**Discussion paper for the second technical meeting ahead of the second plenary of the High-Level Forum on the Future of EU Criminal Justice**

**28 April 2025**

**Topic:** **Reflections on the use of AI to facilitate criminal investigations and proceedings**

|  |
| --- |
| *This document has not been adopted or endorsed by the European Commission. Any views expressed are the preliminary views of the Commission services and may not in any circumstances be regarded as stating an official position of the Commission.* |

**Introduction**

Today, worldwide, various malicious actors widely exploit AI tools to enable a broad range of criminal activities. Criminals increasingly use AI tools to facilitate and expand the scope of crimes such as fraud, misinformation, identity theft, cyber-attacks, generation of illegal content and other illicit activities.

At the same time, authorities have harnessed AI to enhance judicial processes and law enforcement operations in many countries. AI tools can offer significant benefits to justice systems, for example by helping streamline and automate labour-intensive administrative processes, reducing workloads, alleviating case backlogs, and broadly improving efficiency. However, the use of AI in criminal justice also raises concerns. In a global context, AI-powered predictive policing tools have for example been misused to target and detain activists and political dissidents, often under the pretext of counterterrorism. Additionally, AI systems employed in sentencing and parole decisions have faced public criticism for perpetuating racial bias and reinforcing existing inequalities.

It is therefore essential for justice systems to stay abreast of technological developments and harness their potential in a responsible manner which safeguards fundamental rights and is compliant with the relevant EU *acquis*. This is crucial not only for effectively combating technology-enabled forms of cybercrime, but also for enhancing access to justice, improving efficiency, and ensuring the effective, and fair administration of justice. In the EU, the current adoption of AI by courts and prosecution services, however, remains limited. According to the EU Justice Scoreboard 2024, only six Member States use AI in core judicial activities[[1]](#footnote-2).

**Regulatory landscape**

The EU AI Act[[2]](#footnote-3) is the first major comprehensive regulation on the development and use of artificial intelligence in the world setting a global benchmark. One of the key pillars of the AI Act’s risk-based approach - particularly in defining “prohibited” and “high-risk” AI systems - is inter alia the assessment of their potential impact on fundamental rights, health and safety. In February 2025, the Commission published its Guidelines on prohibited AI practices[[3]](#footnote-4), as well as the Guidelines on the definition of an AI system, as defined in the AI Act[[4]](#footnote-5). Among other clarifications, the Guidelines on prohibited AI practices elaborate on the prohibition against law enforcement authorities using AI for making risk assessments of natural persons with the aim of assessing or predicting the likelihood of a person committing a criminal offence, when such assessments *are based solely on profiling or evaluations of personality traits and characteristics*. The guidelines also provide examples of the exception to this prohibition. Specifically, it does not apply to AI systems used to support the human assessment of the involvement of a person in a criminal activity, which is already based on objective and verifiable facts directly linked to a criminal activity. For instance, the use of AI systems in such cases may be permissible where there is a reasonable suspicion of criminal activity, such as the preparation of a terrorist act or the acquisition of weapons.

Similar guidelines on the classification of “high-risk” AI systems are expected in February 2026. In the law enforcement and justice area the AI Act defines several scenarios where the use of AI could present a high risk to the safety, health or fundamental rights of natural persons, which are set out in Annex III thereof. Such AI systems would then be subject to the requirements set by the AI Act to be fulfilled by the provider before the system is placed on the market or used. These include a risk management system, data quality and governance, documentation, transparency, human oversight, accuracy, robustness and cybersecurity that need to be checked in a conformity assessment. Furthermore, deployers of such systems that are bodies governed by public law will also need to carry out a fundamental rights impact assessment, ensure human oversight and establish monitoring. At the same time, the AI Act provides specific exceptions when the use of AI tools even in high-risk contexts is not considered high-risk because they do not pose a significant risk of harm to the health, safety or fundamental rights of natural persons if they are not materially influencing the decision, e.g. when AI is used for the purpose of narrowly defined procedural tasks.

**Opportunities**

In criminal justice, AI tools have the potential to greatly enhance the efficiency and effectiveness of law enforcement and judicial processes - supporting the efforts of public authorities in combating impunity throughout the entire lifecycle from crime prevention and investigation to sentencing and parole hearings. AI tools can provide valuable assistance across a wide spectrum of tasks, from routine administrative functions to complex decision-making. AI can streamline activities such as the anonymisation and pseudonymisation of judicial decisions, facilitate legal research, support self-monitoring mechanisms to detect bias, and promote greater consistency in sentencing. For the analysis of large and diverse datasets AI is particularly effective at rapidly uncovering connections between individuals, legal entities, and events, as well as correlating diverse data sources to reveal both explicit and implicit relationships. In addition, AI can play a crucial role in detecting artificially generated or manipulated content (such as “deepfakes”), thereby supporting forensic investigations and ensuring the integrity and subsequent admissibility of evidence.

Appropriately trained AI systems are even capable of providing recommendations on sentencing, parole and detention decisions, offering additional support for certain risk assessments in judicial deliberations. Such recommendations and assessments should of course always be approached with due caution - with regard to both the quality of the underlying data and potential assumptions inherent to AI models, and in ensuring that a proper secondary human evaluation takes place to avoid overreliance or misplaced trust in the system’s outputs. Thanks to advancements in natural language processing, AI can also deliver high-quality translations and provide real-time interpretation, albeit AI interpretation tools can supplement but not replace human interpreters in criminal proceedings (at least not under the current EU legal standards protecting the right to interpretation).

**Challenges**

The adoption of AI tools by law enforcement authorities and the judiciary presents a range of challenges, spanning from technical and practical issues to concerns related to the protection of fundamental and procedural rights. While the AI Act establishes a robust and comprehensive legal framework grounded in a risk-based approach, there are singular issues and considerations specific to law enforcement and the judiciary that warrant closer examination, particularly in the context of AI use in the area of criminal justice.

Among these challenges is the pressing need to improve AI literacy among judicial staff. It should be noted that Article 4 of the AI Act requires AI literacy for all deployers and providers of AI systems, including in the justice sector. This includes understanding how to use AI tools effectively, while also being aware of their limitations and potential risks (e.g. potential for bias or "AI hallucinations"). As a cognitive phenomenon automation bias remains an inherent risk, even when individuals are aware of its potential influence and final decision-making authority rests with the judiciary. This bias can lead to an undue reliance on AI-generated recommendations, potentially compromising independent judgment and undermining the fairness of decisions.

One rising challenge for judges relates for instance to verifying the reliability and probative value of AI-generated/manipulated evidence. In addition, there is a growing need to build technical expertise in the justice sector in the development, deployment, and adaptation of AI systems. Unlike traditional software, AI tools - particularly those powered by Large Language Models (LLMs) - are not simple plug-and-play solutions. They often require fine-tuning, reinforced training, and continuous evaluation to ensure they remain fit for purpose, deliver accurate and reliable results, and uphold fairness, non-discrimination, and transparency. With the rise of AI-generated and manipulated content, such as deepfakes, there is also an increasing need for advanced tools to help judicial authorities and forensic experts detect and verify the authenticity of digital evidence. These tools are essential to safeguard the integrity of legal proceedings and prevent the misuse of falsified content.

Another key consideration is the need to ensure equality of arms, particularly with respect to defendants and their legal representatives having equitable access to AI tools. In line with the Charter of Fundamental Rights of the EU and the European Convention on Human Rights, this includes their ability to challenge AI-generated/manipulated evidence, scrutinise the underlying methods and data, and employ AI technologies themselves to uncover exculpatory evidence or otherwise support their defence.

**Questions for discussion:**

1. In your view, in the area of criminal justice, what are the key areas and use cases (e.g. legal research, analysis of electronic evidence, anonymisation, risk assessments, etc) where AI tools can deliver significant value?
2. Without prejudice to the need of compliance with the relevant regulatory requirements set out by the AI Act, including on human oversight, are there any specific use cases or application areas where you believe AI tools could fully replace the human element to deliver efficiency gains?
3. What are the main challenges for the deployment and use of AI tools in criminal justice, both from a practical standpoint and in ensuring that fundamental and procedural rights are upheld, rather than adversely affected?
4. Considering the safeguards and rules of the AI Act, are there particular measures that Member States are considering, or stakeholders consider necessary, to address these challenges?
5. Beyond the horizontal framework of the EU AI Act, is there a need for coordinated action or specific soft measures to assist Member States in adopting AI tools in a way that is both effective and compliant with fundamental rights in the criminal justice area?
6. In particular, for AI systems that are not classified as high-risk, the AI Act allows voluntary codes of conduct - could that be of interest to develop at EU level with the relevant public authorities?

1. [EU Justice Scoreboard - European Commission](https://commission.europa.eu/strategy-and-policy/policies/justice-and-fundamental-rights/upholding-rule-law/eu-justice-scoreboard_en) [↑](#footnote-ref-2)
2. [Regulation - EU - 2024/1689 - EN - EUR-Lex](https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng) [↑](#footnote-ref-3)
3. <https://ec.europa.eu/newsroom/dae/redirection/document/112367> [↑](#footnote-ref-4)
4. <https://digital-strategy.ec.europa.eu/en/library/commission-publishes-guidelines-ai-system-definition-facilitate-first-ai-acts-rules-application> [↑](#footnote-ref-5)