# THE RISK OF AI'S FACIAL RECOGNITION IN CRIMINAL PROSECUTION: THE NEED FOR A BALANCE BETWEEN EFFICIENCY AND THE PROTECTION OF HUMAN RIGHTS

Felipe D. Martarelli Fernandes*, Vinícius Garcia Ribeiro Sampaio**, Rafael Khalil Coltro***

SOMMARIO: 1.- Introduction; 2.- AI for Facial Recognition in Prosecution and its Concerns in Brazil; 3.- Facial Recognition by Artificial Intelligence and the Protection of Human and Personality Rights: A Humanistic and Philosophical Analysis; 4.- Conclusions.

## 1.- **Introduction.**

Artificial Intelligence (AI) has been one of the most disruptive technologies of contemporary times, often regarded as the driving force behind a new industrial revolution. Its impact is vast, spanning multiple fields of knowledge and sectors of society, including education, healthcare, and, notably, the legal system. In the legal domain, AI raises a series of critical questions, particularly regarding its application in criminal prosecution and the administration of justice.

Among the main areas where AI is utilized in criminal law, facial recognition stands out as a system that employs algorithms to identify individuals through facial pattern analysis. While this technology promises advancements in public security, its implementation also presents significant risks, especially concerning the protection of fundamental rights such as privacy, the presumption of innocence, and non-discrimination. International experience has demonstrated growing concerns regarding the use of facial recognition systems, leading to various regulatory initiatives and, in some jurisdictions, even outright bans. In the European Union, for instance, the AI Act classifies the use of AI in criminal prosecution as "high risk," requiring its adoption to be subject to strict transparency and legal security standards. In Brazil, the absence of clear regulations and documented cases of AI misidentifications underscore the urgent need for an in-depth debate on the subject.

This study proposes a critical analysis of the use of AI-driven facial recognition in criminal prosecution in Brazil, considering its legal, ethical, and social implications. Through a comparative approach with European legislation, the research aims to assess the challenges and risks associated with this technology, with particular attention to the protection of fundamental rights and the necessity of an appropriate regulatory framework to mitigate its negative impacts.

## 2.- **AI for Facial Recognition in Prosecution and its Concerns in Brazil.**

* Professor in the Law program at Anhembi Morumbi University and FADISP (Brazil). PhD in Constitutional Law from FADISP (Brazil) and the University of Siena (Italy). Master in Fundamental Human Rights. Specialist in Civil Procedure from PUC São Paulo (Brazil). Lawyer in Brazil.

** PhD student in Political and Economic Law at Universidade Presbiteriana Mackenzie (Brazil) and in Law & Social Change at Università degli Studi Roma Tre (Italy, double degree). Master's in Information Society Law and Bachelor of Laws from Centro Universitário das Faculdades Metropolitanas Unidas (Brazil). Lawyer in Brazil.

*** Professor in the Law program at Anhembi Morumbi University (Brazil). PhD student in Political and Economic Law at Universidade Presbiteriana Mackenzie (Brazil) and Legal Sciences at the Università Degli Studi di Firenze (Italy) in double degree under cotutelle. Master's in law at FMU (Brazil). Specialist in Criminal Law and Criminology from PUC Rio Grande do Sul (Brazil). Lawyer in Brazil.

It is practically a consensus that Artificial Intelligence is a revolutionary technology, whose impact – it is no exaggeration to say – is already being envisioned as a new industrial revolution[1]. It can be used in several applications, such as natural language processing (NLP)[2] and machine learning[3], as well as in the most varied sectors of society, such as education[4] and healthcare[5]. The possibilities seem limitless. For Law, the most important thing is to understand the concrete consequences of this innovation, its potential risks, etc., and not necessarily how it actually works – after all, judiciary operators are not experts in technology, and, when necessary, the law allows them to consult specialists (as happens in judicial technical examinations, for example).

First, it is important to understand why artificial intelligence is so valuable. Unlike what its popular name suggests, this technique does not consist of "artificial intelligence", but rather a sophisticated arrangement of mathematical operations. Although it can be understood that the role of this technique is to replicate human intelligence, there are definitions that seem more appropriate to us, such as the proposal of "artificial communication"[6] instead of "artificial intelligence".

Far from being a matter of (excessive) precision, it seems important to us to distance ourselves from the fascination (which manifests either as celebration or as fear) that sacralizes technology, so that we can observe it properly. Yes, there is no doubt that "As the roles of steam engines in the Age of Steam, generators in the Age of Electricity, and computers in the Age of Information, AI is the pillar of technology in the contemporary era and beyond"[7]. However, the fascination with this technology could lead us to try (pointlessly) to restraint it out of fear, or to allow it without control due to passion.

For this particular study, there are two specific applications to be investigated jointly: facial recognition, which has implications for personality rights, as it contains biometric data, images, etc., and algorithmic prediction, a set of statistical operations used to assess the probability of future events, which requires legal sciences to address various problems (phenomena such as "fossilization", "unfalsifiability", "preemptive intervention problem", "self-fulfilling prophecy", etc.[8]).

The use of AI in criminal prosecution is so vital that the AI Act brings important definitions and provisions on this, such as Article 3 (38), which defines "sensitive operational data": "operational data related to activities of prevention, detection, investigation or prosecution of criminal offences,

[1] Y.K. Dwivedi et al., *Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. International Journal of Information Management*, 57 (2021); https://www.sciencedirect.com/science/article/abs/pii/S026840121930917X.

[2] G. Tecuci, *Artificial intelligence. WIREs Comput Stat*, 4 (2012) 168-180; DOI: 10.1002/wics.200. https://wires.onlinelibrary.wiley.com/doi/10.1002/wics.200.

[3] S.M. Mohammad, *Artificial Intelligence in Information Technology* (2020); https://ssrn.com/abstract=3625444; DOI: http://dx.doi.org/10.2139/ssrn.3625444.

[4] L. Chen, P. Chen, Z. Lin, *Artificial Intelligence in Education: A Review. IEEE Access*, 8(2020) 75264-75278;DOI: 10.1109/ACCESS.2020.2988510. Available at: https://ieeexplore.ieee.org/document/9069875.

[5] S. K. Bhattamisra et al., *Artificial Intelligence in Pharmaceutical and Healthcare Research. Big Data Cogn. Comput.*, 7(2023) n. 10; DOI: https://doi.org/10.3390/bdcc7010010.

[6] E. Esposito, *Artificial communication: how algorithms produce social intelligence*. Cambridge, Massachusetts 2022; https://mitpress.mit.edu/9780262046664/artificial-communication/.

[7] Y. Jiang, et al., *Quo vadis artificial intelligence?. Discover Artificial Intelligence*, v. 2 (2022) n. 4; https://link.springer.com/article/10.1007/s44163-022-00022-8.

[8] H. Matsumi, D.J. Solove, *The Prediction Society: AI and the Problems of Forecasting the Future. GWU Legal Studies Research Paper* 58 (2023); *GWU Law School Public Law Research Paper* 58 (2023); Illinois 2025, forthcoming; https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4453869.

the disclosure of which could jeopardise the integrity of criminal proceedings". One of the most important provisions for those interpreting the European AI Act is found in Recital 59:

"Given their role and responsibility, actions by law enforcement authorities involving certain uses of AI systems are characterized by a significant degree of power imbalance and may lead to surveillance, arrest or deprivation of a natural person's liberty as well as other adverse impacts on fundamental rights guaranteed in the Charter. In particular, if the AI system is not trained with high-quality data, does not meet adequate requirements in terms of its performance, its accuracy or robustness, or is not properly designed and tested before being put on the market or otherwise put into service, it may single out people in a discriminatory or otherwise incorrect or unjust manner. Furthermore, the exercise of important procedural fundamental rights, such as the right to an effective remedy and to a fair trial as well as the right of defence and the presumption of innocence, could be hampered, in particular, where such AI systems are not sufficiently transparent, explainable and documented. It is therefore appropriate to classify as high-risk, insofar as their use is permitted under relevant Union and national law, a number of AI systems intended to be used in the law enforcement context where accuracy, reliability and transparency is particularly important to avoid adverse impacts, retain public trust and ensure accountability and effective redress. […]".

In Brazil, for example, a bill is currently being processed, approved by the Senate, which has 244 amendments[9] (highlighting the political disputes surrounding this issue) and is inspired by the European regulation. Similarly, the use of AI for fact investigation is considered "high-risk" when there are risks to individual freedoms, within the framework of the administration of justice. This transatlantic concern does not exist by chance, nor simply through theoretical inference; on the contrary, there are unacceptable practical cases that demonstrate the need for such contingencies. For example, in Brazil, a 23-year-old young man went to the stadium to watch a football match but was mistakenly "recognized" by police with the support of AI – he was not arrested[10].

In the same state (or province), Sergipe, a 31-year-old woman was not as lucky; she was brutally arrested by four police officers while attending a carnival show and even urinated out of fear – in an interview, she said she was "publicly discriminated against for being poor and black". The Governor declared that he would suspend the use of this tool[11].

Despite episodes like these, the state of São Paulo (the largest in Brazil) has been consistently investing in the expansion of AI usage in public security. For example, a report mentions that since the end of 2024, the city hall of São Paulo (the capital of the state, which has the same name and is not responsible for public security – this responsibility relies on the state) helped to arrest nearly 500 criminals[12]. Naturally, this information should warn us, because Brazil is now building a dangerous combination: Artificial Intelligence (with a concerning history of errors) plus police approaches (with

[9] Brasil, Senado Federal, Projeto de Lei nº 2338-2023; https://www25.senado.leg.br/web/atividade/materias/-/materia/157233.

[10] W. Carmo, *Serial errors expose the fragility of facial recognition as a crime-fighting tool. Carta Capital*, april 19th 2024; https://www.cartacapital.com.br/tecnologia/erros-em-serie-expoem-fragilidade-do-reconhecimento-facial-como-ferramenta-de-combate-ao-crime/.

[11] Carmo, *Serial errors* cit.

[12] D. Oliveira, *Smart Sampa: IA da Prefeitura de São Paulo ajudou a prender quase 500 criminosos*. IT Forum, january 22nd 2025; https://itforum.com.br/noticias/smart-sampa-ia-da-prefeitura-sp-prendeu-500-criminosos/.

a history of violence[13]). The situation worsens when we realize that AI errors tend to increase when used for facial recognition of Black people – the primary victims of police violence in Brazil (and in many other countries).

A concerning point is that cases where individuals are mistakenly identified by artificial intelligence suggest that these are not mere errors, but rather flaws in the programming that underpins the systems. In fact, one aspect that must be considered in this issue is how the so-called "machine learning" responsible for developing the implemented technologies will be shaped. The success (or failure) of a particular AI system depends heavily on the information and data provided to the system for feedback. In the case of a system designed for facial recognition via AI, its functionality is only possible through an analysis of a pre-existing database, derived from photographs or video segments. Equipped with such data, the systems perform calculations, measuring "the distance between the eyes or the shape of the nose, producing what is known as a facial pattern".

Thus, although it may be easy to imagine that an AI system designed for identifying people for criminal prosecution purposes could contribute to society, in the specific case of Brazilian society, it doesn't take much to understand the problem that precedes the very technological implementation: racism[14]. Since recognition is based on physical, ethnic, and racial patterns, this can increase the probability of "detection" when it comes to Afro-descendant individuals.

Costa e Kremer[15] recall that, at a conference held in mid-2017, a North American researcher from the Massachusetts Institute of Technology presented how artificial intelligence systems would be susceptible to failures in recognizing faces, especially those of Black and Brown people. In the study, she demonstrates that AI exhibited "low accuracy in identifying the faces of Black women". However, the study shows that this does not happen when visibility is negative, as the technologies prove to be extremely effective at identifying Black individuals for negative results[16]. In this regard, researchers Woodrow Hartzog and Evan Sellinger[17] explain that facial recognition could be a "perfect" tool for state oppression, as it can violate privacy rights and other fundamental guarantees, leading to the persecution of black individuals or other specific ethnic groups, for instance.

All of this suggests that, alongside the increase in public security efficiency (which will achieve more results with less effort), the number of injustices committed in Brazil will likely also rise, a risk clearly outlined in the European regulation but still present among the slow pace of the Brazilian legislative process.

## 3.- **Facial Recognition by Artificial Intelligence and the Protection of Human and Personality Rights: A Humanistic and Philosophical Analysis.**

[13] J. H. French, *Repensando a Violência Policial no Brasil: Desmascarando o Segredo Público da Raça*, in *Revista Tomo* 31 (2017); https://ufs.emnuvens.com.br/tomo/article/view/7648.

[14] About the subject: R. S. Costa e B. Kremer, *Inteligência artificial e discriminação: desafios e perspectivas para a proteção de grupos vulneráveis frente às tecnologias de reconhecimento facial, Revista Brasileira de Direitos Fundamentais & Justiça*, 149/150 (2022).

[15] Id., *Inteligência artificial e discriminação* cit., 158.

[16]J. Buolamwini, *How I Fight Bias in Algorithms,* TED Video, 2017; https://www.ted.com/talks/joy_buolamwini_how_i_m_fighting_bias_in_algorithms, jan.2025.

[17] W. Hartzog, E. Sellinger, *Facial Recognition is the perfect tool for opression. Medium*, (2018); https://medium.com/s/story/facial-recognition-is-the-perfect-tool-for-oppression-bc2a08f0fe66.

It is evident that modern society is shaped by a struggle between a racist and exploitative capitalist system and a hypocritical, dictatorial attempt at communism. In this scenario, new technologies introduced to the market aim to facilitate citizens' lives in certain aspects but impose a high cost that extends beyond financial burdens to mental and humanistic aspects as well.

Zygmunt Bauman exemplifies this social phenomenon by stating that "nowadays, shopping centers tend to be designed with the sudden awakening and rapid extinction of impulses in mind, rather than the inconvenient and prolonged creation and maturation of desires"[18]. This demonstrates that, as a rule, the primary goal of new technologies is not social evolution but rather wealth accumulation for the maintenance of power.

Among various technological advancements, facial recognition by artificial intelligence represents a significant step forward for social evolution, as it aims to provide greater social security and convenience through the speed of individual identification. However, this technology also poses imminent risks to personality rights and human rights. Therefore, considering the collision of constitutional principles, and as well developed by Robert Alexy, the principle of balancing[19] must be applied in this discussion. Unlike legal norms, where a subsequent or specific norm overrides a prior or general one, in conflicts between principles, the rules of necessity, possibility, and strict balancing must be observed. Thus, although a definitive answer can only arise in a concrete case rather than in the abstract, this study will examine both theses to seek a common ground between the arguments presented here. As Alexy states, "the objective of this balancing is to determine which of the interests— which are abstractly at the same level— carries greater weight in the concrete case"[20].

Upon initial analysis, one side of the balancing argument holds that the use of facial recognition by AI violates human and personality rights. Human rights are the minimum rights of any individual simply by virtue of being human. When codified in legal charters, they become fundamental rights. In the same vein, personality rights are reflections of human and fundamental rights, as they protect essential aspects of human existence, such as the right to image, bodily autonomy, privacy, and honor. In this regard, José Joaquim Gomes Canotilho states:

"Personality rights certainly encompass state rights (e.g., citizenship rights), rights over one's own person (right to life, moral and physical integrity, right to privacy), distinctive personality rights (right to personal identity, right to information technology), and many liberty rights (freedom of expression)"[21].

In line with the aforementioned arguments and in alignment with the ideas of Immanuel Kant, Martha Nussbaum asserts that the human being is not merely a means for production or achievement but an end in itself. In other words, individuals are the central focus of protection and support provided by

[18] Z. Bauman, *Amor líquido: Sobre a fragilidade dos laços humanos*, Rio de Janeiro 2004, 14.

[19] When two principles collide —such as when something is prohibited according to one principle but permitted according to another—one of the principles must yield. This does not mean, however, that the yielding principle should be declared invalid or that an exception clause should be introduced into it. In fact, what happens is that one of the principles takes precedence over the other under certain conditions. Under different conditions, the issue of precedence may be resolved in the opposite way. This is what is meant when it is stated that, in concrete cases, principles have different weights and that principles with greater weight take precedence. Conflicts between rules occur within the dimension of validity, while collisions between principles—since only valid principles can collide—occur beyond that dimension, in the dimension of weight. – R. Alexy, *Theorie der Grundrechte*, Frankfurt 2006, 94.

[20] Alexy. *Theorie* cit., 95.

[21] J.J. Canotilho, *Direito Constitucional e Teoria da Constituição*, 4 ed., Coimbra 2000, 390.

legal norms and statutes, and they cannot be suppressed, not even in favor of technological advancement. In this regard, she states, "political theory begins with an abstract idea of basic rights, founded on the combined ideas of dignity (the human being as an end) and sociability"[22]. Thus, if the human being is an end in itself and is protected by personality rights, no technology can, in the name of social security, utilize such information without the necessary safeguards.

Facial recognition by artificial intelligence can be used as a means of capturing personal information without the user's authorization, and such data may be exploited for nefarious purposes by the state or private corporations.

In nearly all major cities worldwide, cameras in public and private spaces already track and record the identified movements of numerous individuals, many of whom are unaware they are being monitored. If such information is not adequately protected, and sensitive data is not handled securely, there is a significant risk of violating a citizen's right to privacy and personal image. Such data can be used to determine an individual's routine, frequency of visits to certain locations, and social interactions. With this information and artificial intelligence, both state and private entities can manipulate the market, encourage consumer behavior, and even facilitate crimes such as thefts and kidnappings.

At this point, it is crucial to highlight that both state actors and private entities that control artificial intelligence systems may use such images for purposes contrary to the will of the data subject:

"It allows us to understand that societies with a state are inherently divided into dominators and dominated [...] whereas stateless societies ignore this division, despite also being regulated by relations of force and domination"[23].

Throughout human history, the struggle for individual freedom against the state has been a continuous battle. All major revolutions have aimed to secure autonomy and liberty for individuals within society. However, when misused, these new technologies can represent a significant regression in the first-generation rights that were so arduously attained. Similarly, private companies can exploit such information to influence and manipulate the economy in various ways, whether by directing consumption, creating highly segmented markets, or even leveraging behavioral patterns to maximize profits. In this regard, Byung-Chul Han explains that:

"In today's financial capitalism, values are radically eliminated. The neoliberal regime introduces an era of exhaustion. Today, the psyche is exploited. Consequently, this new era is accompanied by mental illnesses such as depression and burnout"[24].

As already mentioned, on the other side of the balancing argument, there is a thesis advocating for a broader application of such technologies, as facial recognition has proven to be an extremely useful tool for public security and social organization. Its implementation is already a reality in various countries worldwide (alongside then, in Brazil), bringing significant benefits to the population.

Facial recognition must be employed with proper balancing between the right to privacy and personal image and the right to information and security, following the principles set forth by Bauman, who

[22] M. Nussbaum, *Fronteiras da justiça – deficiência, nacionalidade, pertencimento à espécie*, 2 ed., São Paulo 2020, 45.

[23] H.G. Carnio, *Fronteiras do Direito – analítica da existência e crítica das formas jurídicas*, 1 ed., Belo Horizonte 2021, 29.

[24] B.C. Han, *Psicopolítica – O neoliberalismo e a as novas técnicas de poder. translated to portuguese by Maurício Liesen*, 7 ed. Belo Horizonte 2020, 46.

stated that "security without freedom is slavery, while freedom without security is complete chaos"[25]. In the same vein, Cesare Beccaria asserts that "fortunate are the nations (if any exist) that did not wait for slow revolutions and uncertain vicissitudes to make excessive evil a norm of good, and that, through wise laws, hastened the transition from one to the other"[26]. The monitoring of individuals within a given society is not necessarily something new. Foucault[27], for example, discusses surveillance as part of the process of governance and control, in which individuals are "self-disciplined" due to the internalization of surveillance, leading to a more effective form of social control.

Thus, when used responsibly and with proper legal and ethical safeguards to protect privacy and personal image rights, facial recognition technology can significantly contribute to security, efficiency, and the modernization of urban spaces and social interactions.

## 4.- **Conclusions.**

Foucault already emphasized in his analysis of power and surveillance that, in modern societies, power is no longer exercised solely in a centralized and visible manner, but is dispersed through technologies and institutional practices. In his works, he develops the famous concept of the "panopticon", inspired by the prison model idealized by Jeremy Bentham, where a single guard can observe all prisoners without them knowing when they are being watched. This model, according to Foucault, is an example of how surveillance becomes an effective form of social control, as people begin to behave as if they were constantly being observed, leading to self-discipline and the internalization of power.

When comparing Foucault's ideas with modern facial recognition and monitoring systems that are gradually being proposed in democratic societies (but are already a reality in societies where individual freedoms are limited), it is possible to observe that a facial recognition system through AI would certainly collaborate with the implementation of widespread surveillance and social control. Today, facial recognition technologies represent an advancement in surveillance, allowing individuals to be monitored in real time in both public and private spaces. As in Foucault's panopticon, people may not know when they are being observed, but the fact that they are aware that the technology is available to monitor them can lead them to behave differently.

However, the difference between the panopticon concept and modern facial recognition systems is that while the panopticon was based on centralized physical surveillance, facial recognition systems can be distributed and automated, utilizing vast databases and algorithms to analyze images and identify individuals. Furthermore, AI and machine learning technology constantly improves its recognition capabilities, which implies increasingly effective and invisible surveillance.

On the other hand, the use of facial recognition systems also raises questions about the limits of privacy, civil rights, and social impacts—issues that Foucault might consider as forms of excessive or oppressive control. Furthermore, as also clarified, deeper social problems, such as racism, risk becoming exacerbated, which could be a serious problem since the foundation of machine learning is

[25] S. Bauman, in: https://www.youtube.com/watch?v=POZcBNo-D4A&t=1589s. Access in 14. feb. 2025.

[26] C. Beccaria*, Dos delitos e das penas*, São Paulo 2006,16.

[27] M. Foucault, *Vigiar e Punir: nascimento da prisão*, 42. ed., Petrópolis 2006.

fundamentally flawed from its inception, potentially leading to a learning model based on ideals incompatible with democratic systems and international human rights.

In light of this issue, at least at this early stage, the perception is that the international community shows a certain resistance to the massive implementation of facial recognition technologies, with this technology being banned in various locations around the world. In 2020, the European Commission proposed a temporary ban on the use of facial recognition in public spaces for surveillance purposes, with the suggestion of more comprehensive regulations on artificial intelligence. The Supreme Court of the United Kingdom ruled that the use of facial recognition by the police in London violated privacy and data protection rights. In the U.S., several cities implemented bans on the use of AI for facial recognition, including San Francisco and Boston.

In Brazil, although with significant debates and some hesitancy, in major cities such as São Paulo and Rio de Janeiro, facial recognition technology has been implemented for public space monitoring. In São Paulo, for example, the technology was adopted in the monitoring system of subway stations and public security cameras to identify individuals wanted by the police.

The use of this technology as a public security policy, as seen, is not immune to criticism, particularly with regard to the possible exacerbation of racism, which is undeniably present in Brazilian society. Furthermore, regarding the transparency debate, the regulatory possibilities for AI systems need to engage with other legislation, especially the LGPD (Brazilian General Data Protection Law), which addresses, among other points, the rights of individuals regarding automated data processing.

In view of this discussion, the approach adopted by the European community, for the most part, seems to be the most prudent: caution is necessary. Before mass implementation of AI for facial recognition, it is essential to clearly define how the AI is fed. Furthermore, the following questions need to be addressed: How are these technologies biased? What is the potential social impact of applying such technologies in societies where racial and social issues, like in Brazil, are glaring?

Only after fully understanding these issues will it be possible to assess whether, in fact, the use of such technology aligns with the fundamental principles of social rights, which have been (and are being) constructed in modern Western societies.

**Abstract.-** L'articolo analizza l'uso del riconoscimento facciale nei procedimenti penali brasiliani, in rapporto alle sfide legali ed etiche che ne derivano. Lo studio prende in esame la classificazione di questa tecnologia considerata 'ad alto rischio' e i meccanismi necessari per garantire trasparenza, responsabilità e certezza del diritto. Casi concreti in Brasile hanno evidenziato l'urgente necessità di una regolamentazione adeguata.

Il contributo valuta anche i rischi del potenziale utilizzo di questa tecnologia come strumento di eccessivo controllo sociale. Affinché l'intelligenza artificiale possa essere utilizzata in modo equo ed efficace nei procedimenti penali appare dunque essenziale stabilire normative chiare, meccanismi di controllo efficaci e linee guida per mitigarne gli impatti negativi, garantendo la tutela dei diritti fondamentali dei cittadini.

This paper examines the use of facial recognition in Brazilian criminal prosecution, addressing the legal and ethical challenges involved. From a comparative perspective with European legislation, particularly the AI Act, the study analyzes the classification of this technology as “high risk” and the mechanisms required to ensure transparency, accountability, and legal security. Concrete cases in Brazil, such as wrongful arrests resulting from AI misidentifications, highlight the urgent need for appropriate regulation. The discussion also explores the impact of AI on the phenomenon of “fossilization” and the issue of “self-fulfilling prophecy”, considering the risk of reinforcing discriminatory patterns and the potential use of this technology as a tool for excessive social control. The study concludes that, for AI to be used fairly and effectively in criminal prosecution, it is essential to establish clear regulations, effective oversight mechanisms, and guidelines to mitigate its negative impacts, ensuring the protection of citizens’ fundamental rights.