative technique**. Hacking [14]can be performed under strict conditions, including prior obtainment of warrants and other legal authorizations. Clear guidelines should be provided for police hacking operations, including the types of crimes that can be investigated and the circumstances under which hacking can be used.

2.2 Recommendations for policymakers

Although the fast-growing evolution of AI technologies can benefit society and improve the daily work of LEA officers, policymakers must ensure that they come up with legislative tools that can cope with those constant changes and ensure an effective protection of human rights and individual liberties. For this reason, the following statement was made by one of the participants: "The use of AI tools is necessary, but this use must have clear boundaries which is also why it is necessary to set clear legal and ethical limits."[15.]

Thus, one of the main findings that have been extracted from the recommendations delivered by stakeholders is the need for an enhanced and comprehensible legal framework on the use of AI technologies in policing activities. That need was also showcased by ALIGNER and STARLIGHT in their Policy Recommendations.[16]

1. Data processing for LEAs purposes: a quest for a supplementing framework

One of the main recommendations under the quest for a robust legal framework is to further clarify how data should be stored, and for how long.

Accent has been put on the need to develop clear guidelines and standards for the **collection, storage, restriction and use of biometric data** by LEAs. It is crucial to ensure that these data are regularly reviewed, according to the data accuracy principle and processed both lawfully and proportionally to the purpose they serve. Risks of discriminatory use of data, abuse of power and

[14] This is how 'hacking' was presented on the crowdsourcing platform : *'AI tools can be used by the police for hacking purposes. Hacking by police through AI is aimed at preventing and identifying terrorists and criminals. However, AI has been used also to hack the smartphones of journalists, government officials and human rights activists in several countries. Protests took place in several countries over allegations that the government used AI tools to illegally monitor public figures. These AI tools are seen as an unacceptable form of surveillance, carried out with a lack of transparency, without people being aware of it. The AI tools allowing the police to hack people devices can be seen as a form of oppression rather than a form of protection'.*

[15] Slovak Stakeholder Policy Lab, December 13th 2022.

[16] 5th ALIGNER Public Workshop, June 2023

breach of fundamental human rights and freedoms have been identified as potential issues deriving from an unclear legal framework.

Stakeholders have expressed concerns on the lack of clear regulations regarding data collection, retention, especially with regards to the **duration of data retention**, and the **number of individuals with access** to the data. Legislation must therefore provide **clear limitations on data retention**, and clarify deletion policies in order to minimise the risk of data breaches and misuse. It must clearly define for which purpose data can be collected, and how it can be kept. Once the authorised retention time has come to an end, data should be deleted from the AI system. Considering Directive 680/2016 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, this could be a recommendation for the internal legal orders of the Member States according to Articles 5 (Time-limits for storage and review) and 20 (Data protection by design and by default) of the Directive or for the establishment of guidelines setting concrete criteria for the storage period depending on the processing operations and their purposes .[17]

In sum, stakeholders have recommended a supplementing legal framework regarding data retention and better harmonisation of the AI legislation at the national and EU levels.[18] Furthermore, the applicable legal framework could be developed to establish standards and guidelines for developing and implementing predictive policing algorithms that are designed to ensure fairness and transparency. As, according to the latest Amendments to the AI Act,[19][20] predictive policing is proposed to be among the prohibited AI practices, this recommendation is of limited value. The above would secure a higher level of protection of personal data and enable a more harmonised possibility for judges to intervene regarding the permission for data usage by LEAs.

2. Creation of intermediary and/or oversight bodies

The second trend we have observed as a recommendation, is the **creation of intermediary bodies** to avoid a discriminatory use of AI tools. Stakeholders have recommended the establishment of an independent oversight body to review the use of biometric identification tools and ensure that they are used in compliance with relevant laws, regulations, and ethical standards. Ideally, there should be an independent authority with "technical, organisational, and practical capabilities to assess the

[17] 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 available at : EUR-Lex - 32016L0680 - EN - EUR-Lex (europa.eu)

[18] https://www.fairtrials.org/articles/news/eu-parliament-votes-on-major-ai-law

[19] 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 available at: Texts adopted - Artificial Intelligence Act - Wednesday, 14 June 2023 (europa.eu)

[20] Key EU Parliament votes tomorrow on major AI law including landmark ban on 'predictive policing' and criminal 'prediction' systems - Fair Trials

system's compliance with legal and ethical rules and regulations set by interdisciplinary committees and stakeholders".[21]

More specifically, an independent oversight body is suggested to be created to review and monitor police hacking operations. **Oversight bodies** should also monitor the use of predictive policing algorithms and act as a support to make sure that they are not being used in a discriminatory or prejudicial way.

Finally, **audits** of predictive policing algorithms should be performed by independent bodies, to evaluate their accuracy, fairness, and potential for bias.

All in all, concerns regarding predictive policing, which is envisioned to be among the prohibited AI practices, were raised and accompanied by the urge to establish safeguards, including auditing by oversight bodies.

3. Establishment of lawfulness, transparency and accountability protocols for LEAs

On a general note, it is suggested that rules on the legality of using AI systems by LEAs under specific conditions are implemented. The purposes, specific operations, obligations, requirements under which AI systems are authorised to be used by LEAs, and the administrative offences, and strict internal protocols at the LEA level should be determined by the legislation, primarily at the EU level, to ensure the lawful use of AI systems by LEAs.

In order to ensure the respect of human rights and freedoms, another recurring recommendation that we have observed is the need to make it legally compulsory for LEAs to conduct **thorough reporting regarding their use of data** available. There is also a general trend advising LEAs to publish regular reports on the use of biometric identification tools, including information on the number and types of tools they have used, as well as the purpose for which they are used. LEAs should also report on the accuracy, bias and error rates of the AI technologies, more specifically the predictive policy systems. LEAs should use data that is representative of the communities they serve, and regularly update and audit their data to ensure that they remain accurate and unbiased. In order to prevent authorised access, **encryption** of all data obtained through police hacking operations should be required. The above recommendations are reflecting the principles of lawnfulness, transparency and accountability, which are essential to all AI systems; let alone in the fields of biometric identification, predictive policing and police hacking operations, highlighting the increased worry about those specific operations.

An overall recommendation in this regard could be the establishment of lawfulness, transparency, and accountability protocols at the policy level for LEAs about the use of data by them. Within this framework, AP4AI, as indicated by the STARLIGHT project, could be suggested as a tool to support with the enforcement of the accountability principle.[22]

4. Anti-discrimination; an offline and online framework

[21] Greek Stakeholder Policy Lab, May 25th 2022.

[22] EUROPOL,CENTRIC, Accountability Principles for Artificial Intelligence (AP4AI) in the Internal Security Domain available at : https://www.ap4ai.eu/

One of the participants to the Stakeholder Policy Labs[23] advised that "both public authorities and private entities prohibit artificial intelligence systems that categorise individuals based on biometric data into groups based on ethnic origin, gender, as well as political or sexual orientation or other grounds of discrimination prohibited under Article 21 of the Charter of Fundamental Rights".[24]

Finally, the same participant concluded that "the use of artificial intelligence for the automatic recognition of human features (face or gait, fingerprints, DNA, voice, keyboard typing and other biometric or behavioural signs) in public spaces in any context. Therefore, in order to ensure compliance, according to Article 5 of the proposal, large scale artificial intelligence systems for remote identification should be prohibited in online spaces" by analogy.[25]

2.3 Recommendations for technology developers

Although bias may emerge from human misuse of the AI tools as mentioned above, another way to efficiently fight against a discriminatory utilisation of technologies is to ensure that tools are designed in an ethical manner.

First of all, recommendations were made at multiple occasions as to design AI tools that come in support of decision-making, but do **not replace human beings**.

According to Stakeholder Policy Lab participants, "There should always be human supervision on the whole lifecycle of the AI system; AI systems need to support the decision making, not to make the decisions".[26]

With regards to the design of the AI tools, it was once again broadly recommended that **human supervision and intervention must always remain possible**. The function of tools should be to support the decision-making, however, humans must ultimately remain responsible for the final decision. In other words, the ability for human intervention should be preserved, allowing human oversight and control over the AI system's actions. Over-reliance on technologies presents a significant number of ethical risks and should be avoided.

While AI tools can be an asset, especially to save time on certain tasks that do not need specific skills, it is important to maintain a collaborative approach where AI supports but does not replace human expertise. Indeed, human expertise should remain the most valuable resource when it comes to exercising policing activities.

Although automated processing can be helpful to support the work of LEAs and save a considerable amount of time, it is the role of technical developers to ensure that no automated tool is developed without ensuring the possibility for humans to intervene at any step of the process.

The risk of misuse of AI technology has been broadly pointed out, leading to recommendations on the need to **elaborate ethical tools from the very early design phase**. To avoid any bias based on

[23] Slovak Stakeholder Policy Lab, December 13th 2023
[24] Article 21 - Non-discrimination | European Union Agency for Fundamental Rights (europa.eu)
[25] Slovak Stakeholder Policy Lab, December 13th 2022.
[26] Greek Stakeholder Policy Lab, May 25th 2022.

gender, ethnical background, sociological background, sexual or political orientation, a principle of "non-discriminatory algorithms"[27] should be implemented. This would secure a free-of-bias tool, which does not disproportionately focus on individuals with specific physical characteristics or in certain territories.

In that respect, all tools should be developed including **robust data protection and security measures** to protect (especially) biometric data from unauthorised access, use, or disclose, as from the development phase, inline with the data protection by design and by default approach. [28]

With regards to the development phase, Stakeholder Policy Lab participants have highlighted that: "police officers should be involved in the development of the algorithms, bringing their police expertise to determine what is to be detected and to improve the reliability of the AI tool."[29]

Once the development of the AI tool is advanced enough, **algorithms should be trained** on a diverse and representative dataset, which accurately reflects the population it is meant to serve, and this way avoids discriminatory results. This includes not only demographic data, but also data on crime patterns, socioeconomic factors, and other relevant variables.

It was further proposed by the participants that algorithms must be designed and tested for fairness and transparency, with measures in place to detect and prevent discrimination, such as:

1. Full algorithmic transparency and accessibility for experts and researchers specialised in the field (with a role of auditing, managing the risk and/or oversighting assessment procedures);
2. Rules must ensure display of error percentages, trust regions, etc., so that users/operators, can have some indication as to how much the results can be trusted or not;
3. Results need to support (not to lead) the investigations.

Once a tool is in use, a process for **ongoing evaluation** of algorithms should be put in place to ensure that they remain effective, accurate and fair. Reservations towards predictive policing algorithms have been expressed and the risks of bias they present; a constant evaluation would be required. However, it was expressed that predictive policing algorithms identifying future offenders using personal data should be banned.[30] The increasing volume of negative discussions surrounding predictive policing algorithms, especially regarding perceived discrimination, indicates that citizens are becoming more aware of the potential harms associated with these tools. The negative perception of predictive policing tools is likely due to the fact that they rely on historical data to make predictions about future criminal activity, which may perpetuate biases and discrimination in the law enforcement system. This is in line with the latest Amendments to the AI Act Proposal, according to which predictive policing are proposed as prohibited AI practices.

[27] Italian Stakeholder Policy Lab, April 20th 2023.

[28] 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 available at : EUR-Lex - 32016L0680 - EN - EUR-Lex (europa.eu), Article 20

[29] Spanish Stakeholder Policy Lab, April 27th 2023.

[30] The Verge, EU draft legislation will ban AI for mass biometric surveillance and predictive policing available at : EU draft legislation will ban AI for mass biometric surveillance and predictive policing - The Verge

Privacy-enhancing technologies such as encryption, anonymisation and pseudonymisation should be used to minimise the risks associated with biometric identification. However, for the purposes of this deliverable, it is noted that according to the latest Amendments to the AI Act Proposal, real-time remote biometric identification in public spaces and biometric categorisation systems are anyway proposed as prohibited AI practices. [31]

3 Recommendations empowering citizens

In this section we will shed light on recommendations that can empower citizens to have a better understanding of the reasons why the use of AI tools in police activities is necessary, and how data collection is being processed. At the same time, we will elaborate on the recommendations made by citizens requesting a more important involvement in the decision-making process.

3.1 Awareness-raising

Our results have shown the challenges and concerns presented by the use of AI technologies in the police sector. The impact of AI on fundamental rights and freedoms, such as the right to data protection cannot be denied. To **bring citizens closer** to the decision making process in that field, an effort needs to be made to enhance communication and foster awareness.

In that sense, Stakeholder Policy Lab participants have suggested that: “Transparency and clear communication with the public are vital aspects that should not be overlooked. It is crucial to inform the citizens transparently about the deployment of AI in the surveillance network. By providing clear information, the public can develop a better understanding of the system's capabilities, limitations, and safeguards in place.”

The need for citizens to be **more informed** about the reasons why AI tools are used in policing activities appeared at multiple occasions from the recommendations extracted from our crowdsourcing and Stakeholder Policy Labs activities. Indeed, when it comes to predictive and detection systems, stakeholders expressed their interest in seeing regulations being developed and enforced in a way that promotes and ensures **citizens’ awareness** regarding the existence and specifications of AI systems. This can, for instance, include the definition of any necessary (obligatory) and optional requirements, guidelines and elements. Lastly, informing citizens of the **different steps of the design of AI tools** can nurture a more transparent approach, demonstrating how ethical considerations have been taken into account. To promote awareness and transparency, it was indicatively required that law enforcement agencies publish regular reports on the use of AI (biometric identification) tools, including information on the number and types of tools used, the purposes for which they are used, and the outcomes of their use.

[31] European Parliament, 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, 9.5.2023

In a similar manner, stakeholders have proposed that **awareness campaigns** shall inform citizens on the use of biometric identification tools and their potential impact on privacy and other fundamental rights. The need for public education, under the general AI literacy notion, to increase understanding and awareness of AI tools and their benefits and their risks was highlighted. This is of particular importance for citizens with low engagement and lower technology literacy (e.g. elderly people, or citizens living in remote areas with limited access to technologies).

Clear and accessible information should therefore be available for citizens, defining the scope, purpose, use and limitations of biometric identification tools. Potential risks to their privacy and potential breaches of individual rights should also be disclosed in a simplified manner, as to be understood by citizens regardless of their level of knowledge on AI technologies.

Finally, **academic research** on the purpose and impact of AI in surveillance systems is beneficial to society, as it makes valuable insights available to the public, and can serve as a basis for a collective reflection, which can in turn be used throughout the decision making process.

LEAs' reporting on data processing

One of the key findings with regards to transparency is the recommendation to "inform and be transparent about the proper (appropriate) use of data.".[32] Furthermore, access to personal data that is processed with the assistance of AI technologies by LEAs has been brought up by stakeholders to enhance the trust of citizens in the use of AI in policing activities. From this point of view, it is also supported that whenever a citizen's personal data is used by the police, the citizen should be informed about it timely (automatically, if possible) and the citizen shall have the right to review the details and / or request further information concerning the types of data used, how they have been collected, the reason(s) why they are collected, how they are used, until when they will be kept. Clear and accessible information should be provided to citizens about the purpose and use of -among others- biometric identification tools, and the potential risks to their privacy and other rights.

It was reported that citizens have reservations concerning the processing of biometric identification tools which use, analyse and process personal data. Indeed, there is a general fear that, although necessary to support LEAs in maintaining security, these tools might be discriminatory. Also, citizens express concerns over the use of the results provided by biometric identification tools, in the sense that their data might be used for completely different purposes than the ones they have been initially collected for. At the same time, specific groups or subgroups might be unfairly targeted by biometric identification tools based on their specific characteristics.

In addition, the issue of consent for the processing of biometric data by the LEAs was brought up. This finding was related to the following AI use case as part of the crowdsourcing platform: an AI system with a database of billions of images taken from social media users without their consent/knowledge, which would be used by the Police to compare the face of a person with public images of people from social media and online sources until a match is found.[33] According to this notion, legislation should require LEAs to obtain explicit and informed consent

[32] Spanish Stakeholder Policy Lab, April 27th 2023.
[33] popAI D3.3 "Citizen produced priorities and recommendations for addressing AI in the security domain"

from citizens before collecting, storing or using their biometric data, including the introduction of mechanisms allowing citizens to easily **withdraw their consent** and have their biometric data deleted before the authorised retention time has expired. It was also expressed that in such cases e.g., the remote collection of special categories of personal data (including biometrics), the processing by the LEAs should be automatically denied or conducted depending on the citizen's approval.

This finding is controversial, because according to the LED and the latest draft Amendments to the AI Act Proposal, processing (of special categories) of data is lawful regardless of the consent of the data subject. [34] Further on, as explained in the Article 29 Opinion on the Law Enforcement Directive, "the consent of the data subject can never in itself constitute a legal ground for the processing (of special categories) of data in the context of the Directive.".[35] That is because, resorting to consent for processing (of special categories) of personal data by LEAs could have negative results and lead to violations of fundamental rights and freedoms, since the relationship between LEAs and citizens is characterised by subordination due to the power imbalance between the parties; therefore, consent by citizens cannot be considered to be freely given.[36] Consent as a legal basis for processing of personal data can even lower the level of existing protection, authorising operations which would otherwise be unlawful.

However, the above do not exclude the possibility that the voluntary agreement of the data subjects (or consent/approval) can be required **as an additional safeguard** by national law in cases where processing of special categories of personal data which is particularly intrusive to the data subject takes place, as described in Recitals 35, 37 of the LED.[37] According to this limited interpretation of the recommendation, national laws should - to the extent authorised by LED- **additionally** require the citizens' voluntary agreement in cases of processing of their biometric data.

[34] See 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) Article 9– Processing of special categories of personal data ; 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 available at : EUR-Lex - 32016L0680 - EN - EUR-Lex (europa.eu), Article 10- Processing of special categories of personal data; European Parliament, Artificial Intelligence Act 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 (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD)) available at : https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EN.html Recital 7 a

[35] Article 29 Data Protection Working Party, Opinion on some key issues of the Law Enforcement Directive (EU 2016/680), Adopted on 29 November 2017

[36] FRA, Handbook on European data protection law - 2018 edition,Data protection, privacy and new technologies Data protection, p.143

[37] 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 available at : EUR-Lex - 32016L0680 - EN - EUR-Lex (europa.eu), Recitals 35, 37

The concerns regarding processing of biometrics by LEAs and the function creep, expressed in the form of (additionally) requiring consent from the citizens, are essential to the public dialogue regarding the future regulation of AI. In a similar notion of the EU recognising the controversies around the processing of biometric data, it is noted that, according to the latest proposed Amendments to the AI Act:

- "real-time remote biometric identification in public spaces, "
- "biometric categorisation systems, "
- "AI systems that create or expand facial recognition databases through untargeted scraping of facial images from internet/CCTVs" and
- "AI systems to infer emotions of natural persons in the areas of law enforcement, border management", (in workplace and education)

are proposed to be among the prohibited AI practices.[38]

As consent is not a panacea, the EU legislator should create a framework determining which AI practices would be lawful or unlawful, considering, among others, the best interests of the citizens.

3.2 Citizens' involvement

Although the use of AI technologies in the security domain can be perceived positively by citizens, it remains a sensitive topic that requires finding the right balance between ensuring citizens' protection and respecting rights and privacy. As our results have shown, citizens have identified three primary areas of concerns: discrimination, privacy, and legitimacy. To tackle these concerns, and/or contemplate a holistic approach in the evolution of technologies, it has been recommended to involve citizens at the preliminary stages of the decision making process.

Involving citizens in the decision making processes around the use of AI in surveillance and security operations has been considered as a means to amplify the affected community's voice and build a relationship of trust among citizens and the State. Engaging citizens would help them gain a more thorough understanding of the purpose of AI tools in policing and allow a more inclusive approach where society would be able to express their thoughts and concerns about the use of AI. At the same time, citizens should be given the opportunity to raise their opinions on the limitations in the use of personal data they would like to see determined by the legislation. Creation of discussion platforms involving citizens, technical developers, LEAs and policymakers has been strongly recommended. This can be achieved through the public consultation to take place in the context of algorithmic impact assessment, as proposed in the Draft Convention on AI by the Council of Europe.[39] In this sense, the HRESIA (Human Rights, Ethical and Social Impact Assessment) and the

[38] European Parliament, Artificial Intelligence Act 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 (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD))

[39] Council of Europe Draft Convention on AI, Human Rights, Democracy And Rule Of Law - MIAI (ai-regulation.com)

SIA (Social Impact Assessment) models are of great value as self-assessment tools which take into consideration the public participation and have an interest in societal core values.[40]

To further enhance a safe and ethical use of data, the creation of a central platform, accepting **citizens' reports of incidents** associated with the (potential) violation of rights, freedoms, etc. due to the use of AI by LEAs should be implemented. This would allow citizens to easily report any potential misuse. They would therefore have the possibility to lodge a complaint and object potential unjust decisions that have been made using their personal data.

Involvement of citizens would therefore be a two-fold progress : on the one hand, dialogue with citizens should be encouraged as from the early stages of AI technology developments and over not only ethical questions but also legal measures to be implemented to ensure the lawful and ethical processing of their data. On the other hand, guarantees should be provided to citizens to secure a reporting system whereby they can report any breach or misuse of their data, and easily lodge a complaint whenever they observe a violation of their rights.

4 Conclusions

Although AI technologies can provide a valuable support to LEAs in the exercise of their operational functions, the right balance between authorising the use of AI and protecting human rights and freedoms must be found. Recommendations provided throughout this report reflect the opinions and thoughts of stakeholders who have agreed to take part in our activities and share the outcomes of their reflection.

While the training of LEAs is a recurrent trend that appears to have been broadly agreed upon by the different stakeholders involved in this exercise, using AI in an ethical manner that supports the work of LEAs and yet respects privacy cannot be achieved without a solid legislative framework. Legislation at the EU level must be adapted and further developed to ensure that both access to the tools and data retention are strictly limited in time and scope.

Ensuring a non-discriminatory use of AI tools can be accomplished through appropriate training as mentioned above. However, it is fundamental to have processes in place that guide technical developers in the development of technologies that are completely free of bias. As from the early stage of conception, AI tools which aim at being used by LEAs should constantly be checked, experimented and assessed (social impact and/or impact on rights and fundamental freedoms) to secure a fair and ethical result that does not target citizens according to any discriminatory criteria.

In response to the community building and ecosystem engagement established by the project, the application and use of AI should be completely human-centric and socially oriented. Consequently, to help designing the future of the AI legal framework, informing citizens and collecting their opinion is crucial. Moreover, citizens should have more access to their data, and should be able to request clarification on where, when, how, and why their data are being used.

[40] Elsevier, Computer Law & Security Review, Volume 34, Issue 4, August 2018, Pages 754-772, Alessandro Mantelero, AI and Big Data: A blueprint for a human rights, social and ethical impact assessment

In a nutshell, a holistic approach is recommended, combining rigorous training of LEAs, communication towards citizens, robust regulations and AI systems designed in a way that upholds ethical and legal principles and safeguards societal values.

The recommendations gathered in this deliverable will provide output for the remaining recommendations of WP4 as well as the communication and dissemination of the popAI project findings via WP5 activities.

5 References

Article 29 Data Protection Working Party, Opinion on some key issues of the Law Enforcement Directive (EU 2016/680), Adopted on 29 November 2017

European Union Agency for Fundamental Rights, EU Charter of Fundamental Rights Article 21 - Non-discrimination | European Union Agency for Fundamental Rights (europa.eu)

Council of Europe Draft Convention on AI, Human Rights, Democracy And Rule Of Law - MIAI (ai-regulation.com)

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 available at : EUR-Lex - 32016L0680 - EN - EUR-Lex (europa.eu)

Elsevier, Computer Law & Security Review, Volume 34, Issue 4, August 2018, Pages 754-772, Alessandro Mantelero, AI and Big Data: A blueprint for a human rights, social and ethical impact assessment

European Parliament, Artificial Intelligence Act 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 (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD)) available at : https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EN.html

European Parliament, 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, 9.5.2023 available at : https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/CJ40/DV/2023/05-11/ConsolidatedCA_IMCOLIBE_AI_ACT_EN.pdf

EUROPOL, CENTRIC, Accountability Principles for Artificial Intelligence (AP4AI) in the Internal Security Domain available at : https://www.ap4ai.eu/

FRA, Handbook on European data protection law - 2018 edition,Data protection, privacy and new technologies Data protection

Fifth (5th) ALIGNER Public Workshop, June 2023

High-Level Expert Group on Artificial Intelligence (AI HLEG), Ethics Guidelines for Trustworthy Artificial Intelligence (AI) available online at: https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-

assessment

Key EU Parliament votes tomorrow on major AI law including landmark ban on 'predictive policing' and criminal 'prediction' systems - Fair Trials available at : https://www.fairtrials.org/articles/news/eu-parliament-votes-on-major-ai-law

popAI Crowdsourcing platform available at: popAI (ecas.org)

popAI D2.1 "Functionality taxonomy and emerging practices and trends"

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

popAI D3.2 "Report on citizen discourses and attitudes towards controversies"

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 D3.6 "Photo Competition Results"

popAI Dashboard available at : PopAI Dashboard (citizens.is)

popAI Draft Report Greek Stakeholder Policy Lab, May 25th 2022

popAI Draft Report Italian Stakeholder Policy Lab, April 20th 2023

popAI Draft Report Slovak Stakeholder Policy Lab, December 13th 2022

popAI Draft Report Spanish Stakeholder Policy Lab, April 27th 2023

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) available at: https://eur-lex.europa.eu/eli/reg/2016/679/oj

The Verge, EU draft legislation will ban AI for mass biometric surveillance and predictive policing available at: EU draft legislation will ban AI for mass biometric surveillance and predictive policing - The Verge

6 Annex

6.1 Crowdsourcing platform (example of questions: negative or positive sentiment)

Please indicate your level of agreement with the statements below.
Please bear in mind that there is no right or wrong answer. We want to hear your perception of the case.

In your opinion, AI systems used by the police for biometric identification...

...respect human rights.

Totally disagree 1 2 3 4 5 6 7 Totally agree

...have enough human oversight.

Totally disagree 1 2 3 4 5 6 7 Totally agree

...are accurate.

...are reliable.

Totally disagree 1 2 3 4 5 6 7 Totally agree

...respect privacy.

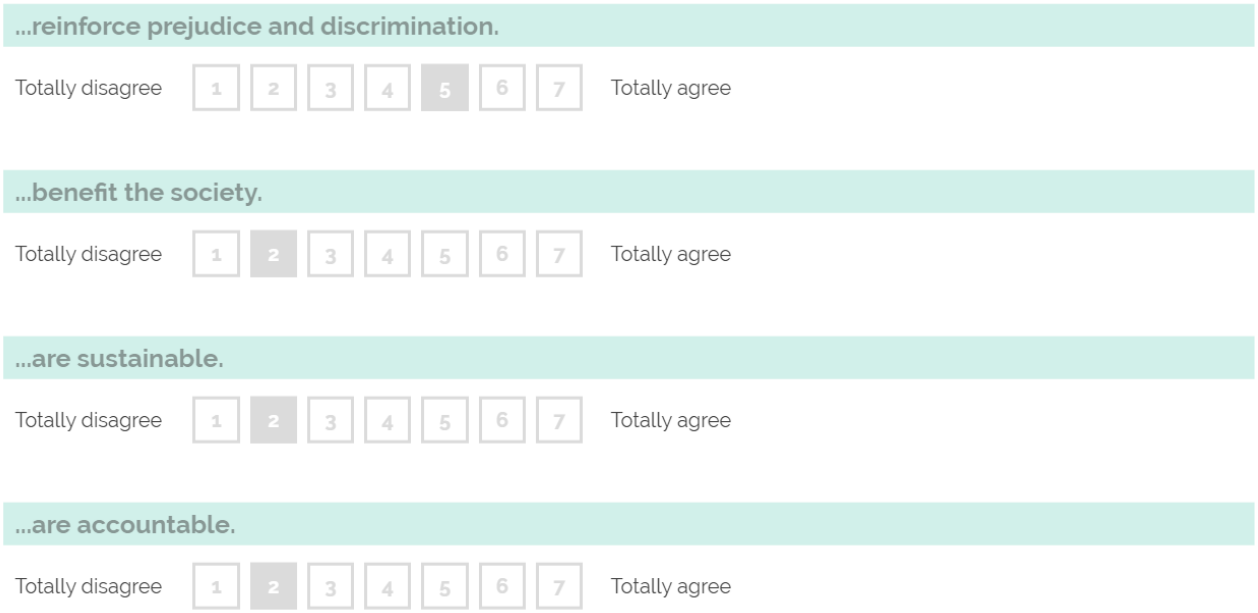
Totally disagree 1 2 3 4 5 6 7 Totally agree

...access to people's data legitimately.

Totally disagree 1 2 3 4 5 6 7 Totally agree

...are used with transparency.

Totally disagree 1 2 3 4 5 6 7 Totally agree



6.2 Crowdsourcing idea generation (example of question: idea generation)

Law enforcement agencies inceasingly use artificial intelligence (AI) systems to prevent crime, based on the promise that AI tools can help in the prediction and anticipation of crime. The AI algorithms used to predict crime rely on historical police data together with other data such as demographic, socio-economic data, as well as real time data from digital devices (mobiles for example). These data are merged together to create models that predict where crime is most likely to occur and also who may commit it. Serious concens have been raised regarding the accuracy of AI in predictive policing. These concerns regard the data that are used and the accuracy of the output. The way data are used might breach data protection laws in terms of privacy and also lead to some erroneous output where people living in area that are socio- economical disadvantaged are more targeted. In this way, the AI system in use reinforces discrimination. This has been considered a serious threat to human rights.

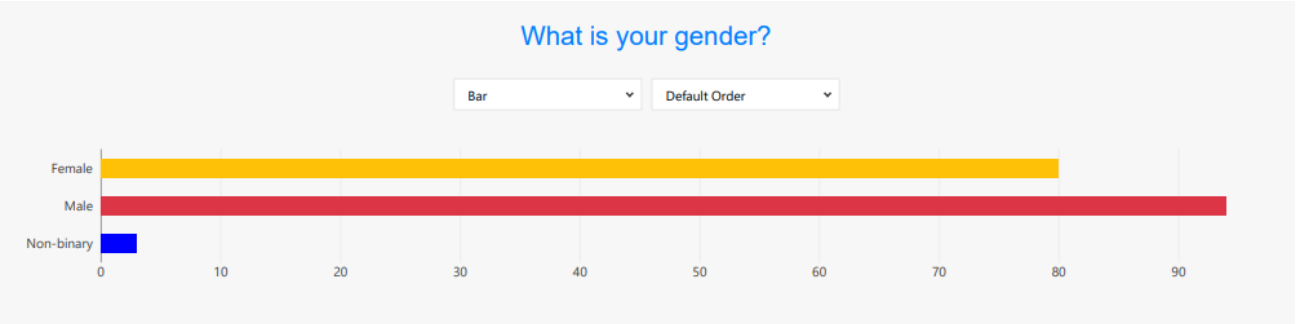
How can we avoid discrimination and prejudice resulting from predictive policing algorithms?

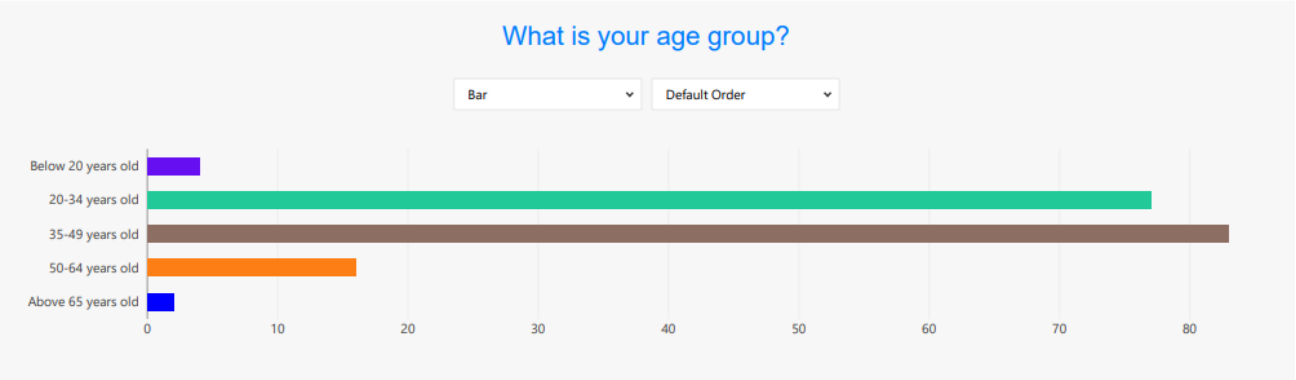
What inclusion and diversity measures should be taken in order to mitigate biases?

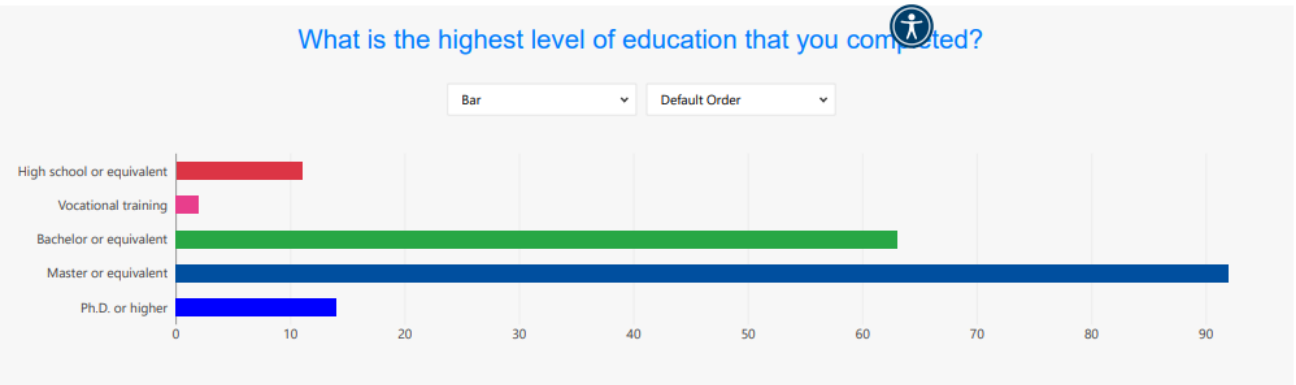
Also in this case being transparent on how these algorithm are trained is essential to avoid any problems. The key rest on the data quality on which these algorithms have been trained on.

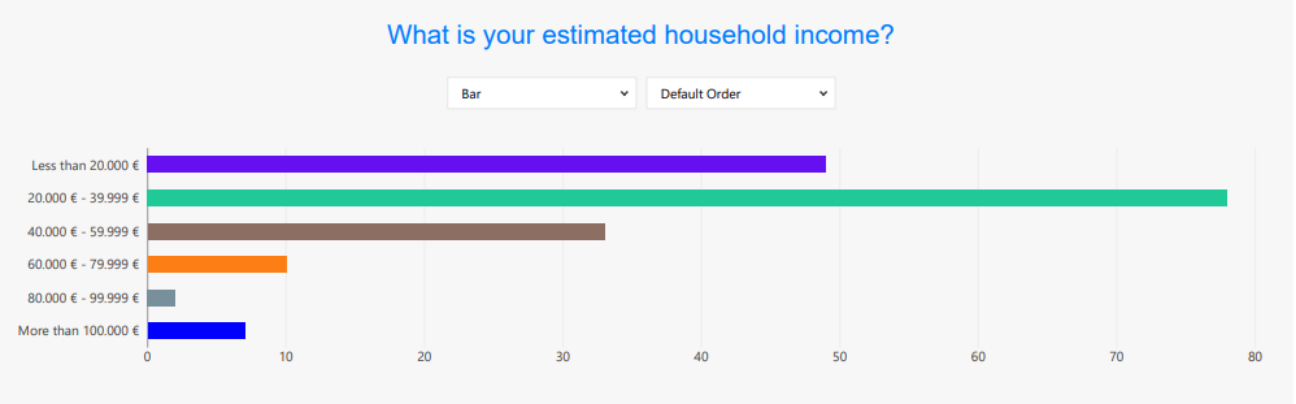
6.3 Crowdsourcing platform participants' statistical data

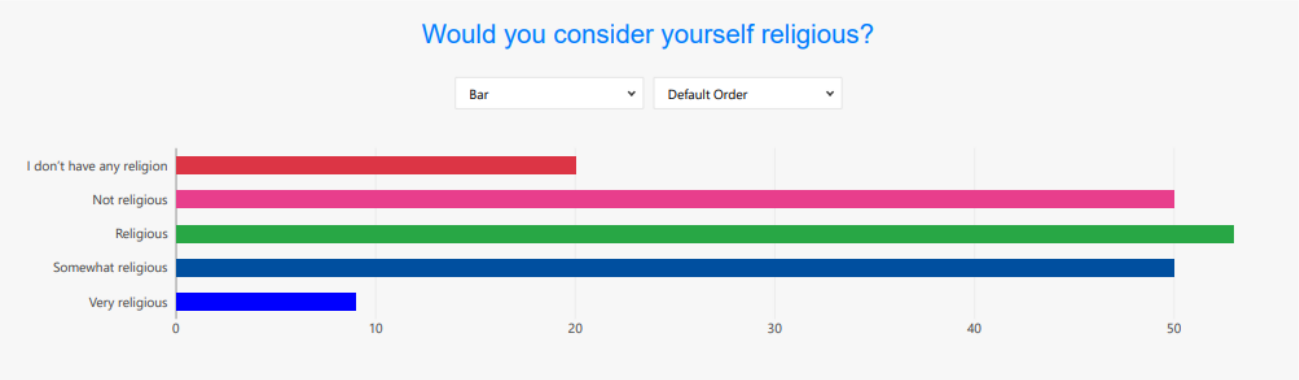
6.3.1 Crowdsourcing platform participants' statistical data; Phase 1 :


What is your gender?
Bar
Default Order
Female
Male
Non-binary
0
10
20
30
40
50
60
70
80
90

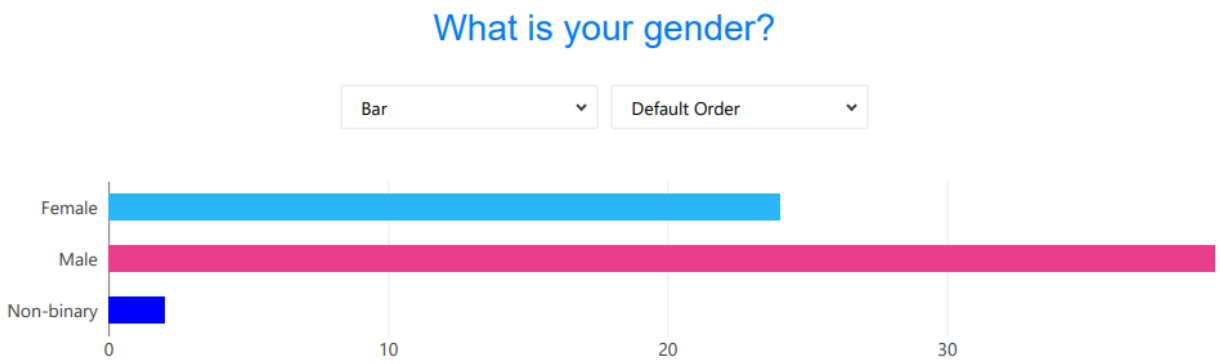

What is your age group?
Bar
Default Order
Below 20 years old
20-34 years old
35-49 years old
50-64 years old
Above 65 years old
0
10
20
30
40
50
60
70
80

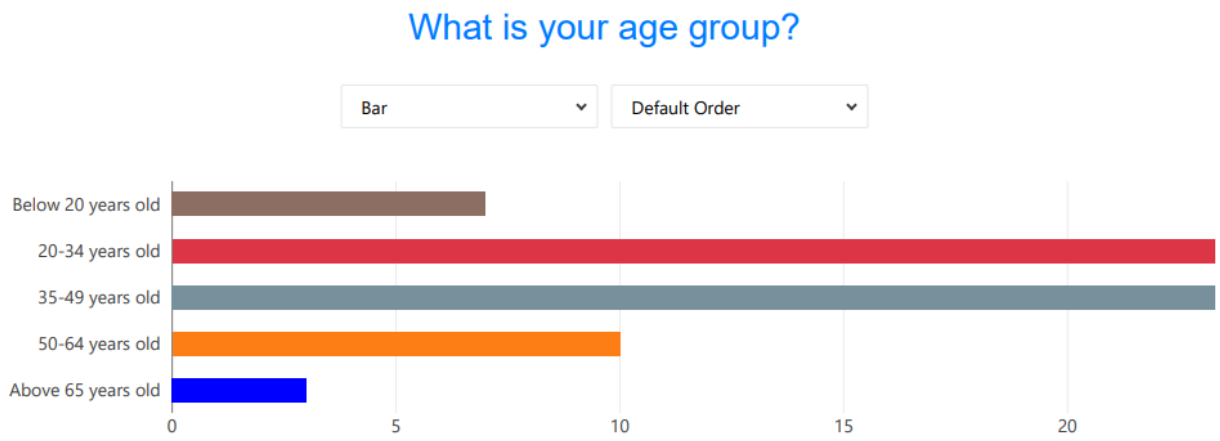

What is the highest level of education that you completed?
Bar
Default Order
High school or equivalent
Vocational training
Bachelor or equivalent
Master or equivalent
Ph.D. or higher
0
10
20
30
40
50
60
70
80
90


What is your estimated household income?
Bar
Default Order
Less than 20.000 €
20.000 € - 39.999 €
40.000 € - 59.999 €
60.000 € - 79.999 €
80.000 € - 99.999 €
More than 100.000 €
0
10
20
30
40
50
60
70
80

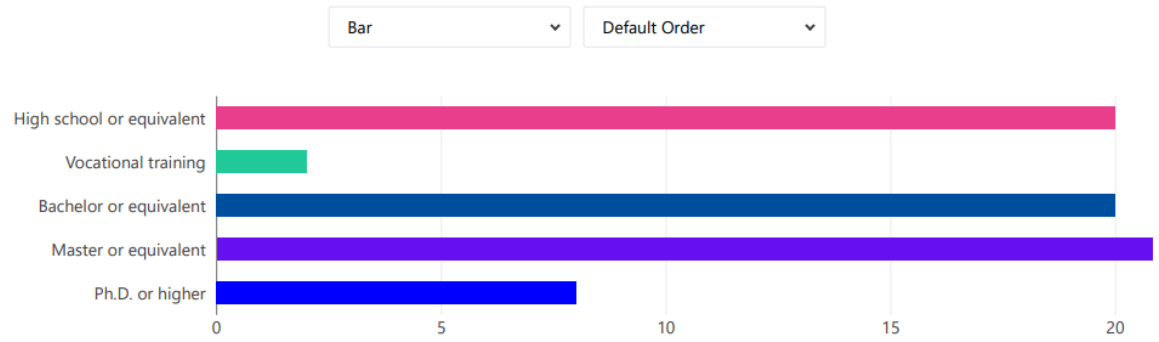


6.3.2 Crowdsourcing platform participants' statistical data; **Phase 2:**

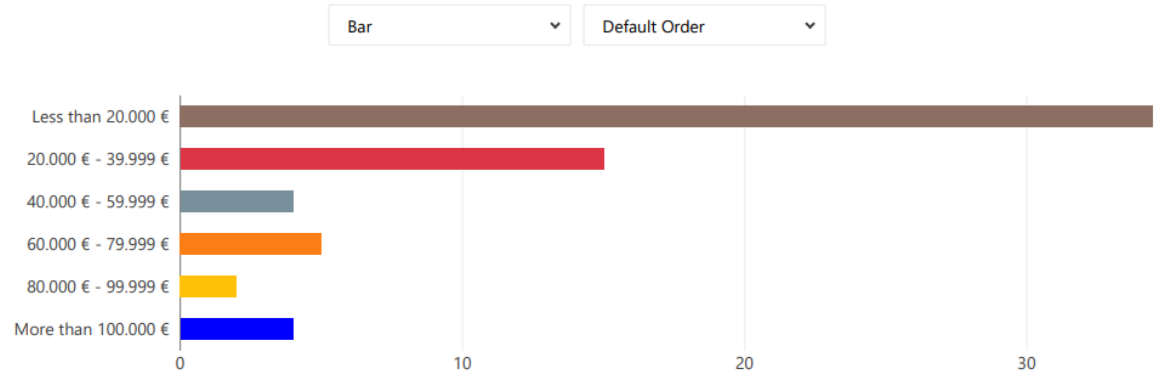




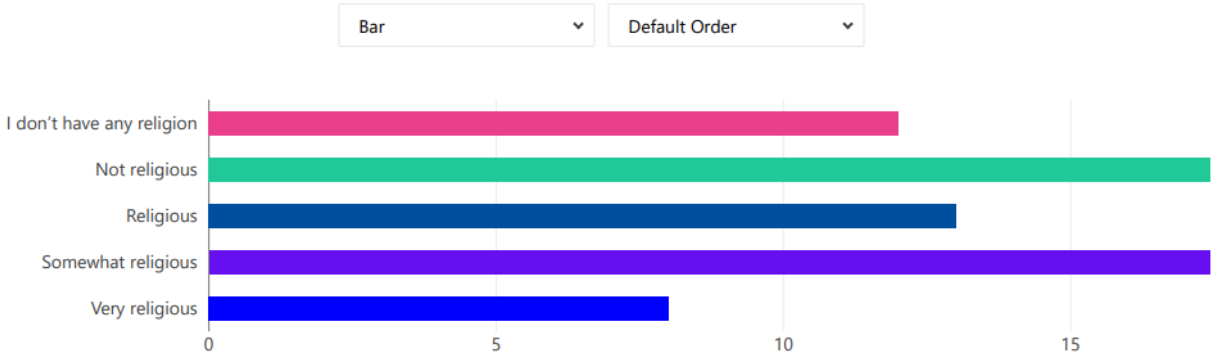
What is the highest level of education that you completed?



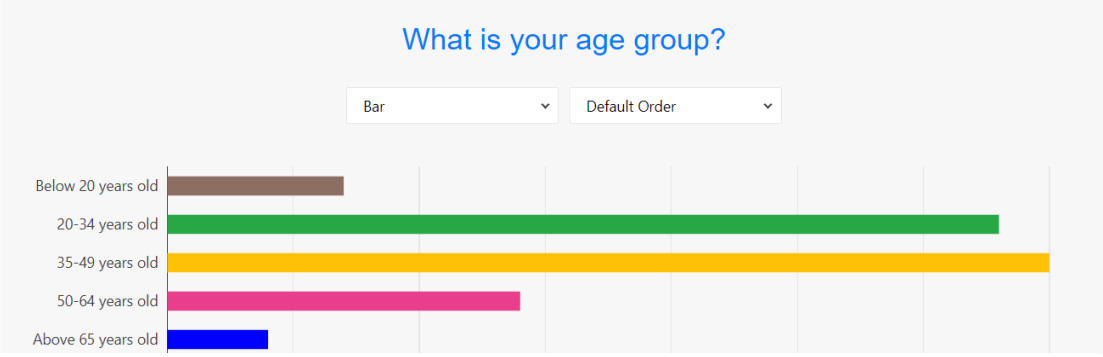
What is your yearly income?

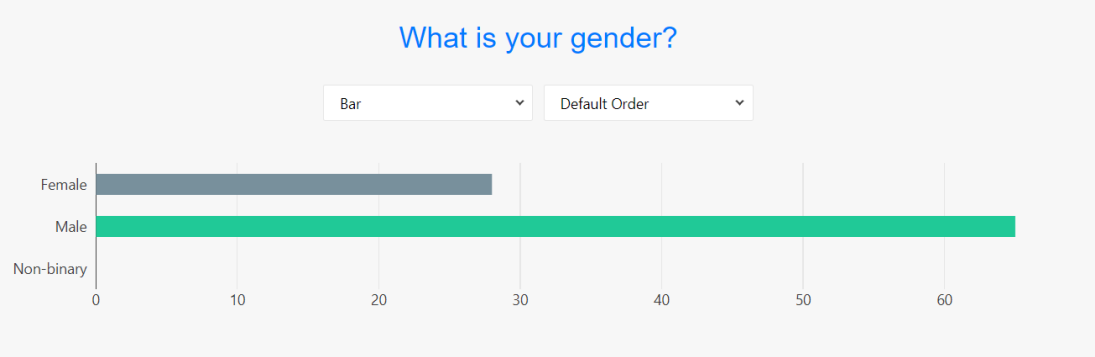


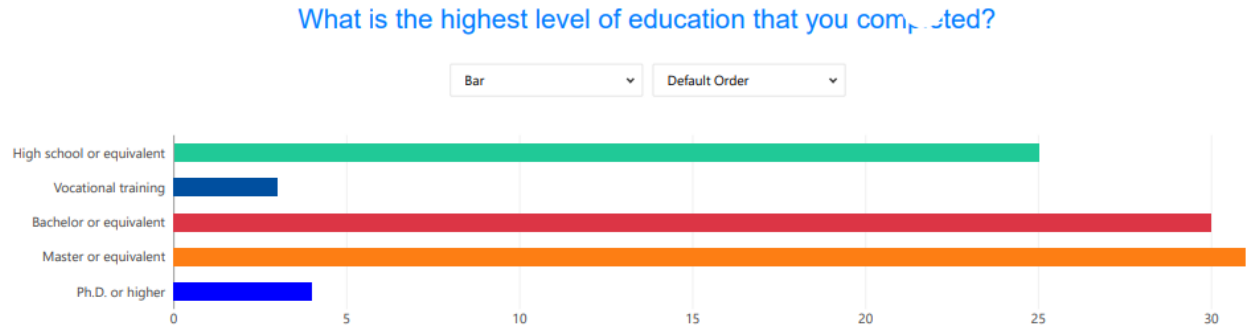
Would you consider yourself religious?

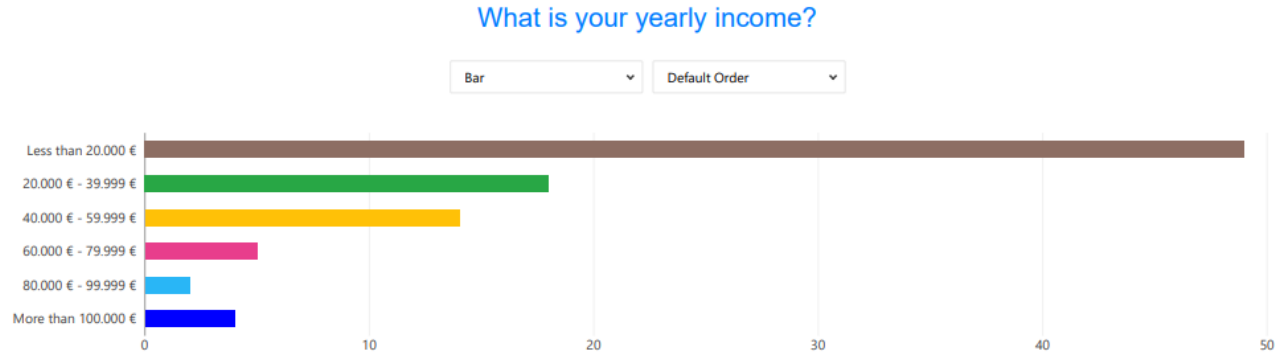


6.3.3 Crowdsourcing platform participants' statistical data; Phase 3:









6.4 Countries of residence of crowdsourcing platform participants

6.5 Example of Policy Labs case studies

Country	Case study 1	Case study 2
Greece	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.	The AI system will detect dangerous driving using video footage from traffic management cameras or other real-time footage
Germany	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	AI to process CSAM material: 1. Various hard disks and data carriers are seized from one suspect

	injured by a knife.	2. Within the framework of international police reporting systems, hundreds of suspicious online contents are reported to the German police every day. 3. All suspicious and seized material is individually visually inspected manually by the officers
Slovakia	AI in support of monitoring the social networks (crime prediction): All suspicious and seized material is individually visually inspected manually by the officers	Use of ethics too box (see Annex 8.3)
Italy	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	Use of ethics toolbox (see Annex 8.3)

	in the video streams generated by the connected camera system.	
Spain	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.	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.

6.6 Policy Labs participants

Greece:

LEAs	14
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
CERTH	1
KEMEA	1
Shadow researcher	3

Germany:

LEAs	8
Logobject	2
Munich Innovation	1
Adesso SE	2

Slovakia:

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
LEAs	27

Italy:

Think Legal	1
Ethic Solution	1
LEAs	8
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

Spain:

LEAs	24
CIDALIA	1
University of Alcala	1
City Council Department	3
Ministry of Interior	1

FUNDACIÓN SECRETARIADO GITANO	1
OBERAXE (government organisation)	1
University Complutense of Madrid	1

6.7 Example of questions raised to the Stakeholder Advisory Board during the popAI Plenary in Rome, Italy

Could you name the fields in which AI tools are essential to protect citizens?

1. Recognition
2. Communication
3. Prediction & Analytics
4. Surveillance
5. Crime Prevention
6. Crime Investigation
7. Cyber Operations
8. Migration, Asylum, Border Control
9. LEAs Training
10. Administration of Justice

Should the use of “real-time” remote biometric identification systems in publicly accessible spaces, be allowed for LEAs?

Yes, in any case for the protection of the citizens.

Yes, in any case of a criminal investigation.

Only in specific cases.

No.

On a general note, can you mention some recommendations, practices, rules and ideas that could foster citizens’ trust of AI used by LEAs?