ALIGNER

First ALIGNER Policy Brief

About ALIGNER

ALIGNER is a European Commission-funded Coordination and Support Action that brings together European actors at the cross section of AI, Law Enforcement, and Policing to collectively identify and discuss needs for paving the way for a more secure Europe in which AI supports law enforcement agencies while simultaneously empowering, benefiting, and protecting the public.

To achieve this, ALIGNER has established a forum for exchange between practitioners, civil society, policymaking, research, and industry to design an AI research and policy roadmap meeting the operational, cooperative, and collaborative needs of police and Law Enforcement.

Duration: 36 months 01/10/2021 – 30/09/2024

Website: www.aligner-h2020.eu

The ALIGNER team

ALIGNER's interdisciplinary team includes three European law enforcement agencies and research scientists and industry experts with focus on Artificial Intelligence, Ethics, and Law. The project is coordinated by Fraunhofer IAIS.

1.1 Recommendations

Based on 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 1 provides a systematic overview of these recommendations. The overview adapts the policy ontology originally developed by popAI [1], 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 [1], while the initial ALIGNER policy recommendations were first published in September 2022 as part of ALIGNER D2.3 [2]. 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.

1.1.1 Recommendation overview

No.	Recommendation	Implementation Levels	Type	Target audiences	Themes / Aims
1	Provide common guidelines and unbiased specialist support to LEAs for the development, procurement, deployment, and use of AI technology.	Regulatory, Organisational		EU Parliament, European Commission, Member State Parliaments, Ministries, 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, 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, LEAs, researchers, industry, and civil society organisations to strengthen citizens' confidence in the use of AI tools by LEAs.	Regulatory, Organisational, Research, Societal		Member States Parliaments, Ministries, 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 LEAs	Diversity, Equality, Social Inclusion
6	Extend and adapt European and national research programmes to better facilitate evidence-based, participatory research into LEA needs regarding AI, the potential implications of the use of AI by 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 1: Overview of policy recommendations

1.1.2 Recommendations in detail

Recommendation 1

Provide common guidelines and unbiased specialist support to 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 [3] highlight the need for and the lack of clear guidelines for Law Enforcement Agencies (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 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 LEAs with choosing, procuring, and integrating AI technologies. On a European level, Europol and its EU Innovation Hub for Internal Security[1] might be the prime target to establish such a centre where 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 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 LEAs.

[1] 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 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[2],[3],[4],[5] 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 [4] (which is based on the MAGNETO[6] 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 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 put

[2] 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
[3] World Economic Forum (WEF) AI Governance framework: https://www.weforum.org/projects/model-ai-governance-framework
[4] NOREA Guiding Principles Trustworthy AI investigation: https://www.norea.nl/uploads/bfile/a344c98a-e334-4cf8-87c4-1b45da3d9bc1
[5] AI Assessment Catalog of Fraunhofer IAIS: https://www.iais.fraunhofer.de/en/research/artificial-intelligence/ai-assessment-catalog.html
[6] https://www.magneto-h2020.eu/

into place (too bureaucratic). A valid approach to alleviate these concerns might be the development of a LEA-specific AI directive (similar to the Law Enforcement Directive [5]).

Regardless of these issues, any legal mechanism on EU and national level related to the use of AI technologies by 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 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 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 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 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 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, LEAs, researchers, industry, and civil society organizations to strengthen citizens' confidence in the use of AI tools by LEAs.

Civil society organisations are often not included in consultations regarding the employment of AI by 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, 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[7] [8], especially in the era of AI. In 2010, the Women Police Officers Network (WPON)[9] 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.[10]

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"*[11]. 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.

[7] Women, U. N. (2021). Handbook on gender-responsive police services for women and girls subject to violence.
[8] 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.
[9] Kekić, D., Đukanović, D., & Tomić, M. Women Police Officers Network (WPON).
[10] This and the following paragraph were first published by popAI in [1].
[11] 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 LEA needs regarding AI, the potential implications of the use of AI by 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"[12], which sparked concerns among Members of European Parliament calling on the European Commission to clarify its purpose[13]. 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.[14]

[12] INDECT (Intelligent Information System Supporting Observation, Searching and Detection for Security of Citizens in Urban Environment), Cordis Project Page.
[13] Euractiv (2011), "MEPs question 'Big Brother' urban observation project".
[14] This paragraph was first published by popAI in [1].

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 LEAs (recommendation 1), assessment frameworks (recommendation 2), an evaluation of the existing legal mechanism as well as their effects on LEA work (recommendation 3), stakeholder engagement techniques in the context of AI technologies for 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 LEAs in an ethical, legal, and societally acceptable way.

1.1.3 References

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

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

[3] 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.

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

[5] 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].

Find out more about ALIGNER'S AI research and policy roadmap

These policy recommendations are an extract from the larger AI research and policy roadmap of the ALIGNER project. You can access all versions of the full document here:

https://aligner-h2020.eu/deliverables/

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