ALIGNER

ALIGNER D5.5

First Update of the Research Roadmap for AI in Support of Law Enforcement and Policing

Deliverable No.	D5.5
Work Package	WP5
Dissemination Level	PU
Author(s)	Daniel Lückerath, Valerie Wischott (Fraunhofer)
Co-Author(s)	Donatella Casaburo (KUL), Lindsay Clutterbuck (CBRNE), Peter Svenmarck, Tommy Westman (FOI)
Contributor(s)	-
Due date	2023-03-31
Actual submission date	2023-03-31
Status	Final
Revision	1.0
Reviewed by (if applicable)	Philip Engström (SPA), Dominic Kelly (CBRNE), Kai Pervölz (Fraunhofer), Christian Rabini (MPD), Adelina Zahirovic (SPA) Ari Basen (SIEAB), Shaban Buzar (SIEAB), Penny Duquenoy (SIEAB), Marco Filippi (SIEAB), Fredrik Heintz (SIEAB), Karl Hertting (LEAAB), Peter Kröjs (LEEAB), Andrius Paskauskas (SIEAB), Oliver Rose (SIEAB)

This document has been prepared in the framework of the European project ALIGNER – Artificial Intelligence Roadmap for Policing and Law Enforcement. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 101020574.

Contact:

info@aligner-h2020.eu
www.aligner-h2020.eu

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement no. 101020574.

Executive Summary

The European Commission-funded Coordination and Support Action *ALIGNER: Artificial Intelligence Roadmap for Policing and Law Enforcement* brings together European actors concerned with Artificial Intelligence (AI), Law Enforcement, and Policing to collectively identify and discuss needs for paving the way for a more secure Europe in which Artificial Intelligence supports law enforcement agencies while simultaneously empowering, benefiting, and protecting the public.

This deliverable presents the second iteration of the research roadmap, a key output not only of work package (WP) 5 "Outreach and Roadmap" but of the whole project. The roadmap compiles all the (intermediate) project results. Specifically, the roadmap

- presents the ALIGNER narratives – visions of potential futures regarding the use of AI by criminals and law enforcement agencies;
- identifies practitioner needs that need to be met to counter (future) criminal use of AI and bring AI into service for law enforcement and policing;
- identifies and assesses AI technologies that can support practitioners under the postulated narratives;
- discusses how AI technologies might aid criminals in future and could lead to new crime patterns;
- identifies and discusses ethical, legal, and organizational implications of the use of AI by law enforcement agencies; and
- gives recommendations to policymakers and researchers on how to address the identified trends to meet the operational, cooperative, and collaborative needs of police and law enforcement agencies (P&LEA) in the context of AI, while acknowledging ethical, and legal implications.

To account for the broad network of actors in the fields of artificial intelligence, law enforcement, and policing, ALIGNER's research roadmap addresses

- LEA, policing, and criminal justice practitioners, including technical staff who are interested in applying, adapting, or co-creating upcoming research trends;
- research programmers and policymakers in local, regional, and national governments and other legislative bodies, who are interested in policy recommendations addressing identified gaps with regard to AI solutions for law enforcement;
- standardisation bodies to advance the unification of models, methods, tools, and data related to the use of AI in law enforcement;
- the research community surrounding artificial intelligence, law enforcement and policing, as well as ethical, legal, and societal assessment; and
- the industry community surrounding artificial intelligence and law enforcement who will receive directions for future developments and business opportunities.

The ALIGNER roadmap is a living document that is iteratively developed, extended, and adapted over the course of two years, starting with the initial publication in September 2022. Subsequent publications will follow every six months.

The work of ALIGNER – and subsequently this roadmap – assumes a vision of the future where AI is a constant criminal threat and a regular tool used by law enforcement agencies. Within this vision,

ALIGNER focuses on a limited number of topical areas with highest relevance for P&LEAs and other actors in the field of law enforcement and AI. To start with, the first iteration of the roadmap (D5.3) focused on the topic of 'Disinformation and Social Manipulation' and the associated challenges and opportunities. Specifically, the initial roadmap provided the description of the first ALIGNER narrative – a description of a potential future scenario of the (mis)use of AI for disinformation and social manipulation; initial practitioner needs and AI technologies – both specific to the narrative and more general; a general overview of ethical and legal implications of the use of AI by P&LEAs; and a first overview of identified research projects in the field of AI.

This second iteration of the roadmap (D5.4) now extends the first iteration with an overview of ongoing EU policy processes both relating to AI in general as well as specifically on the use of AI by P&LEA as well as six initial policy recommendations developed jointly with the EU AI cluster and experts from ALIGNER's advisory boards. These recommendations are:

1. Provide common guidelines and unbiased specialist support to P&LEAs for the development, procurement, deployment, and use of AI technology.
2. Establish unified frameworks for the evaluation of AI tools during development and deployment ensuring their ethical, legal, and societal compliance.
3. Review existing and establish new legal mechanisms to ensure that AI systems and their use are ethical, legal, and societally acceptable.
4. Develop meaningful dialogue between regulators, P&LEAs, researchers, industry, and civil society organisations to strengthen citizens' confidence in the use of AI tools by P&LEAs.
5. Support and invest in the development of guidelines for gender-sensitive and gender-responsive policing in the AI era.
6. Extend and adapt European and national research programmes to better facilitate evidence-based, participatory research into P&LEA needs regarding AI, the potential implications of the use of AI by P&LEA, and potential criminal use of AI.

The majority of the content for this roadmap results from work conducted by individual project partners, an online survey that ran between May and August 2022, four workshops held by ALIGNER with practitioners from law enforcement and policing, research and academia, industry professionals, and policymakers in 2021 and 2022, as well as expert discussions during several research and policy events.

Table of contents

List of Abbreviations

Abbreviation	Meaning
AI	Artificial Intelligence
ECHR	European Convention on Human Rights
P&LEA	Police & law enforcement agency
N	Sample size
WP	Work package
WPON	Women Police Officers Network

1. Introduction

This deliverable has been prepared for the European Commission-funded Coordination and Support Action *ALIGNER: Artificial Intelligence Roadmap for Policing and Law Enforcement*. ALIGNER aims to bring together European actors concerned with artificial intelligence, law enforcement, and policing to collectively identify and discuss needs for paving the way for a more secure Europe in which artificial intelligence supports police and law enforcement agencies while simultaneously empowering, benefiting, and protecting the public. To achieve this, ALIGNER will

(1) facilitate communication and cooperation between actors from law enforcement, policing, policymaking, research, industry, and civil society about the changing dynamics of crime patterns relevant to the use of AI by establishing a workshop series;
(2) identify the capability enhancement needs of European LEAs;
(3) identify, assess, and validate AI technologies with potential for LEA capability enhancement by implementing a technology watch process that includes impact and risk assessments;
(4) identify ethical, societal, and legal implications of the use of AI in law enforcement;
(5) identify means and methods for preventing the criminal use of AI via the development of a taxonomy of AI-supported crime;
(6) identify policy and research needs related to the use of AI in law enforcement by mapping practitioner needs and emerging crime patterns with identified AI technologies; and
(7) employ the gathered insights to incrementally develop and maintain an AI research roadmap.

This deliverable presents the second iteration of the research roadmap, a key output not only of work package 5 “Outreach and Roadmap” but of the whole project. The roadmap compiles all the (intermediate) project results achieved up to now. Specifically, the roadmap

- presents the ALIGNER narratives – visions of potential futures regarding the use of AI by criminals as well as police and law enforcement agencies;
- identifies practitioner needs that need to be met to counter (future) criminal use of AI and bring AI into service for law enforcement and policing;
- identifies and assesses AI technologies that can support practitioners under the postulated narratives;
- discusses how AI technologies might aid criminals in future and could lead to new crime patterns;
- identifies and discusses ethical, legal, societal, and organizational implications of the use of AI by law enforcement agencies; and
- gives recommendations to policymakers and researchers on how to address the identified trends to meet the operational, cooperative, and collaborative needs of police and LEAs in the context of AI, while acknowledging ethical, legal, and societal implications.

To account for the broad network of actors in the fields of artificial intelligence, law enforcement, and policing, ALIGNER’s research roadmap addresses

- LEA, policing, and criminal justice practitioners, including technical staff who are interested in applying, adapting, or co-creating upcoming research trends;
- research programmers and policymakers in local, regional, and national governments and other legislative bodies, who are interested in policy recommendations addressing identified gaps with regard to AI solutions for law enforcement;

- standardisation bodies to advance the unification of models, methods, tools, and data related to the use of AI in law enforcement;
- the research community surrounding artificial intelligence, law enforcement and policing, as well as ethical, legal, and societal assessment; and
- the industry community surrounding artificial intelligence and law enforcement who will receive directions for future developments and business opportunities.

The content of the roadmap results from work conducted by individual project partners, an online survey that ran between May and August 2022, as well as four workshops held by ALIGNER with practitioners from law enforcement and policing, research and academia, industry professionals, and policymakers in 2021 and 2022. In addition, ALIGNER partners participated in expert discussions during several research and policy events. Lastly, ALIGNER intensively exchanged with its sibling projects popAI[1] and STARLIGHT[2] as well as the EU project AP4AI[3], that together with ALIGNER form the AI cluster of EU research projects.

The roadmap is structured as follows: This section continues with giving an overview of the publication timeline of the roadmap and a short description of what is newly included or modified in this iteration of the document. Section 2 then introduces the first ALIGNER narrative, as well as initial practitioner capability enhancement needs and AI technologies. Section 3 continues with an overview of general ethical and legal aspects of the use of AI by law enforcement agencies, before section 4 provides an initial overview of ongoing EU policy processes related to AI use by P&LEA as well as first policy recommendations. The roadmap closes with the initial version of the AI technology catalogue – a more detailed overview of the AI technologies identified for the first narrative and an outlook towards the next iteration of the roadmap. In addition, the annex to the roadmap provides an overview of relevant research projects in the field of AI and more detailed information from ALIGNER's online survey on capability enhancement needs.

1.1 Publishing Timeline

The ALIGNER roadmap is not a fixed document. To account for the rapid developments in the field of AI, the roadmap will be treated as a living document that is iteratively developed, extended, and adapted over the course of two years, starting with the initial publication in September 2022. Table 1 gives an overview of the publication timeline for the roadmap.

[1] https://www.pop-ai.eu/
[2] https://starlight-h2020.eu/
[3] https://www.ap4ai.eu/

Table 1: Publication timeline of the ALIGNER roadmap

When	What
Sep 2022	◆ **For narrative 1:** ○ Description of the narrative ○ Initial set of practitioner needs ○ Initial set of AI technologies ◆ General ethical & legal considerations ◆ Additional general capability enhancement needs ◆ Initial mapping of projects and initiatives
Mar 2023	◆ Identification of relevant policy / research processes / strategies ◆ Initial set of policy recommendations
Sept 2023	◆ **For narrative 1 (if necessary):** ○ Update to set of AI technologies ○ Results of impact assessment for initial set of AI technologies ○ Related challenges & unintended consequences (technical, ethical, legal) ◆ **For narratives 2, 3, 4:** ○ Narrative descriptions ○ Practitioner needs ○ Sets of AI technologies ○ Challenges and potential unintended consequences (technical, ethical, legal) ○ Results of impact assessments for related set of AI technologies ◆ Updated set of policy / research process / strategies and policy recommendations (if necessary) ◆ Initial taxonomy of AI supported crime ◆ Update to project mapping
Mar 2024	◆ **For all narratives:** ○ Final set of AI technologies ○ Updated challenges & consequences (technical, ethical, legal) ○ Updated impact assessments ◆ Preliminary set of cybersecurity requirements ◆ Initial set of desirable approaches to overcome challenges & unintended consequences
Sep 2024	◆ Gap analysis ◆ Final policy recommendations ◆ Final set of cybersecurity requirements ◆ Final impact assessments for all scenarios ◆ Final taxonomy ◆ Final project and initiative mapping

1.2 What's New in This Version?

This second version of the ALIGNER roadmap, published in March 2023, extends the initial version, published in September 2023, with an overview of relevant policy processes relating to AI in general and AI use by P&LEA specifically, as well as an initial set of policy recommendations, based work conducted in ALIGNER as well as in collaboration of the EU AI cluster (section 4). It also includes minor modifications of the publishing timeline (shifting updates for narrative 1 into September 2023) as well as minor updates of the introductory text and the conclusion to account for the second version being published (Executive summary, section 1, section 6).

2. Narratives and Capabilities

The work of ALIGNER – and subsequently this roadmap – assumes a vision of the future where AI is a constant criminal threat and a regular tool used by law enforcement agencies. Within this vision, ALIGNER focuses on a limited number of topical areas with highest relevance for P&LEAs and other actors in the field of law enforcement and AI. These high-interest topics are captured in the form of narratives: high-level descriptions of potential futures, including how AI might be used for criminal behaviour as well as to support LEAs.[4]

The focus of the first narrative was selected based on expert input from ALIGNER's advisory boards and in collaboration with several other research projects. The initial selection was then validated via an online survey that ran between May and August 2022 (see also Annex B). A similar process will be followed for future narratives.

2.1 Narratives and Emerging Crime Patterns

2.1.1 Narrative 1: AI, Disinformation and Social Manipulation

We live in a world where artificial intelligence is a ubiquitous technology, used daily by almost everyone, be it as part of smart household appliances, during the daily commute, as personal assistant, as recommender service, or to support decision making processes. In this world AI is also used by criminal subjects, from isolated individuals, organized criminal networks of different sizes, to state-sponsored malicious entities. At the same time, law enforcement agencies regularly employ AI technologies to prevent, detect and counter criminal activities, find patterns for investigations, and support with their day-to-day work.

One especially active area for criminal activity lies in the domain of disinformation and social manipulation, especially prior to political elections. First, criminals use AI for phishing attacks to gather personal data and identify high-value targets who are subsequently attacked with highly targeted phishing attempts ('tailored phishing or spear phishing'). The goal of these attacks, if successful, is to manipulate or coerce targets to gain unauthorised access to computer networks, e.g., of election campaigns, large research companies, or industry organizations. These phishing attacks may involve online attempts to persuade or trick individuals into divulging passwords or access codes or, if the opportunity arises, using harvested data to subject them to blackmail or coercive threats.

Besides targeted phishing attacks and data harvesting, criminals use artificial intelligence to create and disseminate selective misinformation and specifically created disinformation, apparently emanating from official or well-informed sources. This disinformation uses artificially generated videos, images, text, and sound, including deep fakes of public figures, and is generated by AI-fuelled 'bots'.

To counter these threats, law enforcement agencies also bring AI to bear: They use veracity assessment methods to detect disinformation, then employ deanonymisation techniques like authorship

[4] In the working context of ALIGNER, the overarching vision of the future is also called the 'archetypical scenario'. Within this vision, ALIGNER builds 'scenarios' that specify potential uses of AI by criminals as well as P&LEA and related implications. Each scenario is further fleshed out using 'narratives'. See also ALIGNER D2.2 [1] for additional details.

attribution and the geolocation of images to identify from where the disinformation originated. This is supported by techniques for the detection of synthetic images and videos.

2.2 Practitioner Capability Enhancement Needs

To identify in which areas of law enforcement and policing work artificial intelligence can unfold the most potential and to identify potential barriers for the deployment of AI – other than ethical, legal, and societal, which are discussed in section 3 – the ALIGNER team firstly assessed the current use of AI by P&LEAs, secondly the areas in which practitioners, researchers, and other actors in the field of AI, law enforcement, and policing identify the highest potential of AI, and thirdly where they see the largest challenges when introducing AI. This information was gathered during the ALIGNER workshops as well as via an online survey.[5]

2.2.1 Status quo of AI in law enforcement and policing

When discussing the use of AI with P&LEAs, it becomes evident that at present, AI is not used at all or only to a limited extent in the operative work of P&LEAs. This is supported by the survey results. 17 P&LEA practitioners indicated that AI is currently used to a very little or some extent. Six people do not use it at all in their work, while two people indicated that AI is used greatly in their work (Figure 1)[6]. These results are not surprising as discussions with practitioners showed that many police and law enforcement agencies still grapple with the basic technological demands for the use of AI and the leadership in many P&LEAs needs to be convinced of the fundamental impact AI can and will have on their organisations to support broader use of AI. P&LEAs also indicated in interviews that usually only highly specialized cybercrime units currently employ AI to a great extent, as AI is a prerequesit for their daily work. In other P&LEA departments the use of AI is instead much more dependent on individual motivation of investigators, e.g. an investigator seeking additional specific capacities or someone employed in a research department wanting to examine the use of a novel technology.

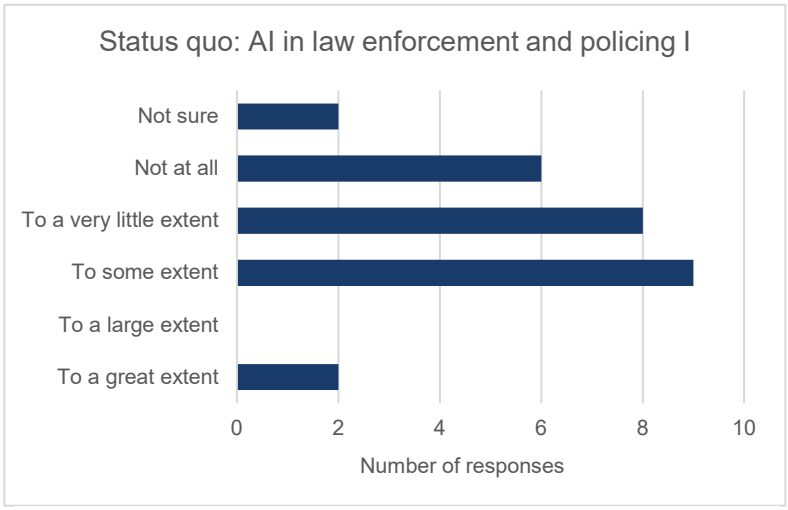


Figure 1: Results of the question “To what extent is AI currently being applied in your work?”

At the same time, a large number of P&LEAs are convinced that AI can enhance existing functions and capabilities or enable the development of new capabilities. However, the extent to which AI has brought benefits varies (Figure 2). Most respondents of ALIGNER’s survey indicated that the functions and capabilities of law enforcement and policing have benefitted to some extent, with fewer respondents

[5] For more information on the survey, please see ANNEX B.
[6] This and the following question (Figure 2) should only be answered by P&LEA. However, the number of responses differs between those two questions. This means that the sample could also include non-practitioners.

indicating that they have benefitted largely or to a great extent. However, no one indicated that functions and capabilities have not improved at all, but at least to a very little extent. From the survey sample, it appears that AI has enabled the development of new functions and capabilities rather than improving existing functions and capabilities.

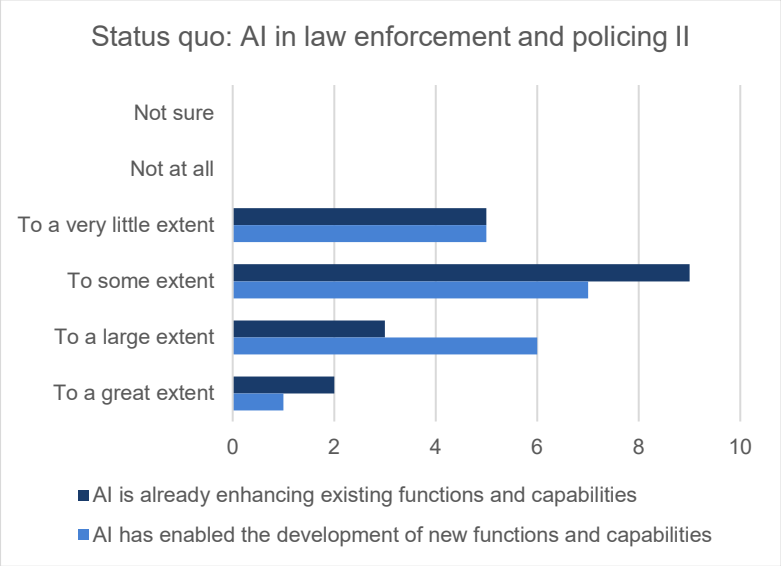


Figure 2: Results of the questions "To what extent do you think the use of AI is enhancing existing/has enabled the development of new functions and capabilities in law enforcement and policing?"

2.2.2 Potentials of AI in law enforcement and policing

Considering that AI is only used to a limited extent by P&LEAs – although if used seems to enable the development of new capabilities – the question arises: Is AI even seen as relevant for P&LEAs by practitioners and other actors in the field? And if so, in which areas of work would AI have the greatest impact? All participants of ALIGNER's workshops hinted at the high relevance of AI for P&LEAs and the survey results support this assessment. Figure 3 shows that AI is generally considered to be highly relevant for law enforcement and policing. Indeed, 95% of the participants stated that it is "relevant" or "very relevant" [7].

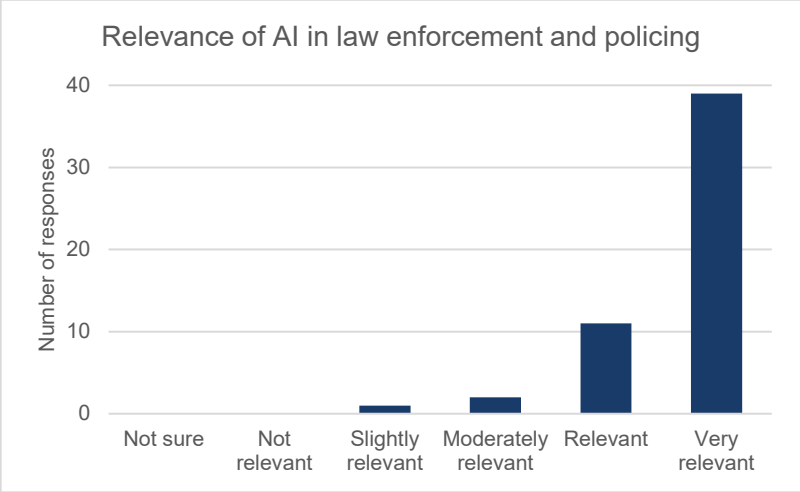


Figure 3: Results of the question "How relevant do you think the use of AI is in law enforcement and policing?

To identify specific work areas in which AI might support P&LEAs, ALIGNER delineated seven different categories of law enforcement and policing capabilities and functions based on working sessions held during the first two ALIGNER workshops (see Figure 4 and ALIGNER D2.2 [1]). During the first two ALIGNER workshops, P&LEA practitioners as well as researchers and industry professionals, unsurprisingly, identified those work areas that are heavily dependent on data as most promising for the application of AI. The survey responses support these results: Participants were asked to rate the extent to which each of the named functions and capabilities could benefit from the use of AI (Figure 4). The highest level of agreement is found in data and information handling processes, where almost

[7] This question and all the following questions were answered by all participants.

90% of participants believe, they could benefit to a large or great extent from the use of AI. This is followed by biometric recognition and identification (83%[8]), digital forensics (81%) and the detection and prevention of crimes and threats within the digital domain (78%). There is less consensus for incident reaction and response (65%), autonomous vehicles, robots, and drones (64%), and the detection and prevention of crimes and threats outside the digital domain (56%).

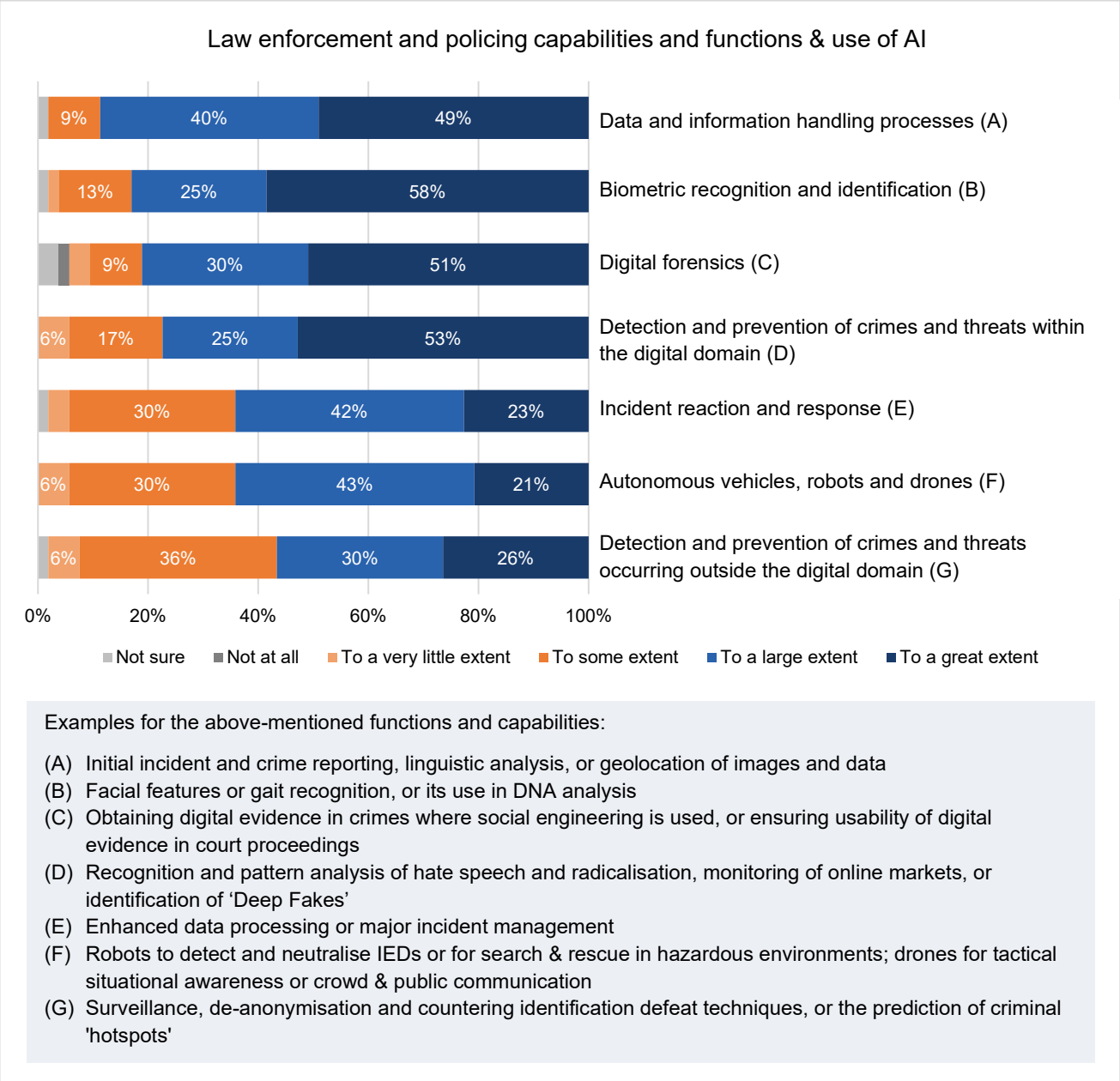


Figure 4: Results of the question "To what extent could the following law enforcement and policing functions and capabilities benefit from the use of AI?"

However, potential does not necessarily imply immediate benefits. Therefore, ALIGNER also asked survey participants to identify work areas where AI could be used immediately to bring about consequential beneficial changes to policing and LEA practice. Such an open question allowed the participants to formulate their views in their own words. The answers obtained were roughly clustered

[8] This and the following numbers in brackets refer to participants who answered with "to a large extent" and "to a great extent".

and prioritised, resulting in the following areas that were mentioned several (≥ 5) times:[9] (i) Data and information handling processes; (ii) Digital forensics; (iii) Prevention of crimes within the digital domain, with a focus particularly on social media analysis; and (iv) Biometric identification. These are in line with the highest ranked work areas that exhibit the highest potential in general.

2.2.3 Challenges of AI in law enforcement and policing

If AI has such a large potential for P&LEAs, why is it not already in broader use? What hinders the deployment of AI at law enforcement and policing institutions? When asked these questions, workshop and survey participants[10] brought up several challenges, which can be broadly categorized into

- **Ethical challenges** related to topics of discrimination, transparency, trust in the AI decision making process, and human oversight;
- **Legal challenges** related to safeguarding of fundamental rights, handling of AI system failures, privacy concerns, or ensuring usability of results from processes that make use of AI in court;
- **Institutional challenges** related to complicated procurement processes or difficulties in transferring promising outputs from research projects into practice; and
- **Technical challenges** related to the need to integrate AI technologies into legacy IT systems, the lack of appropriate training data, or the lack of knowledge and understanding of the technology.

"While AI can enhance capabilities […], this does not mean it is a good use of AI for society."

Figure 5: Quotation of one survey response which outlines the need for tradeoff between usefulness and the ethical and societal questions to explore.

The ethical and legal dimensions of AI in law enforcement and policing were universally regarded as the most important issue, both by workshop and survey participants (see Figure 5 for an example quote). The most obvious issue here is certainly the compliance with fundamental rights, data protection, and privacy regulations, as AI technologies usually require large amounts of data, which can easily result in (unintentional) mass surveillance. Other concerns relate to algorithm bias or the concern that AI is not used in a responsible way, e.g., fairly towards each citizen. In this context, the lack of trust in AI (presumably both among the public and among practitioners in P&LEAs) is mentioned several times as a challenge. Related to these concerns is the aspect that AI, when used by law enforcement and police agencies, must not replace the human brain or human decision making, e.g., in the interpretation of laws, as it is not considered capable of handling "the margin between right and wrong".

In addition to ethical and legal concerns, one reason for the lack of trust could also be the lack of knowledge and understanding of the technology and thus the lack of transparency. Another technical challenge related to the lack of labelled training data for AI. This and the use of "bad quality" or "wrong" training data may then in turn have ethical and legal implications, such as creating algorithm bias.

Other important challenges mentioned are further legal issues, e.g., how to legally handle a failure of an AI system causing any kind of harm, and institutional issues, e.g., the degree of digitisation of law enforcement and policing agencies. In discussions with law enforcement practitioners, the complex procurement practices at public offices, the perceived aversion of top-level hierarchy towards AI

[9] The full list of unclustered and unprioritized answers can be found in Annex B.
[10] The full list of original responses can also be found in Annex B.

systems, and general problems of transferring promising research results into practical use (e.g., because the technology developer does not provide support after the project ends) were also mentioned.

2.2.4 Challenges related to the first narrative

With specific relation to the disinformation and social manipulation narrative, more practical issues arose in individual and small group discussions with law enforcement and policing practitioners. While AI systems to detect “fake news” are already available, it is unclear who should decide on what is a “reliable source” and what is not when employing veracity assessment techniques. Beside the issue of responsibility, there arise also legal issues: when does something legally constitute fake news? And when does the distribution of fake news become a crime? If at all? While some European countries have established legal and organisational instruments to tackle fake news (e.g., Germany, the Czech Republic, Hungary, and France), experts and civil society representatives regularly raise concerns that these efforts might undermine free speech.

Given the speed and penetration power of bot networks, when it comes to distributing disinformation, P&LEAs require tools that (i) can stop the spread of disinformation quickly and effectively, and (ii) can identify the spread of disinformation early – ideally before large scale distribution begins. The latter would require ways to identify and monitor the deployment of bot networks and potentially the early identification of disinformation sources. Related to this challenge is the ethical and legal question of: when does the use and/or deployment of a bot network become a crime? If at all? Or: Is there an ethically and legally acceptable case for employing bot networks? A case for the latter could be made, when considering prebunking – also called inoculation – as a means to counter fake news. Here, someone is purposely confronted with a very small amount of fake news to cause them to defend their position with suitable arguments. The idea is to increase the resilience of people against malicious outside influences. However, this again brings about ethical and legal challenges: who should decide when to employ prebunking techniques? Who keeps oversight of these procedures?

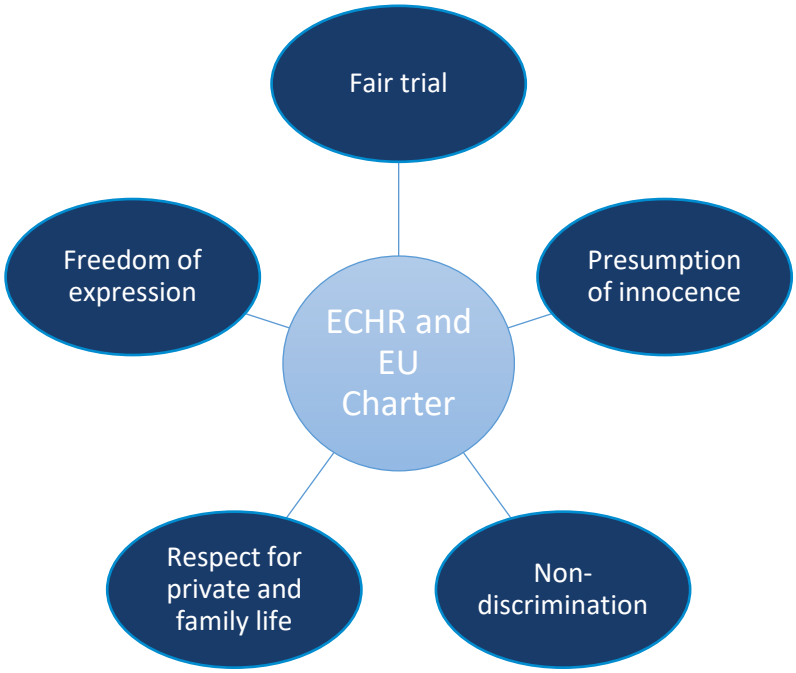
To counter disinformation using deep fakes, P&LEAs would need sufficient resources (personnel, time, money) to deploy and train “counter AI”. Similarly, geolocalisation of images, which becomes relevant to identify disinformation in image form, requires large amounts of labelled data.

3. Challenges and Opportunities

3.1 Ethical and Legal Aspects

While AI-enabled tools can bring clear benefits to the work of P&LEAs, they also raise numerous legal and ethical concerns, as already pointed towards in the previous section. If not properly developed and deployed by P&LEAs, these technologies can significantly harm the fundamental rights of the concerned individuals. For instance, AI-assisted tools used for law enforcement purposes can deliver biased or unexplainable outputs or lead to excessive and indiscriminate surveillance.[11] Therefore, it is crucial to specifically assess the potential risks that may arise from the P&LEAs' use of AI tools, and identify methods and best practices to prevent harm, well before the said tools are developed and deployed in practice.

To date, there is no concrete European legal framework regulating the use of AI tools in the law enforcement field. Nevertheless, many existing pieces of legislation have focused on fundamental rights protection to establish obligations for state authorities that must be observed also by P&LEAs while deploying AI tools.



In the European Union, fundamental rights of individuals are guaranteed and safeguarded by the two major human rights instruments adopted by the Council of Europe and the European Union: the European Convention on Human Rights (ECHR) and the Charter of Fundamental Rights of the EU (the Charter). AI-enabled tools used for law enforcement purposes are susceptible to affecting a multitude of human rights guaranteed by the two instruments, as these rights are closely connected to each other. However, in the present context, particular attention should be paid to: the presumption of innocence and the right to an effective remedy and a fair trial; the right to equality and non-discrimination; the right to respect for private and family life and the right to protection of personal data; and, finally, to freedom of expression and information.
For each of these rights, the relevant provisions of both the ECHR and the EU Charter as well as their further implications are summarized in the tables below. Additionally, the same tables show the potential harmful impact on fundamental rights of LEAs' use of AI-enabled tools, together with some suitable mitigation measures.

[11] See https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing for an example (accessed 2022-09-29)

Presumption of innocence, right to an effective remedy and to a fair trial

Relevant provisions	Articles 6 and 13 ECHR and Articles 47 and 48 EU Charter.
Definition & consequences	Anyone charged with a criminal offence must be presumed innocent until proved guilty according to law. Anyone whose rights and freedoms are violated has the right to an effective remedy before a tribunal. Everyone is entitled to a fair and public hearing within a reasonable time by an independent and impartial tribunal previously established by law, including rights: ◇ to be informed promptly of the nature and cause of the accusation; ◇ to bring their arguments and evidence as well as scrutinise and counteract the evidence presented against them; and ◇ to obtain an adequately reasoned and accessible decision.

AI-related risks	Mitigation measures
Predictive policing tools profile individuals before any crime is committed, potentially obliging the targeted individuals to prove their innocence even in absence of solid evidence against them.	Ensuring human oversight and that factual elements flagged by the AI tool are not considered proven, unless supported by solid evidence.
The opacity of the AI tools may undermine the understanding of the output generated and hide eventual biases, making a decision hard to challenge by the defendant as well as the judge.	Assessment of the accuracy and reliability of the AI tool deployed.
	Prosecution should be able to sufficiently explain the outputs generated by the AI tools used, to allow all relevant parties to challenge the evidence produced.
Unlawful collection and preservation of AI-generated evidence may lead to unreliability and inadmissibility in a criminal proceeding.	Ensuring lawful collection and preservation of chain of custody of AI evidence with appropriate safeguards.

Right to equality and non-discrimination

Relevant provisions	Article 14 ECHR and Articles 20 and 21 EU Charter.
Definition & consequences	Everyone is equal before the law. Any discrimination based on any ground such as sex, race, colour, ethnic or social origin, genetic features, language, religion or belief, political or any other opinion, membership of a national minority, property, birth, disability, age or sexual orientation shall be prohibited. ◇ Everyone should be protected against discriminatory decisions or policies, including automated decision-making based on sensitive data.

AI-related risks	Mitigation measures
The inaccuracy or non-exhaustivity of the criteria used in the design of the algorithm, as well as the poor quality or the existence of biases in the datasets used, may lead AI tools to perpetuate or generate discriminatory outputs.	Enhancing the quality and diversity of the datasets used to feed the AI tools, to avoid biased outputs.
	Avoiding the use of unlabelled datasets, to lower the risk that the new crime patterns or new criminal profiles identified are based on sensitive characteristics of the individuals.
	Expanding the room for human intervention in both the design and deployment stages of the AI tools, to minimise the risks of inaccurate outputs.

Right to respect for private and family life and right to protection of personal data

Relevant provisions	Article 8 ECHR and Articles 7 and 8 EU Charter.
Definition & consequences	Everyone has the right to respect for their private and family life, home and communications. ◇ Self-development without state interference. Everyone has the right to the protection of personal data concerning them. ◇ Personal data must be processed fairly for specified purposes and on a legitimate basis. ◇ Rights of access and rectification. ◇ Independent oversight.

AI-related risks	Mitigation measures
Large datasets including a vast amount of personal and sensitive data may cause a disproportionate interference with privacy and data protection rights.	Where possible, opting for synthetic datasets or anonymised datasets with lowest risks of re-identification.
	Performing a data protection impact assessment, to assess the legality and proportionality of the interference and strict adherence to data protection principles and relevant secondary legislation.
Continuously merging and repurposing different datasets may lead to the development of mass surveillance tools and chilling effects.	Avoiding the repurposing of datasets and limiting their use to the original purpose foreseen during the data collection.

Freedom of expression and information

Relevant provisions	Article 10 ECHR and Article 11 EU Charter.
Definition & consequences	Everyone has the right to freedom of expression, including freedom to hold opinions, communicate and acquire information ◇ State negative obligation not to interfere and positive obligation to facilitate the exercise of the right

AI-related risks	Mitigation measures
AI-enabled surveillance tools may lead to a chilling effect especially for minority groups, who may refrain from expressing their opinions.	Avoiding a targeted use of such tools on minorities and marginalised communities, to not deter them from publicly expressing their opinions.
Data stored and recorded by AI tools may be non-easily accessible for individuals who want to exercise their right to information.	Ensuring the information stored by the AI systems is available, understandable and easily exportable.

4. Policy and Research Recommendations

4.1 Relevant Policy Processes and Frameworks

Two levels of policy need to be considered with respect to the use of AI technology by policing and law enforcement: 1) General policy frameworks on AI technology and its use that address all sectors, and 2) specific policy frameworks directly relating to P&LEA's use of AI.

The most important general policy framework on AI in the EU is the '*Proposal for a Regulation laying down harmonised rules on artificial intelligence'* [2] first published by the European Commission in April 2021 and currently in discussion between the Commission, the Council and the Parliament. Widely known as the **EU Artificial Intelligence Act** (EU AI Act), it proposes the first ever EU legal framework on AI and aims to ensure that AI systems placed and used on the internal European market are safe and respect fundamental rights (see Figure 6). To achieve these objectives, the EU AI Act takes a risk-based approach, categorizing different uses of AI into four levels: Unacceptable Risk, High Risk, Limited Risk and Minimal Risk. While AI systems of unacceptable risk (e.g., manipulative systems or social scoring) are prohibited, AI systems of high risk (e.g., systems for border control management, tools to detect deep fakes, or profiling tools used for crime analytics) are permitted but subject to specific requirements and ex-ante conformity assessments. On the other hand, AI systems of limited risks (e.g., chatbots) are only subject to minimal transparency obligations that ensure that persons interacting with the system are informed that they are interacting with an AI, while AI systems with minimal risk (e.g., video games or spam filters) are not subject to any obligations. Crucially, most AI systems with intended use by P&LEAs, like systems used for polygraphs, for detection of deep fakes, for evaluation of the reliability of evidence, for prediction of the occurrence of a criminal offence, for profiling of persons, and for crime analytics, fall under the high risk category.

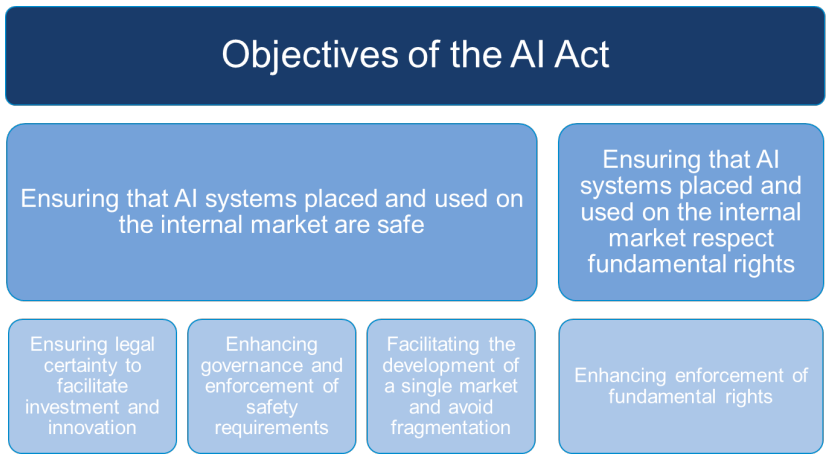


Figure 6: Objectives of the EU AI Act

Another important general policy framework is the final report of the Special Committee on Artificial Intelligence in a Digital Age [3], which was adopted by the European Parliament in May 2022 and aims to establish an artificial intelligence roadmap for up to 2030, with more than 150 policy recommendations on governance, data sharing, digital infrastructure, investment, e-health, e-governance, industry and security.

More specific policy frameworks that directly relate to the use of AI by P&LEA include the study on *'Artificial Intelligence and Law Enforcement: Impact on Fundamental Rights'* [4] commissioned by the EU LIBE Committee in July 2020. It examined the use of AI technology in predictive policing, facial recognition, border security and its use in the wider criminal justice system. The study concluded with a suggestion of six policy recommendations.

Perhaps the most significant EU policy relating to the use of AI for policing and law enforcement, at least until the final passing of the EU AI Act and its introduction of the legal framework, is the Resolution on *'Artificial Intelligence in criminal law and its use by the police and judicial authorities in criminal matters'* [5] passed in October 2021. The resolution stated that *"the relationship between protecting fundamental rights and effective policing must always be an essential element in the discussions on whether and how AI should be used by the law enforcement sector…"* and details the potential ethical and human rights risks associated with police and law enforcement utilization of AI technology, calling for transparency, human centric systems and appropriate governance and legal frameworks. It concludes by calling for *"comprehensive guidelines, recommendations and best practices in order to further specify the criteria and conditions for the development, use and deployment of AI applications and solutions for use by law enforcement and judicial authorities…"* and highlights the need, *"to consider whether specific legislative action on further specifying the criteria and conditions for the development, use and deployment of AI applications and solutions by law enforcement and judicial authorities is needed."*

To summarize, there is a range of EU policy relevant to the use of AI technology by P&LEAs. Soon, these instruments are likely to be joined by binding EU legislation in the form of the EU Artificial Intelligence Act.

4.2 Recommendations

Based on those ongoing policy processes, discussions with experts from policing and law enforcement, research (including ethicists), industry, and policy during ALIGNER workshops in 2021 and 2022, as well as results from research and policy events jointly conducted with the EU AI cluster (ALIGNER, popAI, STARLIGHT, AP4AI), six initial policy recommendations could be derived. Table 2 provides a systematic overview of these recommendations. The overview adapts the policy ontology originally developed by popAI [6], identifying for each recommendation at what level (Societal, Regulatory, Organisational, or Research) a recommendation should be implemented, whether the recommendation is reactively () targeting the current state-of-play or proactively () anticipating new policy actions, who is the target audience for the recommendation, and which themes / aims are addressed by the recommendation. The recommendations are then described in more detail in the remainder of the section.

The ALIGNER project team graciously acknowledges that parts of these recommendations and their detailed descriptions were first published by colleagues from the popAI project in [6], while the initial ALIGNER policy recommendations were first published in September 2022 as part of ALIGNER D2.3 [7]. The ALIGNER and popAI project teams have since worked together to harmonize their recommendations. They presented these harmonized recommendations for the first time at a joint ethics event co-organized between DG Home, ALIGNER, AP4AI, popAI, and STARLIGHT in January 2023. The ALIGNER team has now iterated these recommendations again for publication in the roadmap.

4.2.1 Recommendation overview

No.	Recommendation	Implementation Levels	Type	Target audiences	Themes / Aims
1	Provide common guidelines and unbiased specialist support to P&LEAs for the development, procurement, deployment, and use of AI technology.	Regulatory, Organisational		EU Parliament, European Commission, Member State Parliaments, Ministries, P&LEAs	Fairness, Transparency, Equality, Privacy, Human Rights, Non-Discrimination, Minimize misuse, AI Applicability
2	Establish unified frameworks for the evaluation of AI tools during development and deployment ensuring their ethical, legal, and societal compliance.	Regulatory, Organisational, Research		EC DG Home, EU Parliament, European Commission, Research Institutes, Industry, P&LEAs	Fairness, Transparency, Equality, Privacy, Human Rights, Non-Discrimination, Trustworthy AI
3	Review existing and establish new legal mechanisms to ensure that AI systems and their use are ethical, legal, and societally acceptable.	Regulatory		EU Parliament, European Commission, Member States Parliaments	Fairness, Transparency, Equality, Privacy, Human Rights, Non-Discrimination, Minimize misuse, Trustworthy AI¸ AI Applicability
4	Develop meaningful dialogue between regulators, P&LEAs, researchers, industry, and civil society organisations to strengthen citizens' confidence in the use of AI tools by P&LEAs.	Regulatory, Organisational, Research, Societal		Member States Parliaments, Ministries, P&LEAs, Research Institutes, Industry, Civil Society Organisations	Diversity, Transparency, Social Inclusion, Awareness, Trustworthy AI
5	Support and invest in the development of guidelines for gender-sensitive and gender-responsive policing in the AI era.	Regulatory, Organisational, Societal		EC DG Home, Ministries P&LEAs	Diversity, Equality, Social Inclusion
6	Extend and adapt European and national research programmes to better facilitate evidence-based, participatory research into P&LEA needs regarding AI, the potential implications of the use of AI by P&LEA, and potential criminal use of AI.	Regulatory, Research		European Commission, Ministries / National Funding Agencies, Research Institutes, Civil Society Organisations	Social Inclusion, Trustworthy AI, AI Applicability

Table 2: Overview of policy recommendations

4.2.2 Recommendations in detail

Recommendation 1

Provide common guidelines and unbiased specialist support to P&LEAs for the development, procurement, deployment, and use of AI technology.

Interactions during multiple activities of the EU AI Cluster comprised of ALIGNER, AP4AI, popAI, and STARLIGHT, including exchanges with other projects (see Annex A), survey results (see section 2.2 and Annex B), as well as other research activities [8] highlight the need for and the lack of clear guidelines for P&LEAs regarding the development, procurement, deployment, and use of AI technologies. This includes, first and foremost, guidance on the reliable evaluation of the ethical, legal, and societal implications of the use of AI (see also recommendation 2), supporting effectiveness of AI evaluations by moving away from a black box approach towards explainable AI, as well as target-group-specific training.

A specific issue in the development and deployment of AI relates to data protection and the necessary trade-off between protecting personal, sensitive data and the need for large 'real-world' datasets for training applicable AI models. Specific guidance on how to ensure data protection while simultaneously allowing for training AI models with real-world applicability is very much needed.

However, guidelines alone will not be sufficient. The complex, dynamic, but at the same time highly regulated environment in which P&LEAs operate requires that they have access to unbiased, specialist support during the development, procurement, deployment, and use of AI technologies. To achieve this, the EU and Member States should establish a European network of multidisciplinary trustworthy AI support centres to support P&LEAs with choosing, procuring, and integrating AI technologies. On a European level, Europol and its EU Innovation Hub for Internal Security[12] might be the prime target to establish such a centre where P&LEA can safely test and evaluate AI technologies in clearly defined 'sandboxes'. However, this support centre needs to be complemented by national centres to lower hurdles for engagement (e.g., due to language barriers). Such centres need to be independent entities, funded nationally and not dependent on other funding mechanism, that can then provide a form of external certification for AI technologies, also covering algorithm audits and evaluations of the extent to which systems use "democratic" data in addition to "robust" algorithms.

Critically, these support centres should also act as societal nodes where different actors affected by AI technologies (i.e., civil society organisations) as well as specialists in ethics, law, and AI development engage in discussions with P&LEAs on whether, how, and when to employ which AI technology (see also recommendation 4). For this reason – and to provide a neutral testing ground – these support centres should explicitly not develop AI technologies themselves.

Without such guidance and support there is a high risk of abuse and/or misuse of AI technologies leading to stigmatization, discrimination and potential violence of privacy and human rights. As such it is important that the EU and Member States encourage and support the development of clear guidelines and support structures for the use of AI technologies by P&LEAs.

[12] https://www.europol.europa.eu/operations-services-innovation/innovation-lab/eu-innovation-hub-for-internal-security

Recommendation 2

Establish unified frameworks for the evaluation of AI tools during development and deployment, ensuring their ethical, legal, and societal compliance.

The guidelines and support needed to ensure ethical, legal, and societal compliance, as well as the actual applicability of AI technologies, need to be grounded in evidence-based, unified evaluation frameworks. Given the special role of P&LEAs within society, such assessment frameworks will need to follow a broader approach to impact assessment. As identified by popAI, the literature proposes several AI tool assessment frameworks[13],[14],[15],[16] as well as methods that provide indicators of risks a company might face when adopting an AI tool, while also including mitigation actions and best practices that might be followed. Each of these frameworks includes different guidelines, assessment criteria and mitigation recommendations concerning the adoption of AI. However, most of them focus on the private sector, resulting in a lack of assessment frameworks and clear implementation procedures that provide guidelines, recommendations, and mitigation indicators for the adoption of AI tools in the public sector (see also recommendation 1). The AP4AI Framework for assessing the accountability of AI systems as well as the ALIGNER Fundamental Rights Impact Assessment [9] (which is based on the MAGNETO[17] Ethical Risk Assessment Form) take steps in this direction but need to be further aligned with other frameworks.

Therefore, there is an ongoing need for more extensive research both on the development of such frameworks and the development of the corresponding interdisciplinary assessment measures/metrics. With such frameworks, the adoption of an AI tool can be evaluated against a set of interdisciplinary metrics, developed in an inclusive manner, including the system scope, performance, usability, data used for training and evaluation including ethical processing, human rights impact, as well as ensuring compliance with data protection. Such frameworks should also include specific guidelines on mitigating bias of AI models and datasets.

Recommendation 3

Review existing and establish new legal mechanisms to ensure that AI systems and their use are ethical, legal, and societally acceptable.

Operative guidelines for the development, procurement, deployment, and use of AI technologies, based on evidence-based, unified evaluation frameworks, will need to be flanked by binding legal mechanisms to ensure that these technologies are ethical, legal, and societally acceptable. The EU AI Act is a step in this direction, although based on numerous discussions with representatives from P&LEA, civil society, research, industry, and policy, there remain valid concerns from different actors on its definition of AI (too broad), the exemptions included for high-risk AI technologies (too many), and its affect when

[13] High-Level Expert Group (HLEG) - Assessment List for Trustworthy Artificial Intelligence (ALTAI): https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment
[14] World Economic Forum (WEF) AI Governance framework: https://www.weforum.org/projects/model-ai-governance-framework
[15] NOREA Guiding Principles Trustworthy AI investigation: https://www.norea.nl/uploads/bfile/a344c98a-e334-4cf8-87c4-1b45da3d9bc1
[16] AI Assessment Catalog of Fraunhofer IAIS: https://www.iais.fraunhofer.de/en/research/artificial-intelligence/ai-assessment-catalog.html
[17] https://www.magneto-h2020.eu/

put into place (too bureaucratic). A valid approach to alleviate these concerns might be the development of a P&LEA-specific AI directive (similar to the Law Enforcement Directive [10]).

Regardless of these issues, any legal mechanism on EU and national level related to the use of AI technologies by P&LEAs needs to ensure that there is always a competent and knowledgeable 'human in the loop' if AI technology is used in critical decision-making processes. The nature of the work carried out by P&LEAs, its impact on individuals and on wider society require that AI technology should not directly replace human decision making. Without this safeguard, all the checks and balances that are intrinsic to decision making in P&LEAs cannot occur or are compromised, e.g., the fundamental issues of accountability, explicability, transparency, and compliance with the rule of law. Even if an AI technology does not directly take decisions, but only informs a human operator, the information provided via the AI technology has the potential to influence the decision. As such, it becomes of utmost importance that the data and information on which the AI technology is trained, tested, validated and used is accurate and does not perpetuate existing biases and stereotypes present on society.

Legal mechanisms in the EU and nationally should support the continuous, inclusive, and multidisciplinary monitoring of AI technology across their lifecycle. In particular, EU Member States should invite civil society organisations and create joint working groups, which will check the individual AI technologies used by P&LEAs to highlight potential issues from such usage (a posteriori monitoring and assessment). These joint working groups should also be consulted when designing and developing new AI technologies that will be applied in the future (a priori monitoring and assessment). The purpose is to improve and adapt these technologies appropriately to ensure that they protect citizens' rights. This will support the use of existing technologies, as well as the development of new ones to cover the current needs. This interaction between different actors related to the use of AI technologies by P&LEAs should be continuous (e.g., via the AI support centres suggested in recommendation 2) and should strengthen the involvement of civil society in all stages of the operation of an AI technology (design, implementation, maintenance, upgrade).

To facilitate this interaction, the European Commission and EU Member States need to better promote and ensure citizens' awareness regarding the existence and implementation of an AI technology and enable objection to potential unjust decisions.

Open discussions between different actors related to the use of AI technologies by P&LEAs can support transparency at every stage to minimize the risks of discrimination. In addition, this should also be considered in the procurement of systems, where, for example, the technical specifications must be accepted by civil society organizations and agencies, while monitoring and assessment by representatives of social and other bodies should be foreseen in the system implementation phase.

Recommendation 4

Develop meaningful dialogue between regulators, P&LEAs, researchers, industry, and civil society organizations to strengthen citizens' confidence in the use of AI tools by P&LEAs.

Civil society organisations are often not included in consultations regarding the employment of AI by P&LEAs. Therefore, they express their concerns on emerging risks through announcements and legal actions. This gap is creating tensions that are constantly widening and damage the trust between the involved parties.

To repair the trust issues, civil society organisations should be involved in open dialogues with European and national regulators, P&LEAs, researchers, and industry regarding the employment of AI technologies. The results of such activities would enable European Member States to integrate European regulations (see recommendation 3) into their law, tailoring it to the culture and the specificities that govern their societies. Civil society organisations should be actively involved in the process of designing and implementing AI technologies, as well as in the monitoring of the existing ones. They should also determine the best way to operate these systems to ensure human rights and generate acceptance across citizens.

Recommendation 5

Support and invest in the development of guidelines for gender-sensitive and gender-responsive policing in the AI era.

This recommendation aims at the development of corresponding guidelines for the promotion of gender-sensitive and gender-responsive policing[18] [19], especially in the era of AI. In 2010, the Women Police Officers Network (WPON)[20] was established with the support of Southeast Europe Police Chiefs Association. Its scope was to place gender-sensitive policing at the top of the agenda of police reform and to serve as a platform for knowledge and experience exchange across police services, needs and priorities of policewomen. This network has so far achieved gender-sensitive policing with an emphasis on recruitment, selection, and professional development of women in police services. However, apart from this initiative, it is important in today's developed society to promote and develop appropriate actions and guidelines on the equality of all people in society to ensure no group is disadvantaged over another in its treatment by the police.[21]

This policy recommendation aims at the development of the corresponding guidelines, from the EU and the relevant EU-funded projects, to raise awareness on the position of women in police services and the development and implementation of sustainable solutions for the improvement of recruitment and retention of women personnel and their active involvement in the design and development of AI systems for security purposes. In addition to gender-sensitive policing, the aim is to achieve gender-responsive policing, which means taking into account *"the needs of all parts of the community, women and girls, men and boys including minority or marginalised groups […] to ensure no group is disadvantaged over another in its treatment by the police"*[22]. To achieve both, the suggested guidelines should focus on the empowerment of gender equality in law enforcements with an emphasis on the needs of all parts of the community and facilitate the inclusive design and development of the corresponding AI technologies to ensure that no group is mistreated by the police. Furthermore, these guidelines shall be based on the outcomes of the WPON and the Southeast Europe Police Chiefs Association that proved that the absence of data leads to ineffective policies and legal frameworks, and that it is necessary to include the appropriate information so that gender-sensitive policing can be enhanced.

[18] Women, U. N. (2021). Handbook on gender-responsive police services for women and girls subject to violence.
[19] Bonkat-Jonathan, L., & Ejalonibu, G. L. (2021). A Review of Some Discriminatory Laws against Women and the Need for Legislative-Gender Responsive Actions in Nigeria.
[20] Kekić, D., Đukanović, D., & Tomić, M. Women Police Officers Network (WPON).
[21] This and the following paragraph were first published by popAI in [6].
[22] International Association of Women Police, Gender-responsive policing, https://www.iawp.org/Gender-Responsive-Policing-GRP.

Recommendation 6

Extend and adapt European and national research programmes to better facilitate evidence-based, participatory research into P&LEA needs regarding AI, the potential implications of the use of AI by P&LEA, and potential criminal use of AI.

EU- and nationally funded security projects, and specifically those developing AI driven technologies, have often raised concerns, see for example the FP7 project INDECT "Intelligent information system supporting observation, searching and detection for security of citizens in urban environment"[23], which sparked concerns among Members of European Parliament calling on the European Commission to clarify its purpose[24]. The – sometimes overly restrictive – secrecy of such projects and lack of publicly available information, together with the perceived potentially negative impact on civil liberties and fundamental rights call for new approaches towards accountability. One way to address these issues, while maintaining the required level of security, would be the establishment of specialised interdisciplinary Ethics and Legal Committees that review proposals and ongoing research projects in the security domain on a continuous basis, so as to prevent potentially serious ethical, societal, and legal issues as well as abuse of human rights. Aligned with recommendations 1 and 2 these Committees should have ethical, legal, technical, organisational, and practical capabilities to assess an AI technology's ethical, legal, and societal compliance. This could act as a form of internal certification for research projects in relation to an AI technology's accountability and the ethical, inclusive and secure-by-design AI systems in the course of research and development.

In addition, research conducted in the context of the H2020 project popAI identified the stakeholder groups involved in the research, development, use, and implementation of AI technology, as well as those who promote awareness regarding emerging risks, and push for relevant policies. These different categories of stakeholders should not be seen as "rivals" but rather as key components of a unified ecosystem that co-shape the development and use of AI in the security domain. The identified stakeholders are namely, LEAs, social and humanities research, policy makers, government and public bodies, technologists / data scientists, civil society organizations, national and local authorities, ICT and software companies, and police academies. Mapping EU-funded projects in the security domain, 348 different stakeholders were collated with the majority of stakeholders being ICT and software companies, followed by universities and research organisations. It is recommended that the EC explores ways (i.e., call requirements, specifications) for EU-funded projects to include civil society organisations in the early stages of the AI technology design and development as they are underrepresented in the project consortia, while their voices are very important to preserve privacy and human rights. Likewise, project partners were geographically mapped. The analysis indicated that various European countries such as Albania, Denmark, and Ukraine have been underrepresented to date in EU-funded projects in the security domain. Involvement of partners from underrepresented Member States would enable the inclusion of potentially cultural and geographic differences regarding the needs and acceptance of AI systems. Thus, it is recommended that the EC explores ways (i.e., call requirements, specifications) for EU-funded projects to include underrepresented Member States in the AI design and development.[25]

---

[23] INDECT (Intelligent Information System Supporting Observation, Searching and Detection for Security of Citizens in Urban Environment), Cordis Project Page.
[24] Euractiv (2011), "MEPs question 'Big Brother' urban observation project".
[25] This paragraph was first published by popAI in [6].

Lastly, the implementation of recommendations 1-5 needs to be supported by further AI-specific research in the security domain. This includes the development of guidelines aligned with the needs of P&LEAs (recommendation 1), assessment frameworks (recommendation 2), an evaluation of the existing legal mechanism as well as their effects on P&LEA work (recommendation 3), stakeholder engagement techniques in the context of AI technologies for P&LEAs (recommendation 4), as well as guidelines for gender-sensitive and gender-responsive policing (recommendation 5). This also includes additional research into countering criminal use of AI technologies and employing AI technologies in support of P&LEAs in an ethical, legal, and societally acceptable way.

5. AI Technology Catalogue

This section provides a detailed description and an initial assessment of the AI technologies with relevance for the first ALIGNER narrative. For each technology, a brief description is provided. In addition, each entry provides information on

- **Effectiveness** – A rough estimate in the short-term on effectiveness and performance described in non-technical language.
- **Robustness** – An assessment in the short-term on how robust the technology is for being able to handle counter measures, data quality issues and out-of-distribution examples (examples of a type it has not been trained on).
- **Development** – A mid-term perspective of what the current development efforts are and who are doing it. A general assessment of where the technology is heading within the next few years.
- **Projected future** – Long-term perspective of where this technology may end up a few years from now.
- **TRL** – An assessment of maturity using the simplified Technology Readiness Level scale.
- **Categorisation** – A categorisation of the technology using structured models. A mapping to known classes of technologies indicates capabilities the technology may support.

The assessment uses admiralty code: confirmed, probably true, possibly true, doubtful, improbable, cannot be judged

5.1 Deanonymization – Authorship attribution

Authorship attribution is the task of identifying the author of a given text document within a set of possible candidates.

A set of relevant textual features are used to create a ”fingerprint” of the author. This ”fingerprint” can be matched against a given set of candidates.

Examples of crimes where authorship attribution is important are illegal drug marketing, online threats, and extremism propaganda.

Authorship attribution has shown promising results for e-mails, forum posts, tweets, and blog posts.

Effectiveness (short-term perspective): High

Can reliably find the matching author for a variety of textual content. However, the accuracy is not sufficient to use it as evidence in courts.

Robustness (short-term perspective): Medium

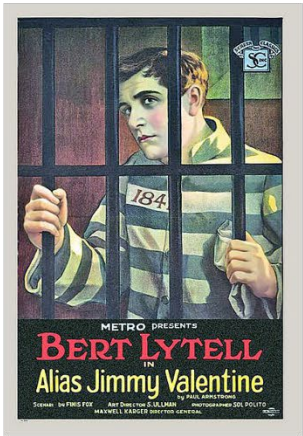
Can only match authors within a set of known candidates. Open authorship attribution where the author may not be among the candidates is much harder. *Probably* requires reasonably sized samples of written texts from all candidates.

Development (medium-term perspective): Active

Active commercial and academic research by many different actors. Recent efforts directly use source material to implicitly learn relevant features. This improves performance considerably compared to previous approaches.

Projected future (long-term perspective): Promising

Future development that directly use source material will improve performance of authorship attribution. Since performance is already very high, the accuracy will be close for acceptance as evidence in courts.



Picture from Wikimedia, Creative Commons 4.0

Currently TRL 4-6

Capabilities:

- Digital forensics
- Prevention and detection
- Reaction and response

Technologies:

- Machine learning
- Classification

5.2 Deanonymization – Geolocalisation of images

Gelocalisation of images is the task of locating where an image was taken on earth when location metadata is missing or is incomplete. The task requires comparison of the target image with millions of images with location metadata to find the corresponding location. Automated tools are necessary for geolocalisation of images since humans perform poorly on this task.

Gelocalisation of images has improved considerably with recent AI techniques that identifies distinguishing features among huge amounts of images. Another trend is that the aerial perspective from publicly available satellite images is increasingly used to supplement ground level images. Recent developments combine the two perspectives for remarkable performance on a city scale.

Effectiveness (short-term perspective): High

Accuracy is highly dependent on the size of the geographic area. For city size areas, one kilometre precision is often possible with sufficient accuracy. Geolocalisation on earth is more difficult, especially of images with few features.

Robustness (short-term perspective): Medium

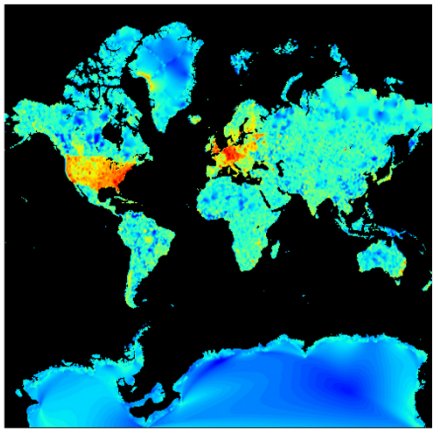
Robustness is highly affected by the trade-off between the geographic area size and accuracy. Huge variation in capture time and weather is also a problem. Information about scene type and context improves robustness.

Development (mid-term perspective): Active

Active commercial and academic research by many different actors. Recent efforts combine ground and satellite images. This improves performance considerably compared to only using ground images.

Projected future (long-term perspective): Promising

Image databases from social media and satellites will increase over time. Future development will increase performance and have high precision for even larger geographic areas.



Picture from Wikimedia, Creative Commons 4.0

Currently TRL 4-6

Capabilities:

- Digital forensics
- Prevention and detection
- Reaction and response

Technologies:

- Machine learning
- Classification

5.3 Veracity assessment – Disinformation detection

Disinformation detection is the task of detecting fraudulent information that is intentionally spread to mislead people. Social media makes it easy to quickly spread large amounts of fraudulent information. The information may even be automatically generated in ways that are difficult for humans to detect. Automatic veracity assessment is necessary since manual assessment is costly and time-consuming.

Automatic veracity assessment consists of identifying claims that require assessment, finding sources that support or refute the claim, and assessing veracity using these sources. Recent efforts directly use source material to implicitly learn relevant features for all stages of veracity assessment. This improves performance considerably compared to previous approaches that only used contextual information (author, place of publication).

Discovery of the underlying intent may require an aggregated judgement from several detections of fraudulent information.

Effectiveness (short-term perspective): Medium

Effectiveness in detection of fraudulent information is *probably* highly context dependent. The effectiveness is high in simple contexts (reviews) and moderate in complex contexts (scientific facts).

Robustness (short-term perspective): Low

Most approaches only use one source (often Twitter). Robust detection of fraudulent information likely requires comparison of multiple sources.

Development (mid-term perspective): Very active

Active commercial and academic research by many different actors. Detection of fraudulent information is if of interest for news agencies, public health, and businesses. Automatic selection of instances to label simplifies creation of datasets.

Projected future (long-term perspective): Hard to assess

Bots that automatically generate fraudulent information will make disinformation detection even more important. *Probable* to be an arms race between generation and detection of fraudulent information.

Picture from Wikimedia, Creative Commons 4.0

Currently TRL 1-3

Capabilities:

- Digital forensics
- Prevention and detection
- Reaction and response

Technologies:

- Machine learning
- Classification

5.4 Detection of synthetic images

Today it is often impossible for a human to tell if an image has been computer generated. Therefore we need tools to aid us in this task that automatically detect synthetic content in visual images.

Most detectors are *probably* only usable on synthetic images that are generated by a specific algorithm. A new detector is *probably* needed for each new generative algorithms. Detectors have to be updated frequently due to the rapid development of generative algorithms. Detectors can be updated using either in-house expertise or by subscription to such a service.

Alternative countermeasures to synthetic images could include strong authentication techniques, such as block chains, which *probably* provide sufficient protection (but only for some cases).

Although detection of synthetic images has its limitations, the detectors will *possibly* succeed against less sophisticated actors who rely on out-of-the-box models (pre-trained and downloadable or available as a service) and are not able to modify them on their own.

Effectiveness (short-term perspective): Medium

Synthetic images can be detected if they have been generated by algorithms that the detector is trained on. Some detectors can detect synthetic images that are generated with unknown algorithms (out-of-distribution images).

Robustness (short-term perspective): Low

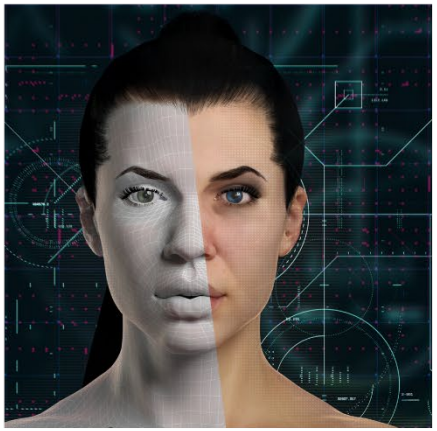
Simple perturbations (cropping, compression, noise) reduce the likelihood of detecting synthetic images. There is currently no countermeasures to such perturbations. Synthetic images can also be tailored to avoid detection by known detectors.

Development (mid-term perspective): Very active

Active commercial and academic research by many different actors. Detection of synthetic images are of interest for news agencies and providers of images/photos. Likely to improve within a few years.

Projected future (long-term perspective): Hard to assess

Probable to be an arms race between generation and detection of synthetic images. AI-based tools are likely the only viable option for automatic detection of synthetic images.



Shutterstock, used with License

Current maturity: TRL 4-6

Capabilities:

- Digital forensics
- Prevention and detection
- Reaction and response

Technologies:

- Machine learning
- Classification

5.5 Detection of synthetic video

Today it could be impossible for a human to tell if a video has been computer generated. Therefore we need tools to aid us in this task that automatically detect synthetic content in video.

A detector may use hand-crafted features, data-driven features, unique ”fingerprints” of a generative algorithm, or artefacts in eye blinking, lip synching, facial landmarks, vocabulary, combinations of word classes or sound frequencies of speech. The features and artefacts are small enough that a human will not necessarily detect them.

Detectors have to be updated frequently due to the rapid development of generative algorithms. Detectors can be updated using either in-house expertise or by subscription to such a service.

Alternative countermeasures to synthetic videos could include strong authentication techniques, such as block chains, which *probably* provide sufficient protection (but only for some cases).

Effectiveness (short-term perspective): Low

Today it is often possible to detect synthetic video with reasonable performance for well known generative algorithms. However, the detectors do not generalise well to synthetic videos that are generated with unknown algorithms.

Robustness (short-term perspective): Low

Simple perturbations (cropping, compression, noise) reduce the likelihood of detecting synthetic videos. There is currently no countermeasures to such perturbations. Changing the generative algorithm will often thwart detection.

Development (mid-term perspective): Very active

Active commercial and academic research by many different actors. Detection of synthetic videos are of interest for news agencies and providers of videos. Likely to improve within a few years.

Projected future (long-term perspective): Hard to assess

Probable to be an arms race between generation and detection of synthetic videos. AI-based tools are likely the only viable option for automatic detection of synthetic videos.

Picture from Wikimedia, Creative Commons 4.0

Currently TRL 1-3

Capabilities:

- Digital forensics
- Prevention and detection
- Reaction and response

Technologies:

- Machine learning
- Classification

6. What comes next?

The next ALIGNER roadmap will be published at the end of September 2023. It will contain an updated set of AI technologies for the first narrative, including their expert impact assessment in terms of technological risks, as well as ethical and legal implications and how to address these. In addition, the next roadmap will also include three new narratives that investigate AI and cybercrime (one focusing on targeting individuals while another narrative focuses on organizations as targets of cybercrimes) as well as AI, robots, drones, and vehicles. Lastly, the next roadmap will look at challenges and potential unintended consequences and provide an initial taxonomy of AI supported crime.

7. References

[1] L. Clutterbuck, “ALIGNER D2.2 Archetypical Scenarios and their Structure,” H2020 ALIGNER, GA no. 101020574, 2022.

[2] European Commission, “Proposal for a Regulation of the European Parliament and of the Council for Laying down harmonized rules on Artificial Intelligence and amending certain union legislative acts, COM(2021) 206 final,” 21 April 2021. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52021PC0206&from=EN. [Accessed 31 03 2023].

[3] European Parliament, “Artifdicial intelligence in a digital age, European Parliament resultion of 3 May 2022, 2020/2266(INI),” 3 May 2022. [Online]. Available: https://www.europarl.europa.eu/doceo/document/TA-9-2022-0140_EN.pdf. [Accessed 31 03 2023].

[4] G. G. Fuster, “Artificial Intelligence and Law Enforcement: Impact on Fundamental Rights,” European union, Brussels, 2020.

[5] European Parliament, “Artificial intelligence in criminal law and its use by the police and judicial authorities in criminal matters , 2020/2016 (INI),” 6 October 2021. [Online]. Available: https://www.europarl.europa.eu/doceo/document/TA-9-2021-0405_EN.pdf. [Accessed 31 03 2023].

[6] Ziouvelou et al., “popAI Policy Brief - 1st Year,” Deliverable D1.6, H2020 popAI, GA no. 101022001, 2022.

[7] L. Clutterbuck, R. Warnes and I. Marsh, “ALIGNER D2.3 Policy recommendations,” H2020 ALIGNER, GA no. 101020574, 2022.

[8] J. Laufs and H. Borrion, “Technological innovation in policing and crime prevention: Practitioner perspectives from London,” *International Journal of Police Science & Management,* pp. 1-20, 2021.

[9] D. Casaburo and I. Marsh, “ALIGNER D4.2 Methods and guidelines for ethical & law assessment,” H2020 ALIGNER, GA no. 101020574, 2023.

[10] European Parliament and Council of the European Union, “Directive 2016/680 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purpose of the prevention, investigation, detection or prosecution of criminal offences,” 4 May 2026. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32016L0680. [Accessed 31 03 2023].

[11] UNSD, “Methodology - Standard country or area codes for statistical use (M49),” [Online]. Available: https://unstats.un.org/unsd/methodology/m49/. [Accessed 14 09 2022].

Annex A: Projects and Initiatives Mapping

Name	Brief Description	Website
AIDA - Artificial Intelligence and advanced Data Analytics for Law Enforcement Agencies	AIDA will develop a Big Data Analysis and Analytics framework equipped with a complete set of effective, efficient and automated data mining and analytics solutions to deal with standardised investigative workflows, extensive content acquisition, information extraction and fusion, knowledge management and enrichment through novel applications of Big Data processing, Machine Learning, AI and predictive and visual analytics. It will do so in a way that ensures societal benefits and consequences are integral part of design and deployment efforts.	https://www.project-aida.eu/
AP4AI – Accountability Principles for Artificial Intelligence in the Internal Security Domain	The AP4AI will create a global framework for AI accountability for policing, security and justice. This framework will be grounded in empirically verified accountability principles for AI as carefully researched and accessible standard, which supports internal security practitioners in implementing AI and Machine Learning tools in an accountable and transparent manner and in line with EU values and fundamental rights.	https://www.ap4ai.eu/
ARCSAR -Arctic and North Atlantic Security and Emergency Preparedness Network	Addresses the Arctic and North-Atlantic (ANA) region, preparing to cope with the security and safety threats that will result from increased commercial activity in the region including traffic through the northern passages, cruise traffic, and offshore oil and gas activity	https://arcsar.eu/
ARESIBO - Augmented Reality Enriched Situation awareness for Border security	The top priorities of ARESIBO will be scientific excellence and technological innovation. It will enhance the current state-of-the-art through technological breakthroughs in Mobile Augmented Reality and Wearables, Robust and Secure Telecommunications, Swarm Robotics and Planning of Context-Aware Autonomous Missions, and Artificial Intelligence (AI), in order to implement user-friendly tools for border and coast guards.	https://www.aresibo.eu/
CC-DRIVER - Understanding the drivers of cybercriminality, and new methods to prevent, investigate and mitigate cybercriminal behaviour	The CC-DRIVER project seeks to understand the drivers of cybercriminality and researches methods to prevent, investigate and mitigate cybercriminal behaviour.	https://www.ccdriver-h2020.com/project

CONNEXIONS - InterCONnected NEXt-Generation Immersive IoT Platform of Crime and Terrorism DetectiON, PredictiON, InvestigatiON, and PreventiON Services	CONNEXIONs aims to develop and demonstrate next-generation detection, prediction, prevention, and investigation services. These services will be based on multidimensional integration and correlation of heterogeneous multimodal data, and delivery of pertinent information to various stakeholders in an interactive manner tailored to their needs, through augmented and virtual reality environments.	https://www.connexions-project.eu/
CREST - Fighting Crime and TerroRism with an IoT-enabled Autonomous Platform based on an Ecosystem of Advanced IntelligEnce, Operations, and InveStigation Technologies	CREST's overall objective is to improve the effectiveness and efficiency of LEAs intelligence, operation, and investigation capabilities, through the automated detection, identification, assessment, fusion, and correlation of evidence acquired from heterogeneous multimodal data streams	https://project-crest.eu/
CYCLOPES Fighting Cybercrime – Law Enforcement Practitioners' Network	CYCLOPES establishes a Europe-wide network to combat cybercrime.	https://cyclopes-project.eu
D4FLY - Detecting Document frauD and iDentity on the fly	The project focuses on enhancing the quality and efficiency of identity verification at border crossings in all modalities: land, air, and sea by providing faster and more secure border control solutions.	https://d4fly.eu/
DARENET Danube river region Resilience Exchange Network	DAREnet is building a dynamic multi-disciplinary community of practitioners, operating in a network of civil protection organisations. The network is supported by a broad range of stakeholders from policy, industry and research. Together they build an interdisciplinary ecosystem to foster synergies, innovation and its uptake across the Danube Region.	www.darenetproject.eu/
DARLENE - Deep AR Law Enforcement Ecosystem	Investigating how cutting-edge augmented reality (AR) technology can be deployed to help law enforcement agencies (LEAs) and first responders make more informed and rapid decisions especially in situations where time is of the essence. The project develops innovative augmented reality (AR) tools that aim to improve situational awareness when responding to criminal and terrorist activities	https://www.darleneproject.eu/
eNOTICE European Network of CBRNE Training Centres	The overall goal of the eNOTICE project is to establish a European network of CBRN training, testing and demonstration centres aiming at enhancing CBRN training capacity for improved preparedness and incident response through increased collaboration between CBRN training centres and practitioners' needs-driven CBRN innovation and research.	https://www.h2020-enotice.eu/

EU-HYBNET Empowering a Pan-European Network to Counter Hybrid Threats	The project is the 1st EU initiative which brings together pan-European practitioners and stakeholders to identify and analyse common challenges, and requirements to counter hybrid threats. It conducts research, highlights innovation initiatives, arranges training events to test innovations and makes recommendations for the uptake, industrialisation and standardisation of these innovations.	https://euhybnet.eu/
EXERTER Security of Explosives pan-European Specialists Network	EXERTER will provide practitioners with the operative knowledge and tools for enhancing the security of our society and to highlight innovative methods, tools and technologies, which can contribute in the fight against terrorism and serious crime. The aim is to help practitioners reach an improved capability, as well as to identify needs within standardisation and industrial development connected to Security of Explosives	www.exerter-h2020.eu
EXFILES - Extract Forensic Information for LEAs from Encrypted SmartPhones	EXFILES will use software exploitation, hardware methods and combined methods to give law enforcement officials the tools and protocols for rapid and consistent data extraction in strict legal contexts.	https://exfiles.eu/
Fire-IN Fire and rescue Innovation Network	EU-wide one-stop shop for Fire-& Rescue Faster and cheaper access to the state-of-the-art Fire & Rescue technology for the whole of Europe	https://fire-in.eu/
FORMOBILE - From mobile phones to court – A complete FORensic investigation chain targeting MOBILE devices	Working in collaboration to create an end-to-end mobile forensic investigation chain, striving to improve digital safety, and security in the EU while respecting fundamental rights.	https://formobile-project.eu/
GRACE - Global Response Against Child Exploitation	GRACE aims to equip European law enforcement agencies with advanced analytical and investigative capabilities to respond to the spread of online child sexual exploitation material.	https://www.grace-fct.eu/
I-LEAD Innovation - Law Enforcement Agencies Dialogue	i-LEAD will build the capacity to monitor the security research and technology market in order to ensure a better matching and uptake of innovations by law enforcement agencies with the overarching aim to make it a sustainable Pan-Europan LEA network.	https://i-lead.eu/
ILEANET Innovation by Law Enforcement Agencies networking	ILEAnet aims to build a sustainable organisational Law Enforcement Agency (LEA) practitioners network focused on research & innovation addressing LEA challenges, together with a community of individuals interested to exchange and collaborate in this area. By encouraging such discussion between practitioners and experts from academia and industry, the project will stimulate LEA capabilities to influence, develop and take up research, development and innovation (RDI) that is useful and usable for LEAs, and thus help them to tackle the major challenges they face.	https://www.ileanet.eu/

iMARS - image Manipulation Attack Resolving Solutions	iMARS improves the operational capacity of passport application officers, border guards and forensic experts by providing Image Morphing and manipulation Attack Detection (MAD) solutions, Document Verification and Fraud Detection (DVFD) solutions, and by providing training, guidelines, share best practices and contribute to standardisation.	https://imars-project.eu/
INCLUDING Innovative Cluster for Radiological and Nuclear Emergencies	INCLUDING seeks to provide a full-fledged and comprehensive training in the RN security sector at European level. Starting from the existing training resources of the Partners in the Consortium, in most cases developed in the framework of EC projects, INCLUDING aims to enhance practical know-how and to boost a European sustainable training and development framework for practitioners in the RN Security sector.	https://including-cluster.eu/
INSPECTr - Intelligence Network and Secure Platform for Evidence Correlation and Transfer	The principle objective of INSPECTr will be to develop a shared intelligent platform and a novel process for gathering, analysing, prioritising and presenting key data to help in the prediction, detection and management of crime in support of multiple agencies at local, national and international level	https://inspectr-project.eu/
iProcurenetNet European Procurer Networking for security research services	iProcureNet aims to create an ecosystem of procurers, prescribers, legal advisors and other key stakeholders of security procurement, to share procurement trends and needs, and open pathways for joint procurement.	https://www.iprocurenet.eu/
LOCARD - Lawful evidence collecting and continuity platform development	automate the collection of digital evidence in any electronic format and medium. Its goal is to provide a comprehensive management approach to handle digital evidence to be presented in a court of law, alleviating many issues of current art and practice	https://locard.eu/
MEDEA Mediterranean practitioners' network	MEDEA is an EU funded Coordination and Support Action project the scope of which is to establish and further develop a regional Network of practitioners and other security related actors in the Mediterranean and the Black Sea region.	https://www.medea-project.eu/
METICOS - A Platform for Monitoring and Prediction of Social Impact and Acceptability of Modern Border Control Technology	developing a platform that integrates information systems and networks of data sources in order to validate the efficiency and users acceptance of border control technologies. The proposed platform will provide metrics and KPIs to authorities and decision-makers, based on a number of independent variables: performance expectancy, effort expectancy, facilitating conditions, physical privacy, accuracy, information privacy, ethical and societal perceptions, securing positive societal impact and maximize border control process efficiency	https://meticos-project.eu/
NO-FEAR Network Of practitioners For Emergency medicAl systems and cRitical care	NO-FEAR will bring together a pan-European and beyond network of emergency medical care practitioners, suppliers, decision and policy makers to collaborate and exchange knowledge, good practices, and lessons learned.	http://no-fearproject.eu/

NOTIONES NetwOrk of an intelligence and security practitiOners with iNdustry and academia actorS	The NOTIONES project gathers actors from 15 European countries to develop European intelligence cooperation in the fight against crime.	https://cordis.europa.eu/project/id/101021853
PEN-CP Pan-European Network of Customs Practitioners	PEN-CP is 'a Novel Customs Innovation Boosting Network and On-line Platform to establish a customs practitioner network which facilitates translating customs security research and innovation ideas and requirements into scalable, viable solutions, technologies, and process improvements	https://www.pen-cp.net/
popAI - A European Positive Sum Approach towards AI tools in support of Law Enforcement and safeguarding privacy and fundamental rights	The core vision of pop AI 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)	https://www.pop-ai.eu/
ROXANE - Real time network, text, and speaker analytics for combating organized crime	ROXANNE collaborates with Law Enforcement Agencies (LEAs), industry and researchers to develop new tools to speed up investigative processes and support LEA decision-making. The end-product will be an advanced technical platform which uses new tools to uncover and track organized criminal networks, underpinned by a strong legal framework.	https://www.roxanne-euproject.org/
STARLIGHT - Sustainable Autonomy and Resilience for LEAs using AI against High priority Threats	Law enforcement agencies' (LEAs) data-rich environments provide the opportunity to adopt artificial intelligence tools and capabilities that improve investigatory practices and limit the criminal misuse of AI. Through STARLIGHT, LEAs will collaboratively develop their autonomy and resilience in the use of AI for tackling major criminal threats.	https://starlight-h2020.eu/
SUSQRA - Protection against improvised explosive devices	SUSQRA aims at the development of an expert system to quantitatively assess the extent of damage caused by improvised explosive devices (IEDs) almost without using experiments.	https://www.emi.fraunhofer.de/en/business-units/security/research/susqra-sprengvorrichtungen-praevention-risikoanalyse.html
TAILOR - Foundations of Trustworthy AI – Integrating Reasoning, Learning and Optimization	Purpose of building the capacity of providing the scientific foundations for Trustworthy AI in Europe by developing a network of research excellence centres leveraging and combining learning, optimization and reasoning.	https://tailor-network.eu/

Annex B: Additional information on the online survey

To expand on the information gathered in ALIGNER's workshops and obtain a more comprehensive picture of the capability enhancement needs of law enforcement and policing agencies, a survey was designed and conducted (Figure 7). The main aim of this survey was to gain an understanding of the capability enhancement needs perceived by those working in the field of law enforcement and policing. A further aim was to explore the potential challenges associated with integrating AI into law enforcement and policing further. The target group included practitioners working in this field as well as other professionals, e.g., from research institutions, who are concerned with the topics of AI, law enforcement, and policing.

Aim
- Understand the capability enhancement needs of law enforcement and policing

Target group
- Practitioners and professionals working in the field of law enforcement and policing

Scope
- 25 questions in closed and open format

Timeframe
- 25 May - 25 August 2022

Figure 7: Survey facts.

The survey consisted of a total of 25 questions, some of which were asked in a closed format with predefined answer options and some in an open format. This mixed approach was chosen to ensure an objective evaluation of the results on the one hand (closed questions) and to give participants the opportunity to address additional aspects on the other (open questions). Data-sensitive and personal questions, such as age or gender, were kept optional if this information was not crucial for gaining knowledge. All other questions were either provided with the option to skip the question or tick the "Not sure" option. This approach was chosen to counteract overload, e.g., in case of misunderstanding or not understanding the question, and to support higher data quality. The survey was open from 25 May 2022 on and responses received by 25 August 2022 were included in the roadmap. A three-months period was therefore set for the collection of survey responses. To gather opinions and experiences from the dedicated target group, a snowball sampling method was used. The survey was disseminated among ALIGNER's advisory board members as well as related projects and their respective networks. Additionally, the link was published on LinkedIn.

It is important to note that the survey results only reflect the opinions of the sample studied and that no conclusions can be drawn for the entire population of interest. Furthermore, the identified capability enhancement needs in which AI could be of use are considered from a one-dimensional perspective that does not take into account all the potential consequences that would result from the application of AI in these areas. The initial collection of challenges in the survey scratches some important issues to consider and provides an impetus to discuss these within society as a whole.

Demographic information on the survey sample

The survey was completed by a total of 53 respondents, of whom 16 (32%) were female and 34 (68 %) were male[26] (Figure 8). The age distribution among the participants was quite balanced, with the largest part of the sample (35%) being between 45 and 54 years old and 20 % representing respectively the age groups 25 to 34 years, 35 to 44 years, and 55 to 64 years. A small proportion of the sample (2%) was in the age groups 18 to 24 years and 65 years and older[27] (see Figure 8 for totals).

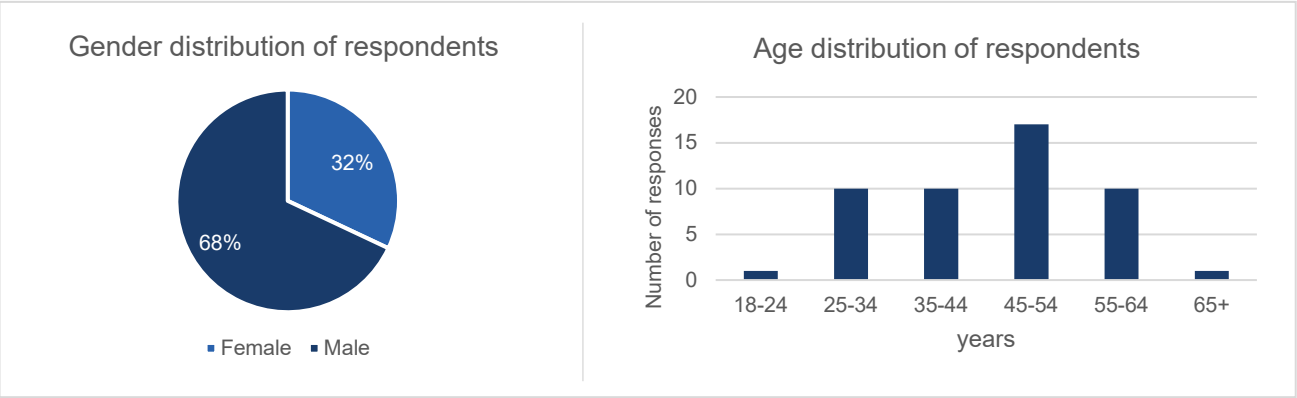


Figure 8: Results of the optional questions “What is your gender?” and “What is your age?”

The sample consisted of 19 people working in law enforcement and policing (as practitioners) and 29 people working in research and academia. Two persons indicated “civil society” and “other” as their work organisation, and one person works in industry (Figure 9). The distribution of countries represented by participants’ work organisations is shown in Figure 10. Most participants (25 persons) were from Southern European countries (Greece, Italy, Kosovo, Portugal, Spain), followed by 14 persons working in Western European countries (Belgium, France, Germany, Netherlands). A proportion of 9 people work in Northern Europe (Estonia, Ireland, Lithuania, Sweden, UK) and 5 persons work in Eastern Europe (Bulgaria, Poland, Slovakia). This loose division into four geographical regions of Europe is based on a methodology of the Statistics Division of the United Nations Secretariat [2].

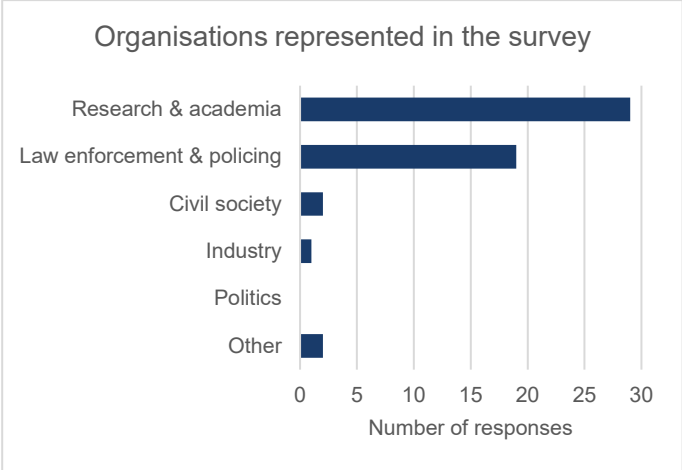


Figure 9: Results of the question “In which type of organization do you work?"

[26] The question about gender was optional, so that n in this question deviates slightly from n total.
[27] The question about age was optional, so that n in this question deviates slightly from n total.

Region	Country	No.
Southern Europe	Italy	7
	Portugal	7
	Greece	6
	Spain	3
	Kosovo	2
	Total	**25**
Western Europe	Germany	7
	Netherlands	4
	Belgium	2
	France	1
	Total	**14**
Northern Europe	Sweden	3
	United Kingdom	3
	Estonia	1
	Ireland	1
	Lithuania	1
	Total	**9**
Eastern Europe	Poland	3
	Bulgaria	1
	Slovakia	1
	Total	**5**

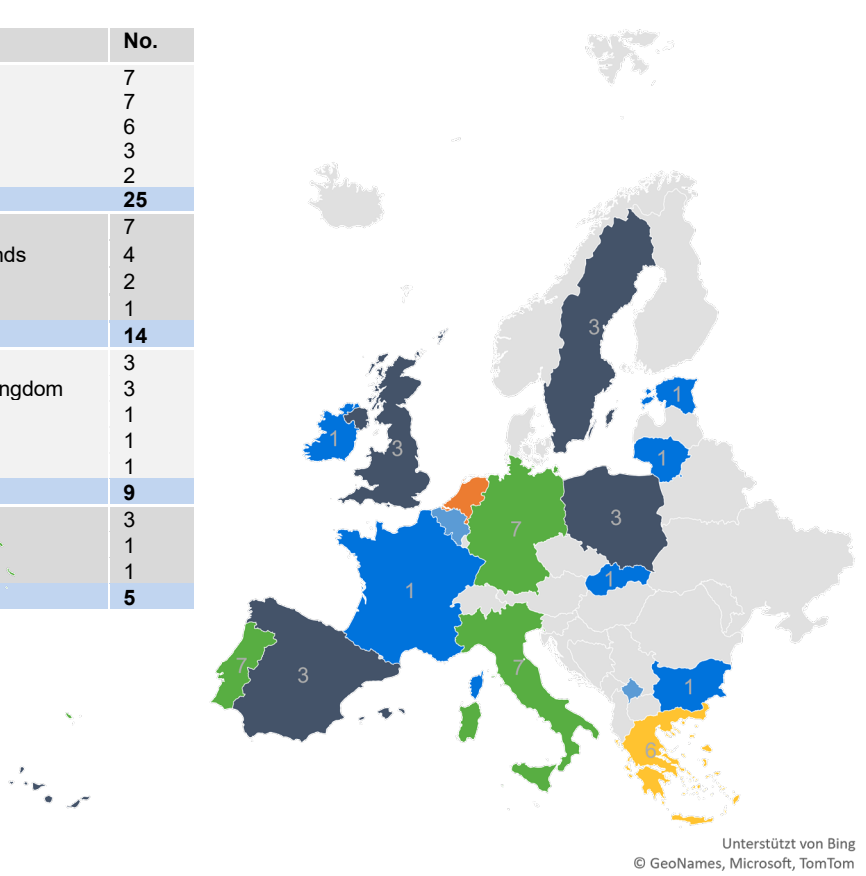


Figure 10: Results of the question “In which country is your organization based?” *Note: Numbers on the map represent number of responses (e.g. “1” = 1 person working in country x, “2” = 2 persons working in country y).*

Unprioritized and categorized answers on potentials and challenges

Table 2 and Table 3 provide the raw data obtained from the survey on the questions “Where do you think AI can be applied immediately and would bring the greatest (immediate) benefit?” and “What do you think are the biggest challenges to introducing AI into law enforcement and policing?”. Answers not given in English have been translated, no other processing of the data has been performed.

Table 3: Original answers to the question "Where do you think AI can be applied immediately and would bring the greatest (immediate) benefit?"

Where do you think AI can be applied immediately and would bring the greatest (immediate) benefit?
DNA analysis, object recognition, automated picture to picture comparison,
Digital forensics
1. Biometric recognition and identification, especially analyzing DNA traces (find similar traces in the data bank). Face recognition works quite well already. 2.Prevention of crimes within digital domains.
Detection and prevention of crimes and threats within the digital domain...
data handling and information handling processes
medicine
Data analytics across multiple police sources/systems
Data analytics, in particular filtering of relevant data
In intelligence and operational efforts
social media analysis and early detection of terrorism-related online crime (e.g. recruitment, propaganda, etc.)
digital forensics
Digital forensics
Data and information handling processes, Digital forensics
Computer Examinations and Phone examination.
Data Management
Cyber crimes
Face recognition (all capabilities of deep learning applied to images). Introducing advanced control tools (like deep reinforcement learning) to devices.
Monitoring of Social networks for detecting hate speech, radicalisation,...
evasion of excise duties
Risk assessment, social media analysis
information handling: making the best use of information that LE already has available (eg, criminal reports)
Digital forensics, passport and ID-related tasks
Fingerprints recognition and matching
Drones used for rescue operations
Detection and prevention of crimes and threats occurring outside the digital domain
Collecting and organization of data
Detection and prevention of crimes and threats within the digital domain
Biometric recognition and identification
Automation of search and data correlation procedures
video surveillance; usage of IoT devices (swarm optimisation)
Mass crime processing and analysis, facial recognition, pattern recognition
predictive policing, automated mapping of crimes and data analysis

AI and its tools can be applied immediately in identifying and predicting threats in cybersecurity processing a large amounts of data increasing both the speed and accuracy of decision-making processes.

Table 4: original answers to the question "What do you think are the biggest challenges to introducing AI into law enforcement and policing?"

What do you think are the biggest challenges to introducing AI into law enforcement and policing?
general level of digitalization of †he LEA, Trustworthiness
tendering process of public bodies, lack of transparency in the result creation process, IT legacy systems
The need for police & law enforcement agencies to understand what problems and challenges they face that technology can assist them with, followed by where and how they can obtain and operate the most suitable and ethically acceptable solutions
people are hesitant to use new technologies and need to have enough trust into the system, transparency about the AI system and how it was created to rule out implementing human biases into the system, to prevent harm by the system/ that the system gets hacked and used against you
Decision makers' lack of knowledge and concerns about AI technologies. Decision makers do not understand the technology.
Data protection.
Challenges are related to the psychology of the individual and groups to adopt disruptive technologies...
legislation and knowledge
Having a clear and precise definition of AI and communicating that definition to public & policy makers
Lack of labelled data for training of AI, experts for labelling are already a scarce resource
To be aware of the ethics and LAW
harmonisation of the regulation across countries; acceptance and training at operational level
privacy issues
Preventing algorithmic bias
data protection
Transparency and explainability (in a wide sense)
Internal safety rules, money, law gaps.
Capacity building. Capacity of understanding and using AI.
AI will only give "hints", the police person is the only one who can decide whether this is relevant or not... Therefore according to me the largest challenges is to make AI useful for human (and gives them elements to improve their work, not replace their brains!)
To use it as the evidence for the court purposes
Operators to understand and accept benefits from AI
1. The human confidence of an AI system to be used in a legal issue. 2. How to legally manage a fail of an AI system that produces damages of any kind.
Trust
GDPR, bad reputation of AI
Crime detection, logistics
staying within legal & ethical boundaries; data governance (internal processes related to data quality, management, standardisation, etc)
Law and ethical principles, defining the exact ways that AI can and cannot be utilised

The biggest challenges are (1) making sure that the tools developed do not infringe privacy and lead to mass surveillance, (2) using AI into law enforcement and policing requires handling uncertainty (for instance in Computer vision) (3) Interpreting laws is subjective and is dependent on the situation, being able to handle this margin between right and wrong is a human trait that is difficult to enforce with an AI
Lack of transparency / human in control.
NLP
While AI can enhance capabilities as given above, this does not mean it is a good use of AI for society.
That it is not in breach of human rights or legislation
Legal framework
Privacy right compliance
the shift to new knowledge
Training of personnel and acquisition of tailored equipment to allow the usage of advanced AI capabilities
Using the AI technologies in a responsible way (e.g., fairly to every citizen).
rule of law - gdpr regulations and data protection issues
Proposals are often made by companies that miss the target of law enforcement or do not have much benefit (e.g. Precobs = making crime forecasts))
protection of privacy, chilling effect, human oversight
The biggest challenge facing the AI into law enforcement and policing is the need to reconcile AI's data with the with the human right to privacy taking into consideration current privacy legislation and culture.