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

D4.3: White paper for Technology Developers

Grant Agreement ID	101022001	**Acronym**	popAI
Project Title	A European Positive Sum Approach towards AI tools in support of Law Enforcement and safeguarding privacy and fundamental rights		
Start Date	01/10/2021	**Duration**	24 Months
Project URL	https://www.pop-ai.eu/		
Document date	30/09/2023		
Nature	R: Document, report	**Dissemination Level**	PU: Public
Author	Panagiotis Douris (KEMEA)		
Contributors	Dimitra Papadaki (KEMEA), Geogia Melenikou (KEMEA), Sofia Segkouli, Anastasios Drosou, George Lazaridis, Vangelis Kopsacheilis (CERTH), Pinelopi Troullinou (TRI), Virginia Bertelli (ETICAS), Xenia Ziouvelou, Marilena Sinni, Dimitris Kyriazanos (NCSRD), Claire Morot-SIr (ECAS)		
Reviewers	Lilian Mitrou (External Ethics Advisor), Sofia Segkouli, Anastasios Drosou, George Lazaridis, Vangelis Kopsacheilis (CERTH), Andreas Ikonomopoulos (NCSRD)		

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

Executive Summary

Nowadays, the advancements and progress in the field Artificial Intelligence (AI) as well as the number and variety of its applications keep increasing. Security lies among the numerous sectors of the potential application of AI under which Law Enforcement Agencies (LEAs) could use AI to assist them in their everyday tasks and operational activities. It is the high performance of the AI algorithms, including their high speed of processing, analysing, automating, visualising data / results, and increased accuracy that LEAs aim at taking advantage of, especially within the context of decision-making. Numerous AI practices for Law Enforcement purposes are considered in principle as high-risk.[1][2] Therefore, the execution, performance, functionalities, and results produced by the corresponding systems need to be carefully and thoroughly assessed, analysed, and examined to ensure they comply with the applicable legislation, legal and ethical requirements. The purpose of this deliverable is to present, detail and analyse several recommendations for and from Technology Developers for the ethical use of AI for LEAs produced in the lifetime of the popAI project. It also aims at contributing to the identification and collection of the best multi-disciplinary practices for the same purpose, together with the related outputs of Tasks 4.1 (Recommendations for and from LEAs and policymakers), T4.2 (Recommendations for and from the Civil Society as presented in D4.2) for the purposes of Task 4.4 as it will be demonstrated in "D4.4 Synthesis: a collection of best multidisciplinary practices". The result of this deliverable is a White Paper in the form of a report, in order to outline the recommendations in a concise and intelligible manner.

[1] "Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS", p. 27, European Commission, URL: https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0001.02/DOC_1&format=PDF, last accessed online on 02/05/2023

[2] Andreas Liebl and Till Klein, "AI Act: Risk Classification of AI Systems from a Practical Perspective", applied AI, URL: https://aai.frb.io/assets/files/AI-Act-Risk-Classification-Study-appliedAI-March-2023.pdf, last accessed online on 02/05/2023

Table of Contents

Table of Figures

List of Terms & Abbreviations

Abbreviation	Definition
AI	Artificial Intelligence
AES	Advanced Encryption Standard
ALTAI	Assessment List for Trustworthy AI
CSA	Coordination and Support Action
CEPOL	European Union Agency for Law Enforcement Training
DoA	Description of Action
EAB	Ethics Advisory Board
EC	European Commission
EU	European Union
FS	Foresight Scenario
GA	Grant Agreement
GDPR	General Data Protection Regulation
ICT	Information and Communications Technology
ID	Identification
IPR	Intellectual Property Rights
IT	Information Technology
LEA	Law Enforcement Agency
LED	Law Enforcement Directive
MS	Member State
NDA	Non-disclosure agreement
QR	Quick Response
PoC	Point of Contact
SAB	Stakeholders’ Advisory Board
SPL	Stakeholder Policy Lab
UUID	Universally Unique IDentifier
USA	United States of America
XAI	Explainable AI

1 Introduction

Since its advent, AI keeps being introduced in more and more fields and expanding its application. Security is one of these sectors and LEAs have also started introducing or would like to use AI to assist them in their operations. AI plays a crucial role in multi-domain operations, which encompass several use cases, such as surveillance, forensics / analytics, communication, prevention and investigation of crime incidents or malicious acts. Considering the large number of such use cases, there is a need for a systematic, complete, and clear organisation of LEA functionalities along with their corresponding relations to AI techniques, data sources and potentials sources of controversies. To meet this need, a Law Enforcement functionality taxonomy has been introduced in deliverable D2.1, aiming at the reflection of the basic aspects of LEA functionality use cases, application area, AI technology used and respective data sources.

Based on the outcomes of WP2 framework and especially of the empirical research of WP3, the main aim of this deliverable is to produce and deliver a set of recommendations with practical value for technology developers, including academia and industry as well as SMEs (considering AI services and product designers), when designing AI systems and related products, developer tools, and processing data. Towards the goal of designing AI tools, accepted, and valued by the civil society and LEAs, specific principles and applicable legal frameworks have been taken into consideration. These include, but are not limited to, the framework identified under WP2 specifically for LEAs, the ALTAI principles, the applicable data protection legislation, and the draft Amendments to the AIA Proposal.[3]

Additionally, the dynamic interaction among legal and technical actors has been made best use of, so as to be in a good position to effectively translate the legal and ethical principles to technical specifications and vice versa. The taxonomy and trends presented in D2.1 as well as the stakeholders' views, as detailed in D3.4, have been taken advantage of. Last but not least, consideration of sibling project (ALIGNER, STARLIGHT) outcomes, have been taken into consideration.

1.1 Aim and Scope

Although the WP4 tasks are orientated towards different groups and except for the relevant differences they exhibit as they seek to produce recommendations addressed to diverse stakeholders, their common goal is to produce recommendations for the ethical use of AI by LEAs. Taking into consideration the nature of popAI project and the context of the type of execution, the GA, and the specific descriptions therein, in conjunction with the existing and developing legal framework, the recommendations aim to illuminate and complement the current applicable legal and ethical frameworks.

The main aim of the present deliverable is to produce a set of recommendations for technology developers associated with the ethical use of AI for LEAs. For a graphical representation of the route to the production of these recommendations, please refer *to Figure 1 below.*

[3] AI HLEG, Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self-assessment, URL: https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment, last accessed online via web browser on 24/7/2023.

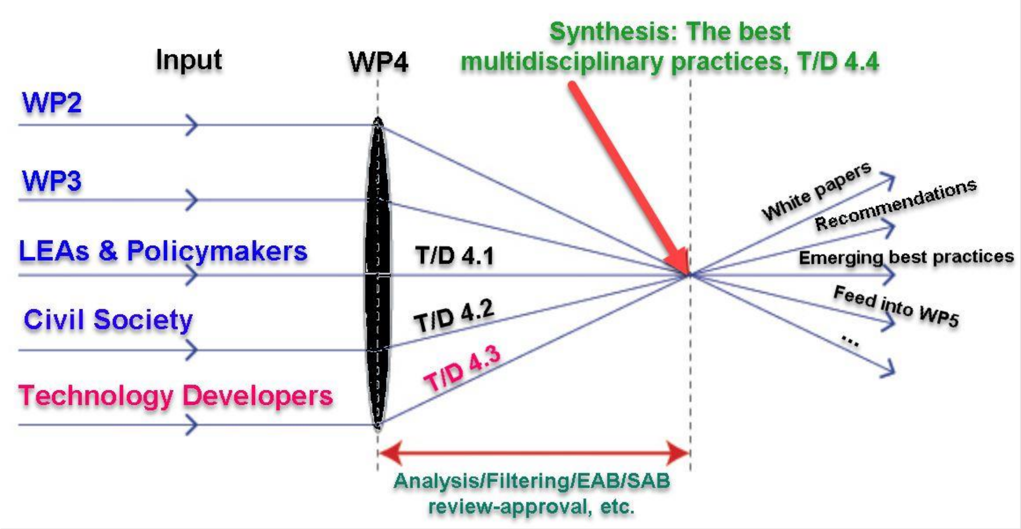


Figure 1. The route to the production of recommendations & further actions

These recommendations are considered valuable for a number of reasons, quoted below:

- The popularity of AI has increased in recent years, as it has been introduced in widely used applications worldwide, such as search engines, social media, popular software, etc. It also finds widespread use in various security related application (see popAI D2.1).[4] In order to face problems caused by this widespread adoption, there is a need for ethical and legal guidelines and provisions, respectively that are specific to the security field.
- Several issues and publicised controversies, which have gained great publicity, including the "Facebook-Cambridge Analytica data scandal", Clearview AI case, Prokid case (see popAI D3.1) have garnered substantial attention.[5] These instances have raised public awareness and shed light on the importance of addressing legal and ethical concerns associated with AI technologies in the security field.
- The growing prominence of AI has resulted in heightened interest, awareness, and engagement from civil society. The public perception and involvement in discussions, surrounding AI ethics have been amplified, highlighting the need for detailed recommendations.
- Law Enforcement is in the spotlight, given that actions by LEAs involving certain uses of AI systems are characterised by a significant degree of power imbalance that are likely to result

[4] D2.1 - Functionality taxonomy and emerging practices and trends", p. 12, NCSRD, Project title: A European Positive Sum Approach towards AI tools in support of Law Enforcement and safeguarding privacy and fundamental rights (popAI), Horizon 2020, Grant Agreement ID: 101022001

[5] "D2.1 - Functionality taxonomy and emerging practices and trends", p. 12, NCSRD, Project title: A European Positive Sum Approach towards AI tools in support of Law Enforcement and safeguarding privacy and fundamental rights (popAI), Horizon 2020, Grant Agreement ID: 101022001

in surveillance, arrest or deprivation of a natural person's liberty as well as other adverse impacts on fundamental rights.

- At this stage, there is an absence of a harmonised regulatory framework on AI. The AI Act is still a work in progress and clear rules and guidelines are needed to support the ethical development and deployment of AI. However, data protection legislation (GDPR and LED) is applicable.
- Striking a balance between the advantages and potential risks of AI is key, hence special emphasis needs to be placed on the provision of recommendations that will aim to the mitigation of potential risks and of adverse impact on fundamental rights.
- Since AI intersects with various domains, including technology, legal science, security, etc., a multi-disciplinary approach is necessitated and therefore the contribution, views, and recommendations from various groups of actors as well as potentially affected groups to tackle ethical considerations effectively. Indicatively, contributions, perspectives, and recommendations from diverse stakeholders, including potentially affected groups are vital in shaping ethics guidelines.
- "Research ethics is based on the explicit European commitment to human rights", so the production and delivery of such recommendations is in line with the will and goals set by the EU.[6] So, such guidance resonates with the goals and intentions set forth by the European Union.

Furthermore, the recommendations contribute to the design of AI-based technologies and tools, which could potentially be accepted and valued by both citizens and LEAs. In addition to the principles for a trustworthy AI set by the EU, the associated developed products and procedures entailed, such as the processing of data, must be secure and transparent in conformity with the existing applicable legal framework (EU Charter of Fundamental Rights, EU data protection legislation) as well as their functions and outcomes need to be comprehensible and interpretable without violating any Intellectual Property Rights (IPR). Moreover, the set of recommendations also aims at informing and further raising the citizens' awareness on the most fundamental characteristics an AI-system and related product must possess, so as to be compliant with all the principles, rules and guidelines associated with the ethical use of AI for LEAs. Additionally, there are also practical guidelines and recommendations for certain actions the users and operators should take or, equally importantly, recommendations against taking some other actions and / or cautions, warnings, etc. The recommendations either have a proactive character or a reactive one, in the sense that they are either associated with actions and advice the users or operators should take or avoid so as not to encounter any issues or with actions and advice they should take or avoid after they have encountered a certain issue, respectively.

Additionally, this deliverable concerns the production and delivery of multi-perspective recommendations, in the sense that their origin is multi-stakeholder and multidisciplinary, seeking

[6] "Ethics for researchers, Facilitating Research Excellence in FP7", p.4, European Commission, URL:https://ec.europa.eu/research/participants/data/ref/fp7/89888/ethics-for-researchers_en.pdf, last accessed online via web browser on 2/5/2023

to cover the diverse needs and requirements of all groups of interest. The main goal of the recommendations concerns the protection and full respect of the principles set by EU for a trustworthy AI, human rights, freedoms, law, ethics, societal as well as environmental values and ethics. Another important aspect that has been seriously considered, while producing the set of recommendations is that careful consideration and special attention to vulnerable groups, non-adults, and minorities should be given for each one of these groups on a group-dedicated basis. Furthermore, inclusiveness has also been considered, ensuring no affected group and no group of interest is excluded, thus further pursuing, and ensuring diversity, multi-disciplinarity and that the identified views are included.

Further to producing recommendations, which achieve the aforementioned goals, some additional goals have also been set. No matter how useful, important, and valued the recommendations could be considered, we would like to confidently argue that their effective dissemination via WP5 and the efficacy of informing the groups of interest are concerned. Moreover, wherever possible, it has been attempted to propose recommendations, which are robust or are associated with robust frameworks and solutions. For example, it has been attempted to produce recommendations that are robust with respect to changes in the legislation framework, ethics, technology, etc. Whenever robustness was not possible, ways to adapt to changes and (ways to make the appropriate) modifications have been suggested with the aim to make our recommendations configurable or customisable.

Concerning the scope of this deliverable, it aims at taking advantage of the knowledge obtained from the literature review of WP2 and the empirical research of WP3 with the broader ecosystem, to create a library of group-specific recommendations appointed to AI technology developers. To this end, exhaustive research has been conducted to cover the identified by popAI fields of AI applications for LEAs.

Furthermore, we captured and elicited the views and information from several types of developers (academia, industry, SMEs), both from inside the consortium and outside of it to validate/evaluate and update the recommendations, as the last methodological step. More specifically, we reached out to developers, and we also asked the partners of the consortium to reach out to different developers. We also involved developers from inside the consortium with relevant knowledge and experience. Moreover, it is via the interaction, collaboration, and exchange of ideas among the involved groups, including but not limited to technology developers that the recommendations produced, are expected to be useful to and usable by LEAs.

Finally, often it is argued that given that the LEAs are governed by law and they do not have divisions, capable of making their own laws, but they enforce the law (i.e., the principle of the separation of powers is effectively enforced), they act and use all technology and products ethically and if they not, there already exist the proper measures, established procedures, etc. to check, monitor, control, audit, judge, and punish them, if needed, and to take all steps necessary to resolve any issues and / or remediate them, etc. However, this is not always the case in practice. Therefore, the need for the production of recommendations is of great importance.

Since the whole deliverable is associated with the ethical use of AI by LEAs, we proceed with brief definitions highlighting the difference between morality and ethics to distinguish between these terms and to avoid any potential confusion, as follows:

- ***Morality***: is often used with reference to an individual's moral standards for themselves. Morality is also defined as something that is personal and normative, whereas ethics is the standards of "good and bad" distinguished by a certain community or social setting.[7]
- ***Ethics***: a system of moral principles and the associated rules of conduct arising from them. Furthermore, ethics constitutes a set of moral principles that determine right or wrong behaviour. The term refers to a person's moral beliefs or principles which govern their conduct. In essence, some differences include the following:
  - Ethics concern how an individual behaves / acts, whereas morals are associated with what they believe. So, morality is often (more) subjective, whereas ethics tend to be more objective.
  - Ethics is based on logic and reason and a widely shared set of established values, while morals can be based on religion, culture, tradition, etc.
  - Ethics also deal with professional conduct, while morals usually deal with personal conduct.

For a graphical representation of ethics versus morality versus law via a Venn diagram,[8] please refer to *Figure 2 below.*

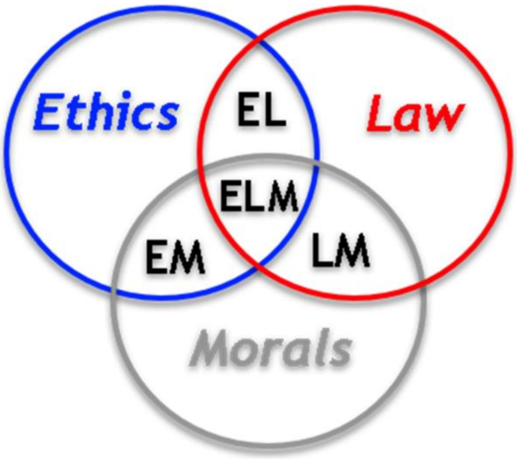


Figure 2. Ethics vs Morality vs Law (Venn diagram)

1.2 Relation with other WPs, Tasks and Deliverables

The dependence of T4.3 "Recommendations for and from technology developers" and, consequently, of this deliverable on other WPs, Tasks and Deliverables can be briefly outlined as follows:

D4.3 received input from:

[7] https://www.britannica.com/story/whats-the-difference-between-morality-and-ethics

[8] More, Trenchard. "On the construction of Venn diagrams." *The Journal of Symbolic Logic* 24.4 (1959): 303-304.

WP1:

- D1.6 “Policy briefs - 1st year”

WP2:

- D2.1 “Functionality taxonomy and emerging practices and trends”
- D2.2 “Legal casework taxonomy: emerging trends and scenarios”
- D2.3 “The controversies and risks that will shape AI in the next 20 years”
- D2.4 “Ethical frameworks for the use of AI by LEAs”
- D2.5 “Practical ethics toolbox for the use of AI by LEAs”

WP3:

- D3.1 “Map of AI in policing innovation ecosystem and stakeholders”
- 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”
- D3.5: “Foresight Scenarios for AI in Policing”
- D3.6: “Photo Competition Results”

WP4:

- D4.2: “White Paper for Civil Society”
- D4.1: “White Paper for LEAs and policymakers”

D4.3 provides output to:

WP1:

- D1.7: “Policy briefs - 2nd year”

WP4:

- D4.3: “White Paper for LEAs and policymakers”
- D4.4: “Synthesis: a collection of the best multidisciplinary practices”

WP5:

- D5.2 “Final community building and ecosystem engagement activities plan”
- D5.6: “Communication & Dissemination plan – final”
- D5.7: “Sustainability and exploitation plan”
- D5.8 “popAI roadmaps”

We can refer to the aforementioned dependencies as explicit, in the sense that they can be considered the most direct ones, whereas all other dependencies can be referred to as “implicit”, as we can assume the dependencies on them are (more) indirect. For a graphical representation of the interdependencies between WP4, its Tasks and Deliverables and other WPs, Tasks and Deliverables, please refer to *Figure 3*.

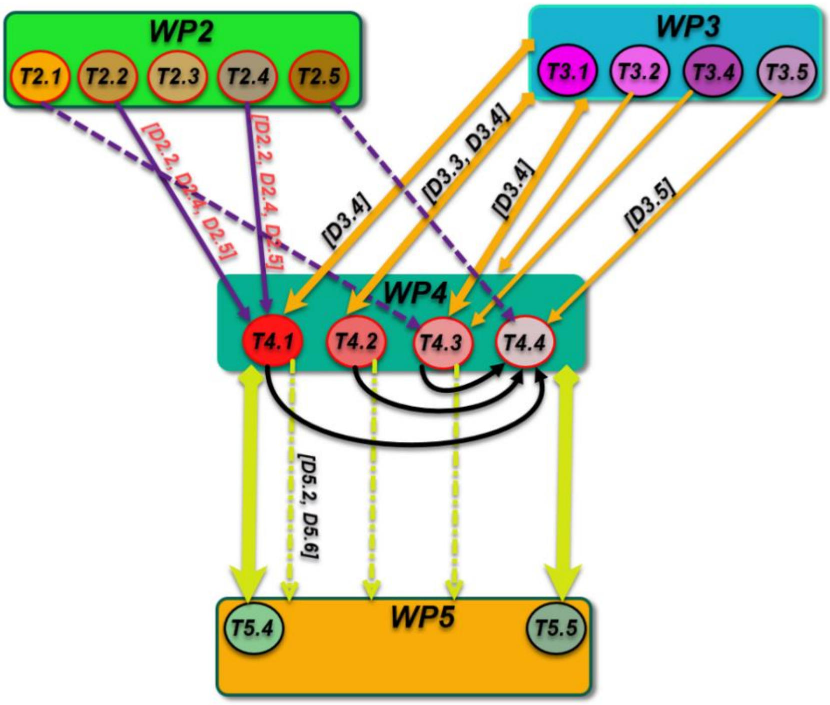


Figure 3. Interdependence of WP4 with other WPs

Essentially, within the same WP, there exist dependencies between different deliverables and tasks (i.e., the term "self-dependency" is associated with the interrelation between tasks and deliverables of the same WP). For instance, the recommendations for technology developers can stem from another group, such as the policymakers, the citizens as well as from other developers, too (possibly from a different entity, such as SMEs, academia, etc.). The self-dependencies cannot be discarded or ignored, as different entities in the same group may borrow useful ideas, concepts, methodologies, strategies, etc. from other ones and lend some of theirs to those. Additionally, through the exchange of ideas and practices between them, the emergence of best practices and lessons learnt are possible.

An implicit dependence is on WP1, on "D1.6 Policy briefs – 1st Year" which provides input to WP4, while WP4 provides output to "D1.7 Policy briefs – 2nd Year".

1.3 Structure of the Deliverable

The rest of the deliverable is structured as follows: *Section 1* provides the Introduction to the deliverable, *Section 2* presents the methodological approach adopted for the production and delivery of recommendations to technology developers, including the relation to other WPs and deliverables. The guidelines and criteria set as well as the sources of information, input, and data are mentioned, too. *Section 3* details the recommendations to technology developers[9]. *Section 4* is associated with the evaluation of the outputs and results of this deliverable and the corresponding Task (i.e., Task 4.3). *Section 5* discusses some potential shortcomings of the approach followed and proposes future extensions and improvements. *Section 6* concludes the deliverable.

2 Methodology

2.1 Objectives

The overall procedure defined, adopted, and followed towards the production and delivery of the "Recommendations for and from Technology Developers" constitutes an approach, which is in turn, based on a methodology, which seeks to achieve specific goals set, such as the following:

- Remain compliant with the GA and the milestones set therein.
- Include and / or consider the outputs of all the associated tasks, deliverables, workshops, outputs from other "sister" / relevant projects etc., and thus satisfy all the appropriate interdependencies between tasks and deliverables related to this one.
- Ensure the produced results are in line with the existing ethical and legal frameworks, including the personal data protection framework and the ALTAI principles.
- Take into consideration the latest developments in the AI Act Proposal (EU) and the Draft Convention on Artificial Intelligence, Human Rights, Democracy, and the Rule of Law (CoE).
- Ensure the results of this deliverable are valid, useful and of practical use for the LEAs and the technology developers (i.e., for all target groups).
- Produce multidisciplinary results and cover the needs of LEAs and the target groups under consideration from all different perspectives of interest.
- (Cross-)validate the results generated to ensure their validity from different perspectives and according to the needs of LEAs and the target groups of interest.
- Present the recommendations in an attractive and comprehensive way for audiences and detail their usefulness in an unambiguous reasonable manner.
- Discuss the robustness of the approach, potential limitations as well as ways to increase the former and reduce the latter.
- Ensure that as many fields and cases of interest as possible have been covered, including the (fields of) application(s) of LEAs, the principles set by EU for a trustworthy AI, human rights, freedoms, law, ethics, societal as well as environmental values and ethics, society needs, etc.
- Effectively capture and elicit(ate) the target groups' input and resolve any (potential) contradictory answers / feedback received.

[9] "D2.1 - Functionality taxonomy and emerging practices and trends", p. 12, NCSRD, Project title: A European Positive Sum Approach towards AI tools in support of Law Enforcement and safeguarding privacy and fundamental rights (popAI), Horizon 2020, Grant Agreement ID: 101022001

- Present, deliver, disseminate, and communicate the produced recommendations appropriately in order they are informative, clear and they spur the targeted audiences' interest.
- Feed the recommendations and useful output into "D4.4 - Synthesis: a collection of the best multidisciplinary practices" and the dissemination activities appointed to citizens, policymakers, LEAs, and the industry under WP5.

Noteworthily, based on what has already been presented above concerning the methodological approach, the task of effectively combining the heterogeneous inputs from all the diverse groups of interest, while satisfying the heterogeneous needs and demands, is challenging. These outputs serve as a presentation of the project results to the public and groups of interest.

2.2 Methodological approach

WP4 applies a combination of doctrinal and empirical research, in order to answer the question of what the emerging best practices or recommendations for the ethical use of AI would be. D4.3, in particular, provides for the recommendations for the ethical use of AI by LEAs, appointed to the technology developers, as described in the GA. On the one hand, one of the sources where it draws the theoretical framework from, is WP2 "Security AI in the next 20 years: trends, practices and risks", while following up on the latest developments regarding the ethics principles (especially the ALTAI principles) and the regulation of AI-related issues at the European level (most importantly the AI Act Proposal and the Draft Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law), to ensure that it stays up-to-date, due to the numerous and essential latest changes in the forthcoming legal landscape.[10] Furthermore, it makes use of WP3 'Empirical Knowledge Collection and Management Framework' results, with a focus, among others, on the Policy Labs of Task 3.4 to which stakeholders including LEAs and policymakers, civil society representatives and technology developers participated. In addition to the above sources, auxiliary sources were utilised to inform the recommendations, such as the D1.6 Policy briefs. Furthermore, questionnaires based on WP2 and WP3 taxonomies, functionalities, and controversies have been developed under WP4. The questionnaires were appointed to LEAs and policymakers of the Consortium as well as to limited externals, and serve WP4 as an assistant tool to support, update, crosscheck and evaluate the emerging results regarding recommendations. Lastly, it takes into consideration the feedback of popAI SAB and EAB to inform the recommendations and the sibling projects' (ALIGNER and STARLIGHT) proposals on the issue of ethical AI for law enforcement purposes.

The rationale behind this approach is that it studies the relation between the existing and forthcoming ethical and legal frameworks and the opinions of the interested stakeholders with various backgrounds (LEAs, policymakers, civil society, technology developers), in order to seek potential solutions or balance exercises to the quest of ethical AI for LEAs. However, it sets the ALTAI principles as a minimum threshold to classify the popAI findings as emerging best practices or

[10] News, European Parliament, EU AI Act: first regulation on artificial intelligence, https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence, last accessed online via web browser on 23/07/2023.

recommendations. Therefore, the role of the ethics principles and emerging legal frameworks are twofold: they serve both as a source and as a threshold to filter the identified findings, in the sense that among all identified suggestions, those which are in accordance with or not contrary to the principles, are classified as emerging best practices or recommendations. The operational and organisational structure of LEAs and their hierarchical chain of command as well as the need for following a reporting framework, which satisfies their needs and which is compliant with their reporting framework, standards, templates, etc. have also been taken into consideration for the production of recommendations. Indicatively, these characteristics pose certain constraints and dictate several design characteristics as well as the development process. For example, any AI systems and related products with reporting capabilities need to be aligned with the aforementioned characteristics and must not violate the constraints posed and therefore, so must the associated recommendations do.

Taking into account the nature of the terms “recommendations” or “emerging best practices”, their purpose in the present deliverable is to illuminate the existing concerns regarding AI for Law Enforcement purposes, specify the obligations of technology developers and complement the legal framework by serving as a practical guide that will help technology developers create and provide to the LEAs ethically and legally compliant AI systems.

The methodology followed and enforced can be divided into three (3) main phases: a. Collect / Elicit existing data, b. Analyse new data / input, c. Produce new and/or Update and/or discard elements of the recommendations. Towards the direction of eliciting / collecting input, output and outcome, the following sources have been utilised to produce recommendations for and from technology developers:

- popAI deliverables
- literature, bibliography
- EU and CoE draft legislation, ethics guidelines
- popAI workshops
- popAI Policy Labs
- popAI crowdsourcing platform
- popAI Consortium meetings (e.g., plenary meetings)
- popAI, STARLIGHT, ALIGNER workshops and information related to the projects’ deliverables
- WP4 questionnaires
- SAB, EAB feedback

2.3 Process, Guidelines, Criteria, Constraints

The cycle of the production of recommendations for / from technology developers for the ethical use of AI from LEAs can be broken down into the following steps (assuming there is new information or input available, i.e., the cycle of production is not complete):

- ***Step1*** - Collect input and information to create entries of recommendations based on the sources mentioned under Sections 1.2 and 2.2.
- ***Step2*** - Analyse the collected input and information, and decide which ones classify as emerging best practices or recommendations. The recommendations are grouped based on

their thematic categorisation representing the main trends. Those which are in accordance or not contrary to the applicable ethics principles, and especially the ALTAI principles and the operational characteristics of the LEAs classify as emerging best practices/ recommendations.

- ***Step3*** - Produce new and / or update/modify or discard identified recommendations. A part of the procedure is to request that the EAB / SAB check the recommendations of Step 3, and depending on their feedback, go back to step 3 to update, correct, modify, or discard certain elements of the recommendations.

Evidently, if there is not any new input or information available, the cycle is considered complete and thus the whole process is terminated. The aforementioned cycle is presented in *Figure 4 below*.

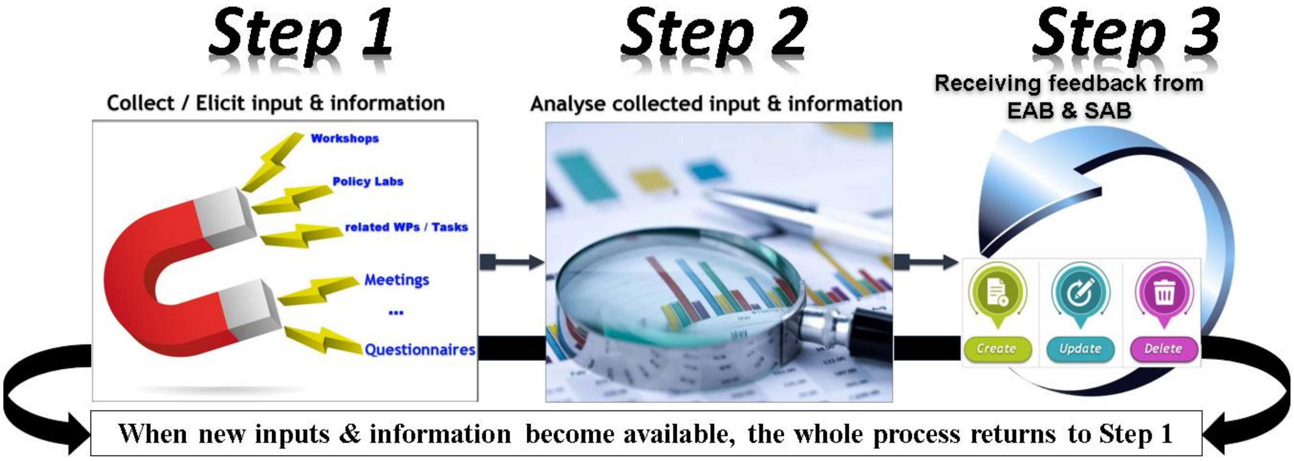

Step 1
Collect / Elicit input & information
Workshops
Policy Labs
related WPs / Tasks
Meetings
...
Questionnaires
Step 2
Analyse collected input & information
Step 3
Receiving feedback from EAB & SAB
Create
Update
Delete
When new inputs & information become available, the whole process returns to Step 1

Figure 4. Cycle of the production of recommendations

2.4 Sources of Input

In the course of the project execution, the following activities and actions which have been carried out served as the **main** sources for the present deliverable among those mentioned under Sections 1.2 and 2.2:

- ***PopAI literature review and research*** which has been conducted throughout the execution of the project in addition to WP2,[11] to identify templates for white papers[12] and to keep up with the latest developments regarding the AI Proposal as elaborated in D4.1. [13]

- ***PopAI stakeholder Policy Labs*** (SPLs)***:*** Within the context of "T3.4 - Engaging LEAs and relevant experts through policy labs", several PLs have been organized (see also D3.4),[14] where LEAs have been engaged, together with technology developers, legal experts,

[11] popAI D2.1 - Functionality taxonomy and emerging practices and trends", D2.5 - Practical ethics toolbox for the use of AI by LEAs"

[12] Brussels, 19.2.2020 COM(2020) 65 final WHITE PAPER On Artificial Intelligence - A European approach to excellence and trust available at : https://commission.europa.eu/system/files/2020-02/commission-white-paper-artificial-intelligence-feb2020_en.pdf

[13] popAI D4.1 – White Paper for LEAs under 3.1

[14] popAI D3.4 – "Stakeholder attitudes, priorities, and recommendations for addressing AI in the security domain in practice [M20], ECAS

policymakers and civil society representatives, from which valuable outputs have been extracted.

- ***PopAI Crowdsourcing platform:*** Based on "T3.3 – Crowdsourcing stakeholder attitudes and pro-active solutions ideation", the crowdsourcing methodology has been applied to achieve the citizens' active engagement in order to understand their perceptions on the use of AI by LEAs. Lawful, ethical social sensing (listening) has also been employed and analysis as well as post-processing of the information collected has illuminated different dimensions of social sensing. For more information on the data, see D3.3.
- **Multi-Disciplinary Foresight Scenarios:** The creation of FSs, for the co-creation of which, multidisciplinary groups have collaborated, has assisted the production of recommendations for technology developers as well as their validation/evaluation. As a methodology, FSs are increasingly used to support policy making and decision-making, in general. Although the development of future scenarios was pursued, the parameter "accuracy", based on the present data and information has been considered in this context.[15]
- ***PopAI Student Photo and Caption Competition:*** Within the context of "T3.6 - Engaging New Citizens through student photo and caption competition" (see also D3.6*)*, a competition was organised and managed by CERTH and disseminated through a campaign hosted in the project platform.[16] Universities were supported to administer an open call for students to reflect with a photograph and short narrative caption on the ethical issues related to different AI and policing controversies. The main aim was to reach new audiences, unfold the complex reflections on ethical concerns of AI policing data and the provision of a rich qualitative data source for understanding emergent and future concerns, which recommended guidelines around the use of AI use by LEAs. The results of this Task has also been taken into consideration and contributed to the production of recommendations within the context of this deliverable.

For a further and detailed description of the data regarding the participants in the above empirical activities, we refer the reader to popAI WP3 'Empirical Knowledge Collection and Management Framework'.

An additional step towards the formulation of the recommendations, especially in order to evaluate the identified ones, was to develop questionnaires based on D2.1 and D3.1 taxonomies and controversies, to distribute them to the task contributors (inside the Consortium) who would also send them to external technology developers, as described in Section 4.2 of the present deliverable. The same procedure was followed for D4.1, while the questionnaires were answered by the LEAs and policymakers of the Consortium and certain externals, as described therein. The above procedure is depicted in *Figure 5*, while a template of the questions for T4.3 is found in ANNEX B. D4.2 was covered by the dedicated to it D3.3 task with the respective questionnaires of the crowdsourcing platform.

[15] popAI D3.5 – "Foresight Scenarios for AI in Policing", TRI, especially p.13-15

[16] popAI D3.6 – "Photo Competition Results", CERTH

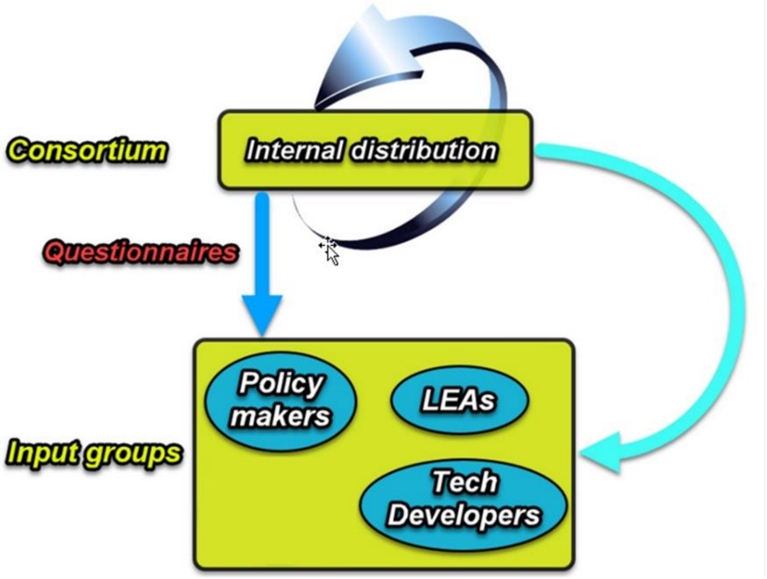


Figure 5. Input collection from the groups of interest

The appropriate information sheets and informed consent forms for T4.3 questionnaires, were drafted by KEMEA (see *Figure 6*, *Figure 7*, *Figure 8* and *Figure 9* in the *ANNEX*), for participants outside the Consortium , while personal data has been properly anonymised whenever necessary.

As a last step, the consultation with the EAB to guide us in the initial steps of drafting the recommendations, along with initial feedback requested by the SAB during the popAI plenary in Rome were taken into account. Some of the input collected includes highlighting the importance of the human oversight principle, that cases of fight against terrorism and criminal investigation could be the exceptions where the use of more intrusive AI tools may be justified, that the adoption of AI technologies may require effort, costs and time, especially for SMEs, and that LEAS could be more outward looking when it comes to the adoption of AI tools. The SAB mentioned that it would be important for LEAs, before using a new AI tool, to publish an open paper, sharing quantitative and qualitative data about the challenges that they may face.

The feedback of the SAB has been also requested for the present deliverable and pending; so, the overall SAB feedback is to be incorporated into the synthesis of D4.4.

In addition, for the purposes of the present deliverable, the opinion of the EAB Chair has been considered via their deliverable review and consultation throughout the project. As far as the number of additional experts who contributed to the evaluation of the recommendations are concerned, questionnaires (see section 4.2) were sent to technology developers within and outside of the popAI Consortium. Regarding the experts within the Consortium, it was completed by the Head of Software Development Department of Hellenic Police (1), an NGO officer of ECAS (1), a technology developer of TRI (1), a post-doc researcher and two technology developers of CERTH (3), one (1) machine-learning expert of NCSRD as well as an external data scientist of TRI London (1). Moreover, concerning the experts external to the Consortium, all members of the SAB gave their valuable input, and a total of eight (8) additional external experts from Ubitech (SME), Vicomtech (SME), Bavarian Police (LEA), CEA (RTO) and EUROPOL (LEA/Policy). The above constitutes an actively involved group of twenty-three (23) experts in addition to experts' opinions provided or extracted through popAI Policy Labs and workshops.

3 Recommendations to and from Technology Developers

It is possible to categorise the produced recommendations with respect to different types of categories. For instance, it is possible to categorise them according to the following types of categories:

- The functionality categorisation identified in D2.1, which is the top tier of the taxonomy, as mentioned therein, i.e., Recognition, Communication, Prediction & Analytics, Surveillance.
- The area of application in law enforcement, which constitutes the second tier according to D2.1,[17] such as crime prevention, crime investigation, cyber operations, migration, asylum, border control, LEA training, administration & Justice.
- Compliance with the principles for a trustworthy AI set by the EU, that is: Human agency and oversight, Technical Robustness and safety, Privacy and data governance, Transparency, Diversity, non-discrimination and fairness, Societal and environmental well-being, and Accountability.
- The actions associated with the production cycle and the processes of AI systems and related products, e.g., design, development, data processing.

Additionally, the activities of T4.3, associated with this deliverable are interlinked with the corresponding ones of T1.6 in that the principles, life cycle, and phase are linked to the AI policy ontology created as part of D1.6 Policy Briefs,[18] which have also been taken into consideration for the creation of recommendations.

For the purposes of the current deliverable the recommendations for technology developers are presented according to the production cycle of AI systems, starting with the recommendations at the stage of the design, followed with the stage of development and those suggested during data processing, as in the T4.3 description, and concluding with a set of horizontal or general recommendations.

3.1 Recommendations to / from technology developers concerning the design of AI systems, tools and related products

The set of recommendations generated are described in detail below:

[17] "D2.1 - Functionality taxonomy and emerging practices and trends", p. 12, NCSRD, Project title: A European Positive Sum Approach towards AI tools in support of Law Enforcement and safeguarding privacy and fundamental rights (popAI), Horizon 2020, Grant Agreement ID: 101022001

[18] "D1.6 – Policy briefs – 1st Year", NCSRD, Project title: A European Positive Sum Approach towards AI tools in support of Law Enforcement and safeguarding privacy and fundamental rights (popAI), Horizon 2020, Grant Agreement ID: 101022001

- ***Ensure both the problem definition as well as all the goals of the problem solving are clearly defined and detailed. (Purpose limitation).***

In short, the intended purpose(s) of the AI system, tool and related product need to be clear. The term “intended purpose” refers to the specific use for which an AI system is designed. This includes the context and conditions under which the system is intended to be used. The intended purpose should be clear to users in the instructions and documentation associated with the AI system, and it should provide guidance on its appropriate application and usage. Moreover, any relevant limitations and / or restrictions related to the application and / or usage of it need to be made available, too.

- ***AI systems for law enforcement need to be human-centred***

For the AI systems, tools, and related products to be human-centred, the LEA and citizen involvement is crucial as analysed per category:

- *Involvement of LEAs*: Human-centred design implies placing the needs, preferences, and experiences of the end-users -LEAs- at the forefront of the design process. It involves actively LEAs throughout the design stages, seeking their input and feedback, and iterating on designs based on their insights. Furthermore, starting from the design phase, the operational needs of the LEAs must be effectively captured and translated into technological ones and each one of the latter as well as any individual and collective functionality needs to be mapped to the ethical framework in place.
- *Citizen involvement*: Human-centred design for technologies in law enforcement need to take into consideration the perspective of citizens because it ensures that the solutions developed are not only effective from an operational standpoint but also respectful of the rights, values, needs and expectations of the individuals being served.

In general, to avoid any unexpected effects arising from the nature of the AI itself with respect to the ethical framework enforced, it is recommended that some checkpoints be in place to prevent any violations, unethical actions, unexpected behaviour, etc. In essence, these checkpoints can check and ensure that every step is compliant with, and does not violate any citizens’ rights, ethics, and legal frameworks, etc.

- ***The AI systems and related products and services need to be “ethical by design” to the maximum extent possible***

The “ethical use of AI for LEAs” needs to be ensured from the design phase. It is an ongoing and dynamic process, which must be enforced during all phases of development as well as upon deployment and actual usage, too. To build technologies that are ethical by design, the developer team needs to integrate ethical considerations and principles into the very foundation of a project or technology. This involves proactively addressing ethical concerns from the outset, embedding safeguards, and ensuring responsible practices throughout the entire development lifecycle. In this way, ethical considerations become an integral and proactive part of decision-making processes. Towards the same direction, the technology developers could consider increasing the adoption of a suitably modified "zero trust security model" (or as close as possible to it) for the design of AI systems, tools, and related products.[19]

The practical ethics toolbox developed within the context of “T2.4 - From ethical frameworks to ethics in practice”, also serves as a useful training tool. It can assist LEAs in getting familiarised with practical aspects associated with the ethical use of AI and it can also be useful to technology developers, as they can refer to it to design AI systems, tools and related products that can be easily understood by LEAs.[20]

➢ ***Risk management by design***

Planning the implementation and continuous update of a risk management system throughout the whole lifecycle of the AI systems, tools and related products is strongly recommended (which could include a risk assessment and/or an impact assessment). The risk management system should include the following components:

- The identification, estimation, and evaluation of the risks to health, safety, fundamental rights, and democracy
- The evaluation of the risks after the system is out in the market
- The outline of concrete and detailed mitigation measures
- The users’ training
- The testing of the technology
- The evaluation of the impact on the groups affected with a strong emphasis on vulnerable groups

The use and inclusion of control and monitoring mechanisms for AI systems, tools and related products needs to be appropriately planned from the design phase. If and whenever possible, the inclusion / placement of additional control and monitoring mechanisms in the systems (after agreeing on this with, and after informing the chain of command of LEAs, as needed) is recommended to further ensure the users / operators use the systems ethically.

3.2 Recommendations to / from technology developers concerning the development of AI systems, tools and related products

[19] He, Yuanhang, et al. "A survey on zero trust architecture: Challenges and future trends." Wireless Communications and Mobile Computing 2022 (2022)

[20] “D2.5 - Practical ethics toolbox for the use of AI by LEAs”, ERI, Project title: A European Positive Sum Approach towards AI tools in support of Law Enforcement and safeguarding privacy and fundamental rights (popAI), Horizon 2020, Grant Agreement ID: 101022001

- ***Every single process / task / module / functionality of AI systems, tools and related products needs to be developed and function in full respect of the law, ethics requirements and guidelines of the MSs and fully respect the EU applicable framework*.**

This recommendation concerns all types of technology developers and LEAs. The principle of the primacy (also referred to as "precedence" or "supremacy") of EU law must be always respected by all MSs unless any of the latter have negotiated "opt outs" or exemptions.[21] For this purpose, an interdisciplinary collaboration and open discussion are necessary for the software to be accepted, so technology developers / SMEs / academia need to communicate their intensions, plans and strategy to the legal partners, ethics experts, and regulatory bodies. It is also possible that MS-specific versions of an AI system/ tool or related technology (e.g. software) will be required to account for these differences per MS, unless it is not technically feasible. These details also need to be included in the relevant documentation.

- ***In order to adhere to the requirement of data fairness and inclusivity, and to avoid any under-representation of certain groups in society and / or any sort of polarisation, racism, etc., it may be desired to hardcode or force limits to be monitored during the evolution of the algorithm and / or in the results produced***

It is possible that the AI system, tool, and related product (to be) developed, could exhibit a behaviour with characteristics of polarisation, prejudice, and bias against certain groups in society, especially vulnerable groups, or minorities. For instance, the underlying algorithm may deduce that poor neighbourhoods are more likely to be involved in certain criminal activities, or that people with a certain colour on their skin may be more or less likely to do certain actions than others, etc. Moreover, some neighbourhoods, where there is increased (or reduced) policing may be considered to be more or less likely to experience higher criminality and this may be amplified. (e.g., the output of such an algorithm be fed back into the same algorithm as input). Specifically, if police forces are by default or preventively accumulated in poorer neighborhoods, it is likely that they record more incidents compared to neighborhoods where police forces are less, correlating thus higher criminality rates at poor neighborhoods. This could reinforce a "feedback loop", according to which data regarding poor neighborhoods would be reported and according to it more police forces would be sent to the field.

To avoid effects similar to these, it is recommended that some statistical limits (and / or parameter values, more generally) are hardcoded or input to the algorithm as thresholds which will be monitored during the evolution of the algorithm and checked against previously or latterly generated results. For instance, if the algorithm systematically correlates specific characteristics (e.g. protected grounds for discrimination) with incidents/ criminal activity, it should be checked for bias, which could then require the input of fair data.

The values of these parameters will be updated and enriched when the regulations / legislation is updated and enriched or changed, too. Frequent audits, checks, reviews, and reporting both on

[21] European Parliament, STUDY Requested by the JURI committee, The primacy of European Law, available at: [The primacy of European Union law (europa.eu)]

demand and scheduled as well on the output and results produced by the AI system, tool and related product are recommended as well.

The **involvement of potentially affected vulnerable groups** in the design phase is also recommended. This inclusive approach is expected to further ensure that the design process considers the unique needs, challenges, and experiences of these groups, resulting in more inclusive and effective solutions.

- ***For the implementation / use / testing of the AI-based systems sandboxes in protected environments / settings need to be developed***.

To avoid exposure of the AI-based systems under consideration sandboxed, secured, protected environments and settings need to be used and applied. Especially during the testing procedure, the use of such systems and settings is strongly recommended to decrease the risks entailed. The developers can then choose to progressively expose these systems to real-world conditions to approach the actual conditions in the operational environments of interest. Examples of such sandboxes could be testing innovative AI systems under national or European research programmes. The establishment of regulatory sandboxes, and specifically at least one national regulatory AI sandbox per MS, is also prescribed by the draft AI Act Proposal, so that development, testing and validation of innovative AI systems is conducted under oversight before these systems are put into the market or into service. [22]

- ***The TRL of the AI systems, tools and related products for LEAs must be nine (9)*** [23]

The AI systems, tools, and related products to be put in the market and used by LEAs need to attain a high level of maturity that makes these systems suitable, safe, secure, and stable enough for use in the intended operational environments of LEAs. To this end, there is no room for experimentation, especially considering the possibility that humans and / or the security, safety of the environment may be at stake, as a consequence.

- ***The entity developing the AI systems, tools and related products for LEAs needs to outline the specific limitations of these systems.***

Apart from detailing the capabilities, the functionalities, and the operation instructions, it is important for the entity developing the AI systems, tools, and related products for LEAs to detail the limitations of the systems. These need to be reported via the appropriate documentation. This documentation may not be officially available in order to ensure that no third party will be aware of

[22] 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))(1) ,Article 53

[23] Commission Decision C(2014)4995, HORIZON 2020 – WORK PROGRAMME 2014-2015 General Annexes Page 1 of 1 Extract from Part 19 - G. Technology readiness levels (TRL), https://ec.europa.eu/research/participants/data/ref/h2020/wp/2014_2015/annexes/h2020-wp1415-annex-g-trl_en.pdf

the limitations, vulnerabilities, shortcomings, bugs, risks, etc., but it needs to be available to the agreed points of contact and possibly to the end users and operators. The fields of application as well as certain limitations associated with their applicability need to be mentioned, too.

- ***All sorts of limitations, weaknesses, bugs, shortcomings, risks, etc. and requirements for installation, smooth operation and functionalities of AI systems, tools and related products need to be detailed and reported via the appropriate documentation***.

If applicable, any sort of limitation, weakness, bug, vulnerability, etc. needs to be outlined in the appropriate documentation. Given the AI system, tool and related product under consideration is directed to LEAs, these shortcomings should not be disclosed to any third party to avoid the possibility that these third parties can potentially take advantage of these shortcomings and attempt to exploit them. Furthermore, the AI systems, tools and related products will be installed on pre-existing infrastructure. Therefore, in addition to the basic installation instructions and system requirements, the (in)compatibility with other components needs to be mentioned. Moreover, after identifying the weakest / most vulnerable spots of the AI systems, tools, and related products (as far as the ethical use and security, etc. is concerned), the technology developers can group and outline them appropriately and share them with the right PoC(s) of LEAs only. For instance, if / when the AI system or any subsystem connects to a network, that network needs to be a trusted one and / or meet the necessary requirements, according to the security regulations, safety measures, etc. set in the manuals and relevant documentation.

- ***The results and reports need to be secured and locked (or even encrypted) to prevent anyone from altering / corrupting them***.

This recommendation aims at ensuring that the results and / or reports output by the AI systems, tools and related products will not be modified, either willingly or unwillingly. To this end, there exist a few measures that can be taken towards the direction of ensuring that the data will not be modified and towards the direction of holding accountable those who have modified the data. For instance, the data / results can be locked and secured by using a strong and complex enough password and / or strong encryption algorithm. Furthermore, the data / results can also be mined and / or certifiable / verifiable through appropriate means (e.g., Quick Response (QR) codes, use of security stamps, etc.) and the IDentification (ID) of the end user / operator and / or that of his / her terminal or similar (e.g., the Universally Unique IDentifier (UUID) of their device, timestamp, electronic signature, etc.).

- ***Anonymity and encouragement of operators / users, stakeholders, developers, designers, and all possibly involved, interested, affected parties and groups who wish to report anything associated with the (ethical) use of AI by LEAs, needs to be ensured and protected by law***.

All the involved parties need to be assured that their reports will stay anonymous and / or that they will not face any adverse consequences because they decided to report issues regarding the implementation of the appropriate steps, principles, procedures, or any unjustness or unfairness. Of course, any false accusations that are proved to be made in an attempt to cause harm and / or damage someone's reputation directly or indirectly, etc. need to be treated according to applicable legal procedures depending on the subject's interests.

- ***Personnel who are involved in the development of AI systems, tools and related products for LEAs need to possess security clearances and / or other certifications, and sign the appropriate NDAs, etc.***

Specifically, the personnel need to hold a security clearance, which is at least at the security level the AI systems, tools and related products are associated with. In cases where the personnel are occupied with a specific component or part of the system, then they need to hold a security clearance which is at least at the security level the functionality and purpose of the respective component or part is associated with. In cases where the AI systems, tools and related products are intended to be used transnationally or internationally (e.g., in EU and in the United States of America (USA)), the involved personnel may be required to hold security clearances accepted by all the corresponding continents, countries, states, etc., according to the respective regulations and legislation.

The personnel should also possess certifications confirming they possess the appropriate level of experience and knowledge to be capable of carrying out the development of the AI systems, tools and related products. These certifications need to concern the task of the development itself as well as the field of AI, which their development tasks encompass. It would also be useful for them to possess certifications ideally or at least proven knowledge and experience with the ethical framework and / or legislation their tasks concern (or are linked to).

Additionally, they may need to sign the appropriate NDAs and be informed about the legal consequences in case they violate any of the terms of the NDAs and to be accountable to the degree they are responsible.

The possible types of certifications, the training strategy as well as the relevant specifications fall beyond the scope of this deliverable.

3.3 Recommendations to / from technology developers as to the processing of data

- ***Whenever personal data of the data subjects are processed, the necessary information according to the applicable provisions shall be provided to the data subjects via -among others- technical means. Along with the information regarding the processing operations, and the data controllers and processors, information regarding the data sources, the providers and the algorithmic models used need to be officially disclosed to the data subjects.***

According to Article 13(1) of the Law Enforcement Directive, the following specific information that shall be made available to the data subjects as a minimum:[24]

- the identity and the contact details of the controller,
- the contact details of the data protection officer,
- the purposes of the processing,
- the right to lodge a complaint with a supervisory authority and the contact details of the supervisory authority,
- the right to request from the controller access to and rectification or erasure of personal data and restriction of processing of the personal data concerning the data subject;

In specific cases specified in Article 13(2) LED, the following information, should be additionally provided to the data subjects:

- the legal basis for the processing and the period for which the personal data will be stored, or the criteria used to determine that period,
- the categories of recipients of the personal data, including in third countries or international organisations further information shall be also provided.[25]

As LED provides the above as the minimum information required, it is suggested that in addition, information about the providers, the data sources and the algorithmic models used is provided to the data subjects when their personal data is processed via AI systems for law enforcement purposes.

The data subjects need to be informed about the above via appropriate means. For this purpose, the technology developers need to develop, set up, and incorporate the appropriate functionalities needed to ensure the process of the disclosure to the data subjects is performed appropriately, safely, securely, and according to the legislation in force and the LEAs' needs. Such information could be indicatively provided automatically to the data subjects through technical means. Even when there is a technical capability for the data subjects' to be informed automatically, they should be also able to exercise the rest of their data protection rights throughout technical means. In case data subjects are informed via technical means, it must be ensured that the least possible amount of the subjects' data for the purposes needed, are processed, by default and that the appropriate level of security measures is implemented to safeguard the rights and freedoms of the data subjects.

In cases where the national legislation additionally requires the agreement/consent of the data subjects, as in Recital 35 of the LED (DNA tests in criminal investigations or the monitoring of his or her location with electronic tags for the execution of criminal penalties), or as in Recital 37 of LED (processing of sensitive personal data in relation to fundamental rights and freedoms, e.g. revealing

[24] 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, Article 13

[25] 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, Article 13

racial or ethnic origin, that is particularly intrusive to the data subject), [26] it is further suggested that information sheets and consent forms are provided to the data subjects customised to the specific processing operation.

- ***Any sort of (potentially) sensitive data stored or transferred must be encrypted***

Depending on the application, the most appropriate algorithm can be chosen. For instance, AES-256 can be used in cases where a truly secure and trustworthy encryption is needed. This algorithm is also quantum resistant and is considered the encryption algorithm of choice for governments, financial institutions as well as security-conscious enterprises around the globe. Furthermore, the communications, which employ end-to-end encryption offer enhanced security, preventing third parties from accessing data while they are transferred from one end system or device to another. The type of encryption needs to be adequately strong, so as to deem the respective processes and the corresponding systems secure.

- ***Data anonymisation - pseudonymisation***

The choice between anomymisation, pseudonymisation, or use of personal data depends on the purpose and context of the processing. In case when data anonymisation is preferable but not feasible, there shall be a justification, explaining why it was not possible to ensure data anonymisation and how privacy for the data subjects is protected.[27] The exact methods and techniques applied towards achieving pseudonymisation shall be mentioned and the relevant risk assessment needs to be carried out and reported to the appropriate PoCs.

- ***Logging of all actions associated with logging into/out of the system, request / change of elevation rights and user roles / data from inside and outside the system, generation of results, reports, ID of operator / user and terminal (or PC or portable device or similar), date and time and / or duration of action (if applicable), crashes, (possible) security incidents, any (potential) anomaly, critical event, and generally any serious deviations from the normal operation is necessary, crucial and needs to be conducted continuously, saved and stored securely, protected, checked and monitored frequently***.

Any report generated must clearly state which results were produced automatically (i.e., without human intervention and / or supervision and / or grant of permission) and clearly mention during which stage, who (e.g., operator / user ID) and how they have intervened or granted their permission, etc.). Additionally, the technology developers can introduce / implement (additional) proactive steps to monitor the behaviour of the AI tools, periodically, on demand and when specific triggers are

[26] 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 Recitals (35) and (37)

[27] Moretón, Alvaro, and Ariadna Jaramillo. "Anonymisation and re-identification risk for voice data." *Eur. Data Prot. L. Rev.* 7 (2021): 274

present, they can send the appropriate warnings and / or alerts and / or reports to the appropriate PoCs, users, operators, etc., as required.

- ***Proper alerting, warnings, error codes, actions associated with privacy / personal data processing, sharing, transferring, etc. need to be included and satisfy several standards and possess / exhibit specific characteristics, depending on the application.***

It is recommended that the most important events are linked to the appropriate alerts, signals, warnings, error codes, etc. For instance, the more serious the event, the higher its visibility must be (e.g., the setup and incorporation of proper alerting, warnings, error codes, actions associated with privacy / personal data processing, etc.). Furthermore, a coupling of these alerts, warnings, etc. with visuals, such as colour codes is also recommended so that they are easier and simpler to distinguish. Furthermore, proper signalling/alerting and protection, measures, etc. against any attempt to bypass or deactivate any of the alerts or to disable colour codes is recommended. Alerting and reporting is recommended to be always directed to at least two distinct people (and there should be an immediate replacement of any one of them in case of absence).

- ***The inclusion of representatives and experts stemming from policymakers, LEAs, citizens, technology developers (e.g., from academia, SMEs, industry, etc.), legal experts, ethics experts, and relevant stakeholders needs to be demanded / guaranteed, starting from the design phase, and continued throughout the whole cycle from development to deployment, including during the processing of data***.

Thorough discussion and close collaboration between the technology developers and LEAs is necessary throughout the whole procedure of the development. Moreover, since the goal is for the whole society, LEAs, policymakers, and technology developers to benefit, it makes sense to require that the needs and demands of all parties be communicated clearly and openly. In a sense, this is also linked to ensuring equality / inclusion and representation of all groups and parties of interest. Furthermore, effective, efficient, and clear communication need to be ensured among developers to avoid loss of crucial information or miscommunication and to make sure everyone is aware of how others' tools / components / modules, etc. function.

3.4 Horizontal recommendations

The recommendations presented throughout this section are those that either do not fall within the scope of the categories or those that fall within the scope of all of them (in the sense that they can be considered as more general). It is also noticed that the following recommendations are mostly deployment-oriented, however, they require that the technical experts design and develop the AI systems in a manner according to which, they can be deployed as explained in the following paragraphs. Additionally, it is worth mentioning that the majority of them has emerged from and identified thanks to the outputs and outcomes of the related tasks and deliverables of WP2.[28]

- ***Human supervision needs to be ensured (i.e., with the aid of technical/technological means) during the whole lifecycle of the AI system (ideally) and the final decision must be made by humans***.

Humans need to supervise and monitor all processes, activities, data, and results. They also need to be the final decision-makers because they and / or the developers are accountable, and they and / or the developers will also be held responsible in case something goes wrong and /or if any damage is caused.

- ***System explainability regarding outcomes and recommendations needs to be ensured. It is important that the system can indicate the key parameters used, regarding a specific result / outcome***

Explainable AI (XAI) is highly recommended to be deployed. In general, it is desired that the end users / operators are aware of the values of the key parameters, so that they can be aware of how the system outputs specific results. Moreover, even if it is not easy to correlate the outputs / results with the specific inputs, this can be done or found later, as long as the values of the parameters of interest have been stored, together with the values of the associated outputs / results. Therefore, the need for logging and storing a number of important parameter values, in general is also recommended.

- ***The AI systems, tools and related products for LEAs need to possess reporting capabilities, in line with, and in full respect of the reporting procedures of the LEAs, the systems will be used by***.

Most of the users / operators (if not all of them) will need to report their findings to their chain of command, as needed. These reports need to be in line with the reporting procedures of the corresponding LEA. For instance, specific templates need to be used, the report needs to have an appropriate format, there may be specific stamps and / or signatures of some officers, users / operators, and there may also exist timestamps and / or security characterisations, codes, etc. Moreover, for the findings to be of any use in the court of law, they also need to be presented through the appropriate reports so as to ensure their validity, originality, and lawfulness.

- ***The AI systems and related software may need to offer the capability for the grant of permission from the right number and rank of persons needed (e.g., at least four eyes rule) for certain processes / procedures, etc.***

[28] D2.1 - Functionality taxonomy and emerging practices and trends, D2.5 - Practical ethics toolbox for the use of AI by LEAs

This must be compliant and in line with the procedures, operational requirements, and regulations enforced in the LEAs' operational environments.

- ***A list of at least the most essential qualifications an operator needs, in order to use AI system, tools and related software, products, modules, etc. should be drafted, preferably with reference to the respective training modules, certification(s) (if/where applicable), including the period of validity (i.e., expiration) per certification***:

The formulation of a list of the necessary on the one hand, and desirable qualifications, on the other, including certifications, education, trainings, curricula (e.g., based on CEPOL methodology, learning / training strategy) [29] can be of great value. Such a list could also be incorporated into the operations manual of an AI tool by the providers and made available to the LEAs. Lessons learnt and the emergence of best practices are also anticipated within the same context.

- ***The AI systems and related software to be used by LEAs need to offer the capability to LEAs to stop / pause them timely and review / audit the parts of the process of interest***.

This could offer LEAs the opportunity to check if something is wrong, while the AI systems and related software is executed or even to confirm that everything functions as intended. It could also allow them to confine a problem before expanding any further. For example, the LEAs may suspect the results are prejudiced against a specific race or a certain minority, or that the system has proceeded without asking for a confirmation, although it was needed or it has not asked for a subject's concern, etc.

- ***Recovery / remediation / issue-problem handling steps need to be clearly mentioned, detailed and easy to follow as much as possible from the operator / user***.

These actions need to be documented in a detailed procedure, preferably in a step-by-step manner. The documents containing them need to be kept at specific, secure places and made available to the appropriate personnel. The chain of hierarchy needs to be informed with a specific order, incident handling reports need to be updated accordingly, together with the disaster recovery plan, the business continuity plan, and the relevant procedures.

- ***Within the context of accountability, the provider of an AI system to LEAs needs to give the persons' required details (e.g., contact details) as PoCs, together with the technological field they are responsible in (e.g., data processing, backend, etc.), so that the LEAs can contact / consult them if necessary***.

It is argued that it is not enough to provide general contact details, as the AI systems used by LEAs may potentially expose citizens' personal data and / or violate their rights. Therefore, more contact details may be needed (i.e., from more people) for each one of the most important functionalities

[29] CEPOL, European Union Agency for Law Enforcement Training, https://www.cepol.europa.eu/

present and at least one (1) additional contact per functionality to account for the main responsible contact’s absence.

- ***The technical support of AI systems needs to be made necessary and considered as an integral part of these systems and thus formally written in the appropriate documentation and mutually agreed upon among the entity / entities and the LEAs or their appointed legal representatives***.

According to the risks present, based on the type and functionalities of the AI system under consideration, the technical, IT, ICT support, etc. need to be ensured and agreed upon officially, e.g., by signing the appropriate agreements. The reason is that the AI systems cannot be left without maintenance and support, as the possibility exists that they may start malfunctioning.

- ***Frequent support and checks of the infrastructure and equipment - and minimum necessary requirements ensuring safe functioning and operation, etc.***

This extends beyond the most directly associated recommendations, but it is still required, as security breaches in the digital infrastructure could also have adverse effects on the ethical use of AI by LEAs. Therefore, there need to be frequent audit reports with signatures of the personnel (e.g., the appropriate personnel may need to check if certain functionalities and / or equipment perform as expected). So, this recommendation is associated with ensuring the ethical use of AI by LEAs from a wider perspective. It also considers the coupling of safety and security to ensure the ethical use of AI by LEAs and can include the incorporation of self-checks and automations in these systems.

- ***Frequent, scheduled as well as on-demand, automated and / or manual backups, relevant capabilities and recovery steps need to be in place and detailed in the appropriate documentation.***

For instance, some attacks or specific security issues may disrupt the functionalities and allow malicious actors to steal (sensitive) data or could result in LEAs losing control of the operation of AI systems and related products. Therefore, the use of backups (both automated and manual ones) frequently is recommended.

- ***The AI system and related software (to be) used by LEAs needs to be certified with respect to their technical characteristics in relation to security, safety, ethics, compliance with law of the target MS and EU.***

Given the AI system and related software (to be) used by LEAs includes functionalities that could potentially lead to the violation of ethics frameworks in place and / or law or citizens’ personal rights, this software needs to be certified with respect to the functionalities related to these potential risks. These certifications need to come from reliable entities at least mutually agreed upon, and accepted from the entities involved in the development, installation, deployment, and use of the AI systems of interest or according to the regulations and rules, if applicable.

- ***The AI systems, software and associated components, parts, etc. need to be reviewed, audited by trusted third parties and updated frequently within deadlines, agreed upon and stated clearly and formally in all necessary documentation***.

The AI systems need to be reviewed periodically and updated frequently to prevent them from becoming outdated and / or (potentially) insecure and / or unstable. For this purpose, the developers are recommended to introduce alerts, which will be categorised with respect to the severity and urgency of the associated update, which is required to be carried out. The most significant alerts can also be represented with a specific, distinguishable colour, the size of the associated message should be bigger as compared to a less important / urgent alert and appear on a location of the screen that will be detected by the user / operator as fast and as easily as possible.

Furthermore, AI systems need to be monitored, to undergo frequent reviews, checks, testing, and audits to ensure they function as expected. Scheduled and not scheduled audits and checks can also be carried out by trusted third parties (i.e., externals). The findings of the audits need to be reported together with the auditors' recommendations and ratings.

- ***Whenever possible, the civil society needs to be informed about the introduction and / or use of an AI system through official channels (at least two distinct channels) or at least two independent bodies, officially representing society***.

Technology developers need to work closely with the LEAs to develop, set up and, incorporate the appropriate functionalities so as to inform properly and timely the (potentially) affected people. For instance, there are occasions where some citizens' data may be used. In these occasions, the citizens need to be informed appropriately. The requirement for the existence of at least two (2) distinct official channels or at least two (2) independent bodies is associated with the enforcement of multivocality and the relevant virtues of democracy. For the same purpose, public versions of the associated documentation need to be available to the civil society.

4 Evaluation of Recommendations

Aiming at the production of recommendations within the context of this deliverable as well as results, certain steps have been followed to pursue the establishment of valid outputs and outcomes. These are desired to be useful, practical, and reproducible by the target groups as well as civil society. Therefore, towards the direction of (cross-)validating/evaluating the produced results and recommendations, the following practices have been adopted:

4.1 Synergies with sibling projects

As a further step towards the evaluation, synergies with ALIGNER, STARLIGHT and through STARLIGHT with AP4AI were established.[30] The collaboration with these projects has been proved beneficial for the production of recommendations as well as for popAI overall. Experts and stakeholders from these projects have been invited and participated in workshops organised within the context of popAI, and through their involvement, the opportunity to exchange findings with regards to recommendations was identified. One of the ALIGNER policy recommendations which relates to the current deliverable is to ensure that LEAs always have knowledgeable and competent human-in-the-loop utilising AI tools assisting them in decision making, while the collaboration among ethics and legal experts, technical experts as part of the regulatory sandboxes was suggested in the Interim Policy Recommendations of STARLIGHT.[31] For further information regarding the participants to the popAI Workshops , we refer the reader to the popAI WP5 'Dissemination, Communications and Sustainable Community Engagement' and as regards the joint activities, to the dissemination WPs of ALIGNER and STARLIGHT projects as well.

For more information on the issue, please also check popAI D4.4.

4.2 T4.3 Questionnaires

Questionnaires have been distributed to the partners of the consortium according to the process visualised in *Figure 5* above, along with the consent forms and information sheets for externals to the Consortium (see ***ANNEX A***)], which are considered experts in the topics of interest and their contribution has been taken into consideration, too, as far as the production of recommendations is concerned. A copy of the questionnaires developed and distributed within the context of T4.3 can be found in ***ANNEX B*** – Questionnaires of T4.3(see *Figure 10* through *Figure 18).* To sum up the answers to the questions in the questionnaire, 30% of the participants work in SMEs, 25% in RTOs – academia, 20% in NGOs and the rest 25% in "Other - Technology Development". Furthermore, 40% answered there is legislation enforced in their country, concerning the design and development of AI systems and related products for LEAs, whereas 60% answered "No". Moreover, 80% of them have at least one (1) specialised department in their entity that has the expertise to deal with AI-related issues, while the rest 20% do not have any. In addition, 70% of them are currently developing / designing AI-enabled technologies or tools for LEAs, and the rest 30 % are not. Interestingly, all the participants highlighted the need for continuous (re-)training of LEAs on AI and the need for informing society and raising their awareness. All the detailed answers (i.e., the ones answering the questions, which needed longer responses) have been taken into consideration and incorporated into the produced recommendations, where applicable and appropriate.

[30] H2020-SU-AI-2020, SU-AI01-2020 - Developing a research roadmap regarding Artificial Intelligence in support of Law Enforcement, Project title: An AI roadmap for law enforcement agencies (ALIGNER), Grant Agreement ID: 101020574, URL: https://cordis.europa.eu/project/id/101020574, last accessed online via web browser on 2/5/2023 ; H2020-SU-AI-2020, SU-AI02-2020 - Secure and resilient Artificial Intelligence technologies, tools and solutions in support of Law Enforcement and citizen protection, cybersecurity operations and prevention and protection against adversarial Artificial Intelligence, Project title: Sustainable Autonomy and Resilience for LEAs using AI against High priority Threats (STARLIGHT), URL: https://cordis.europa.eu/project/id/101021797, last accessed online via web browser on 2/5/2023; https://www.ap4ai.eu/, last accessed online via web browser on 2/5/2023.

[31] 5th ALIGNER Public Workshop, June 2023

4.3 Evaluation by experts

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

Finally, a dedicated report will be drafted by the External Ethics Advisor of the popAI project chairing the EAB which will include her opinion on the recommendations provided in the present deliverable as well as in D4.1 “White Paper for LEAs” and D4.2 “White Paper for Civil Society”.

The SAB will provide their feedback to be incorporated into popAI D4.4, as below.

4.4 Evaluation through the last deliverable of WP4

The last deliverable of WP4 is D4.4 “Synthesis: a collection of the best multidisciplinary practices” for the ethical use of AI by LEAs and will collect and examine the recommendations included in the present deliverable D4.1, D4.2 and D4.3 and, ultimately, evaluate them in order to present them as the best multidisciplinary practices emerging from interdependent and collaborative work of people with different specialties and experiences. Within D4.4, the feedback of the EAB and the SAB will be incorporated.

To ensure the validity of our results we have checked the reliability of our sources of information and data. Moreover, our research includes checks as to the following types of validity: face validity, content validity, internal validity, external validity, statistical conclusion validity and criterion-related validity with respect to credibility, authenticity, criticality, and integrity.[32]

5 Discussion

This section presents some of the challenges faced, possible extensions of this work, and what the future holds as far as the (ethical) use of AI is concerned.

Firstly, one of the difficulties sometimes faced was the inherently contradictory groups and their interests. Sometimes, the LEAs kept expressing the need for wider and more extensive access to data, while citizens kept asking for less access to their data and for better control and monitoring of the authorities that would like to use them, stricter, more complex, and secure protocols, and continuous training of LEAs. The LEAs also mentioned that the adoption of stricter, more complex protocols and measures and the need for their continuous training could deem the AI systems and related products only accessible to a few LEAs who have received training. Moreover, LEAs have argued the possibility that they may end up trying to fight against truly advanced technology used by criminals using old and ineffective means.

[32] Chase, S., C. Mandle, and R. Whittemore. "Validity in qualitative research." Qualitative Health Research 11.4 (2001): 522-537; Creswell, John W., and Cheryl N. Poth. Qualitative inquiry and research design: Choosing among five approaches. Sage publications, 2016.

6 Conclusion

The rapid and increasing involvement of AI in the daily tasks and operations of LEAs is currently starting to gain attention. Therefore, the need for LEAs to be informed, trained, organisationally prepared prior to the use of AI tools to be ethically and legally compliant and trusted by citizens is of vital importance. In this deliverable, the aim and scope, the strategy and the methodology followed have been analysed as well as the sources of information, the involved groups, and the approach to elicit(ate) and analyse the inputs and information needed. The results are the recommendations to / from technology developers for the ethical use of AI by LEAs. They were indicatively categorised as: recommendations regarding the stage of design of AI systems, the development of AI systems, the processing of data and horizontal recommendations for technology developers regarding the ethical use of AI by LEAs, which are of a more general nature. The present recommendations along with the related outputs of Task 4.1 (Recommendations for and from LEAs and Policymakers as presented in D4.1) and Task 4.2 (Recommendations for and from the Civil Society as presented in D4.2), will form a set of multidisciplinary best practices that will be presented in D4.4 “Synthesis: a collection of best multidisciplinary practices” and feedback by the EAB and SAB members will be incorporated.

7 References

AI HLEG, Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self-assessment, URL: https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment, last accessed online via web browser on 24/7/2023

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))(1) : https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EN.html last accessed online via web browser on 09/08/2023

Andreas Liebl and Till Klein, AI Act: Risk Classification of AI Systems from a Practical Perspective, applied AI, URL: https://aai.frb.io/assets/files/AI-Act-Risk-Classification-Study-appliedAI-March-2023.pdf, last accessed online on 2/5/2023

AP4AI https://www.ap4ai.eu/ , last accessed online via web browser on 2/5/2023.

Brussels, 19.2.2020 COM(2020) 65 final WHITE PAPER On Artificial Intelligence - A European approach to excellence and trust available at : https://commission.europa.eu/system/files/2020-02/commission-white-paper-artificial-intelligence-feb2020_en.pdf

CEPOL, European Union Agency for Law Enforcement Training, https://www.cepol.europa.eu/

Chase, S., C. Mandle, and R. Whittemore. "Validity in qualitative research." Qualitative Health Research 11.4 (2001): 522-537

Creswell, John W., and Cheryl N. Poth. Qualitative inquiry and research design: Choosing among five approaches. Sage publications, 2016.

Commission Decision C(2014)4995, HORIZON 2020 – WORK PROGRAMME 2014-2015 General Annexes Page 1 of 1 Extract from Part 19 - G. Technology readiness levels (TRL), https://ec.europa.eu/research/participants/data/ref/h2020/wp/2014_2015/annexes/h2020-wp1415-annex-g-trl_en.pdf

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

Ethics for researchers, Facilitating Research Excellence in FP7, p.4, European Commission, URL:https://ec.europa.eu/research/participants/data/ref/fp7/89888/ethics-for-researchers_en.pdf, last accessed online via web browser on 2/5/2023

European Parliament, STUDY Requested by the JURI committee, The primacy of European Law, available at: The primacy of European Union law (europa.eu)

H2020-SU-AI-2020, SU-AI01-2020 - Developing a research roadmap regarding Artificial Intelligence in support of Law Enforcement, Project title: An AI roadmap for law enforcement agencies (ALIGNER), Grant Agreement ID: 101020574, URL: https://cordis.europa.eu/project/id/101020574, last accessed online via web browser on 2/5/2023

H2020-SU-AI-2020, SU-AI02-2020 - Secure and resilient Artificial Intelligence technologies, tools and solutions in support of Law Enforcement and citizen protection, cybersecurity operations and prevention and protection against adversarial Artificial Intelligence, Project title: Sustainable Autonomy and Resilience for LEAs using AI against High priority Threats (STARLIGHT), URL: https://cordis.europa.eu/project/id/101021797, last accessed online via web browser on 2/5/2023
More, Trenchard, On the construction of Venn diagrams, *The Journal of Symbolic Logic* 24.4 (1959): 303-304

He, Yuanhang, et al. "A survey on zero trust architecture: Challenges and future trends." Wireless Communications and Mobile Computing 2022 (2022)

Moretón, Alvaro, and Ariadna Jaramillo. "Anonymisation and re-identification risk for voice data." *Eur. Data Prot. L. Rev.* 7 (2021): 274

News, European Parliament, EU AI Act: first regulation on artificial intelligence, https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence, last accessed online via web browser on 23/07/2023

popAI D1.6 – Policy briefs – 1st Year

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

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

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

popAI D3.5 – Foresight Scenarios for AI in Policing

popAI D3.6 – Photo Competition Results

popAI D4.1 – White Paper for LEAs

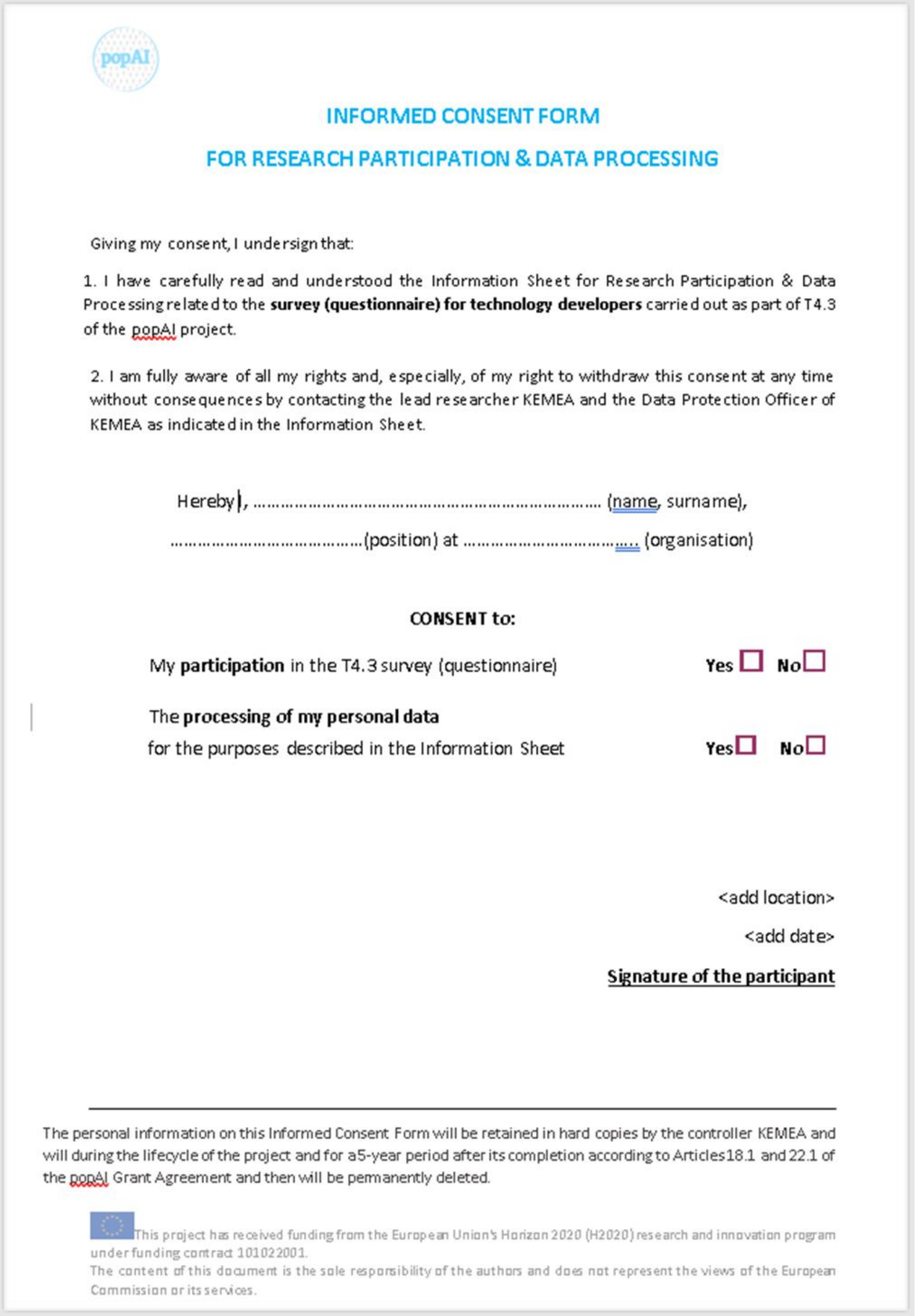
Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS, p. 27, European Commission, URL: https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0001.02/DOC_1&format=PDF, last accessed online on 2/5/2023

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

ANNEXES

ANNEX A – Informed Consent Form & Information Sheet

Informed consent form for research participation & data processing within the context of answering the questionnaires for this deliverable (see *Figure 6*)



INFORMED CONSENT FORM

FOR RESEARCH PARTICIPATION & DATA PROCESSING

Giving my consent, I undersign that:

1. I have carefully read and understood the Information Sheet for Research Participation & Data Processing related to the **survey (questionnaire) for technology developers** carried out as part of T4.3 of the popAI project.

2. I am fully aware of all my rights and, especially, of my right to withdraw this consent at any time without consequences by contacting the lead researcher KEMEA and the Data Protection Officer of KEMEA as indicated in the Information Sheet.

Hereby I, ……………………………………………………………………… (name, surname),

…………………………………….(position) at ……………………………….... (organisation)

CONSENT to:

My **participation** in the T4.3 survey (questionnaire)	**Yes** □	**No** □
The **processing of my personal data** for the purposes described in the Information Sheet	**Yes** □	**No** □

<add location>

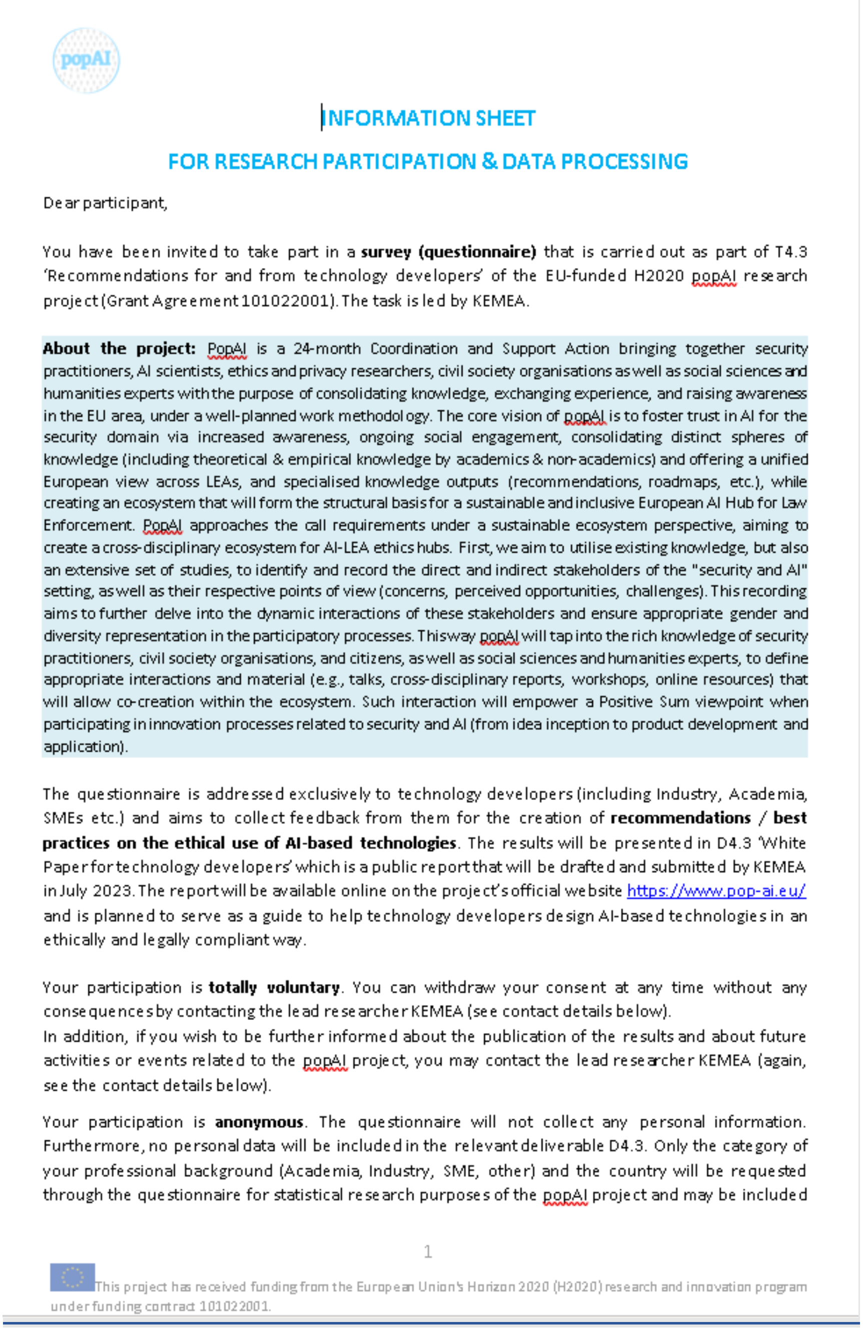
<add date>

Signature of the participant

The personal information on this Informed Consent Form will be retained in hard copies by the controller KEMEA and will during the lifecycle of the project and for a5-year period after its completion according to Articles18.1 and 22.1 of the popAI Grant Agreement and then will be pemanently deleted.

This project has received funding from the European Union's Horizon 2020 (H2020) research and innovation program under funding contract 101022001.
The content of this document is the sole responsibility of the authors and does not represent the views of the European Commission or its services.

Figure 6. Sample informed consent form - questionnaires for this deliverable



INFORMATION SHEET

FOR RESEARCH PARTICIPATION & DATA PROCESSING

Dear participant,

You have been invited to take part in a **survey (questionnaire)** that is carried out as part of T4.3 'Recommendations for and from technology developers' of the EU-funded H2020 popAI research project (Grant Agreement 101022001). The task is led by KEMEA.

About the project: PopAI is a 24-month Coordination and Support Action bringing together security practitioners, AI scientists, ethics and privacy researchers, civil society organisations as well as social sciences and humanities experts with the purpose of consolidating knowledge, exchanging experience, and raising awareness in the EU area, under a well-planned work methodology. The core vision of popAI is to foster trust in AI for the security domain via increased awareness, ongoing social engagement, consolidating distinct spheres of knowledge (including theoretical & empirical knowledge by academics & non-academics) and offering a unified European view across LEAs, and specialised knowledge outputs (recommendations, roadmaps, etc.), while creating an ecosystem that will form the structural basis for a sustainable and inclusive European AI Hub for Law Enforcement. PopAI approaches the call requirements under a sustainable ecosystem perspective, aiming to create a cross-disciplinary ecosystem for AI-LEA ethics hubs. First, we aim to utilise existing knowledge, but also an extensive set of studies, to identify and record the direct and indirect stakeholders of the "security and AI" setting, as well as their respective points of view (concerns, perceived opportunities, challenges). This recording aims to further delve into the dynamic interactions of these stakeholders and ensure appropriate gender and diversity representation in the participatory processes. This way popAI will tap into the rich knowledge of security practitioners, civil society organisations, and citizens, as well as social sciences and humanities experts, to define appropriate interactions and material (e.g., talks, cross-disciplinary reports, workshops, online resources) that will allow co-creation within the ecosystem. Such interaction will empower a Positive Sum viewpoint when participating in innovation processes related to security and AI (from idea inception to product development and application).

The questionnaire is addressed exclusively to technology developers (including Industry, Academia, SMEs etc.) and aims to collect feedback from them for the creation of **recommendations / best practices on the ethical use of AI-based technologies**. The results will be presented in D4.3 'White Paper for technology developers' which is a public report that will be drafted and submitted by KEMEA in July 2023. The report will be available online on the project's official website https://www.pop-ai.eu/ and is planned to serve as a guide to help technology developers design AI-based technologies in an ethically and legally compliant way.

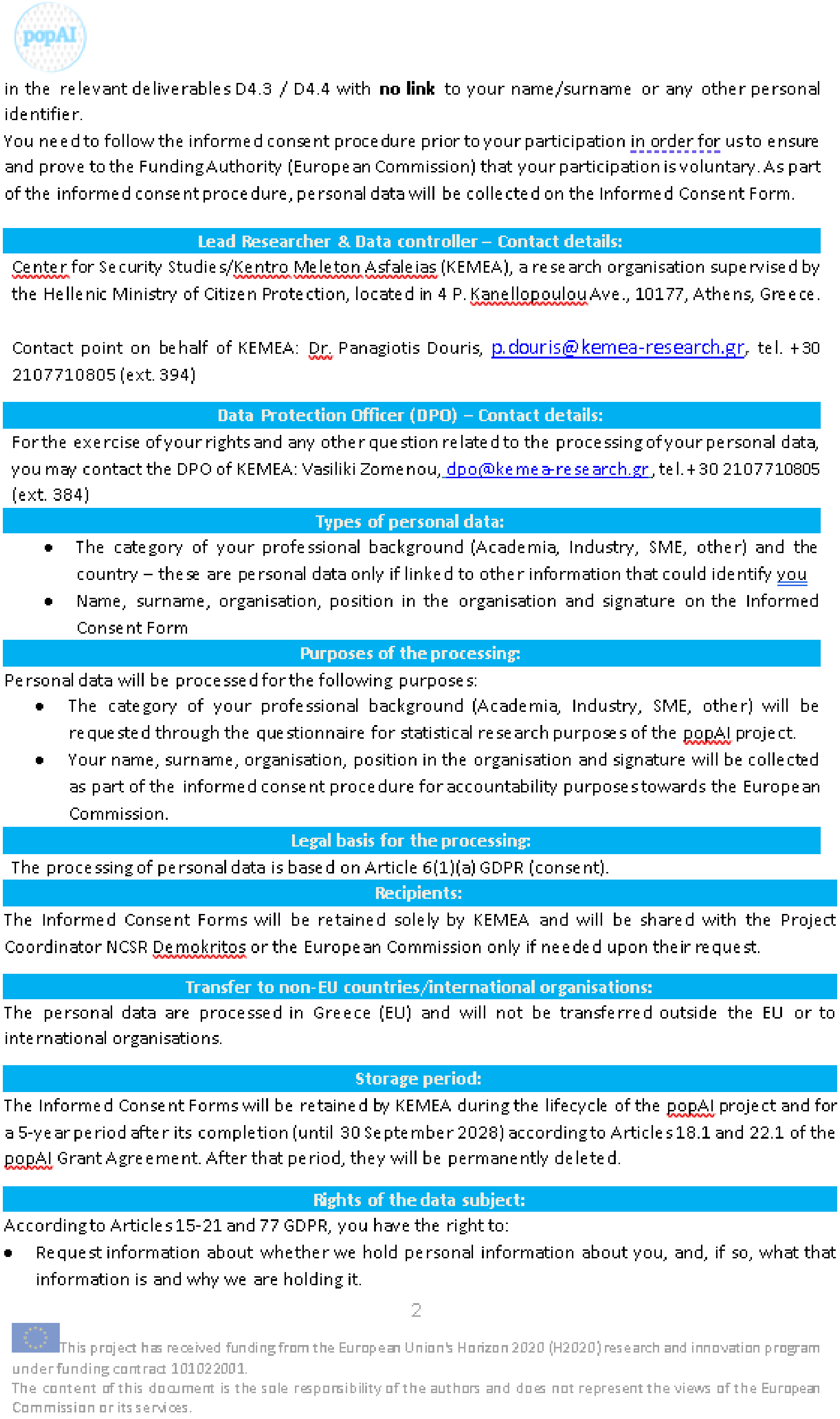
Your participation is **totally voluntary**. You can withdraw your consent at any time without any consequences by contacting the lead researcher KEMEA (see contact details below).
In addition, if you wish to be further informed about the publication of the results and about future activities or events related to the popAI project, you may contact the lead researcher KEMEA (again, see the contact details below).

Your participation is **anonymous**. The questionnaire will not collect any personal information. Furthermore, no personal data will be included in the relevant deliverable D4.3. Only the category of your professional background (Academia, Industry, SME, other) and the country will be requested through the questionnaire for statistical research purposes of the popAI project and may be included

1

This project has received funding from the European Union's Horizon 2020 (H2020) research and innovation program under funding contract 101022001.

Figure 7. Sample information sheet for research participation & data processing – questionnaire for this deliverable (1/3)



in the relevant deliverables D4.3 / D4.4 with **no link** to your name/surname or any other personal identifier.

You need to follow the informed consent procedure prior to your participation in order for us to ensure and prove to the Funding Authority (European Commission) that your participation is voluntary. As part of the informed consent procedure, personal data will be collected on the Informed Consent Form.

Lead Researcher & Data controller – Contact details:

Center for Security Studies/Kentro Meleton Asfaleias (KEMEA), a research organisation supervised by the Hellenic Ministry of Citizen Protection, located in 4 P. Kanellopoulou Ave., 10177, Athens, Greece.

Contact point on behalf of KEMEA: Dr. Panagiotis Douris, p.douris@kemea-research.gr, tel. +30 2107710805 (ext. 394)

Data Protection Officer (DPO) – Contact details:

For the exercise of your rights and any other question related to the processing of your personal data, you may contact the DPO of KEMEA: Vasiliki Zomenou, dpo@kemea-research.gr, tel. +30 2107710805 (ext. 384)

Types of personal data:

- The category of your professional background (Academia, Industry, SME, other) and the country – these are personal data only if linked to other information that could identify you
- Name, surname, organisation, position in the organisation and signature on the Informed Consent Form

Purposes of the processing:

Personal data will be processed for the following purposes:

- The category of your professional background (Academia, Industry, SME, other) will be requested through the questionnaire for statistical research purposes of the popAI project.
- Your name, surname, organisation, position in the organisation and signature will be collected as part of the informed consent procedure for accountability purposes towards the European Commission.

Legal basis for the processing:

The processing of personal data is based on Article 6(1)(a) GDPR (consent).

Recipients:

The Informed Consent Forms will be retained solely by KEMEA and will be shared with the Project Coordinator NCSR Demokritos or the European Commission only if needed upon their request.

Transfer to non-EU countries/international organisations:

The personal data are processed in Greece (EU) and will not be transferred outside the EU or to international organisations.

Storage period:

The Informed Consent Forms will be retained by KEMEA during the lifecycle of the popAI project and for a 5-year period after its completion (until 30 September 2028) according to Articles 18.1 and 22.1 of the popAI Grant Agreement. After that period, they will be permanently deleted.

Rights of the data subject:

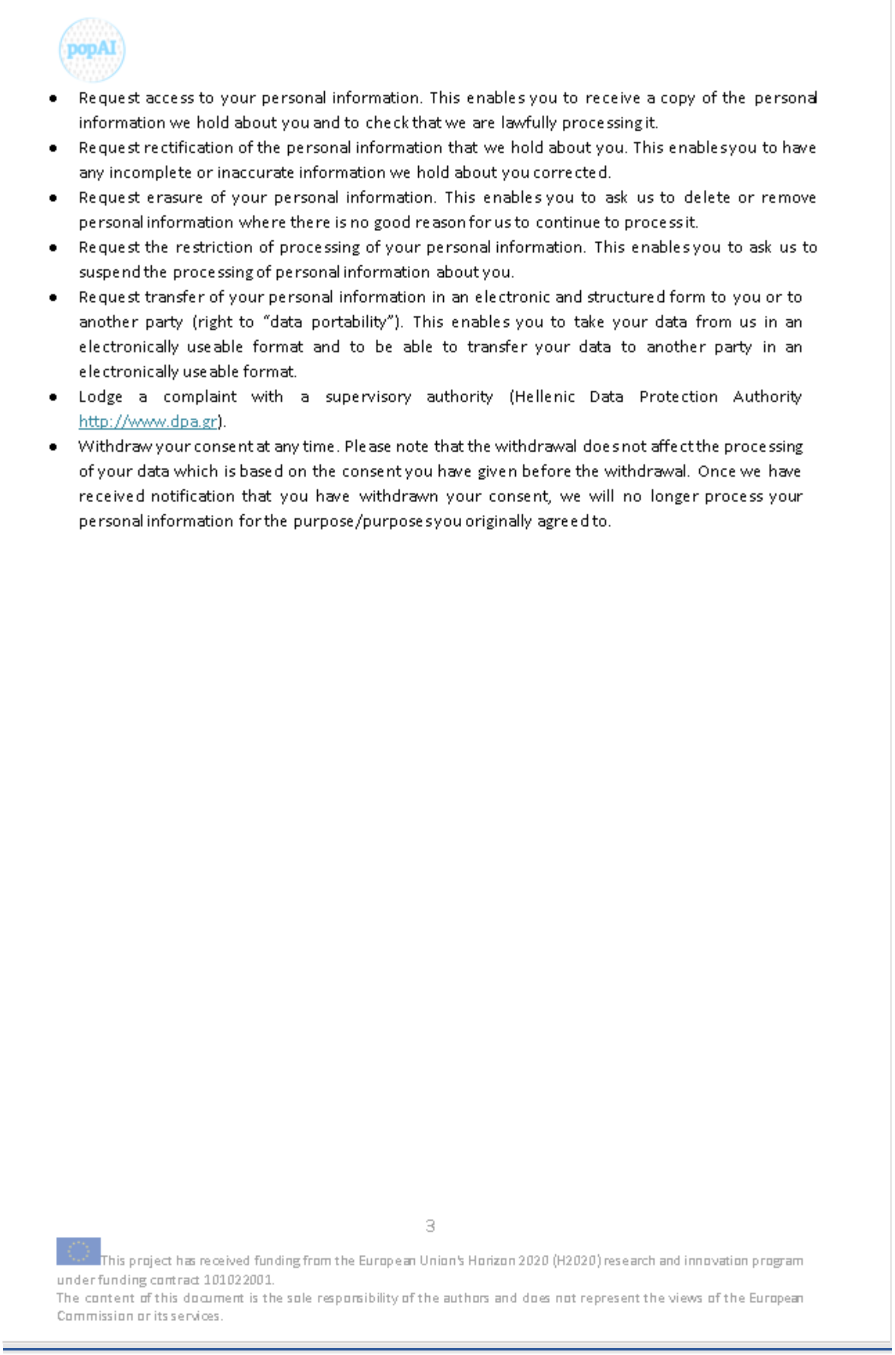
According to Articles 15-21 and 77 GDPR, you have the right to:

- Request information about whether we hold personal information about you, and, if so, what that information is and why we are holding it.

2

This project has received funding from the European Union's Horizon 2020 (H2020) research and innovation program under funding contract 101022001.
The content of this document is the sole responsibility of the authors and does not represent the views of the European Commission or its services.

Figure 8. Sample information sheet for research participation & data processing - questionnaire for this deliverable (2/3)

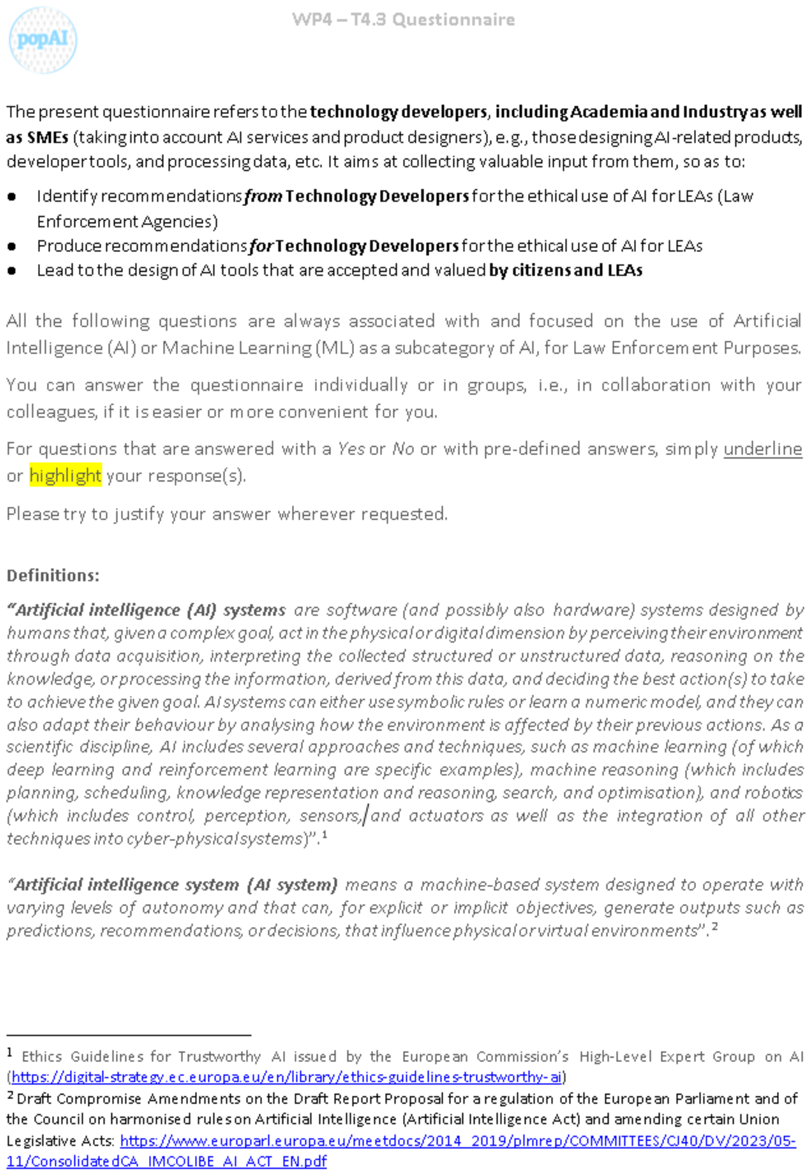


- Request access to your personal information. This enables you to receive a copy of the personal information we hold about you and to check that we are lawfully processing it.
- Request rectification of the personal information that we hold about you. This enables you to have any incomplete or inaccurate information we hold about you corrected.
- Request erasure of your personal information. This enables you to ask us to delete or remove personal information where there is no good reason for us to continue to process it.
- Request the restriction of processing of your personal information. This enables you to ask us to suspend the processing of personal information about you.
- Request transfer of your personal information in an electronic and structured form to you or to another party (right to "data portability"). This enables you to take your data from us in an electronically useable format and to be able to transfer your data to another party in an electronically useable format.
- Lodge a complaint with a supervisory authority (Hellenic Data Protection Authority http://www.dpa.gr).
- Withdraw your consent at any time. Please note that the withdrawal does not affect the processing of your data which is based on the consent you have given before the withdrawal. Once we have received notification that you have withdrawn your consent, we will no longer process your personal information for the purpose/purposes you originally agreed to.

This project has received funding from the European Union's Horizon 2020 (H2020) research and innovation program under funding contract 101022001.
The content of this document is the sole responsibility of the authors and does not represent the views of the European Commission or its services.

Figure 9. Sample information sheet for research participation & data processing - questionnaire for this deliverable (3/3)

ANNEX B – Questionnaires of T4.3



WP4 – T4.3 Questionnaire

The present questionnaire refers to the **technology developers, including Academia and Industry as well as SMEs** (taking into account AI services and product designers), e.g., those designing AI-related products, developer tools, and processing data, etc. It aims at collecting valuable input from them, so as to:

- Identify recommendations ***from*** **Technology Developers** for the ethical use of AI for LEAs (Law Enforcement Agencies)
- Produce recommendations ***for*** **Technology Developers** for the ethical use of AI for LEAs
- Lead to the design of AI tools that are accepted and valued **by citizens and LEAs**

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

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

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

Please try to justify your answer wherever requested.

Definitions:

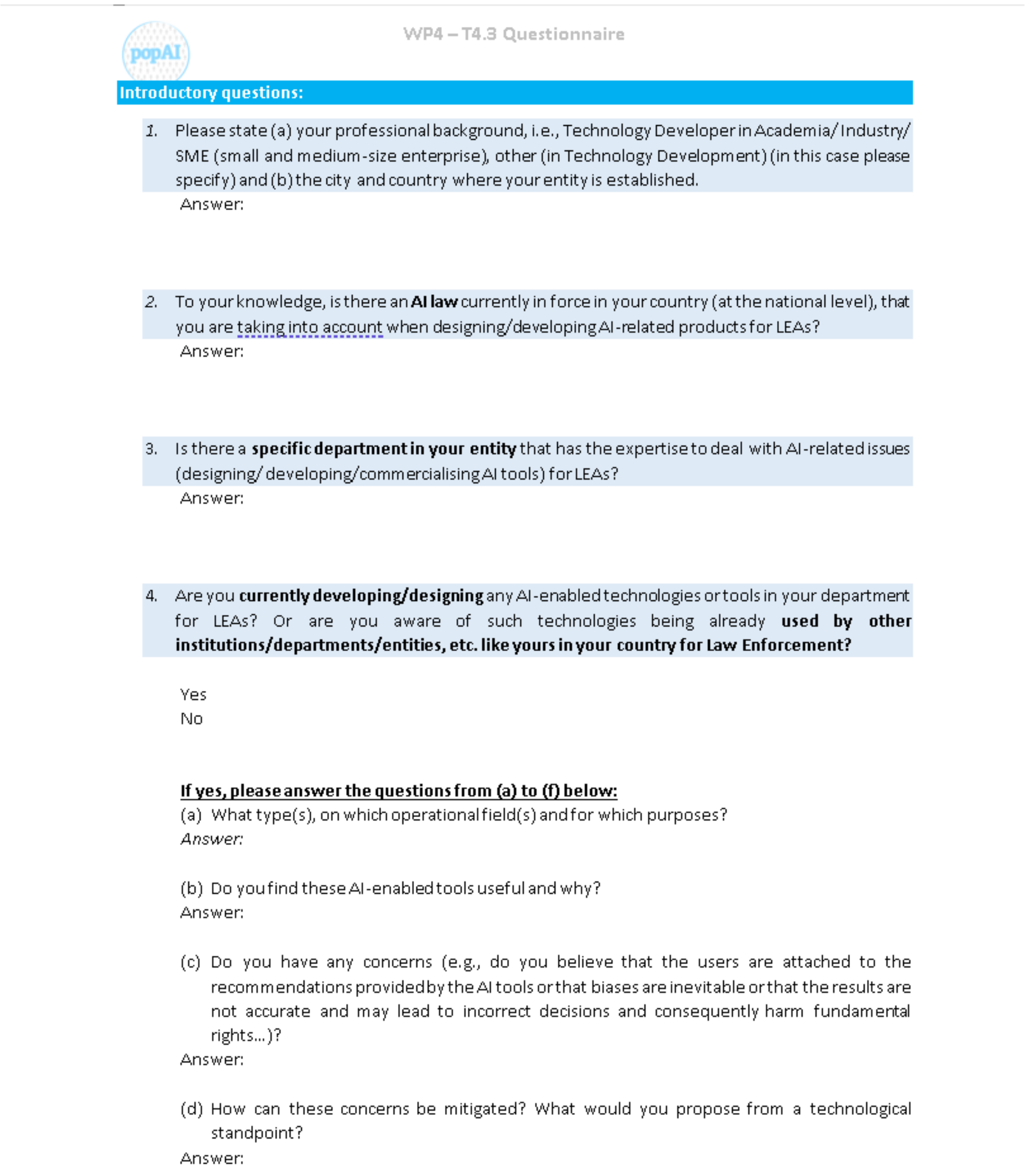
"Artificial intelligence (AI) systems *are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data, and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions. As a scientific discipline, AI includes several approaches and techniques, such as machine learning (of which deep learning and reinforcement learning are specific examples), machine reasoning (which includes planning, scheduling, knowledge representation and reasoning, search, and optimisation), and robotics (which includes control, perception, sensors, and actuators as well as the integration of all other techniques into cyber-physical systems)".*[1]

"Artificial intelligence system (AI system) *means a machine-based system designed to operate with varying levels of autonomy and that can, for explicit or implicit objectives, generate outputs such as predictions, recommendations, or decisions, that influence physical or virtual environments".*[2]

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

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

Figure 10. Questionnaire used in T4.3 (1/9)



WP4 – T4.3 Questionnaire

Introductory questions:

1. Please state (a) your professional background, i.e., Technology Developer in Academia/ Industry/ SME (small and medium-size enterprise), other (in Technology Development) (in this case please specify) and (b) the city and country where your entity is established.

Answer:

2. To your knowledge, is there an **AI law** currently in force in your country (at the national level), that you are taking into account when designing/developing AI-related products for LEAs?

Answer:

3. Is there a **specific department in your entity** that has the expertise to deal with AI-related issues (designing/ developing/commercialising AI tools) for LEAs?

Answer:

4. Are you **currently developing/designing** any AI-enabled technologies or tools in your department for LEAs? Or are you aware of such technologies being already **used by other institutions/departments/entities, etc. like yours in your country for Law Enforcement?**

Yes
No

If yes, please answer the questions from (a) to (f) below:

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

Answer:

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

Answer:

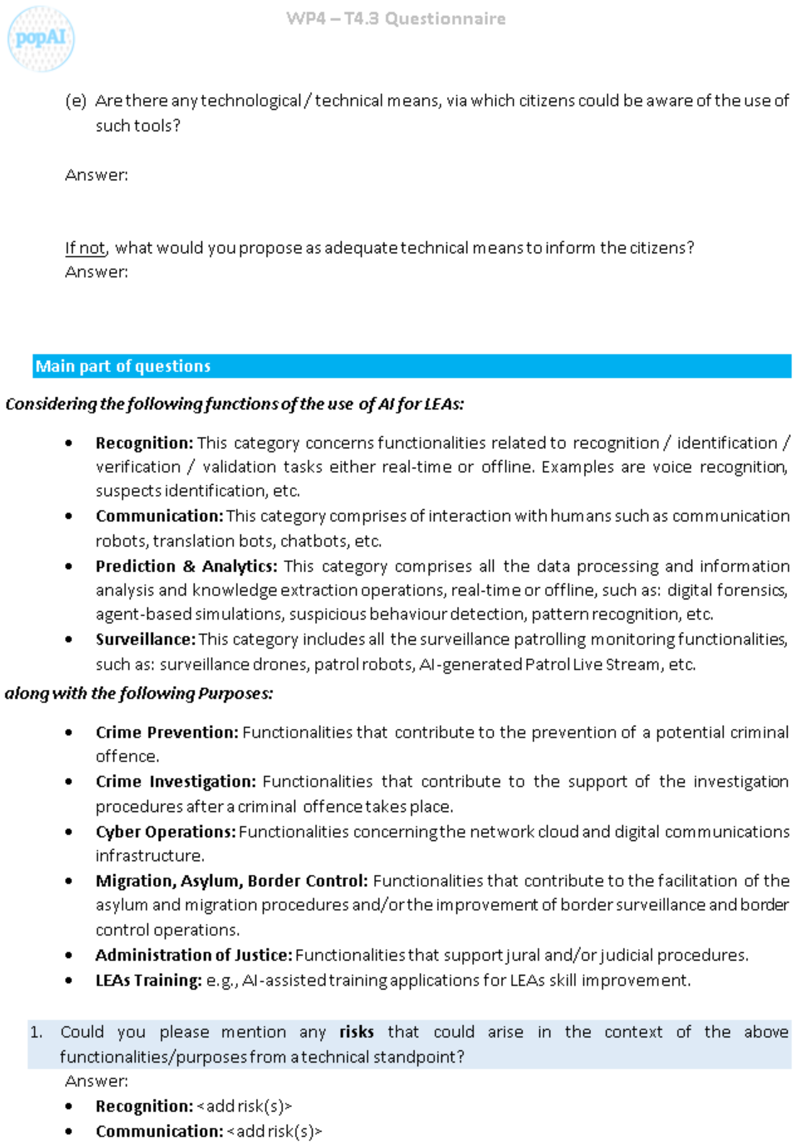
(c) Do you have any concerns (e.g., do you believe that the users are attached to the recommendations provided by the AI tools or that biases are inevitable or that the results are not accurate and may lead to incorrect decisions and consequently harm fundamental rights…)?

Answer:

(d) How can these concerns be mitigated? What would you propose from a technological standpoint?

Answer:

Figure 11. Questionnaire used in T4.3 (2/9)



WP4 – T4.3 Questionnaire

(e) Are there any technological / technical means, via which citizens could be aware of the use of such tools?

Answer:

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

Main part of questions

Considering the following functions of the use of AI for LEAs:

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

along with the following Purposes:

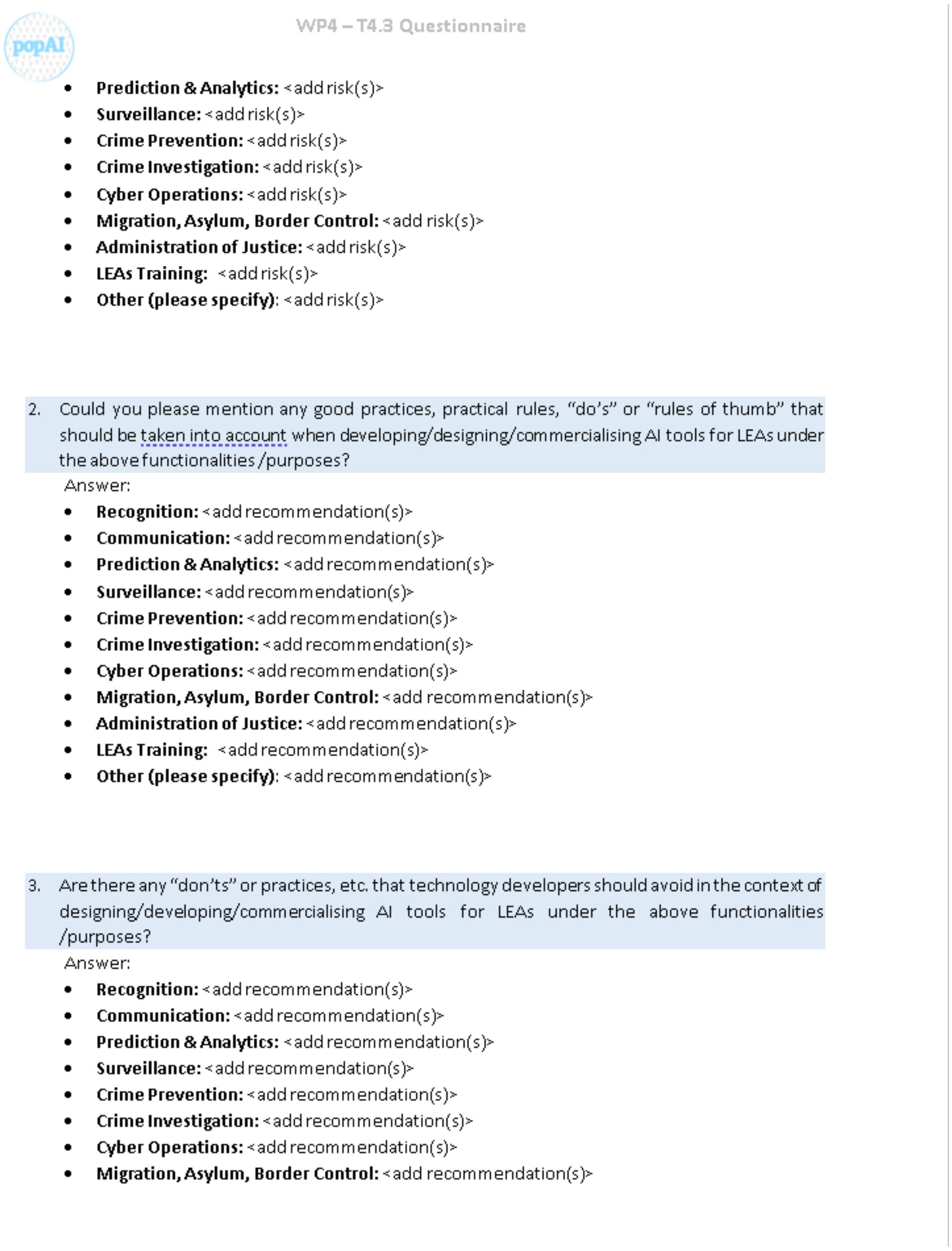
- **Crime Prevention:** Functionalities that contribute to the prevention of a potential criminal offence.
- **Crime Investigation:** Functionalities that contribute to the support of the investigation procedures after a criminal offence takes place.
- **Cyber Operations:** Functionalities concerning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

1. Could you please mention any **risks** that could arise in the context of the above functionalities/purposes from a technical standpoint?

Answer:

- **Recognition:** <add risk(s)>
- **Communication:** <add risk(s)>

Figure 12. Questionnaire used in T4.3 (3/9)



WP4 – T4.3 Questionnaire

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

2. Could you please mention any good practices, practical rules, "do's" or "rules of thumb" that should be taken into account when developing/designing/commercialising AI tools for LEAs under the above functionalities /purposes?

Answer:

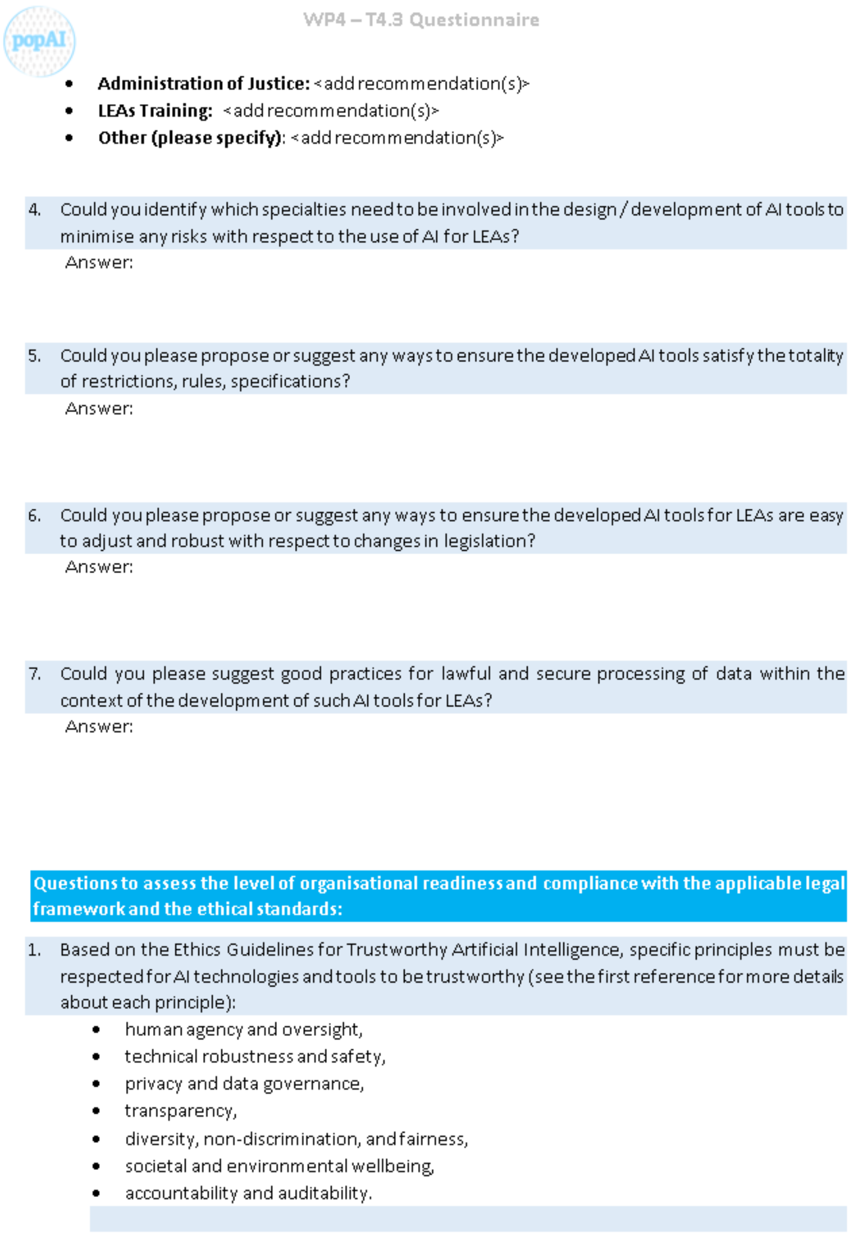
- **Recognition:** <add recommendation(s)>
- **Communication:** <add recommendation(s)>
- **Prediction & Analytics:** <add recommendation(s)>
- **Surveillance:** <add recommendation(s)>
- **Crime Prevention:** <add recommendation(s)>
- **Crime Investigation:** <add recommendation(s)>
- **Cyber Operations:**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

3. Are there any "don'ts" or practices, etc. that technology developers should avoid in the context of designing/developing/commercialising AI tools for LEAs under the above functionalities /purposes?

Answer:

- **Recognition:** <add recommendation(s)>
- **Communication:** <add recommendation(s)>
- **Prediction & Analytics:** <add recommendation(s)>
- **Surveillance:** <add recommendation(s)>
- **Crime Prevention:** <add recommendation(s)>
- **Crime Investigation:** <add recommendation(s)>
- **Cyber Operations:** <add recommendation(s)>
- **Migration, Asylum, Border Control:** <add recommendation(s)>

Figure 13. Questionnaire used in T4.3 (4/9)



WP4 – T4.3 Questionnaire

- **Administration of Justice:** <add recommendation(s)>
- **LEAs Training:** <add recommendation(s)>
- **Other (please specify):** <add recommendation(s)>

4. Could you identify which specialties need to be involved in the design / development of AI tools to minimise any risks with respect to the use of AI for LEAs?

   Answer:

5. Could you please propose or suggest any ways to ensure the developed AI tools satisfy the totality of restrictions, rules, specifications?

   Answer:

6. Could you please propose or suggest any ways to ensure the developed AI tools for LEAs are easy to adjust and robust with respect to changes in legislation?

   Answer:

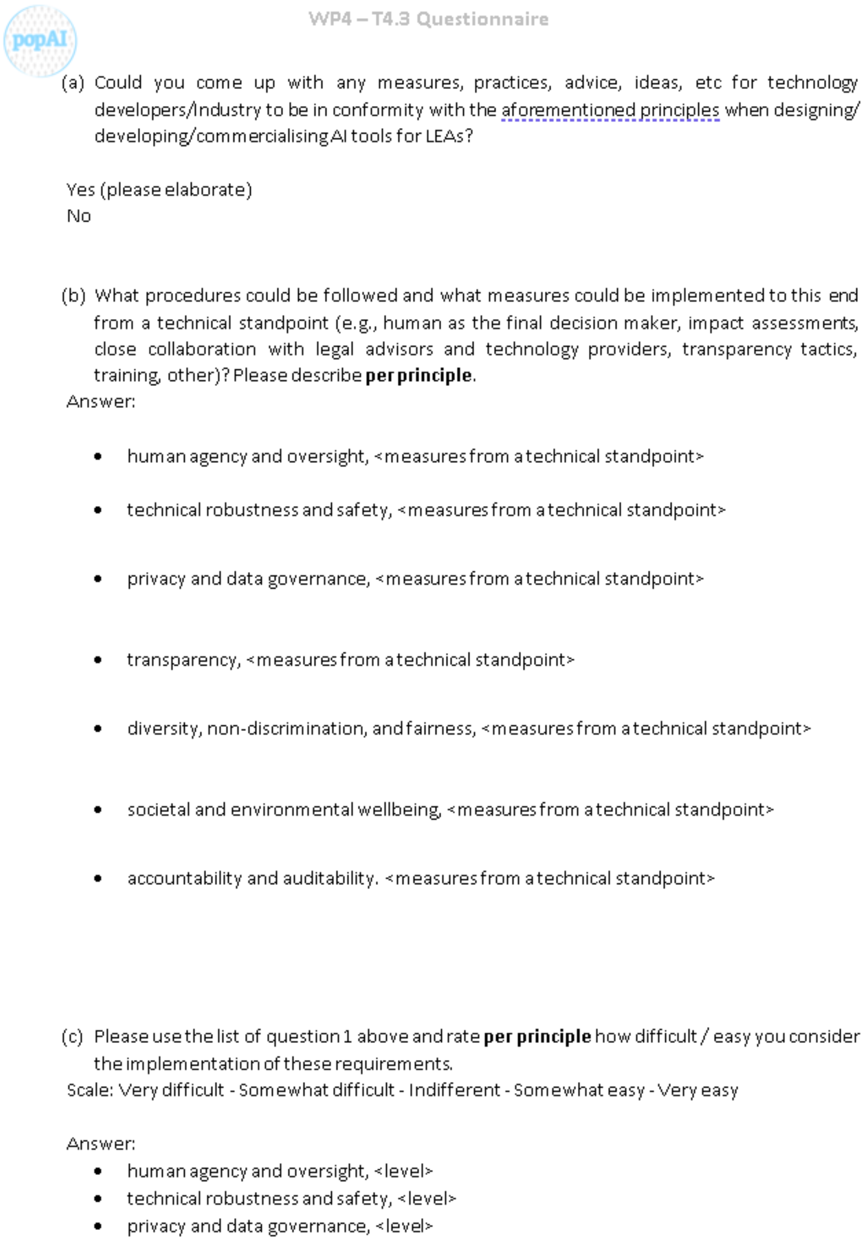
7. Could you please suggest good practices for lawful and secure processing of data within the context of the development of such AI tools for LEAs?

   Answer:

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

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

Figure 14. Questionnaire used in T4.3 (5/9)



WP4 – T4.3 Questionnaire

(a) Could you come up with any measures, practices, advice, ideas, etc for technology developers/Industry to be in conformity with the aforementioned principles when designing/ developing/commercialising AI tools for LEAs?

Yes (please elaborate)
No

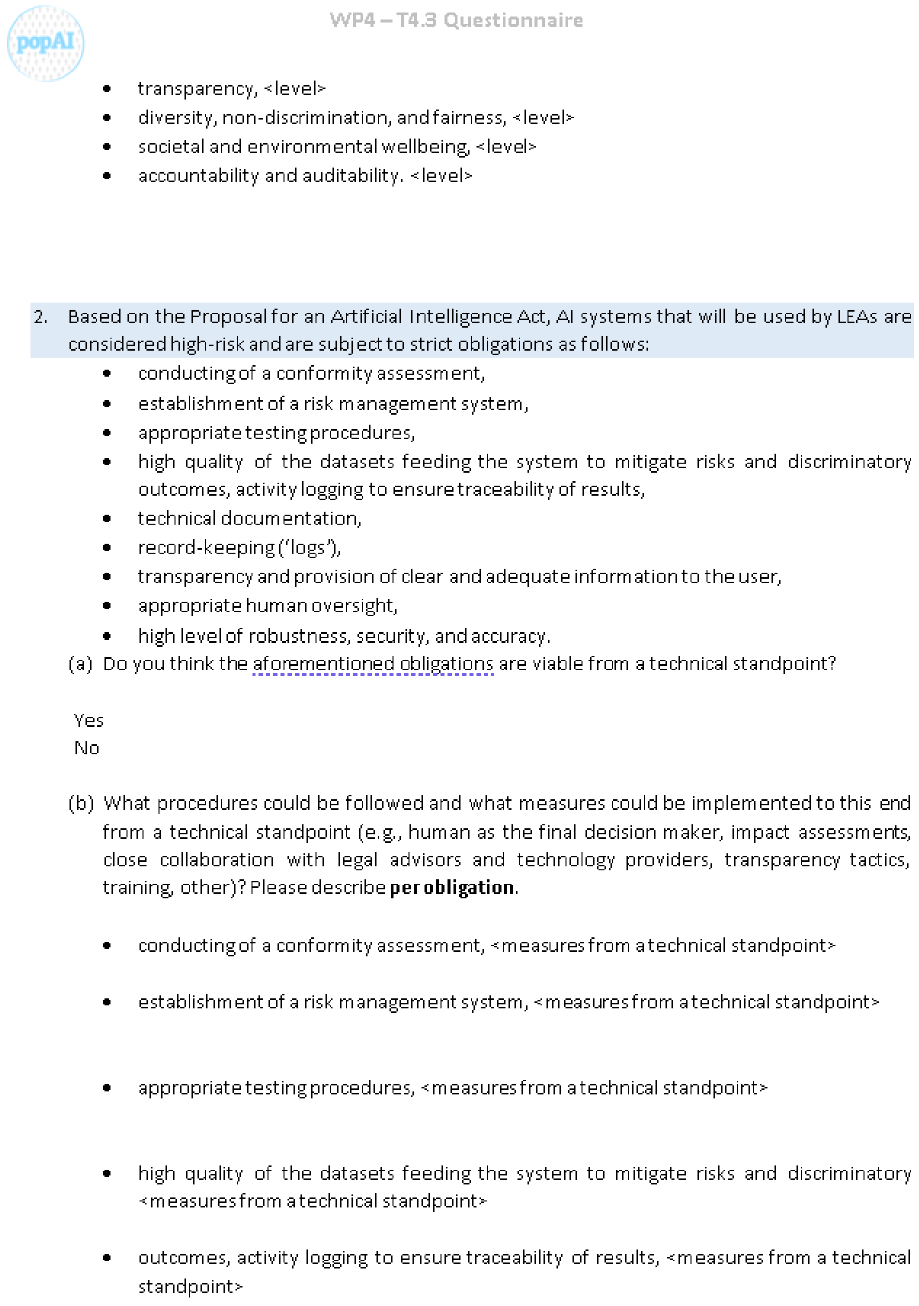
(b) What procedures could be followed and what measures could be implemented to this end from a technical standpoint (e.g., human as the final decision maker, impact assessments, close collaboration with legal advisors and technology providers, transparency tactics, training, other)? Please describe **per principle**.
Answer:

- human agency and oversight, <measures from a technical standpoint>
- technical robustness and safety, <measures from a technical standpoint>
- privacy and data governance, <measures from a technical standpoint>
- transparency, <measures from a technical standpoint>
- diversity, non-discrimination, and fairness, <measures from a technical standpoint>
- societal and environmental wellbeing, <measures from a technical standpoint>
- accountability and auditability. <measures from a technical standpoint>

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

Answer:
- human agency and oversight, <level>
- technical robustness and safety, <level>
- privacy and data governance, <level>

Figure 15. Questionnaire used in T4.3 (6/9)



WP4 – T4.3 Questionnaire

- transparency, <level>
- diversity, non-discrimination, and fairness, <level>
- societal and environmental wellbeing, <level>
- accountability and auditability. <level>

2. Based on the Proposal for an Artificial Intelligence Act, AI systems that will be used by LEAs are considered high-risk and are subject to strict obligations as follows:
   - conducting of a conformity assessment,
   - establishment of a risk management system,
   - appropriate testing procedures,
   - high quality of the datasets feeding the system to mitigate risks and discriminatory outcomes, activity logging to ensure traceability of results,
   - technical documentation,
   - record-keeping ('logs'),
   - transparency and provision of clear and adequate information to the user,
   - appropriate human oversight,
   - high level of robustness, security, and accuracy.

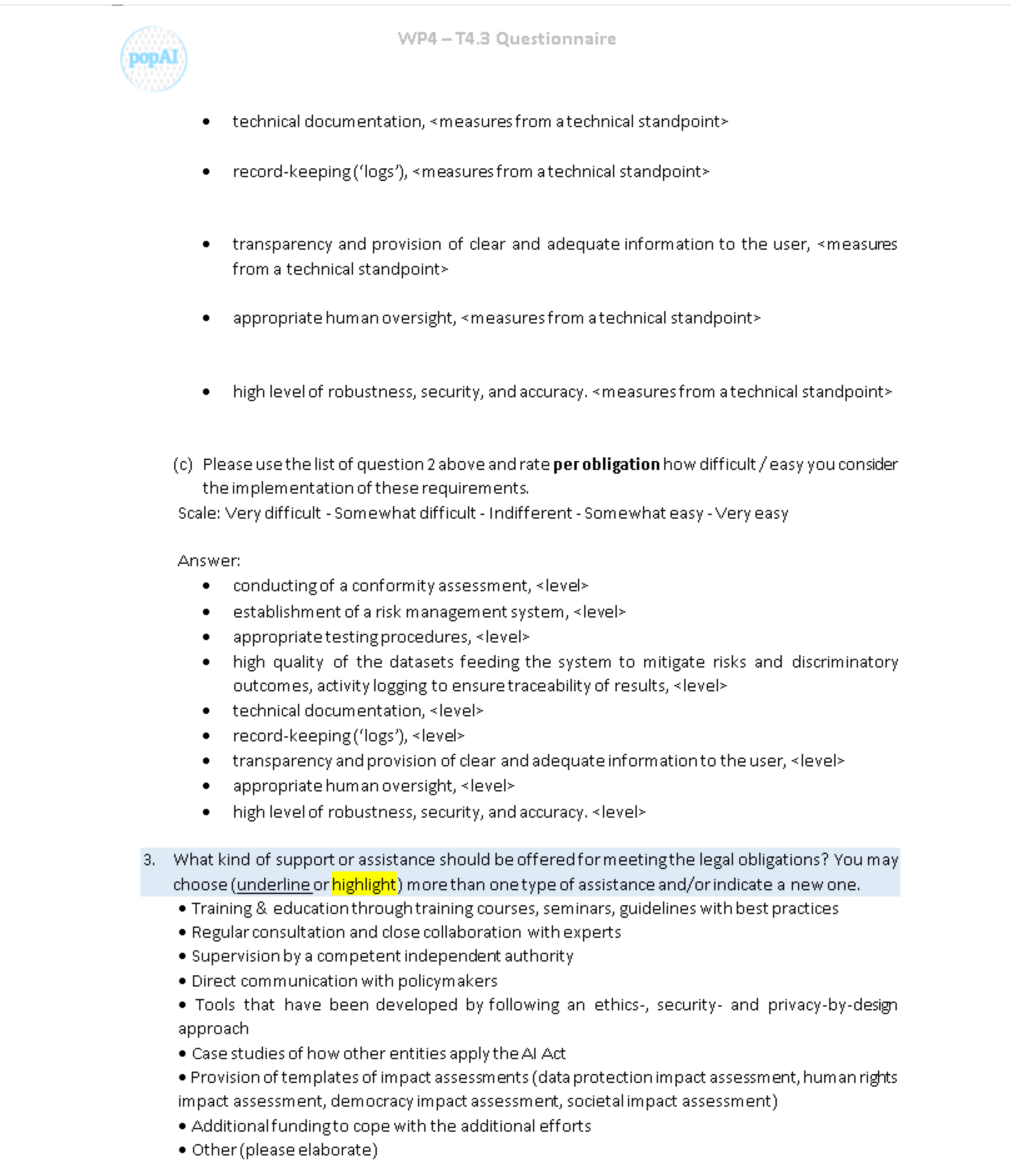
(a) Do you think the aforementioned obligations are viable from a technical standpoint?

Yes
No

(b) What procedures could be followed and what measures could be implemented to this end from a technical standpoint (e.g., human as the final decision maker, impact assessments, close collaboration with legal advisors and technology providers, transparency tactics, training, other)? Please describe **per obligation**.

- conducting of a conformity assessment, <measures from a technical standpoint>
- establishment of a risk management system, <measures from a technical standpoint>
- appropriate testing procedures, <measures from a technical standpoint>
- high quality of the datasets feeding the system to mitigate risks and discriminatory <measures from a technical standpoint>
- outcomes, activity logging to ensure traceability of results, <measures from a technical standpoint>

Figure 16. Questionnaire used in T4.3 (7/9)



WP4 – T4.3 Questionnaire

- technical documentation, <measures from a technical standpoint>
- record-keeping ('logs'), <measures from a technical standpoint>
- transparency and provision of clear and adequate information to the user, <measures from a technical standpoint>
- appropriate human oversight, <measures from a technical standpoint>
- high level of robustness, security, and accuracy. <measures from a technical standpoint>

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

Scale: Very difficult - Somewhat difficult - Indifferent - Somewhat easy - Very easy

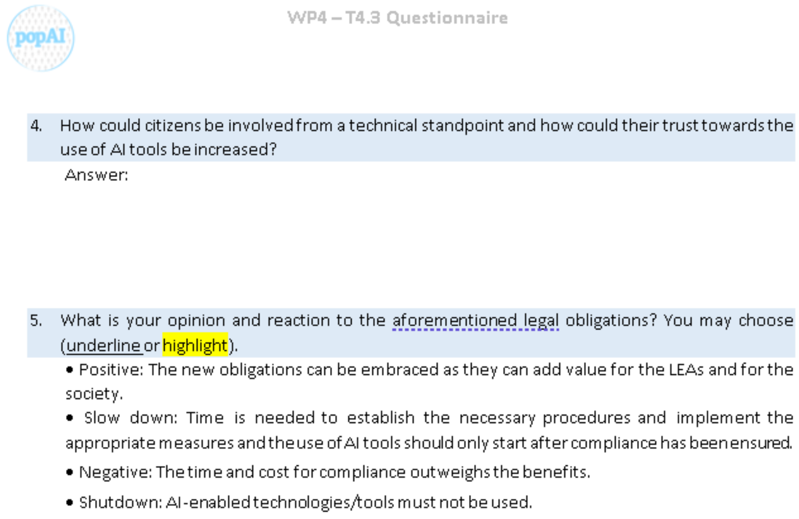
Answer:

- conducting of a conformity assessment, <level>
- establishment of a risk management system, <level>
- appropriate testing procedures, <level>
- high quality of the datasets feeding the system to mitigate risks and discriminatory outcomes, activity logging to ensure traceability of results, <level>
- technical documentation, <level>
- record-keeping ('logs'), <level>
- transparency and provision of clear and adequate information to the user, <level>
- appropriate human oversight, <level>
- high level of robustness, security, and accuracy. <level>

3. What kind of support or assistance should be offered for meeting the legal obligations? You may choose (underline or highlight) more than one type of assistance and/or indicate a new one.

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

Figure 17. Questionnaire used in T4.3 (8/9)



WP4 – T4.3 Questionnaire

4. How could citizens be involved from a technical standpoint and how could their trust towards the use of AI tools be increased?

   Answer:

5. What is your opinion and reaction to the aforementioned legal obligations? You may choose (underline or highlight).

   - Positive: The new obligations can be embraced as they can add value for the LEAs and for the society.
   - Slow down: Time is needed to establish the necessary procedures and implement the appropriate measures and the use of AI tools should only start after compliance has been ensured.
   - Negative: The time and cost for compliance outweighs the benefits.
   - Shutdown: AI-enabled technologies/tools must not be used.

Figure 18. Questionnaire used in T4.3 (9/9)