

THE IMPACT OF AI ON THE PRISON SYSTEM: THE RE-EDUCATION OF PRISONERS

Isabela Domingos¹

the Universidade Federal de Santa Catarina

Nicoletta Castellano²

“La Sapienza”, University of Rome

DOI: <https://doi.org//10.62140/IDNC710773640>

Summary: 1. Introduction 2. The Recommendation of the Council of Europe CM/REC(2024)5 2.1 The Recommendation and its implications 2.2 The Italian Penitentiary system 2.3 AI in Finnish prisons 3. Risks of Artificial Intelligence: between predictive efficiency and ethical challenges 4. Concluding remarks

Abstract: This contribution revolves around the analysis of the Recommendation of the Committee of Ministers of the Council of Europe of 09 October 2024, CM/Rec(2024)5, in which for the first time this body questioned the effects, both beneficial and pathological, of a gradual introduction of artificial intelligence systems in the prison world.

First of all, these are seen as methods that do not replace human beings, but rather as means of enhancing human activities, precisely because they can have a positive impact on the control and surveillance of inmates, by assisting prison staff.

The Recommendation deals with two topics that, in the writer's opinion, are particularly pregnant:

¹ PhD Student in Administrative law at the Universidade Federal de Santa Catarina, Brazil.

² PhD Student in Constitutional law at “La Sapienza”, University of Rome, Italy.

- (i) the concept of cognitive BIAS applied to artificial intelligence;
- ii) the role that artificial intelligence can play in the resocialisation of offenders after release.

It is precisely on this second point that the author's attention will focus, as she has always been interested in the constitutional face of punishment in her own reference system, the Italian one, in which the third paragraph of Article 27 of the Constitution stands out clearly.

In spite of the risks that AI may entail in its very wide field of application, however, the Committee of Ministers highlights how technology can easily put prisoners in touch with their families and their social context of reference, through the use of suitably prepared platforms.

Moreover, AI techniques can also be used to identify a risk of recidivism in the commission of offences: this too, however, must take place in the light of a balanced and in any case human control, since human relationship and assessment cannot be easily replaced by a machine without 'human' reasoning, but only technical reasoning.

Furthermore, the dissertation here focuses on the practical application of artificial intelligence systems in European prisons: a striking example is the Finnish case, a country that was the first to conceptualize, design, and implement artificial intelligence assistants to support the lives of prisoners, but above all, to aid in their re-education and social reintegration.

Introduction

This dissertation, focused on the CM/REC(2024)5 Recommendation of the Committee of Ministers of the Council of Europe and on the practical and theoretical implications that artificial intelligence, with its biases, can have in the field of criminal surveillance, starts from a fundamental observation: if it is true that punishment must rehabilitate, artificial intelligence, in the intentions of the international legislator, must also be constitutionalized, tamed by human norms. The risk that artificial intelligence may stray from the human purposes for which it was created is indeed present and denying it would be a great demonstration of negligence on the part of legal practitioners.

The essay, written by two authors, primarily consists of two parts: the first part, of an explanatory and programmatic nature, written by Dr. Nicoletta Castellano, focuses mainly on the connection between the international legislation produced by the Council of Europe and Italian legislation, which finds its rationale in the third paragraph of Article 27 of the Constitution (chapter 2).

On the other hand, the second part deals with the ethical implications of AI applied to the prison system, such as algorithmic visions and their risks for human rights, was written by Dr. Isabela Domingos (chapter 3).

2. The Recommendation of the Council of Europe CM/REC(2024)5 and the consequences on European legal systems

2.1 The Recommendation and its implementation

The Council of Europe, in the composition of the Committee of Ministers, has shown considerable interest in the issue of the reintegration of convicts after the expiration of their sentence, dedicating a Recommendation on the subject in October 2024³.

The Council of Europe, in the composition of the Committee of Ministers, has shown considerable interest in the issue of the reintegration of

³ This recommendation was undoubtedly prompted by EU acts, as well as other previous acts of the Council of Europe, including, for example, inter alios, the AI Act (at the EU level) and other recommendations, including CM/REC(2021)8. From a jurisprudential perspective, the Strasbourg Court dealt with a case of interference between AI and procedural-criminal rules in *Nuh Uzgun and others v. Turkey* (2022), involving fourteen Turkish citizens who had, for various reasons, attempted to participate in the military coup of July 2016 in Turkey. For the Strasbourg Judges, the violation of the ECHR was based on Article 8, concerning the secrecy of communications and correspondence, as the individuals involved had been subjected to unlawful harassment while deprived of their personal liberty, as their correspondence with their defense attorneys had been violated.

convicts after the expiration of their sentence, dedicating a Recommendation on the subject in October 2024⁴.

The CM/REC(2024)5 focuses on a central goal⁵, that is to guide the legislation of member states toward harmonizing regulations on AI (Artificial Intelligence)⁶ and the rehabilitative effect of criminal sanctions⁷.

First and foremost, the Recommendation emphasizes, from its outset, that criminal sanctions are an exercise of public power; as an expression of public power, interferences with human rights, human dignity, and privacy are risky and undoubtedly present⁸.

An extremely debated issue that carries with it most of the practical risks is that of BIAS⁹. Bias is defined as follows: “It is a systematic pattern of deviation from the norm or from rationality in judgmental mental processes. In psychology, it refers to a tendency to create one's own subjective reality, not necessarily corresponding to the evidence, developed based on the

⁴ The preparatory work for this Recommendation originated from the Annual Conference on Penological Cooperation (PC-CP) held on November 22-23, 2021, in Strasbourg, where the first guiding criteria for drafting a document on the subject were developed.

⁵ “*Recommends that governments of member States:*

- *be guided in their legislation, criminal policy and practice by the principles and rules contained in the appendix to this recommendation;*

- *ensure that this recommendation and its explanatory memorandum are translated and disseminated as widely as possible and, more specifically, among judicial authorities, prosecution, police, prison, probation and juvenile justice services, as well as among private companies which design and provide AI and related digital technologies in the framework of the criminal justice system”.*

⁶ The definition of Artificial Intelligence used in CM/REC(2024)5 was taken from CM/REC(2021)8, and is as follows: “(AI) means a system that is either software-based or embedded in hardware devices, and that displays intelligent behaviour by, inter alia, collecting and processing data, analysing, and interpreting its environment, and by taking action, with some degree of autonomy, to achieve specific goals.”

⁷ Article 3 of the ECHR, also a product of the Council of Europe, states that torture, as well as inhuman and degrading punishment or treatment, are prohibited.

⁸ FORMISANO, A. *L'impatto dell'intelligenza artificiale in ambito giudiziario sui diritti fondamentali*, in *Federalismi.it*, fasc. 22/2024, pp. 112 ss.

⁹ GIACOMO, L. D. *Bias dell' algoritmo: responsabilità e innovazione nell'AI*. Available at: <https://www.diritto.it/bias-algoritmo-responsabilita-ai-innovazione/> Accessed on 20 feb. 2025.

interpretation of available information”, even if not logically or semantically connected, which leads to a misjudgment or a lack of objectivity in judgment¹⁰.

Cognitive biases are forms of evolved mental behavior: some represent forms of adaptation, as they lead to more effective actions in certain contexts or allow for faster decision-making when most needed; others arise from the lack of appropriate mental mechanisms or from the incorrect application of a mechanism that would otherwise be positive in other circumstances. This phenomenon is studied by cognitive sciences and social psychology¹¹.

The etymology of this word is uncertain, as it likely derives from the English "bias,"¹² which properly means "oblique," "deviant," indicating the term itself refers to deviation from the expected outcome under impartial, objective, and neutral conditions.

The drafters of the Recommendation also do not forget how fast the arms race in Artificial Intelligence is progressing: even the pace followed in this direction moves at inhuman, robotic speeds, not entirely human.

If the speed is not human, the use of AI mechanisms must be human: therefore, in this regard, the discussion revolves around “*ethical use of Artificial Intelligence*”¹³, companies involved in scientific-engineering progress in the field of AI, and which similarly advance in the research and practical development of artificial intelligence tools, must keep in mind that in a complex system like

¹⁰ STATE OF MIND. *Bias Euristiche*. Available at: <<https://www.stateofmind.it/bias/>>. Accessed on 20 feb. 2025.

¹¹ STATE OF MIND. *Bias Euristiche*. Available at: <<https://www.stateofmind.it/bias/>>. Accessed on 20 feb. 2025.

¹² CAMBRIDGE. *Bias*. Available at: <<https://dictionary.cambridge.org/dictionary/english-portuguese/bias>>. Accessed on 20 feb. 2025.

¹³ By "ethics" in the context of artificial intelligence, we refer to all those areas that require a significant and non-reductive involvement of ethically evaluable concepts, such as: making decisions, respecting privacy, and the surveillance of incarcerated individuals who are therefore deprived of their personal freedom.

criminal sanctioning, they are dealing with human subjects, not just biometric and statistical data.

The conditions of detention can certainly be improved with an ethical use of artificial intelligence, as it acknowledges and demonstrates that new horizons exist which, until recently, could not even have been remotely imagined¹⁴.

Perhaps hoped for.

Like all products of the technological age, Artificial Intelligence—which now takes on the most varied forms—risks becoming an ambivalent tool, given the undeniable implications that a negative and pretextual use of the tool can have¹⁵.

2.2 The Italian Penitentiary system

With a brief introduction to this point, we now move to the heart of the matter, linking the theme of the application of artificial intelligence to human processes with the central theme in the Recommendation of the rehabilitation of the convict.

From the perspective of the author, the correspondence with the constitutional principles—certainly of a continental nature—of the reference legal system, namely the Italian system, is significant, as it seems that, from the viewpoint of the European body (which is, of course, not to be confused with the European Union), the guiding principle is the same as that which underpins the Italian legal system: rehabilitation.

¹⁴ MAGGIOLINI, P.; FUGINI, M. *Le sfide etiche dell'intelligenza artificiale*, Ledizioni, 2025.

¹⁵ RINALDI, L. *Intelligenza artificiale, diritti e doveri nella Costituzione italiana*, in DPCE Online, fascicolo 1/2022, pp. 201 ss.

Article 27, paragraph 3, of the Italian Constitution¹⁶ firmly expresses the principle that the penalty should primarily aim at the rehabilitation of the convict, while Article 3 of the ECHR¹⁷ that the penalty should not consist of inhuman or degrading treatment¹⁸.

Although more strictly penal theories have spread in the field¹⁹, over the decades, regarding the actual value of criminal sanctions, there is no doubt that, through the lens of a constitutionalist, one can only embrace the rehabilitative and social reintegration value of the penalty.²⁰

Indeed, in addition to the more technical theories on the matter, including the deterrent and retributive theories, one must not overlook the concept that criminal law must be integrated within the framework of

¹⁶ "Criminal responsibility is personal.

The defendant is not considered guilty until the final conviction.

Punishments cannot consist of treatments that are contrary to human dignity and must aim at the rehabilitation of the convict.

The death penalty is not permitted." (Article 27 of the Constitution)

¹⁷ "Nobody shall be subjected to torture or to inhuman or degrading punishment or treatment." (Article 3 of the ECHR)

¹⁸The concept of "prisoners' rights" is not explicitly codified in Article 3 of the European Convention on Human Rights; it is a term coined by case law from both the Convention and national and European doctrines, both conventional and EU-based. The principle, associated with the concept of penitentiary law, was primarily established in cases such as *Campbell and Fell v. the United Kingdom*, *X v. the Federal Republic of Germany*, *Ilse Koch v. Austria*, *Kotalla v. the Netherlands*, based on the succinct provision in Article 3 of the ECHR.

In these rulings, the so-called theory of implicit freedom was hypothesized and realized, according to which the set of rights theorized for man in the broadest sense has been made available even to those deprived of their personal liberty. For further clarification, see: GORI, A. *Condizioni di detenzione nelle carceri europee tra diritto dei detenuti e deroghe all'art. 3 della Convenzione*, in *Rivista ADIR*, UniFi, 2015, cap. 3.

¹⁹ DOLCINI, E. *La pena nell'ordinamento italiano, tra repressione e prevenzione*, in *Rivista Italiana di Diritto e Procedura Penale*, fasc. 2/2021, pp. 383 ss.

²⁰ The author believes that the more strictly technical theories on this point are particularly relevant in the field of strict criminal law, both substantive and procedural. However, from the perspective of a constitutional law dissertation, the rehabilitative function of punishment must always be considered to avoid moving toward paths that are not properly constitutionally oriented.

constitutional law, even if only through the theory of resolving contradictions and the hierarchical pyramid of sources.

If it is true, therefore, that the exercise of public power in the function of criminal punishment must be constitutionally oriented, so too must the use of effective but potentially dangerous tools be guided by solid constitutional mechanisms that are "in step with the times," meaning updated and capable of adapting flexibly to new socio-legal needs.

The Recommendation that is the subject of this research effectively illustrates how the primary goal to be pursued at the national level by public authorities is to provide a human judgment for something that is not human: public and judicial authorities must nonetheless oversee the use of Artificial Intelligence to ensure its use is conscious and ethically oriented.

In this case, in the opinion of the author, "ethics" is not meant in a morally charged sense: this term refers to a constitutional guide for addressing the issue, as the Recommendation was made within a forum, the Council of Europe, which has historically dealt with issues related to human rights.

The close link between the Council of Europe and the European Court of Human Rights is both its symbol and demonstration.

The Strasbourg Court has consistently addressed issues related to the rehabilitation of convicts, with or without the use of artificial intelligence: among these, see the case *Viola v. Italy*, regarding the unconstitutionality of the so-called "life sentence without parole," established by Article 4-bis of Italian Law No. 354/1975, known as the Penitentiary System²¹.

²¹ Marcello Viola had been sentenced by the Court of Assizes of Palmi in the late 1990s for the crimes of mafia-style criminal association, murder, kidnapping, and illegal possession of firearms; he had been under the regime of the so-called "hard prison" (Article 41-bis of the Penitentiary Law) since 2000. He had repeatedly requested special leave permits, which were never granted to him as a non-collaborative inmate with the justice system, since this was the only criterion used to practically assess the convict's "redemption." However, in 2019, the European Court of Human Rights denied this rigid presumptive automaticity, stating that cooperation with justice

From the literal and then critical exegesis of the text of the Recommendation, it is clear that the primary hope expressed by the European legislator is to apply the principle of legality even to mechanisms that, by nature, evade human treatment, which in Italian terms could be translated as a conscious use of the tool of the law reserve²².

2.3 AI in Finnish prisons

Finland represents one of the main examples of implementing artificial intelligence practices for the benefit of the incarcerated population. First of all, Scandinavian countries, in general, have always stood out for providing the best living conditions for prisoners²³; in these prison contexts, social reintegration is conceived as the main goal of imprisonment, which is clearly evident in the architecture of the prisons as well as the modernity and cleanliness – perhaps trivially – with which the prisoners' cells are treated.

The artificial intelligence systems adopted in Finnish prisons have been tested since 2018 under the "Smart Prison" project, aimed at implementing and improving the living conditions of Finnish prisoners by

cannot be the sole indicator of practical and concrete repentance for the purpose of granting special leave permits.

²² The principle of legality, expressed in the form of the law reserve – which, in Article 13 of the Constitution, also includes jurisdiction, in the sense that personal freedom can only be restricted in light of a reasoned act by the judicial authority – has been extensively discussed in doctrinal terms within procedural law. The main risk involved is the jeopardizing of decision-making mechanisms that, since the origins of law, have belonged to the human sphere: the judicial one. Among all, see GALLONE, G. *Riserva di umanità, intelligenza artificiale e funzione giurisdizionale alla luce dell'LA Act. Considerazioni (e qualche proposta) attorno al processo amministrativo che verrà*, in *Rivista Judicium*, Pacini Giuridica, 2024.

²³ OSSERVATORIO BALCANI E CAUCASO. *La Danimarca e le prigioni off-shore*. Available at: <https://www.balcanicaucaso.org/arec/Kosovo/La-Danimarca-e-le-prigioni-off-shore-215757> Accessed on 20 mar. 2025.

introducing systems for digitalizing requests and expressing needs within the prison environment²⁴.

Relevant examples include the establishment of email addresses for prisoners to communicate freely and independently with their loved ones, as well as access to online libraries and university courses that can be followed remotely.

The first “Smart Prison” project in Finland was established in Hämeenlinna, a women’s prison, where each cell was equipped with an artificial intelligence system named “Doris,” which allowed inmates to communicate within the prison system smoothly and easily, and partially with the outside world. The purpose of this installation was to simplify and improve the living conditions of individuals already suffering from the deprivation of personal liberty. The second “Smart Prison,” established in Eastern Finland, was Pyhäselkä in December 2022, and the third, Kylmäkoski, was designed in the autumn of 2023.

In addition to the daily use of artificial intelligence, which helps improve the daily lives of prisoners, AI has been used and is still employed to enhance and implement judicial systems for the so-called “probation service.” This refers to algorithms designed to create post-sentence rehabilitation systems tailored to individual prisoners, with the goal of avoiding unsuitable and ineffective treatment programs for reintegrating the offender into society after the detention period.

The use of artificial intelligence, not necessarily in its most recent forms, but through digital video conferencing platforms, has also been employed in various Italian university centers, allowing prisoners to obtain a university degree while maintaining their incarceration status.

²⁴ PUOLAKKA, P. *Smart prison and digital innovations in prison and probation service of Finland*, in *Rassegna Italiana di Criminologia*, fasc. 2/2023, pp. 153 ss.

An example of this is the Penitentiary University Center established at the "Rebibbia" Correctional Facility in Rome by the University of "La Sapienza." The Roman university has agreed with the Rebibbia prison to start a law course (among others) that enables prisoners to reintegrate socially and professionally once they have served their sentence.

3. Risks of Artificial Intelligence: Between predictive efficiency and ethical challenges

Societies are constantly changing, especially in the post-Covidh19 movement with the implementation of information and communication technology (ICT) in labor relations, public education, transparency in public-private contracts, and the judicial system.

This research addresses the use of Artificial Intelligence in the criminal justice system, specifically to identify the risks and benefits for society. Innovation must be in line with the public interest, with a focus on human dignity. This requires improving security in the collection of sensitive data, the efficiency of the administration of justice and the sustainability of policies aimed at reintegrating convicts into society.

The recent EU regulation 2024/1689 understands that Artificial Intelligence models generate systemic risks, so it aims to encourage the development of reliable and safe systems throughout the European Union (EU) single market in both the public and private sectors, guaranteeing the health and safety of EU citizens, as well as respect for fundamental rights²⁵

²⁵ EUR-Lex. *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance)*. Available at: <https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng>. Accessed on: 20 feb. 2025.

As for risk-based rules, it identifies the need to place rules on artificial intelligence and the prohibition of AI practices with a view to the transparency of AI systems through risk management and the supervision of these technologies before and after they enter the market. On the other hand, there is a need to encourage innovation for the development of small and medium-sized enterprises, for the economic promotion of the local community²⁶.

Among the risks pointed out, there is a need for education and professional training in digital literacy so that citizens can identify the risks of AI. On the other hand, there is also a need to include marginalized groups in technical and vocational courses for the development of emerging technologies aimed at inclusion and citizenship to reduce the discriminatory biases and bring greater plurality to innovations in justice systems.

To ensure safety in the use of artificial intelligence, it is essential to strengthen systems for monitoring and evaluating its results. This process should seek to identify and understand practices that promote social benefits. However, if it is found that the technology acts as an obstacle to the protection of fundamental rights guaranteed by the Constitution of the Italian Republic of 1947, its application must be carefully suspended²⁷.

The European Union (EU) has important documents on the impact of artificial intelligence (AI) on the criminal justice system. The Directive of the European Parliament and of the Council on the adaptation of the rules of non-contractual civil liability to artificial intelligence of August 28, 2022²⁸,

²⁶ EUR-Lex. *Regras para uma inteligência artificial de confiança na UE*. Available at: <https://eur-lex.europa.eu/PT/legal-content/summary/rules-for-trustworthy-artificial-intelligence-in-the-eu.html?fromSummary=31> Accessed on: 20 mar. 2025.

²⁷ SENATO IT. *Constitution of the Italian Republic*. Available at: https://www.senato.it/documenti/repository/istituzione/costituzione_inglese.pdf. Accessed on: 20 feb. 2025.

²⁸ EUR-Lex. *Diretiva do Parlamento Europeu e do Conselho relativa à adaptação das regras de responsabilidade civil extracontratual à inteligência artificial (Diretiva Responsabilidade da IA)*. Available at: <https://eur->

states that the opacity of AI systems makes it difficult to apply liability rules in the civil, administrative and criminal spheres. In this sense, the lack of regulation or inefficient regulation leads to legal uncertainty and high costs in dealing with victims.

In this sense, the application of AI in criminal law by police and judicial authorities in criminal cases is the focus of the report "Artificial Intelligence in Criminal Law" (A9-0232-2021)²⁹. The resolution identifies the trade-offs in the use of artificial intelligence, i.e. the risks and benefits. AI can make data collection more agile and efficient, bringing benefits to the European community by developing new policies whose goal is security and the protection of human dignity. However, it is necessary to establish rules and standards to ensure that AI follows ethical paths in its design, implementation, and use. Therefore, it requires constant monitoring and evaluation of the technologies used in the administration of justice to protect fundamental rights.

A critical aspect of the use of artificial intelligence (AI) in the prison system is algorithmic bias — machine learning based on discrimination based on race, ethnicity, gender, disability, and other marginalized groups. Pedro Domingos believes that there is a “master algorithm” capable of learning from any type of data, improving itself through association and making decisions³⁰. The challenge is to control the risks, such as data manipulation, violation of fundamental rights, including large-scale decision-making.

Predictive analytics uses historical data to train models that identify patterns for criminal detection algorithms that may include predictive

lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:52022PC0496. Accessed on: 20 feb. 2025.

²⁹ EUROPARL. *Relatório sobre a inteligência artificial no direito penal e a sua utilização pelas autoridades policiais e judiciárias em casos penais*. Available at: https://www.europarl.europa.eu/doceo/document/A-9-2021-0232_PT.html. Accessed on: 20 feb. 2025.

³⁰ About machine learning, see: DOMINGOS, P. *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*. New York: Basic Books, 2015.

variables³¹. This means learning from the response variables, making associations and using the input values to make decisions. Berk understands the potential risks of feeding algorithms that violate human rights. The risk lies in the quality and management of the data they feed, as well as how the changes are translated into actual decisions³², a fact that requires AI transparency for systems that are aligned with current legal and ethical guidelines.

Malek points out that machines inherit the cognitive biases of those who feed them. In a societal context that tolerates discrimination, discriminatory decisions involving the guarantee of freedom can reinforce prejudices based on color, race, socio-economic status and gender. However, it cannot be said that all algorithms built internationally with unconstitutional code are unconstitutional³³. The implementation of AI tools in the public sector is a complex issue, as it has a significant impact on marginalized groups.

Algorithmic discrimination generated by artificial intelligence generates financial, physical and psychological damage, and therefore affects human dignity, especially if biases are inserted that reinforce prejudices and inequalities³⁴, whose poor data management triggers low social trust in artificial intelligence systems.

Based on the results of this research, it can be seen that the mitigation of risks related to AI requires a holistic approach, ranging from robust regulation by the European Union (EU), as well as internal regulations

³¹ Prediction is the use of historical data to train models that identify patterns. See: MAHESH, Batta. Machine learning algorithms-a review. *International Journal of Science and Research (IJSR)*. [Internet], v. 9, n. 1, p. 381-386, 2020.

³² BERK, Richard A. "Artificial intelligence, predictive policing, and risk assessment for law enforcement." *Annual Review of Criminology* 4.1, p. 209-237, 2021.

³³ Prediction is the use of historical data to train models that identify patterns. See: MALEK, M. A. Criminal courts' artificial intelligence: the way it reinforces bias and discrimination. *AI Ethics* 2, 233-245, 2022.

³⁴ MANNES, A. Governance, Risk, and Artificial Intelligence. *AI Magazine*, 41(1), p. 61-69, 2020. <https://doi.org/10.1609/aimag.v41i1.5200>

to implement an ethical culture in the use of AI, the establishment of security standards, the protection of human rights and the transparency of systems, especially when used by the criminal justice system.

Concluding remarks

The use of predictive algorithms is associated with narrow artificial intelligence, sophisticated tools that use data and associations to try to anticipate behavior and criminal events. On the other hand, research shows the risks of biased algorithms that contaminate the predictive variables, perpetuating existing prejudices in society and widening inequalities.

The problem is not the use of AI in the criminal justice system, but the transparency of understanding how data is generated and transformed into decisions, what the evaluation criteria are, and whether they reinforce discriminatory practices that can lead to unfair and unequal decisions. For this reason, AI systems must be constantly evaluated to prevent abuse and manipulation, which requires strong regulation and greater societal oversight when the technology is used in the public sector, given its impact on the protection of fundamental rights.

Moreover, it is not a given that the results of artificial intelligence applied in the criminal and penitentiary fields, areas that are at the limit of constitutional criteria compliance, can also be positive. On the contrary, given the technological contribution of today's times for the development of new suitable treatment programs, it is desirable that technical expertise in the criminal and penitentiary fields merge with information technology and engineering skills to achieve humanly positive and acceptable objectives.

References

BERK, R. A. "Artificial intelligence, predictive policing, and risk assessment for law enforcement." *Annual Review of Criminology* 4.1, p. 209-237, 2021.

CAMBRIDGE. *Bias*. Available at: <<https://dictionary.cambridge.org/dictionary/english-portuguese/bias>>. Accessed on 20 feb. 2025.

DOLCINI, E. *La pena nell'ordinamento italiano, tra repressione e prevenzione*, in Rivista Italiana di Diritto e Procedura Penale, fasc. 2/2021, pp. 383 ss.

DOMINGOS, I. M. Los retos de la inteligencia artificial para la protección de los derechos humanos. *Diálogos en DH. Tecnologías de la información*. 1ed. Toluca: CDHEM, 2022, p. 79-86.

DOMINGOS, P. *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World* New York: Basic Books, 2015.

EUR-Lex. *Diretiva do Parlamento Europeu e do Conselho relativa à adaptação das regras de responsabilidade civil extracontratual à inteligência artificial (Diretiva Responsabilidade da LA)*. Available at: <https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:52022PC0496>. Accessed on: 20 feb. 2025.

EUR-Lex. *Regras para uma inteligência artificial de confiança na UE*. Available at: <https://eur-lex.europa.eu/PT/legal-content/summary/rules-for-trustworthy-artificial-intelligence-in-the-eu.html?fromSummary=31> Accessed on: 20 mar. 2025.

EUR-Lex. *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance)*. Available at: <https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng>. Accessed on: 20 feb. 2025.

EUROPARL. *Relatório sobre a inteligência artificial no direito penal e a sua utilização pelas autoridades policiais e judiciais em casos penais*. Available at: https://www.europarl.europa.eu/doceo/document/A-9-2021-0232_PT.html. Accessed on: 20 feb. 2025.

FORMISANO, A. *L'impatto dell'intelligenza artificiale in ambito giudiziario sui diritti fondamentali*, in Federalismi.it, fasc. 22/2024, pp. 112 ss.

GALLONE, G. *Riserva di umanità, intelligenza artificiale e funzione giurisdizionale alla luce dell'IA Act. Considerazioni (e qualche proposta) attorno al processo amministrativo che verrà*, in *Rivista Judicium*, Pacini Giuridica, 2024.

GIACOMO, L. D. *Bias dell'algorithm: responsabilità e innovazione nell'AI*. Available at: <https://www.diritto.it/bias-algoritmo-responsabilita-ai-innovazione/> Accessed on 20 feb. 2025.

GORI, A. *Condizioni di detenzione nelle carceri europee tra diritto dei detenuti e deroghe all'art. 3 della Convenzione*, in *Rivista ADIR*, UniFi, 2015, cap. 3.

MAGGIOLINI, P.; FUGINI, M. *Le sfide etiche dell'intelligenza artificiale*, Ledizioni, 2025.

MAHESH, B. Machine learning algorithms-a review. *International Journal of Science and Research (IJSR)*. [Internet], v. 9, n. 1, p. 381-386, 2020.

MALEK, M. A. Criminal courts' artificial intelligence: the way it reinforces bias and discrimination. *AI Ethics* **2**, 233–245, 2022.

MANNES, A. Governance, Risk, and Artificial Intelligence. *AI Magazine*, 41(1), p. 61-69, 2020. <https://doi.org/10.1609/aimag.v41i1.5200>.

OBSSERVATORIO BALCANI E CAUCASO. *La Danimarca e le prigioni off-shore*. Available at: <https://www.balcanicaucaso.org/aree/Kosovo/La-Danimarca-e-le-prigioni-off-shore-215757> Accessed on 20 mar. 2025.

PUOLAKKA, P. *Smart prison and digital innovations in prison and probation service of Finland*, in *Rassegna Italiana di Criminologia*, fasc. 2/2023, pp. 153 ss.

RINALDI, L. *Intelligenza artificiale, diritti e doveri nella Costituzione italiana*, in *DPCE Online*, fascicolo 1/2022, pp. 201 ss.

SENATO IT. *Constitution of the Italian Republic*. Available at: https://www.senato.it/documenti/repository/istituzione/costituzione_inglese.pdf. Accessed on: 20 feb. 2025.

STATE OF MIND. *Bias Euristiche*. Available at: <https://www.stateofmind.it/bias/>. Accessed on 20 feb. 2025.