

Artificial Intelligence – a Vector for Crime and a Tool for Carrying Out Criminal Justice

*By Carmina-Elena Tolbaru**

This study aims to provide a systematic and comprehensive analysis of the role of artificial intelligence (AI) in both criminal activity and the criminal justice system. It highlights the dual nature of AI's technological advancement: on one hand, its negative aspect, as it enhances the capabilities and efficiency of criminal operations; and on the other hand, its positive potential in strengthening cybersecurity measures. AI can serve as a powerful instrument for criminal conduct, which in turn necessitates its deployment in predictive policing and the more effective attribution and investigation of crimes. Artificial intelligence technology has an essential role in everyday life, in all areas of economic-social life, being the obvious beneficial effect on the whole culture; especially, however, we need to be aware of the negative effects in the field of cybercrime, as AI can help criminals become more sophisticated and scalable, consequently more difficult to detect. Therefore, the rapid progress of information and communication technology is likely to constitute a mechanism for criminal actors to facilitate the commission of crimes, not only by expanding and changing the inherent nature of existing threats, but also by creating new threats. Without well-enforced national and international legislation, cyber security is seriously affected and the management of criminal activities becomes increasingly difficult to achieve, which is a significant concern for law enforcement authorities. Using artificial intelligence in investigations is likely to improve and speed up criminal justice work; In fact, this tool is becoming increasingly useful for detecting crime and preventing illegal activities. But on the other hand, through the ability of artificial intelligence to quickly analyse large amounts of data, there is a flagrant violation of the rights and freedoms of individuals, which is a real challenge for criminal justice. In this respect, it is therefore important to develop a legal framework to ensure the responsible use of AI.

Keywords: *Artificial intelligence; Crime; Criminal justice; Cyber security; Legal framework; Responsibility.*

Introduction

The transformation of society generated by the processes of globalisation and digitalisation is deeply interconnected with the evolution of the criminal phenomenon, influencing both its level and its dynamics, depending on the nature and extent of criminal manifestations. The emergence of artificial intelligence, first in the field of computer science, as a theoretical concept, and then its incidence at a general level,

*Associate Professor PhD, Faculty of Economic Sciences and Law, National University of Science and Technology Politehnica Bucharest, Pitești University Centre, Romania.
Email: carmina.tolbaru@upb.ro

in all areas of economic and social life, has generated new challenges for law enforcement authorities and for legal systems at a global level. This technological advance has facilitated the expansion of transnational crime and the emergence of sophisticated forms of criminality, artificial intelligence today representing a fundamental pillar of the global digital transformation.

Undoubtedly, artificial intelligence (AI) plays an essential role in contemporary society, being integrated into a wide range of fields, from commerce, education and healthcare, to strategic sectors such as national security and criminal justice¹. Through its ability to process large volumes of data and generate quick and efficient decisions, AI contributes significantly to optimizing processes and increasing institutional performance². However, the benefits brought by AI are accompanied by notable risks, especially in the context where the technology is misused for criminal activities or implemented in systems that may endanger the fundamental rights and freedoms of the individual³. In this regard, the new EU AI Regulation aims to ensure that artificial intelligence (AI) systems are developed and used responsibly, thus highlighting the concern for appropriate regulation that ensures a balance between innovation and the protection of democratic values⁴. AI is therefore a dual technology that, on the one hand, acts as a catalyst for cybercrime and new forms of fraud, and on the other hand, offers sophisticated means of preventing, investigating and solving crimes⁵. Thus, in the contemporary legal context, AI becomes both a challenge for authorities and an indispensable tool for strengthening criminal justice.

Studies have shown that areas such as predictive crime analysis, facial recognition or automated detection of financial anomalies demonstrate the potential of AI to revolutionise law enforcement and judicial efficiency⁶. At the same time, deepfakes, automated attacks based on machine learning or indiscriminate surveillance raise urgent questions about regulation, accountability and personal data protection⁷.

In this study, we aim to investigate the dual nature of artificial intelligence in relation to the phenomenon of crime and criminal justice activity. The approach aims to outline an integrated perspective on current issues, such as: the ways in which artificial intelligence is exploited in the commission of modern crimes; its role in the prevention, identification and resolution of criminal acts; the risks and ethical and legal limits associated with the use of AI in judicial systems; as well as the need for a responsible and regulated use of these emerging technologies.

¹Brynjolfsson & McAfee (2014) at 23-24.

²Wang, Kung & Byrd (2018) at 3-13.

³Brundage, Avin, Clark, Toner, Eckerskey, Garfinkel & Amodel (2018); Popescu-Ljungholm & Tolbaru (2025) at 47-62.

⁴European Parliament and Council. (2024). *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (EU AI Act)*. *Official Journal of the European Union*, L 189, 1–257. The AI Act aims to foster responsible artificial intelligence development and deployment in the EU.

⁵For a balanced analysis of the dual nature of artificial intelligence in enabling and combating cybercrime and crypto-asset fraud, see Tolbaru (2023) at 151-156.

⁶Regarding the potential of predictive technologies to enhance policing efficiency—provided that a strong democratic and ethical framework is in place, see Van Brakel (2021) at 104–118. See also Lynch (2024) for a regulatory analysis of facial recognition technology in policing and security contexts, and Duong, Le & Hoang (2023).

⁷Wright (2020) at 147-148; Brundage, Avin, Wang, Krueger, Hadfield & Belfield (2023) at 10-12.

Research Methods used

The realisation of this study was based on a mixed methodological approach, predominantly qualitative, with an emphasis on documentary analysis, normative interpretation and critical assessment of the impact of artificial intelligence (AI) on crime and criminal justice. The descriptive-analytical method was used to identify and interpret the main phenomena, risks and opportunities related to the use of AI in the criminal field, combining both conceptual and legal analysis (to define the notions of AI, digital crime, criminal liability), case study (to highlight practical applicability in various legal systems) and public policy analysis (to evaluate regulations and emerging legislative directions). The work is also based on a series of specific methods such as: documentary analysis, comparative analysis, normative method, inductive and deductive method. Thus, the documentary analysis aimed at the study and systematic review of specialised literature, namely doctrinal studies, research reports, relevant legislation (e.g. "AI Act" - EU) and case law. To highlight the differences in regulation and practice regarding the application of AI in criminal justice in various international legislations, such as the USA, China, EU and Romania, we used the comparative method which thus allowed us to highlight strengths, risks and common trends. Also, in the study, legal interpretation techniques (literal, systematic and teleological) were used to analyse the compatibility of the use of AI with fundamental rights and the principles of fair trial, using the normative method in this regard. Last but not least, the inductive and deductive methods were applied, the inductive method for drawing general conclusions based on case studies and concrete examples, and the deductive method for formulating hypotheses regarding the impact of AI on the justice system and testing them through doctrinal analysis and judicial practice. It should be taken into account that regulations in the field of AI are rapidly evolving, which is why this study focuses on the current legal context, but which needs revalidation in the future.

Results and Discussion

Using AI in Criminal Activities

The rapid evolution of artificial intelligence (AI) has generated multiple benefits for society, but has also provided extensive opportunities for criminal actors, who exploit these technologies maliciously. AI's ability to automate complex tasks, produce deepfakes and circumvent conventional cybersecurity systems has led to its increased use in criminal activities⁸. The reports, which analysed the potential security threat landscape, proposed ways to better predict, prevent and mitigate these risks, consisting of four high-level recommendations, namely⁹:

⁸For an in-depth analysis of the evolution of cyber threats in the European context and the EU's institutional and legal response mechanisms, see Tolbaru (2024) at 79-86.

⁹Brundage, Avin, Clark, Toner, Eckerskey, Garfinkel & Amodel (2018) at 50-51.

- close collaboration between policymakers and technical researchers to investigate, prevent and mitigate potential malicious uses of AI;
- raising awareness among AI researchers and engineers of the dual-use nature of their work, allowing misuse issues to influence research priorities and norms, and proactively reaching out to relevant actors when malicious applications are foreseeable;
- identifying best practices in research areas, with more mature methods for addressing dual-use concerns, such as cybersecurity, and imported where appropriate into AI;
- actively pursuing a broader range of stakeholders and subject matter experts involved in discussions on these challenges.

The process of globalisation, in tandem with accelerated digitalisation, has contributed to the profound transformation of the nature and scale of the criminal phenomenon. The global interconnection of economic systems and technological progress have facilitated not only the expansion of transnational crime, but also the emergence of increasingly sophisticated forms of criminality, such as computer crime, automated fraud or cyber espionage¹⁰.

Over the past two decades, AI has moved from being a niche technology to a key element of the global digital transformation, in sectors such as health, education, public administration and, especially, law enforcement. Thus, according to the OECD report, this expansion of the use of AI has generated both significant opportunities for increasing institutional efficiency (through decision automation, data analysis, productivity, etc.), as well as critical challenges related to surveillance, data protection and use in law enforcement, in the field of preventing and combating crime (predictive policing, digital recidivism assessment)¹¹. In this context, Regulation (EU) 2024/1689 on artificial intelligence, adopted by the European Parliament and the Council, provides a harmonised legal framework for the use of high-risk AI technologies, including in the field of criminal justice, setting clear requirements on transparency, accountability and the protection of fundamental rights¹².

According to a Europol report, AI-based technologies are increasingly being used for surveillance, predictive analytics, facial recognition and identifying suspicious behaviour¹³. In parallel with the legitimate uses of artificial intelligence (AI), the technology is increasingly being exploited by malicious actors to carry out complex criminal activities, especially in cyberspace. One area where this phenomenon is becoming increasingly visible is that of AI-assisted cyberattacks, which allow for both automation and personalisation of cyberattacks, depending on the target¹⁴. The use of machine learning algorithms facilitates the identification of vulnerabilities in IT infrastructures, accelerating the processes of penetration and compromise of systems¹⁵. Thus, AI is used to automate phishing attacks, analyse system vulnerabilities,

¹⁰United Nations Office on Drugs and Crime (UNDOC) (2021); Di Nicola (2022); Shelley (2019) at 223-228.

¹¹OECD (2019) at 5-6.

¹²Regulation (EU) 2024/1689 (2024).

¹³Europol (2023a) at 9-12.

¹⁴Schmitt & Flechais (2024).

¹⁵Brundage, Avin, Clark, Toner, Eckersley, Garfinkel & Amodei (2018) at 6-7.

and evade detection by antivirus software. For example, generative adversarial networks (GANs) can create personalised fraudulent emails that easily fool security filters¹⁶.

At the same time, AI is used to manipulate information through deepfakes, namely synthetically generated but extremely realistic audio-video content, capable of misleading public opinion, compromising reputations or even affecting democratic processes such as elections, through disinformation campaigns. These technologies are based on generative neural networks (Generative Adversarial Networks – GANs), capable of producing images, voices and video sequences that are indistinguishable from authentic content, which thus raises serious information security and ethical issues¹⁷.

Research explicitly shows how AI technologies can facilitate more sophisticated, personalised, hard-to-attribute, and scalable hacking or social engineering attacks. Increasingly, neural networks are being used to intrude into computer systems. AI-enabled cyber threats, including malware and automated attacks, can learn and replicate the behaviours of legitimate users, successfully bypassing traditional detection systems¹⁸. This adaptability makes AI an extremely versatile and dangerous tool in the hands of those who use it for criminal purposes. Banking fraud has become increasingly sophisticated with the adoption of generative artificial intelligence technologies: from automatic document generation and audio/video phishing, to simulated transactions on blockchain platforms and manipulations orchestrated by AI chatbots¹⁹. Not infrequently, investors become targets of "pump-and-dump" campaigns promoted through automated messages and directives designed based on online behaviour. In this context, at a global level there is a need for a reconfiguration and adaptation of the framework for fiscal prevention and supervision, in the sphere of AI-assisted financial crime.

Against the backdrop of these developments, the urgent need for international regulations and the strengthening of institutional capacities in the field of cybersecurity has emerged, so as to limit abuses and protect democratic values and the rule of law. This objective is incorporated into the Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, adopted by the Council of Europe and opened for signature on 5 September 2024, the first binding international treaty on AI governance.

¹⁶Eze & Shamir (2024).

¹⁷Chesney & Citron (2019) at 1760–1765; Maras & Alexandrou (2019) at 255–256;

¹⁸Brundage, Avin, Clark, Toner, Eckersley, Garfinkel & Amodei (2018) at 50-51; Goodfellow, Pouget-Abadie, Mirza, Xu, Warde-Farley, Ozair, Courville & Bengio (2014).

¹⁹Papasavva, Johnson, Lowther, Lundrigan, Mariconiti, Markovska & Tuptuk (2025).

*Challenges in AI-assisted Crime Detection and Prevention*Algorithmic Opacity and Source Detection

One of the most controversial aspects of artificial intelligence, especially in the case of deep learning-based systems, is algorithmic opacity. These systems often operate as “black boxes” and their decisions remain opaque and unexplained. Thus, while they can provide very accurate results, the internal decision-making mechanism remains inaccessible to both users and developers²⁰.

The problem becomes even more acute in the context of investigations into cyber attacks, especially when these attacks are geographically distributed or are carried out through anonymisation and obfuscation technologies, such as VPN networks, Tor, Monero, dark-web or blockchain protocols²¹. In such cases, identifying the real source of the attack is complicated, not only because of technological complexity, but also because of the impossibility of clearly assigning responsibility when decisions are made by non-transparent autonomous systems.

The use of artificial intelligence (AI) in justice and automated decision-making processes raises significant issues of transparency and accountability, especially in the context of “black box” algorithmic systems. According to specialised literature, the lack of the possibility to audit and understand the logic of algorithmic decisions undermines the foundations of the rule of law, endangering the right to a fair trial and the principle of decisional transparency. This opacity can seriously affect the legitimacy of the use of AI in sensitive areas, such as criminal justice, where decisions must not only be correct, but also understandable for the parties involved. In the absence of clear explainability mechanisms, individuals cannot challenge or understand decisions that directly concern them, which runs counter to democratic requirements and human rights norms²². The right to a fair trial implies the right to challenge AI decisions, to receive clear notifications and opportunities to be heard – regardless of whether the decision is automated or artificially assisted. Thus, AI should not operate in a “black box” mode, but should provide interpretable results through a “glass box” system, so that decision-makers (and affected individuals) understand the factors, their weight and the limitations of the model²³. Thus, it is essential to develop explainable and accountable AI models that allow for the auditability of decisions, in accordance with the requirements of international regulations, such as the GDPR.

Legislative Deficiencies in the Criminal Regulation of Artificial Intelligence

One of the most pressing obstacles to regulating the use of artificial intelligence (AI) in a criminal context has been, until recently, the absence of a specific legal framework adapted to the new technological realities. Currently, this regulatory gap is partially remedied by the adoption of international and European legal instruments, such as Regulation (EU) 2024/1689 on artificial intelligence (AI Act) and the Council of Europe Framework Convention on AI, Human Rights, Democracy and the Rule

²⁰Binns (2020); Burrell (2016); Wachter, Mittelstadt & Floridi (2017a).

²¹Faber (2025).

²²Wachter, Mittelstadt & Floridi (2017a).

²³Garrett (2025).

of Law (2024)²⁴. However, in the domestic criminal law of most states, including Romania, there are still no detailed provisions that explicitly regulate the use of AI in the commission of crimes, which generates a significant regulatory vacuum²⁵.

In particular, the issue of attributing criminal liability in the context of AI remains in an uncertain legal area. If an AI system generates an illegal action without direct and immediate human intervention, the determination of guilt becomes deeply problematic. If in classic scenarios the human agent is the active subject of the crime, in the case of AI, questions arise regarding the possible culpability of the developers, users or the entity implementing the system²⁶. This dilemma is exacerbated by the technological specificity of AI, especially machine learning systems, which can adapt their behaviour autonomously based on training data²⁷.

Furthermore, the evolving and opaque nature of AI – often described as a “black box” – accentuates the impossibility of establishing with certainty criminal intent or imputability. In the absence of clear legal norms on algorithmic transparency, decision-making auditability and the delimitation of responsibility, judicial systems face major obstacles in applying the principle of legality of criminalisation, but also in ensuring a criminal response that respects the requirements of fairness, predictability and proportionality²⁸.

Although Regulation (EU) 2024/1689 (the “AI Act”) establishes strict obligations regarding transparency, control and governance of artificial intelligence systems, it does not regulate the attribution of criminal liability in the event that AI acts autonomously. The issues of imputability, intention and causality remain unresolved in criminal law, which highlights the need for national legislative intervention. Therefore, current criminal law is called upon to evolve in parallel with the new European regulations, in order to adequately respond to the challenges raised by the use of artificial intelligence in the field of crime.

Scalability of Attacks

Recent developments in artificial intelligence have profoundly reshaped the risk landscape in cyberspace, facilitating an unprecedented escalation of digital attacks. Systems using machine learning and generative neural networks enable the generation of automated, adaptable, and extremely fast cyberattacks²⁹. Unlike traditional attacks, which required constant planning and human intervention, automated AI systems can identify vulnerabilities, exploit security breaches, and spread malware autonomously, without direct supervision.

Cyberattacks facilitated by artificial intelligence have the potential to simultaneously affect tens of thousands of users and systems in an extremely short time frame, which amplifies the impact and complicates the intervention. Europol highlights that malicious actors can automate essential stages of an attack – identifying vulnerabilities, generating malicious content and disseminating it through scalable

²⁴Regulation (EU) 2024/1689; Council of Europe (2024).

²⁵Brundage, Avin, Wang, Krueger, Hadfield & Belfield (2023).

²⁶Garrett (2025).

²⁷Wachter, Mittelstadt & Floridi (2017).

²⁸Veale & Edwards (2018) at 398–404; Brundage, Avin, Wang, Krueger, Hadfield & Belfield (2023).

²⁹Gupta, Akiri, Aryal, Parker & Praharaj (2023) at 80218-80244.

phishing campaigns. AI also enables digital identity hijacking through deepfakes, infiltrating government networks and chain-compromising interconnected systems, including by exploiting third-party providers³⁰.

The rapid development of artificial intelligence is putting increasing pressure on law enforcement and cybersecurity structures to adapt to an increasingly complex and constantly changing digital threat. According to Europol, law enforcement forces face significant limitations in their ability to detect, attribute and counter attacks generated by autonomous AI systems, particularly in the absence of advanced analytical infrastructures and a clear legal framework on liability for the autonomous actions of intelligent technologies³¹. Furthermore, real-time detection of AI-initiated attacks requires advanced cyber threat intelligence technologies, supported by specific learning algorithms, in a process that inevitably leads to a technological race between attackers and defenders.

In this context, it is essential that international regulations and national cybersecurity strategies include specific mechanisms for monitoring, preventing and sanctioning AI-assisted attacks, as well as measures to strengthen transnational cooperation in the digital field.

Artificial Intelligence as a Tool for ensuring Criminal Justice

Applying AI in the Investigation Process

The implementation of artificial intelligence in criminal investigation has generated substantial transformations in the identification, analysis and prevention of crimes. The Europol report shows that OSINT³² and SOCMINT³³ tools enhanced with machine learning algorithms can process structured and unstructured data simultaneously, providing essential information for investigations on the spot. These capabilities allow for the rapid and automated detection of criminal patterns in data volumes that are far too large for traditional methods³⁴. This technological transformation contributes not only to the efficiency of criminal prosecution operations, but also to the strengthening of prevention capacities, although it also raises a series of legal and ethical issues.

³⁰Europol (2023a) at 9; Europol. (2023b); Europol Innovation Lab (2022);

³¹Europol (2023a) at 12-14.

³²OSINT (Open Source Intelligence) refers to the collection of information legally collected from public sources – such as websites, official databases, press articles or online forums – and used for analytical, investigative or security purposes. In the context of law enforcement, this data is processed using artificial intelligence technologies to identify patterns, connections and potential risks.

³³SOCMINT (Social Media Intelligence) is a subcategory of OSINT that refers to the systematic collection and analysis of data from social media platforms – such as Facebook, Twitter/X, Instagram, TikTok, LinkedIn etc. This form of intelligence allows law enforcement agencies to identify suspicious behaviours, criminal networks, or radicalisation trends, with the support of machine learning algorithms and semantic analysis.

³⁴Europol (2023a) at 9-13.

Predictive Crime Analysis

Predictive crime analysis is one of the most hotly debated applications of artificial intelligence in the criminal justice field. Systems such as PredPol and CompStat are already being implemented by police agencies in the United States and some European countries to predict the likely location and type of future crimes, using historical data on criminal activity. These platforms identify high-risk areas – known as “hot spots” – and facilitate the strategic allocation of patrols and other law enforcement resources, with the aim of proactively preventing the commission of criminal acts³⁵.

However, the benefits of these instruments are accompanied by significant risks that cannot be ignored, namely³⁶:

- AI "learns" from the data it trains algorithms on and tends to replicate the same patterns of disproportionate surveillance; predictive systems can direct more police resources to already heavily patrolled neighbourhoods;
- many algorithms used (e.g. PredPol) are "black boxes", and the authorities cannot clearly explain how the predictions are generated, which complicates the right to challenge of the affected persons;
- excessive interventions, based on algorithmic assumptions, can affect the right to life, to freedom of movement, the right to protection against discrimination;

Thus, according to existing research, predictive systems can perpetuate or even amplify systemic biases existing in historical data, leading to disproportionate surveillance of certain already vulnerable communities.

In conclusion, although predictive analytics offers a clear operational advantage, its responsible use requires a rigorous ethical framework and constant auditing of the data entered and the results generated.

Facial Recognition and Biometric Technologies

Another direction of application of AI in criminal investigation is the biometric identification of suspects, through technologies such as facial recognition, fingerprint analysis or gait analysis³⁷. Platforms like Clearview AI or Face++ allow authorities to identify and locate suspicious individuals by correlating surveillance camera images with social media profiles or official databases³⁸.

Although these technologies increase the efficiency of identifying perpetrators, they raise serious questions regarding the right to privacy, informed consent and the risk of identification errors³⁹. There are concerns about the misuse of these tools in contexts of mass surveillance, which requires the adoption of clear standards on the proportionality and legality of technological interventions.

³⁵Joh (2020) at 1140–1142.

³⁶Lum & Isaac (2016) at 14–16.

³⁷Zeng, Lu, Sun & Tian (2019).

³⁸Dul (2022) at 1-24.

³⁹Chan (2022) at 1-2.

Automatic Detection of Suspicious Behaviour

Artificial intelligence is increasingly present in modern intelligent video surveillance systems, where anomaly detection algorithms automatically identify behaviours that deviate from usual patterns — such as restless movements, prolonged standing, dropping objects, or other unnatural actions⁴⁰. The use of anomaly detection algorithms plays a key role in reducing the risk of attacks, especially in public transport and crowded spaces, to identify potentially dangerous activities or antisocial behaviour in advance for rapid intervention. This type of analysis is also valuable in the context of preventing acts of terrorism or mass events. However, automated behaviour detection involves the risk of false positive interpretations and disproportionate reactions by authorities. These systems may react to legitimate but unusual behaviour, amplifying the risk of abusive profiling and unjustified interventions. Consequently, the implementation of these technologies must be accompanied by human assessment protocols, periodic reviews of algorithmic performance and robust procedural safeguards⁴¹.

Artificial Intelligence in the Judicial Process: Evaluation, Decision and Fairness

Artificial intelligence (AI) has increasingly penetrated the judicial systems of various countries, bringing with it the concern for a more efficient, coherent and data-driven justice. Although AI cannot and should not replace human decision-making in a trial process, it is increasingly used in auxiliary activities of the act of justice. Among the most significant applications are the assessment of the risk of recidivism, support in decision-making and the identification of judicial errors.

Assessing the Risk of Recidivism through Artificial Intelligence in the Judicial Process

One of the most controversial algorithmic tools used in the criminal justice system is COMPAS (Correctional Offender Management Profiling for Alternative Sanctions), a software program that calculates recidivism risk scores, which are frequently used by American courts in decisions about parole, pretrial detention, or sentence length. These scores are based on factors such as a defendant's criminal history, age, employment, and social background⁴².

Although the practical utility of these tools is obvious in terms of streamlining the decision-making process, critics warn about the lack of transparency of the algorithms and the risk of perpetuating systemic prejudices, as historical data may contain biases related to race, socio-economic status or gender⁴³. Additionally, studies show that AI's accuracy in predicting recidivism is not significantly higher than educated popular judgment⁴⁴, questioning the superiority of algorithms over traditional assessments.

⁴⁰Olabi, Favor & Ramon (2025).

⁴¹Wachter, Mittelstadt & Floridi (2017b) at 77–78.

⁴²Dressel & Farid (2018) at 2–3.

⁴³Angwin, Larson, Mattu & Kirchner (2016) at 4–5.

⁴⁴Dressel & Farid (2018) at 5–6.

Assistance in Judicial Decision-making

Another area of application of AI in the judiciary is decision support for judges and clerks, through automated case law analysis, generation of suggestions on judicial practice, and assisted drafting of judgments. In the People's Republic of China, the judiciary uses advanced AI applications to standardise court decisions, reduce drafting errors, and prevent inconsistencies between similar decisions. Such applications, integrated into judicial decision support systems, allow for the rapid extraction of precedents, automatic classification of cases and the suggestion of a legal classification based on the patterns identified in the judicial database. Specifically, the 206 System uses analysis of criminal records to suggest legal classifications and sentence lengths in line with previous practice⁴⁵.

However, this trend raises fundamental questions about the autonomy of the judge and the risk of decision-making authority being diluted in favour of automated algorithmic logic. Experts emphasise that AI must remain an advisory tool, not a normative one, because justice cannot be reduced to the simple statistical uniformity of decisions⁴⁶.

Detection of Judicial Errors

AI is also being used to identify systemic errors or inequalities in sentencing, by automatically detecting abnormal statistical patterns. Algorithms can highlight cases where there are unjustified discrepancies between sentences for similar crimes, suggesting the possibility of judicial errors or even institutional discrimination⁴⁷.

This automated audit capability of the judicial system can contribute to greater transparency, accountability and fairness, provided that AI is used in a way that respects the fundamental rights of litigants and is subject to constant human oversight⁴⁸.

The application of AI in the judiciary has the potential to fundamentally transform the way risks are assessed, decisions are made and dysfunctions are corrected. However, its use must be accompanied by clear regulatory measures, algorithmic auditing, decisional transparency and procedural guarantees, so that the technology is an ally of justice, and not a factor of rigid standardisation or amplification of existing inequities.

Advantages and Limitations of using AI in Criminal Justice

The use of artificial intelligence (AI) in criminal justice is a rapidly expanding field that promises significant structural reforms in the way investigations, criminal trials and the execution of sentences are managed. Although the technology brings

⁴⁵For a more detailed explanation and a better understanding of how artificial intelligence is used in assisted judgment drafting, automatic case classification, precedent identification, and the operation of the "206 System" in suggesting legal qualifications and sentencing ranges, see Li (2025) and Tahura & Selvadurai (2022).

⁴⁶For a detailed analysis of the HART (Harm Assessment Risk Tool) algorithmic model, used by Durham Constabulary to assess the risk of reoffending, as well as the legal and ethical implications of applying algorithmic proportionality in operational decision-making, see Oswald, Grace, Urwin & Barnes (2018) at 223–250.

⁴⁷For a detailed academic examination of how machine learning can detect unjustified sentencing disparities and systemic errors by comparing similar cases, see Ryberg (2023) at 447–462.

⁴⁸Zerilli, Knott, Maclaurin & Gavaghan (2019) at 661–683. For a discussion on the limits of the right to explanation in the context of automated decisions, see Wachter, Mittelstadt & Floridi (2017b) at 90–92.

undeniable benefits in terms of efficiency, standardisation and prevention, its implementation also raises numerous ethical, legal and technological challenges. Thus, a critical analysis of the advantages and limitations of the use of AI is essential for shaping a balanced framework for its integration into the criminal justice system.

Advantages of using AI in Criminal Justice

1. *Efficiency and procedural speed.* One of the most obvious advantages of AI is the ability to manage large volumes of data in real time, which leads to increased efficiency of judicial and police activities. AI can quickly analyse evidence in files, extract relevant case law and formulate legal suggestions, thus reducing the time needed to resolve cases. According to recent research, the use of AI in law enforcement and courts has contributed to increasing procedural efficiency and reducing the workload of judicial staff⁴⁹. For example, as we have shown, in China, artificial intelligence systems are already used to automate the drafting of decisions and for legal assistance in courts, through solutions such as the "206 System", which supports judges in analysing cases and formulating motivations⁵⁰.

2. *Support in criminal decision-making.* AI provides support in assessing the risk of recidivism, determining the proportionality of the sanction and analysing aggravating or mitigating circumstances, through algorithmic models such as COMPAS. These tools contribute to standardizing decisions and avoiding arbitrary differences between courts⁵¹.

3. *Detection of judicial errors and systemic behaviours.* Algorithms can detect statistical anomalies in sentencing, suggesting possible errors or discriminatory practices. This capability can be used to audit the judicial system and promote more fair and transparent criminal justice⁵².

4. *Crime prevention and strategic resource allocation.* Through predictive data analysis, AI allows for the anticipation of high-crime risk areas, supporting law enforcement in the preventive management of the criminal phenomenon. Thus, police resources can be directed more efficiently based on the predictions generated.

Limitations of the use of AI in Criminal Justice

1. *Algorithmic bias and systemic discrimination.* One of the most criticised aspects of AI in justice is the risk of perpetuating or amplifying biases already existing in the historical data used to train the algorithms. Studies have shown that risk assessment systems like COMPAS can discriminate based on race or socio-economic status⁵³, affecting the principle of equality before the law.

2. *Lack of transparency and explainability of decisions.* Many AI models operate as "black boxes", meaning they provide results without providing intelligible

⁴⁹For empirical evidence demonstrating a ~40% reduction in trial duration and ~55% reduction in error rates in AI-assisted criminal case proceedings, see Chen (2025).

⁵⁰For a detailed analysis of how Chinese courts use AI to assist in judgment drafting and legal decision-making support, including the deployment of tools such as the 206 System, see Li (2025).

⁵¹Dressel & Farid (2018) at 6–7.

⁵²Ryberg (2025) at 203–220.

⁵³Angwin, Larson, Mattu & Kirchner (2016) at 4–5.

explanations of the internal decision-making process. This opacity affects the right of litigants to a fair trial, as they cannot effectively challenge an automated decision⁵⁴.

3. *Legal liability issues.* In cases where AI produces erroneous decisions or contributes to judicial abuses, legal liability becomes difficult to delimit: it is not clear whether liability belongs to the programmer, the user institution or the system itself. This ambiguity affects the principle of legality and imputability, fundamental to criminal law—issues directly addressed by the AI Act (2024), complemented by proposals such as the Artificial Intelligence Liability Directive (AILD), designed to provide legal clarity in cases of harm caused by AI⁵⁵.

4. *Reducing the human role and the risk of excessive delegation.* Excessive reliance on AI in the criminal process can lead to a technocratisation of the act of justice, in which decisions are made automatically, without human nuance⁵⁶. This phenomenon contradicts the personal and deliberative nature of the judicial act.

In conclusion, AI has the potential to revolutionise criminal justice, but only if implemented with accountability, transparency, and robust human oversight. It is essential that the use of this technology is accompanied by clear legal frameworks, independent algorithmic audits, and effective mechanisms to challenge automated decisions. In the absence of these safeguards, AI risks becoming not an instrument of justice, but one of systemic injustice.

Ethical and Legal Aspects in the use of Artificial Intelligence in Criminal Justice

The use of artificial intelligence (AI) in criminal and law enforcement systems raises a number of complex challenges, both from an ethical and legal perspective, as these technologies directly interact with fundamental rights and freedoms, such as the right to privacy, to a fair trial and to protection against discrimination⁵⁷. AI, particularly in criminal justice, is not just a technological tool, but also an assisted decision-maker, the results of which can seriously influence human destinies. Therefore, the analysis of these aspects must include considerations related to accountability, transparency, fairness and the protection of personal data.

Legal Liability of AI-assisted decisions

Who is responsible for a wrong decision generated by AI? This is one of the most pressing legal questions. Traditionally, legal responsibility is clearly attributed to the person making the decision; in the case of AI, however, decisions can be influenced by an opaque model trained on historical data sets⁵⁸. This raises questions regarding the principle of legality and imputability, which are essential in criminal law.

⁵⁴Wachter, Mittelstadt, & Floridi (2017b) at 77–78.

⁵⁵For a legal discussion on the ambiguity surrounding liability for harm caused by AI systems—including whether programmers, deploying entities or the systems themselves should be held accountable—see the AI Act (EU 2024/1689) and the proposed AI Liability Directive.

⁵⁶Nikpayam, Kremer & de Véricourt (2024).

⁵⁷Popescu (2014) at 127.

⁵⁸Santoni & Mecacci (2021) at 1057–1084.

Since artificial intelligence does not have legal personality, it cannot be held directly liable under criminal or civil law. Instead, liability for harm caused by the use of AI systems can be attributed to the human or institutional actors involved – developers, system providers, operators or user public entities. According to Regulation (EU) 2024/1689 on artificial intelligence (AI Act), these categories of actors have distinct obligations of compliance, human oversight and traceability, in particular in the case of high-risk systems⁵⁹. However, in the absence of clear rules on the distribution of liability in cases of harm, the traceability of automated decisions remains problematic, generating potential legal accountability gaps and affecting the fundamental rights of litigants.

Transparency and Explainability of Algorithmic Decisions

AI systems used in criminal proceedings must be sufficiently transparent to allow for the verification of decision-making logic. This requirement is essential for respecting the right to a fair trial.

Most machine learning systems, especially deep learning, operate as “black boxes,” meaning they provide results without the decision-making process being clearly traceable or explained. As such, they are extremely difficult to interpret even for specialists⁶⁰. This algorithmic opacity, or lack of explainability, affects the right to a fair trial, as the defendant or his lawyer cannot challenge an automated decision without understanding the reasoning behind it. Thus, the lack of a “right to explanation” affects the principles of procedural transparency enshrined in democratic justice systems.

Fairness and the Risk of Algorithmic Discrimination

Another critical aspect is the discriminatory potential of AI, which can perpetuate or even accentuate existing inequalities in the criminal justice system. Algorithmic bias is a recognised problem in AI systems, especially when they are trained on historical data reflecting racial, social, or gender discrimination. Thus, recidivism risk assessment systems, such as COMPAS, have been criticised for racial bias, as the algorithms were based on historical data reflecting systemic disparities⁶¹. These issues demonstrate that AI is not neutral, but reflects values, biases and past decisions, which necessitates the need for rigorous mechanisms for ad-hoc audit, correction and evaluation. In criminal justice, this can lead to excessive surveillance of certain groups, erroneous arrests or disproportionate sentences⁶².

⁵⁹For a legal framework assigning obligations of traceability, accountability, and human oversight to AI providers and users, especially for high-risk systems, see Regulation (EU) 2024/1689 (AI Act).

⁶⁰Şahin, Arslan & Özdemir (2024) at 859–965.

⁶¹Angwin, Larson, Mattu & Kirchner (2016) at 4-5.

⁶²For empirical and ethical discussions illustrating how AI reflects historical bias and can lead to disproportionate policing and sentencing—thus necessitating robust audit mechanisms—see Ho, Yi-Jen (Ian), Jabr, W. & Y. Zhang (2023). Bharati (2024) and Zilka, Fogliato, Hron, Butcher, Ashurst & Weller (2023).

Right to Privacy and Data Protection

The use of AI in criminal justice involves the processing of a considerable amount of sensitive data – such as criminal records, behavioural data, biometrics or social connections – which are protected by the GDPR (Articles 9 and 10) and the AI Act (EU 2024/1689)⁶³. According to these regulations, processing must be lawful, proportionate and accompanied by appropriate safeguards, in particular when automated systems influence decisions that may profoundly affect the fundamental rights of data subjects.

Technologies such as facial recognition in public spaces or predictive surveillance raise serious privacy concerns, especially in the absence of explicit consent and a solid legal basis. According to the AI Act and the European Data Protection Board (EDB) guidelines, these systems must comply with the principles of legality, necessity, proportionality and data minimisation. In the absence of solid safeguards, there is a risk of widespread algorithmic surveillance, incompatible with European fundamental rights rules⁶⁴.

Inefficiency of the current Legal Framework

Currently, most criminal and criminal procedural legislation does not contain clear rules on the use of AI. The European Union proposed an Artificial Intelligence Act in 2021 to regulate high-risk AI applications, including those in the judiciary, thus adopting Regulation (EU) 2024/1689, also known as the EU AI Act, in force since 1 August 2024, which imposes clear obligations for high-risk AI systems⁶⁵. However, the definition of AI systems is still considered vague, which generates uncertainty for implementers and authorities.

Artificial intelligence in criminal justice brings major opportunities for system reform, but also profound ethical and legal dilemmas that cannot be ignored. For technology to serve justice and not subvert it, it is essential that AI is governed by principles of transparency, accountability, fairness and respect for fundamental rights.

The need for a clear and comprehensive legal framework that regulates the rights, obligations and limitations of the use of AI in the criminal justice system is imperative to avoid abuses, and the legal and institutional architecture must be constantly adapted to respond to these challenges in a responsible and sustainable manner⁶⁶.

⁶³For a detailed analysis of the alignment—or potential conflict—between the human oversight requirement under Article 14 of the AI Act and the automated decision safeguards under Article 22 GDPR, see Sarra (2025) at 45–62.

⁶⁴European Data Protection Board (2021) clearly states that the use of facial recognition in public spaces entails a major risk to privacy and must be strictly limited. See also Article 5 of the AI Act (Regulation (EU) 2024/1689), which prohibits certain uses of AI, including real-time biometric surveillance in public spaces, except under narrowly defined circumstances. For a detailed discussion on the risks of opaque automated decision-making and its impact on fundamental rights, including the right to privacy and individual freedom, see Wachter, Mittelstadt & Floridi (2017b) at 76–99.

⁶⁵Regulation (EU) 2024/1689 (“EU AI Act”), article 6 and article 16.

⁶⁶For an analysis of the evolving legal and regulatory complexities surrounding AI—from international governance to EU policy alignments—see Ori Igwe (2024) at 737–764.

In Romania, the application of AI in criminal justice is still in an exploratory phase. There are doctrinal positions and legal analyses that signal the potential use of facial recognition, crime analysis algorithms, but there is no specific legislative framework or strategic plan dedicated to the application of AI in the criminal system⁶⁷.

The studies proposed by the National Institute of Magistracy and the Ministry of Justice on the automation of deliberations and case sorting show increased interest, but also justified caution regarding the ethical implications⁶⁸.

Conclusions

Through this study, we aimed to provide a critical analysis, based on recent scientific evidence, of the impact of artificial intelligence on the criminal field – both as an emerging challenge and as a unique opportunity for justice reform. We have therefore shown that artificial intelligence represents a tool with immense potential in the field of public security, but also an emerging vector of modern crime. Without a doubt, the use of artificial intelligence in predictive analysis, surveillance, fraud detection and criminal investigation can bring efficiency and precision in combating crime. At the same time, AI technologies can be exploited for illicit purposes, generating significant risks related to cyber attacks, information manipulation and abusive surveillance.

According to international case studies, it has been observed that the level of integration of AI in the field of crime is variable, influenced by political, economic and institutional factors. While some states prioritise innovation, others insist on strict regulations to protect fundamental rights. So, AI is neither beneficial nor evil in itself: its value depends on the normative, ethical and institutional context in which it is implemented. A coherent framework, based on human rights, transparency and responsible innovation, is essential for AI to serve security without sacrificing freedom.

As a concluding direction of this study, we believe that future research should focus on strengthening a responsible and equitable approach to the use of artificial intelligence in criminal justice. In this regard, it is necessary, first of all, to develop ethical, transparent and periodically audited algorithms, capable of ensuring respect for fundamental rights and avoiding the amplification of systemic inequalities. It is also essential to critically assess the impact of AI technologies on vulnerable groups, in order to prevent discrimination and social exclusion. Another priority direction aims to analyse the effectiveness of governance mechanisms and the legal accountability framework in the use of AI in the criminal field, in particular with regard to attributing responsibility for algorithmic decisions. At the same time, comparative studies between legal and institutional regimes in different countries could contribute to identifying best practices in regulating the use of AI. Finally, the role of international cooperation becomes crucial in the face of cross-border challenges

⁶⁷For a critical overview of Romania's initial engagement with artificial intelligence in criminal justice and the absence of a coherent regulatory framework see Stănilă (2020).

⁶⁸See the National Institute of Magistracy (2023), *Study on the Applicability of Artificial Intelligence in Romanian Criminal Justice*, for an analysis of the technological risks and opportunities associated with automated decision-making in the national legal context.

generated by AI-assisted crime, and it is necessary to articulate effective global mechanisms for the prevention, detection and sanctioning of this emerging phenomenon. The proposed directions for regulation and practical application can contribute to a balanced and responsible use of AI, for the benefit of a modern, efficient and humane criminal justice.

References

- Angwin, J., Larson, J., Mattu, S. & L. Kirchner (2016). 'Machine bias: There's software used across the country to predict future criminals. And it's biased against blacks' in *ProPublica*. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.
- Bharati, R.K. (2024). 'Ethical Implications of AI in Criminal Justice: Balancing Efficiency and Due Process' in *Research Review International Journal of Multidisciplinary* 9(7): 93–105. <https://doi.org/10.31305/rrijm.2024.v09.n07.014>.
- Binns, R. (2020). 'Human judgement in algorithmic loops: Decision-making, fairness, accountability, and transparency' in *AI & Society*, 35(2). <http://dx.doi.org/10.2139/ssrn.3452030>.
- Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B. & D. Amodei (2018). *The malicious use of artificial intelligence: Forecasting, prevention, and mitigation*, 50–51. <https://doi.org/10.48550/arXiv.1802.07228>.
- Brundage, M., Avin, S., Wang, J., Krueger, G., Hadfield, G. & H. Belfield (2023). 'Responsible artificial intelligence: Managing risk and regulation in an evolving landscape' in *AI & Society*. 1802.07228.
- Brynjolfsson, E. & A. McAfee (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. W. W. Norton & Company.
- Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1). <https://doi.org/10.1177/2053951715622512>.
- Chen, Q. (2025). 'Improving the trial efficiency of criminal cases with the assistance of artificial intelligence' in *Discover Artificial Intelligence* 5, 110. <https://doi.org/10.1007/s44163-025-00353-2>.
- Chan, J. (2022). *Facial Recognition Technology and Ethical Issues*. Wellington Faculty of Engineering Ethics and Sustainability Symposium, 1. <https://doi.org/10.26686/wfeess.vi.7647>.
- Chesney, R. & D.K. Citron (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. *California Law Review*, 107(6), 1753–1819. <https://doi.org/10.2139/ssrn.3213954>.
- Chris, E., Olabiyi, W., Favor, B. & T. Ramon (2025). *AI-powered surveillance systems and anomaly detection*. [Preprint]. ResearchGate. https://www.researchgate.net/publication/387933428_AI-powered_surveillance_systems_and_anomaly_detection.
- Di Nicola, A. (2022). *Towards digital organized crime and digital sociology of organized crime. Trends in Organized Crime*. Springer. <https://doi.org/10.1007/s12117-022-09457-y>.
- Dressel, J. & H. Farid (2018). 'The accuracy, fairness, and limits of predicting recidivism' in *Science Advances*, 4(1), eaao5580. <https://doi.org/10.1126/sciadv.aao5580>.
- Dul, C. (2022). 'Facial Recognition Technology vs Privacy: The Case of Clearview AI' in *Queen Mary Law Journal*, 2022(1):1–24. <https://qmro.qmul.ac.uk/xmlui/handle/123456789/80559>.

- Duong, H.-T., Le, V.-T. & Hoang, V.T. (2023). 'Deep Learning-Based Anomaly Detection in Video Surveillance: A Survey' in *Sensors*, 23(11), 5024. <https://doi.org/10.3390/s23115024>.
- Eze, C.S. & L. Shamir (2024). *Analysis and prevention of AI-based phishing email attacks*. *arXiv*. <https://doi.org/10.48550/arXiv.2405.05435>.
- Farber, S. (2025). 'The evolving nexus of cybercrime and terrorism: A systematic review of convergence and policy implications' in *Security Journal*, 38(1), art. 29. <https://doi.org/10.1057/s41284-025-00471-7>.
- Garrett, B.L. (2025). *Artificial Intelligence and Procedural Due Process*. *Duke Law School Public Law & Legal Theory Series*, No. 2025-15. SSRN. <https://doi.org/10.2139/ssrn.5093989>.
- Goodfellow, I.J, Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. & Y. Bengio, Y. (2014). 'Generative Adversarial Nets' in *Advances in Neural Information Processing Systems*, 27. <https://arxiv.org/abs/1406.2661>.
- Gupta, M., Akiri, C., Aryal, K., Parker, E. & L. Praharaj (2023). *From ChatGPT to ThreatGPT: Impact of Generative AI in Cybersecurity and Privacy*. 80218-80244. IEEE Access. <https://doi.org/10.1109/ACCESS.2023.3300381>.
- Ho, Yi-Jen (Ian), Jabr, W. & Y. Zhang (2023). 'AI Enforcement: Examining the Impact of AI on Judicial Fairness and Public Safety' in SSRN Electronic Journal. <http://dx.doi.org/10.2139/ssrn.4533047>.
- Igwe, O. (2024). 'Artificial Intelligence: A Twenty First Century International Regulatory Challenge' in *Athens Journal of Law* 10(4):737–764. <https://doi.org/10.30958/ajl.10-4-19>.
- Joh, E.E. (2020). 'Artificial intelligence and policing: First questions' in *Seattle University Law Review* 41(4):1139–1154. <https://digitalcommons.law.seattleu.edu/sulr/vol41/iss4/7/>.
- Li, Z. (2025). 'AI and Human Judges in Chinese Courts' (rev. 6 June 2025) in SSRN. <https://doi.org/10.2139/ssrn.5235753>.
- Lum, K. & W. Isaac (2016). 'To predict and serve?' in *Significance* 13(5):14–19. <https://doi.org/10.1111/j.1740-9713.2016.00960.x>.
- Lynch, N. (2024). 'Facial recognition technology in policing and security—Case studies in regulation' in *Laws*, MDPI, 13(3):1-14., 35. <https://doi.org/10.3390/laws13030035>.
- Maras, M.-H. & A. Alexandrou (2019). 'Determining authenticity of video evidence in the age of artificial intelligence and in the wake of Deepfake videos' in *International Journal of Evidence & Proof* 23(3):255–262. <https://doi.org/10.1177/1365712718807226>.
- National Institute of Magistracy. (2023). *Studiu privind aplicabilitatea inteligenței artificiale în justiția penală română [in English: Study on the Applicability of Artificial Intelligence in Romanian Criminal Justice]*. Bucharest: NIM.
- Nikpayam, H., Kremer, M. & F. de Véricourt (2024). *When Delegating AI-Assisted Decisions Drives AI Over-Reliance*. SSRN. <https://ssrn.com/abstract=4966186> or <http://dx.doi.org/10.2139/ssrn.4966186>.
- Oswald, M., Grace, J., Urwin, S. & G.C. Barnes (2018). 'Algorithmic risk assessment policing models: Lessons from the Durham HART model and 'Experimental' proportionality' in *Information & Communications Technology Law*. 27(2):223–250. <https://doi.org/10.1080/13600834.2018.1458455>.
- Papasavva, A., Johnson, S., Lowther, E., Lundrigan, S., Mariconti, E., Markovska, A. & N. Tuptuk (2025). 'Applications of AI-Based Models for Online Fraud Detection and Analysis: A Systematic Literature Review' in *Crime Science* 14: 7. <https://doi.org/10.1186/s40163-025-00248-8>.
- Popescu-Ljungholm, D. & Tolbaru, C.-E. (2025). 'The impact of artificial intelligence on fundamental human rights in eu countries. Examination of academic studies and the positions of non-governmental organisations regarding the ethical and legal risks of AI' in *Lex et Scientia no. XXXII*, vol.1: 47-62.

- Popescu, D. (2014). *Protectia juridica a drepturilor omului/ Legal protection of human rights*. Craiova: Sitech.
- Ryberg, J. (2023). 'Sentencing disparity and artificial intelligence' in *Journal of Value Inquiry* 57(3):447–462. <https://doi.org/10.1007/s10790-021-09835-9>.
- Ryberg, J. (2025). 'Criminal Sentencing and Artificial Intelligence: What is the Input Problem?' in *Criminal Law, Philosophy* 19:203–220. <https://doi.org/10.1007/s11572-024-09739-2>.
- Santoni de Sio, F. & G. Mecacci (2021). 'Four Responsibility Gaps with Artificial Intelligence: Why They Matter and How to Address Them' in *Philosophy & Technology* 34(4):1057–1084. DOI: 10.1007/s13347-021-00450-x.
- Sarra, C. (2025). 'Artificial Intelligence in Decision-making: A Test of Consistency between the "EU AI Act" and the "General Data Protection Regulation"' *Athens Journal of Law* 11(1):45–62. <https://doi.org/10.30958/ajl.11-1-3>.
- Stănilă, L.M. (2020). *Inteligența artificială, drept penal și sistemul de justiție penală: Amintiri despre viitor*. Universul Juridic.
- Schmitt, M. & I. Flechais (2024). 'Digital deception: generative artificial intelligence in social engineering and phishing' in *Artificial Intelligence Review* 57:324. <https://doi.org/10.1007/s10462-024-10973-2>.
- Shelley, L. (2019). The globalization of crime. In M. Natarajan (Ed.), *International and transnational crime and justice*, 223–228. Cambridge University Press. <https://doi.org/10.1017/9781108597296.037>.
- Şahin, E., Arslan, N.N. & D. Özdemir (2024). 'Unlocking the black box: an in-depth review on interpretability, explainability, and reliability in deep learning' in *Neural Computing and Applications* 37:859–965. <https://doi.org/10.1007/s00521-024-10437-2>.
- Tahura, U. S. & N. Selvadurai (2022). 'The use of AI in judicial decision-making: the example of China' in *International Journal of Law, Ethics and Technology* 2022(3):1–20. <https://doi.org/10.55574/PYEB5374>.
- Tolbaru, C. (2024). 'Cyber threats in the European space. Security measures at the level of the European Union' in *Bulletin of the Transilvania University of Braşov, Series VII: Social Sciences, Law* 17(66):79-86. <https://doi.org/10.31926/but.ssl.2024.17.66.3.9>.
- Tolbaru, C.-E. (2023). Considerations on Combating Money Laundry in the Field of Crypto-Assets, at European Union Level. *RAIS Conference Proceedings*, August 3-4, 2023: 151-156. <https://doi.org/10.5281/zenodo.8310135>.
- Van Brakel, R.E. (2021). 'Rethinking predictive policing: Towards a holistic framework of democratic algorithmic surveillance' in M. Schuilenburg & Peeters, R. (Eds.), *Algorithmic Societies: Power, Knowledge and Technology in the Age of Algorithms*, 104–118. Routledge. 9780367204310_proofs_0.pdf.
- Veale, M. & L. Edwards (2018). 'Clarity, surprises, and further questions in the Article 29 Working Party draft guidance on automated decision-making and profiling' in *Computer Law & Security Review* 34(2):398–404. <https://doi.org/10.1016/j.clsr.2017.12.002>.
- Wachter, S., Mittelstadt, B. & L. Floridi (2017a). 'Transparent, explainable, and accountable AI for robotics' in *Science Robotics* 2(6), eaap6962. DOI: 10.1126/scirobotics.aan6080.
- Wachter, S., Mittelstadt, B. & L. Floridi (2017b). 'Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation' in *International Data Privacy Law* 7(2):76–99. <https://doi.org/10.1093/idpl/ixp005>.
- Wang, Y., Kung, L. & T.A. Byrd (2018). 'Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations' in *Technological Forecasting & Social Change* 126:3–13. <https://doi.org/10.1016/j.techfore.2015.12.019>.
- Wright, D. (2020). 'AI and the future of crime prevention: Ethical challenges and legal dilemmas' in *Journal of Ethics and Information Technology* 22(2):145–158. <https://doi.org/10.1007/s10676-020-09542-1>.

- Zeng, Y., Lu, E., Sun, Y. & R. Tian (2019). 'Responsible Facial Recognition and Beyond'. arXiv. <https://doi.org/10.48550/arXiv.1909.12935>.
- Zerilli, J., Knott, A., Maclaurin, J. & C. Gavaghan (2019). 'Transparency in algorithmic and human decision-making: Is there a double standard?' in *Philosophy & Technology* 32(4): 661–683. <https://doi.org/10.1007/s13347-018-0330-6>.
- Zilka, M., Fogliato, R., Hron, J., Butcher, Ashurst, C. & A. Weller (2023). *The progression of disparities within the criminal justice system: Differential enforcement and risk assessment instruments*. arXiv. <https://doi.org/10.48550/arXiv.2305.07575>.

Legal Instruments

- Council of Europe. (2024). *Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*. CETS 225 - Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law.
- Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (EU AI Act). *Official Journal of the European Union*, L 189, 1–257.
- European Commission. (2022). Proposal for an Artificial Intelligence Liability Directive (AILD). AI Liability Directive.
- European Commission. (2021). *Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act)*. COM(2021) 206 final.
- European Data Protection Board. (2021). *Guidelines 05/2021 on the use of facial recognition technology in the area of law enforcement*. https://edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-052021-use-facial-recognition-technology-law_en
- Europol. (2023a). *AI and Policing: The Benefits and Challenges of Artificial Intelligence for Law Enforcement*. The Hague: Europol. <https://www.europol.europa.eu/cms/sites/default/files/documents/AI-and-policing.pdf>
- Europol. (2023b). *Internet Organised Crime Threat Assessment (IOCTA)*. The Hague: Europol.
- Europol Innovation Lab. (2022). *Facing Reality? Law Enforcement and the Challenge of Deepfakes*. The Hague: Europol. https://www.europol.europa.eu/cms/sites/default/files/documents/Europol_Innovation_Lab_Facing_Reality_Law_Enforcement_And_The_Challenge_Of_Deepfakes.pdf.
- OECD - Organisation for Economic Co-operation and Development. (2019). *Artificial Intelligence in Society: Emerging Trends and Challenges*. OECD Publishing. <https://doi.org/10.1787/eedfee77-en>.
- United Nations Office on Drugs and Crime (UNODC). (2021). *The Use of Artificial Intelligence to Combat Criminal Activities*. Vienna: UNODC.