

# Artificial intelligence in public security and criminal justice systems in Latin America

Due process-based analysis



Supported by **moz://a**

## About Fair Trials

**Fair Trials is the global criminal justice watchdog, campaigning for fairness, equality and justice. The organisation monitors criminal justice systems with a team of experts and has been tracking the increasing use of artificial intelligence in this field since 2020. This report was prepared by Luis Eliud Tapia, Legal Consultant, under the supervision of Verónica Hinestroza, Senior Legal Advisor.**

## Contact

**Verónica Hinestroza**

**Senior Legal Advisor**

Veronica.Hinestroza@fairtrials.net

 @fairtrials  @fairtrials  Fair Trials

First published by Fair Trials in December 2024. © Fair Trials 2024.

“Fair Trials” includes Fair Trials International, Fair Trials Europe, and Fair Trials Americas. Fair Trials International is a registered charity (no. 1134586) and in 2010 was incorporated with limited liability in England and Wales (No. 7135273), and is based at 5 Castle Road, London, NW1 8PR. In May 2014, Fair Trials International founded Fair Trials Europe, which is a registered public foundation in Belgium (registered number 0552.688.677). In 2018 we founded Fair Trials Americas, which is a registered 501(c)(3) public charity in the United States of America (No DLN17053243307017). We were initially founded in 1992 with the name “Fair Trials Abroad”.

## Acknowledgments

**Fair Trials gratefully acknowledges the support of the Mozilla Foundation for the study underlying this report. We also express our sincere thanks to the partners who organised the International Congress on Artificial Intelligence, Law, and Democracy in Mexico City in October 2023: Universidad Iberoamericana Ciudad de México, Konrad Adenauer Foundation, and the Center for Constitutional Studies of the Supreme Court of Justice of the Nation in Mexico. We are indebted to the congress speakers whose insights contributed to this report's findings and to Claudia Pajas Tinoco, Fair Trials intern, for her valuable research support.**

<b>1. Introduction</b>	<b>5</b>
<b>2. Executive summary</b>	<b>7</b>
Introduction	7
Europa	7
Latin America	7
Challenges and risks	9
Conclusions	9
Recommendations	9
<b>3. Basic concepts</b>	<b>11</b>
<b>4. Artificial intelligence in public security and criminal justice systems in Europe</b>	<b>13</b>
<b>5. Preliminary conclusions</b>	<b>15</b>
<b>7. Artificial intelligence and emerging challenges to due process in Latin America</b>	<b>23</b>
<b>8. Conclusions</b>	<b>27</b>
<b>9. Recommendations</b>	<b>28</b>

# 1. Introduction

The use of artificial intelligence (AI)-based systems in public security and the criminal justice system in Latin America responds to a global trend, but also finds an explanation in dynamics specific to the region. In general, in Latin America, the levels of violence and impunity lead the population to put pressure on the authorities. Governments tend to respond to this pressure with measures that increase both direct protection against crime and the perception of security. However, seeking to integrate new technologies to respond to legitimate concerns does not exempt states from guaranteeing human rights in general, especially those related to due process.

Civil society organisations around the world have documented the human rights impacts of artificial intelligence-based technology, including facial recognition. The human rights that are primarily vulnerable to such technology are privacy, equality, and non-discrimination.

However, when studying the justice system as a whole, artificial intelligence can impact other human rights, such as personal liberty, the presumption of innocence, and due process.

This report analyses the use of artificial intelligence-based systems in the justice system under the umbrella of due process, a developing field of study. The report takes as a precedent some experiences documented by Fair Trials in European countries and then focuses on Latin America. The sources analysed for the regional analysis include inputs from academia, civil society, international organisations, and courts of justice.

Additionally, the report draws on the discussions that took place between October 23 and 24, 2023, at the International Congress on Artificial Intelligence, Law and Democracy, organised in Mexico City by Fair Trials, the Universidad Iberoamericana Ciudad de México, the Konrad Adenauer Foundation and the Center for Constitutional Studies of the Supreme Court of Justice of the Nation in Mexico. During 2 days, 5 panels discussed the opportunities and challenges of artificial intelligence, with a focus on security, justice, discrimination, inequality, and democracy.

Verónica Hinestroza from Fair Trials, Marc Rotenberg from the Center for Artificial Intelligence and Digital Police (CAIDP), Priscilla Ruiz from Article 19, and Pablo Nunes from Panóptico participated in the round table on security and artificial intelligence. The panel on justice and artificial intelligence included Carolina Villadiego from the International Commission of Jurists (ICJ), Grecia Macías from the Network in Defense of Digital Rights (R3D), Carlos Guzmán, Criminal Magistrate of the Superior Court of Bogotá and Luis Tapia from Fair Trials.

The report's central theme is to reaffirm that the safeguards of fair trial or due process are not waivable, especially in the face of the use of technologies based on artificial intelligence. Nor does the efficient use of resources justify relativizing the minimum guarantees of due process, whose origin sought to limit the arbitrary exercise of public power.

It should be noted, however, that judiciaries, prosecutors' offices, and security institutions are using AI for other purposes, such as judicial management, the development of case law search engines, and organising the information they collect.<sup>1</sup>

<sup>1</sup> Carolina Villadiego, Panel on Justice and Artificial Intelligence at the International Congress on Artificial Intelligence, Law and Democracy organised by Fair Trials, the Universidad Iberoamericana Ciudad de México and the Konrad Adenauer Foundation, October 2023.

In this sense, the report recognises that it is essential to identify and differentiate the possible uses of artificial intelligence in the justice system to avoid disqualifying the technology a priori. The report takes into account that AI can be useful for the functioning of the judiciary, and generates, in turn, a warning against the so-called technological solutionism,<sup>1</sup> stressing that it is imperative to develop mechanisms to regulate its use with a human rights approach, avoiding adopting this technology uncritically sacrificing the rights of individuals.

Therefore, the report seeks to identify both risks and good practices related to the use of artificial intelligence by public security authorities, criminal justice systems, and judiciaries in Latin America and Europe. In addition, it seeks to reflect on the reconfiguration of due process guarantees in the key of artificial intelligence.

It is hoped that the report will be an input that nurtures a conversation necessary for guaranteeing human dignity from security and human rights, given the increasing use of artificial intelligence by law enforcement and justice system authorities.

---

<sup>1</sup> Luis Fernando García and Grecia Macías, La nueva cara de la discriminación: el impacto diferenciado de la tecnología de reconocimiento facial en grupos vulnerables in SCJN, Discriminación, Problemas Contemporáneos, Ricardo Latapie Aldana (Coord.), pp. 109-110.

## 2. Executive summary

### Introduction

The report analyses the use of artificial intelligence (AI) in public security and the criminal justice system in Latin America, highlighting the need to guarantee human rights, especially due process. It starts from documented experiences in Europe, and recovers, among other debates, discussions from the International Congress "Artificial Intelligence, Law and Democracy" held at the Universidad Iberoamericana on October 23 and 24, 2023.

Based on the above, the experiences of Argentina, Brazil, Chile, Colombia, El Salvador, Mexico, and Uruguay regarding the use of AI in public security, in the criminal justice system, and in the judiciary are presented.

### Key human rights findings from the experience of using some AI systems.

#### Europa

##### 1) Public security

- a. **RADAR-iTE (Germany).** Violence risk assessment with a focus on potential Islamist terrorists, showing bias against Muslims.
- b. **ProKid (Netherlands).** Prediction of crimes committed by children and adolescents, with a high rate of errors and biases.
- c. **NDAS (UK).** Risk assessment showing racial discrimination.
- d. **HART (UK).** Prediction of recidivism with limited accuracy and use of discriminatory data.

##### 2) Criminal justice and prison systems

- a. **Cassandra (Ukraine).** Recidivism risk assessment, with lack of data transparency impacting the right to access information and due process.
- b. **RisCanvi (Spain).** Evaluation of early release with difficulties in appealing decisions influenced by system design.

#### Latin America

##### 1) Public security

- a. **Facial Recognition.** Extensive use with problems of bias and false positives affecting vulnerable populations.
  - i. **OI (Brazil).** Facial recognition technology makes it easier for the police to justify their allegedly illegal actions and has a racially discriminatory bias towards people of African descent due to the area where the cameras are installed.
  - ii. **Proximity Police (Mexico).** Surveillance cameras discriminate against low-income and LGBTTQ+ people. Incorporating facial recognition into

surveillance cameras can reproduce people's biases and facilitate discrimination.

- iii. **Calle Segura (Chile).** The use of facial recognition in drones is opaque, highly fallible and institutionalises biases associated with social class and skin colour.
  - iv. **Facial recognition (El Salvador).** The use of facial recognition by governments accused of committing serious human rights abuses can be highly counterproductive against stigmatised and discriminated populations.
- b. Crime prediction.** Limited and cautious implementation showing biases and errors in predictions.
- i. **Urban Crime Predictive System (Chile).** A system that sought to predict in which areas crimes will be committed based on algorithms. The system can make it easier for the police to harass people because they consider them suspicious.
  - ii. **Empirical study of criminal behaviour identification (Chile).** Research focused on the possibility of predicting the risk of recidivism using algorithms. In more than 30% of the cases, the predictions were wrong. In addition, it was identified that the system may encourage the arrest of people based on their background and not for behaviours committed.
  - iii. **Crime prediction in intermediate cities (Colombia).** Crime prediction models can have discriminatory effects against populations or people living in specific areas. In addition, such systems make errors that are compounded when there is little information, as in intermediate cities such as Bucaramanga.
  - iv. **PredPol (Uruguay).** A tool for predicting where crimes will be committed. A pilot program found that the system showed no better results than traditional technologies.

**2) Judiciary and Criminal Justice.** The use of artificial intelligence in the criminal justice system is under exploration or has been discarded due to its risks. In the general work of the justice system, some AI-based systems are being implemented to support the elaboration of sentences.

- a. **Evaluation of evidence and sentences (Mexico).** The possibility of using artificial intelligence in sentencing has been explored in the journal of the National Institute of Penal Sciences but has not been implemented to date.
- b. **PRiSMA (Colombia).** Development of an automated learning-based recidivism risk assessment system with concerns about the impact on the right to equality and non-discrimination, due process, the right to defence, and judicial impartiality.
- c. **Projeto Sócrates (Brazil).** A system based on artificial intelligence that seeks to reduce the time to issue sentences by analysing previous sentences that are similar. However, it runs the risk of automating historical injustices in the judicial system.
- d. **PROMETEA (Argentina).** System based on artificial intelligence to automate

the preparation of legal opinions. Implied a dramatic reduction in the resolution time of some cases. 90 minutes to 1 minute in contracting documents. The system incorporates controls based on human rights.

- e. **PretorIA (Colombia).** Tool based on artificial intelligence that helps to pre-select the tutelas to be reviewed. In Colombia, around 600,000 guardianships are received per year. Although it is a system that automates activities, the decision is made by a person in order to reduce risks to human rights.

## Challenges and risks

- **Personal Liberty:** Risk of arbitrary restrictions on the freedom of movement of persons who have not been proven to have committed any crime.
- **Due Process:** Affectation of the presumption of innocence, right of defence, right to appeal the ruling, and judicial independence and impartiality due to the fact that, among other things, it is not known how the decision is made, from which databases the information comes from, who finances the systems based on artificial intelligence.
- **Equality and non-discrimination:** Technology based on artificial intelligence automates the discriminatory biases that characterise public security and criminal justice system work.

## Conclusions

The use of AI in justice and security in Europe has shown a high cost to human rights, with significant biases and errors. In Latin America, the use is more limited but presents similar risks. A balanced approach is needed that neither condemns nor uncritically adopts AI, always ensuring the protection of human rights, especially due process.

The importance of a regulated and conscious approach to the risks that artificial intelligence may present in the field of public security and justice is emphasised. This approach always promotes the protection of individuals' human rights, especially the right to due process.

There are some favourable opportunities for the use of AI in justice systems, as exemplified by the cases of Argentina and Colombia, as long as transparency, accountability, and respect for human rights are privileged.

## Recommendations

### 1) Prohibitions:

- a) Use of mass surveillance technologies such as facial recognition.
- b) Use of crime prediction systems.

### 2) Transparency and accountability:

- a) Report on the acquisition and use of AI systems.
- b) Develop due process safeguards applicable to AI systems.

### 3) Best practices:

- a) Conduct pilot tests to evaluate performance and risks before implementing the systems.
- b) Implement external audits of AI technologies.

**4) Controlled opportunities:**

- a) Use of AI systems with adequate safeguards, following the examples of Argentina and Colombia.

## 3. Basic concepts

- **Artificial intelligence**

According to the United Nations (UN), Artificial Intelligence (AI) is defined as a "constellation of processes and technologies that enable computers to supplement or replace specific tasks that would otherwise be performed by humans, such as decision-making and problem-solving".<sup>1</sup>

The academic of the University of Buenos Aires, Juan Corvalán, in the Treatise on Artificial Intelligence and Law, refers to 20 definitions of artificial intelligence.<sup>2</sup> This reflects that the concept of AI is under constant debate and under construction.<sup>3</sup>

The organisation Access Now notes that AI is considered more a field of study than a thing and that it can be split into subfields of study such as machine learning and robotics.<sup>4</sup>

An additional definition of AI refers to it as "a set of processes and techniques that enable a computational system to perform statistical calculations capable of identifying patterns and, from them, predicting behaviours."<sup>5</sup> However, AI "is far from having a critical or empathic capability".<sup>6</sup>

Artificial intelligence describes then, the possibility of systems attempting to mimic human thinking by learning and using the generalizations we use to make decisions.<sup>7</sup> However, artificial intelligence is a paradox because it rests on the unfulfilled desire to emulate human cognitive processes.<sup>8</sup>

\*In this report, the umbrella term artificial intelligence will be used generically, understanding it as a discipline that encompasses different forms of automation and uses algorithms. Although it is recognised that there are different concepts, such as soft artificial intelligence, hard artificial intelligence, generative artificial intelligence, and general artificial intelligence, depending on the approach from which they are analysed.

While technical clarity matters, the key issue is to identify the extent to which technology is allowed to replace activities that a person would normally do, such as drafting a sentence.

<sup>1</sup> Res. 73/348 of the UN General Assembly cited in Juan Corvalán et. al. *Inteligencia artificial: bases conceptuales para una aproximación interdisciplinar* in Juan Corvalán, *Tratado de Inteligencia Artificial y Derecho*, T. 1.

<sup>2</sup> Juan Corvalán et. al. *Inteligencia artificial: bases conceptuales para una aproximación interdisciplinar* in Juan Corvalán, *Tratado de Inteligencia Artificial y Derecho*, T. 1.

<sup>3</sup> Access Now. *Human Rights in the Age of Artificial Intelligence*. p. 8.

<sup>4</sup> Id.

<sup>5</sup> Jamila Venturini. *Artificial intelligence beyond polarization*. April, 2024. Available at <https://www.derechosdigitales.org/23564/inteligencia-artificial-mas-alla-de-la-polarizacion/>

<sup>6</sup> Id.

<sup>7</sup> Jordi Nieva. *Inteligencia artificial y proceso judicial*, Marcial Pons, 2018, p. 20.

<sup>8</sup> Fernando Miró Linares quoted in Álvaro Vízcaíno Zamora. *Video surveillance and artificial intelligence: between utopia and dystopia*. p. 30.

- **Algorithms**

An algorithm is key to defining artificial intelligence. Algorithm refers to a set of instructions to perform tasks.<sup>1</sup> It is a model created by people that helps to represent reality.<sup>2</sup> Advances in technology have allowed algorithmic models to become increasingly complex. In the case of artificial intelligence, an algorithm contains the instructions that allow a computer to learn from the environment and perform tasks.<sup>3</sup> A cooking recipe and the procedure for changing a flat tire are examples of algorithms.<sup>4</sup>

- **Machine Learning**

Machine learning (ML<sup>5</sup>) allows machines to learn automatically using patterns and inferences instead of receiving instructions from a person. ML techniques instruct machines to achieve an outcome by providing numerous instances of correct outcomes.<sup>6</sup> The more they train, the more experience they gain and the better they can replicate it in the future.<sup>7</sup>

- **Big data**

Data sets that are too large or complex to be analysed by traditional programs. The increasing availability of big data, thanks to society's growing use of the Internet and rapid improvements in computing power has enabled significant advances in AI over the past 10 years.<sup>8</sup>

- **Facial recognition**

Facial recognition technology uses a person's biometric identification to map the facial features contained in a photograph or video and compares them with information obtained from a face bank to find a match.

- **Automated decision systems**

Procedures in which decisions are initially delegated, in whole or in part, to another person or corporate body which, in turn, executes automated decision-making models for the development of the action.<sup>9</sup>

<sup>1</sup> UNESCO. Global Toolkit on AI and the Rule of Law for the Judiciary. 2023, p. 28. Available at [https://unesdoc.unesco.org/ark:/48223/pf0000387331\\_spa](https://unesdoc.unesco.org/ark:/48223/pf0000387331_spa).

<sup>2</sup> Luis Fernando García and Grecia Macías, The New Face of Discrimination: el impacto diferenciado de la tecnología de reconocimiento facial en grupos vulnerables in SCJN, Discriminación, Problemas Contemporáneos, Ricardo Latapie Aldana (Coord.), p. 111.

<sup>3</sup> UNESCO. Toolkit... p. 28.

<sup>4</sup> Juan Corvalán et. al. Inteligencia artificial: bases conceptuales para una aproximación interdisciplinaria in Juan Corvalán, Tratado de Inteligencia Artificial y Derecho, T. 1.

<sup>5</sup> Machine Learning.

<sup>6</sup> UNESCO. Toolkit... p. 21.

<sup>7</sup> Luis Fernando García and Grecia Macías, The New Face of Discrimination, Op. cit.

<sup>8</sup> Access Now. Human Rights in the Age of Artificial Intelligence, p. 8.

<sup>9</sup> Patricio Velasco Fuentes and Jamila Venturini. Automated decisions in the civil service in Latin America. A comparative approach to its application in Brazil, Chile, Colombia and Paraguay. Derechos Digitales, March 2021. p. 5.

## 4. Artificial intelligence in public security and criminal justice systems in Europe

In September 2021, Fair Trials published the report *Automating Injustice: The Use of Artificial Intelligence & Automated Decision-Making Systems in Criminal Justice System in Europe*.<sup>1</sup> In this report, Fair Trials analysed 13 artificial intelligence or automated decision-making systems that have been implemented in the justice systems in Germany, Spain, the Netherlands, England and Wales, Italy, and Ukraine. What follows is a brief account of some of the systems that do not include facial recognition and their main findings. The presentation is divided between AI systems that were used for public security work and AI systems used by the judiciary or prison authorities.

### i. Crime prediction

- *RADAR-iTE. Germany*

In 2017, the German Federal Criminal Investigation Office (BKA) developed RADAR-iTE, a violence risk analysis tool intended to assess the risk of "potentially destructive criminals" committing acts of Islamist terrorism.

According to the German police, using this system, law enforcement resources are more specifically targeted to individuals identified as being at high risk of committing a violent act. RADAR-iTE uses police information on a person's "observable behaviour" that is collected by social workers. As much information as possible is collected about events in the person's life. From this, the tool classifies people as "high," "conspicuous" or "moderate" risk.

Fair Trials concluded that RADAR-iTE's design and objective were intended to assess the risk of Muslim individuals. From the information available online, it appears that such a risk analysis tool continues to be used.<sup>2</sup>

- *ProKid. Netherlands*

In the Netherlands in 2011, ProKid was developed and used by the police to assess the risk of children under the age of 12 committing crimes. Although ProKid was presented as a crime prediction method, it actually only predicted the probability of a child being linked to a crime in police records. ProKid's algorithms were fed from 2 Dutch police databases. These databases contained information on children who had had any kind of contact with the police either as suspects, victims, or witnesses of crimes. When ProKid assessed that a child posed a risk there could be consequences for him and his family. They could even be separated from their family. ProKid's analysis also resulted in criminal records for their parents.

However, ProKid was documented as making errors. The Dutch Ministry of Security and Justice requested an evaluation of ProKid, which reviewed 2,444 risk analyses and found only 1,542 to be correct. That is, more than one-third of the children were assigned the wrong risk level.

<sup>1</sup> Fair Trials, "Automating Injustice: The use of artificial intelligence & automated decision-making systems in criminal justice in Europe", October 2021. [https://www.fairtrials.org/app/uploads/2021/11/Automating\\_Injustice.pdf](https://www.fairtrials.org/app/uploads/2021/11/Automating_Injustice.pdf)

<sup>2</sup> See The International Centre for Counter-Terrorism (ICCT) <https://www.icct.nl/sites/default/files/2024-02/Female%20Jihadis%20Facing%20Justice%20Chapter%204.pdf>

There is no information that the system is still in use, although new versions such as ProKid 23 have been developed.

- *NDAS. United Kingdom*

In 2016, different police institutions in the United Kingdom created the National Data Analytics Database (NDAS) to assess risks and predict crimes. This machine learning-based tool performs behavioural analysis and predictive modelling to make predictions about the possible future actions of specific individuals. These tools are used to influence police actions to prevent crime.

The use of NDAS has led to almost ten times more people of African descent being stopped and searched by the police than white people in the UK and Wales.

In 2021, the year the Fair Trials report was launched, West Midlands Police stated that they intended to use data from public health, education, social security, and local authority offices to feed the predictive tools. The tool continued to be used at least until 2021.

- *HART, United Kingdom*

The *Harm Assessment Risk Tool* (HART) is also used in the United Kingdom. This is an automated system acquired by Durham Police in 2017 that promises to predict whether a crime suspect will commit a crime in the future. Based on the score assigned by HART, authorities decide whether to subject individuals to criminal prosecution or integrate it into a rehabilitation program.

The information that feeds into HART comes from criminal records, a person's age, gender, and place of residence. From an investigation by Big Brother Watch, a human rights organisation based in the UK, it was discovered that the police hired a consumer classification company to feed HART. Big Brother Watch documented that the company, called Mosaic, included overtly discriminatory and stereotypical criteria to classify consumers, such as their Asian ancestry, low income, working class, or even their name. The variable provided by Mosaic had a high level of influence in classifying the person as high, medium, or low risk. Durham police discontinued the use of Mosaic when it was exposed. However, at least until 2021, HART continued to be used.

Finally, the accuracy of HART, calculated from the number of individuals who actually relapsed, was 62% and the high-risk predictions were only accurate in 52% of cases.

## **b. Criminal justice system**

### **i. Evaluation of precautionary measures**

- *Cassandra. Ukraine*

In 2020, the Ukrainian Ministry of Justice announced that an artificial intelligence-based risk assessment system had been developed. It was said that Cassandra would automate the preparation of reports for issuing precautionary measures and for sentencing. Specifically, Cassandra analyses the risk of recidivism of persons who have committed crimes. These reports are used to support judicial decisions regarding the release of detainees.

Cassandra's core objective is to generate a prediction of a person's likelihood of committing a new crime on a scale of 0 to 97. At the time of publication of the Fair Trials

report, there was no public information on the data feeding Cassandra to perform the risk assessments.

## ii. Early release

- *RisCanvi. Spain*

In 2009, Spain implemented RisCanvi, an algorithm-based risk assessment system used to support the decision as to whether detainees can receive early release treatment. RisCanvi is used by the prison system, but it is worth reviewing how it works. In 2018, 15 thousand assessments were conducted using RisCanvi.

Risk factors used by RisCanvi include criminal history, the person's age at the time of the offense, length of sentence, time in prison, escape attempts, as well as disciplinary reports. RisCanvi also uses information about the person's educational level. For example, low mental ability, severe mental damage, or attempts at self-harm. It also uses the family's criminal history, whether they have criminal or anti-social friends, sexual promiscuity, and drug and alcohol abuse.

The system does not 100% make the decision on early release because it is accompanied by interviews conducted by prison staff. The results of the interviews are entered into RisCanvi, and from that, the system predicts whether the risk is low, medium, or high. A committee decides whether to accept RisCanvi's decision or to reject it with justification.

However, according to the Fair Trials report, it is very difficult for a person deprived of liberty to appeal RisCanvi's decision because the ruling judges consider that the decision is made by the prison authorities even though it is highly influenced by RisCanvi.

Despite criticism, the system is still in use to this day.<sup>1</sup>

## 5. Preliminary conclusions

The use of different automated decision systems in security and justice-related issues in Europe has revealed the criminalization and discrimination of people based on their ethnicity, income, or even place of residence. Furthermore, evaluations have shown that in general, the systems make such serious errors that their effectiveness in supporting crime-fighting and relevant judicial decisions is questionable.

In addition, due to the lack of transparency and access to information about the databases that feed the systems and about how decisions are made, people cannot defend themselves once they are analysed by an automated decision-making system.

Several of the artificial intelligence systems used in Europe in tasks related to public security and justice have been reformed based on research carried out by civil society. Therefore, it is essential to continue the critical approach to their evolution and operation, avoiding falling into the myth of technosolutionism.

<sup>1</sup> El País. The algorithm that evaluates the dangerousness of Catalan prisoners works in a "random" way July 2024, See <https://elpais.com/tecnologia/2024-07-03/el-algoritmo-que-evalua-la-peligrosidad-de-los-presos-catalanes-funciona-de-forma-azarosa.html>

Europe's experience shows that the use of automation and algorithms can deepen the discrimination of vulnerable populations if the information that will feed the systems is not taken care of. Therefore, regulation, transparency, accountability, and the possibility of judicial review of decisions made by artificial intelligence systems are required.

## 6. Artificial intelligence in public security tasks and in the administration of criminal justice in Latin America

Focusing on Latin America, this section follows the logic of the previous one to present experiences in the use of artificial intelligence-based systems in public security tasks and in the administration of justice and criminal justice in Latin America. This division is not homogeneous, but, as a general rule, AI systems for public security tasks are acquired and implemented by law enforcement institutions. Meanwhile, AI systems for the administration of criminal justice are acquired and used by judiciary branches.

### a. Public security

First, the use of artificial intelligence tools for facial recognition and crime prediction is discussed. In the area of security, the technology that has proliferated the most is facial recognition. Brazil is the country with the highest overall reported use of surveillance technologies.<sup>1</sup> However, other countries are also reviewed.

#### i. Facial recognition

- *Facial recognition (Oi). Brazil.*

In a study conducted by the human rights organisation O Panóptico, it was reported that the military police of Rio de Janeiro implemented a pilot program of facial recognition since 2019 as a solution to deal with criminality. The program was implemented with the intervention of the company Oi.<sup>2</sup>

However, according to internal military police documents, the facial recognition system was designed to support the police's claim of innocence in the face of possible allegations of wrongdoing. In other words, the evidence produced by the facial recognition technology would only be used if it matched the police version of the arrest.<sup>3</sup>

The cameras were placed at strategic points in the city of Rio de Janeiro. According to the documents analysed by O Panóptico, there is a racially discriminatory bias towards people of African descent, as well as gender, resulting in false positives at the time of their arrest.<sup>4</sup>

- *Proximity policing. Mexico*

<sup>1</sup> Instituto Igarapé, Implementação De Tecnologias De Vigilância No Brasil e na América Latina, Artigo Estratégico 58, Novembro de 2022, p. 2. Available at <https://igarape.org.br/wp-content/uploads/2022/12/Implementacao-de-tecnologias-de-vigilancia-no-brasil-e-na-america-latina.pdf>.

<sup>2</sup> Nunes, Pablo A Rio of cameras with selective eyes: the use of facial recognition by the Rio de Janeiro state police / Pablo Nunes, Mariah Rafaela Silva, Samuel R. de Oliveira. - Rio de Janeiro : CESeC, 2022, p. 7. Available at <https://drive.google.com/file/d/1ofwajPJR9EAyQ48cuMG8qLsYCf0VEuDu/view>.

<sup>3</sup> Nunes, Pablo A Rio of cameras with selective eyes: the use of facial recognition by the Rio de Janeiro state police, Op. cit.

<sup>4</sup> Id.

The analysis of artificial intelligence in Mexico in security tasks requires reviewing the use of video surveillance cameras (CVV). In 2018, the National Institute of Statistics, Geography and Informatics (INEGI) released that there were 53,949 CVVs in Mexico.<sup>1</sup> Meanwhile, the government of Mexico City reported in 2020 that it had 15,310 CVVs.<sup>2</sup>

In 2016, an academic investigation documented that individuals operating video surveillance systems were following construction workers moving into residential areas in the State of Mexico. They also followed people from the LGBTTTQ+ community in a commercial plaza in Mexico City. Those operating the cameras assessed that construction workers could enter a building to steal because it was unusual to see people walking in these residential areas of the city. In the case of people of sexual diversity, the objective of following them was to monitor whether they kissed or hugged. If they were detected, the security forces asked them to behave differently or invited them to leave the place.<sup>3</sup>

Video surveillance cameras are not only operated by people, they also incorporate facial recognition technology. In this way, artificial intelligence applied to facial recognition allows police officers to make decisions related to public security, i.e., arresting people. In Mexico, arrests based on the use of facial recognition technology have been reported in Aguascalientes<sup>4</sup> and in Coahuila.<sup>5</sup>

In the face of this, it has been documented that face recognition technology reproduces the biases of people labelling the training data of this technology.<sup>6</sup> In addition, these systems make errors due to external factors such as lighting conditions, image resolution, or the angle of the face.<sup>7</sup>

Also in Mexico, in 2020, the Citizen Security Directorate of the municipality of Nezahualcóyotl, Microsoft, and Analytikis created the Proximity Policing Model with the objective of predicting where a crime may occur. To do so, they developed an artificial intelligence program with analytical and predictive capabilities that use data collected through surveillance and public service systems to provide intelligence and information to facilitate tactical and strategic decision-making. According to CodinRights, this can present problems of stigmatization, discrimination, and criminalization of the most vulnerable people.<sup>8</sup>

- *Calle Segura. Chile.*

In March 2019, the then-president of Chile, Sebastián Piñera, unveiled the "Mobile Telesurveillance System as part of the "Calle Segura" (safe street) plan. Said system was announced to use unmanned aircraft, i.e. drones, in the fight against crime. The drones would be equipped with facial recognition technology. According to the information, it was planned to implement the monitoring in the Metropolitan Region and then extend it to the whole country.

<sup>1</sup> Álvaro Vizcaíno Zamora. Video surveillance and artificial intelligence: between utopia and dystopia. p. 10.

<sup>2</sup> Ibid., pp. 10 and 11.

<sup>3</sup> Ibid., pp. 14 and 15.

<sup>4</sup> H. Municipality of Aguascalientes 2019-2021. Bulletin no.1161 "New Security Cameras Positive Results". Consultation on July 7, 2021 at: <https://www.ags.gob.mx/cont.aspx?p=6253> cited in Luis García and Grecia Macías. La nueva cara de la discriminación: el impacto diferenciado de la tecnología de reconocimiento facial en grupos vulnerables en SCJN. Discriminación, Problemas Contemporáneos, Ricardo Latapie, Coord., 2022, Mexico. p. 124.

<sup>5</sup> Luis García and Grecia Macías. La nueva cara de la discriminación: el impacto diferenciado de la tecnología de reconocimiento facial en grupos vulnerables en SCJN. Discriminación, Problemas Contemporáneos, Ricardo Latapie, Coord., 2022, Mexico. p. 125.

<sup>6</sup> Ibid., p. 139.

<sup>7</sup> Ibid., p. 124.

<sup>8</sup> Available at <https://notmy.ai/es/project-item/modelo-de-policia-de-proximidad-es/>

In response, a group of human rights organisations spoke out against the use of this technology. Some of the reasons for rejecting the use of facial recognition technology were that it is highly fallible, it facilitates the institutionalization of biases associated with social class and skin colour, and it lacks transparency and accountability mechanisms.<sup>1</sup>

- *Facial recognition. El Salvador*

In El Salvador, the government of Nayib Bukele announced the acquisition of surveillance cameras with facial recognition capabilities as part of its security policy. The cameras were purchased with funds from a loan from the Central American Bank for Economic Integration (CABEI).<sup>2</sup>

Servicio Social Pasionista, a Salvadoran human rights organisation, expressed concern about the opacity with which this type of technology is being acquired and operated. In addition, because the security policy implemented in El Salvador has been characterised by serious human rights abuses, they fear that the use of these systems will be counterproductive, especially against typically stigmatised and discriminated populations.<sup>3</sup>

## ii. Crime prediction

In Chile, Colombia, and Uruguay, crime prediction systems have been implemented or pilot-tested. In some cases, the technology's use has been announced, but there is no accurate information on whether the governments are using it.

- *Urban Crime Predictive System. Chile*

The Urban Crime Predictive System was developed in 2017 by the University of Chile<sup>4</sup> together with the Department of Criminal Analysis (DAC) of Carabineros de Chile. It was implemented to predict areas of higher risk of crime occurrence to efficiently direct police patrolling.<sup>5</sup>

According to the technical report, the aim was to develop a predictive crime system based on mathematical algorithms in order to carry out focused patrols and to reinforce "the effectiveness of the criminal prosecution system."

The system was analysed by researcher Josefina Buschmann for Derechos Digitales América Latina. Some of the conclusions of the analysis are the following.

There is no such thing as neutral or objective data. Police data are used and their context of production must be analysed. The context of production may be unjust or arbitrary arrests or over/underestimated complaints. Such a context may bias the database and reproduce discriminatory practices that lead to intensified surveillance and preventive identity checks in certain areas.<sup>6</sup>

<sup>1</sup> Datos protegidos, Pronunciamiento "Estamos en contra de la vigilancia masiva en los espacios públicos del "Sistema de televigilancia móvil", April 8-2019. This publication was subscribed by other NGO's S. <https://datosprotegidos.org/estamos-en-contra-de-la-vigilancia-masiva/>

<sup>2</sup> SSPAS, Uso de tecnología en las políticas de seguridad en El Salvador, los riesgos en los derechos humanos. 2021, pp. 23, 25 and 32.

<sup>3</sup> Ibid., see conclusions.

<sup>4</sup> Specifically, by the Center for Analysis and Modeling in Security (CEAMOS).

<sup>5</sup> Josefina Buschmann, Predictive System of Urban Crime: Algorithmic Production of Surveillance and Control Zones in the City, Digital Rights, Artificial Intelligence and Inclusion in Latin America. December 2021. P. 6, Available at [https://www.derechosdigitales.org/wp-content/uploads/02\\_Informe-Chile-Sistema-Predictivo-del-Delito-Urbano\\_ES\\_28042022.pdf](https://www.derechosdigitales.org/wp-content/uploads/02_Informe-Chile-Sistema-Predictivo-del-Delito-Urbano_ES_28042022.pdf)

<sup>6</sup> Ibid., pp. 41 and 42.

In the face of these problems, there is no protocol for an external, civilian audit of police data collection practices.<sup>1</sup>

One of the central points of the assessment states that the "problem with integrating police data into an automated [...] system [...] is that it rearticulates them as scientifically valid data."<sup>2</sup>

Carabineros, the Chilean police, did not consider that the model could impact people's rights because it does not use personal data or identify exact locations. However, Josefina Buschmann points out that the surveillance zones created from the system may imply changes in police behaviour. These changes may include harassment of people who have not committed a crime but whose behaviour is considered suspicious.<sup>3</sup>

- *Empirical study on the identification of criminal behaviour. Chile*

In 2018, the Pontificia Universidad Católica de Chile published research<sup>4</sup> related to the use of automated tools to analyse and predict criminal behaviour and help police reduce crime. The study attempted to predict the criminal behaviour of people who had been previously arrested.<sup>5</sup> The proposed prediction model was not applied.

The research analysed almost 800 thousand arrests made by police officers between 2009 and 2018 and sociodemographic information about the offenders. From this, it estimated a person's chances of recidivism. Some elements that were taken into account were: previous arrests and personal background, including the number of arrests of family members.

Thirty-seven percent of the predictions made by the model that a person would be re-arrested were wrong.<sup>6</sup> The erroneous predictions were labelled as false positives and were acknowledged to carry a high social cost. The research states that crime is simply difficult to predict. It further states that prediction can be improved by adding attributes such as arrest history or biographical indicators such as education or the person's assets.<sup>7</sup>

The study asks key ethical and legal questions. For example, to what extent can the state use personal information for purposes that citizens have not been informed of or consented to? Can the record of previous arrests or other similar information be used as a basis for criminal prevention and investigation actions?

Thus, the research recognises that this type of system can encourage the detention of people based on their personal or family background and not for behaviours they have actually committed. In addition, it could affect the rights of people in situations of vulnerability.

- *Colombia Police. Colombia*

In 2019, the government of Colombia unveiled the "Colombia Policías" project<sup>8</sup>. The objective is to anticipate crime in Bogotá, predicting when, where, who, and why crime occurs through a predictive mathematical program that describes homicides, property crimes involving force or threat of force, personal injuries, and the dynamics behind the

<sup>1</sup> Ibid., p. 42.

<sup>2</sup> Id.

<sup>3</sup> Josefina Buschmann, Sistema Predictivo del Delito Urbano, Op. cit., p. 47.

<sup>4</sup> Van 'T Wout et. al. Big data for the identification of criminal behavior, in Proposals for Chile, Pontificia Universidad Católica de Chile, 2018, Available at: [https://politicaspUBLICAS.uc.cl/content/uploads/2019/03/Libro-completo-en-PDF\\_final-6.pdf](https://politicaspUBLICAS.uc.cl/content/uploads/2019/03/Libro-completo-en-PDF_final-6.pdf)

<sup>5</sup> Ibid., p. 51.

<sup>6</sup> Ibid., p. 65.

<sup>7</sup> Ibid., p. 66.

<sup>8</sup> El Espectador, "El matemático que quiere predecir los crímenes de Bogotá", 24 Jun 2019. <https://www.elespectador.com/ciencia/el-matematico-que-quiere-predecir-los-crimenes-de-bogota-article-867606/>

public perception of security. The program was developed by the District Secretariat of Security, Coexistence, and Justice.

According to CodinRights mapping,<sup>1</sup> there are concerns about the possible stigmatization, discrimination, and criminalization of the most vulnerable people.

- *Crime prediction in intermediate cities. Colombia*

Research published in September 2022 analysed the use of artificial intelligence-based technology to predict crime in an intermediate city, Bucaramanga, located in the Department of Santander in Colombia. The research estimated the predictive capacity of crime in a city for which there is little data and few capabilities compared to major cities such as Bogota.<sup>2</sup>

The research developed a pilot crime prediction model that was not applied. Its data source was the Statistical, Criminal, Contraventional, and Operational Information System of the National Police (SIEDCO). The level of accuracy obtained in the prediction was between 50% and 60%, which means that about half of the predictions made by the model were true crimes.<sup>3</sup>

Research recognises the existence of biases and discrimination in predictive models and invites consideration of these factors prior to implementation.<sup>4</sup>

The conclusions point out that crime prediction models can have discriminatory effects on certain population groups or areas of the city. Therefore, it recommends implementing predictive models in parallel with other efforts.<sup>5</sup>

- *PredPol. Uruguay*

In 2013, the Ministry of Interior of Uruguay acquired the PredPol software, which emerged from a joint research project of the Los Angeles Police Department and the University of California.<sup>6</sup>

The tool promised to predict where a crime would occur. (in spaces of 150 square-meters). PredPol was implemented between 2014 and 2017.<sup>7</sup> PredPol made predictions based on the system that brings together public security information throughout Uruguay.

The analysis of the tool's operation concluded that not all crimes are reported, and the police bases, as expected, are built from what individuals report. This implies, for example, that some crimes are reported more than others. It was also identified that police presence in an area may increase the likelihood of a crime being identified or reported.

Once the evaluation process was completed, the Ministry of the Interior observed that the areas where PredPol was used did not show better results than those where the static system developed by its own technicians was used. For this reason, it was decided to discontinue the use of PredPol.<sup>8</sup>

<sup>1</sup> CodinRights <https://notmy.ai/es/mapeo-de-proyectos/>

<sup>2</sup> Juan-David Galvez-Ferreira et. al. Predicting crime in intermediate cities: a machine learning model in Bucaramanga, Colombia.

<sup>3</sup> Ibid., p. 94.

<sup>4</sup> Ibid., p. 85.

<sup>5</sup> Ibid., pp. 95 and 96.

<sup>6</sup> WWF. Algorithms and artificial intelligence in Latin America. A study of implementations by governments in Argentina and Uruguay, 2018, p. 27.

<sup>7</sup> Id.

<sup>8</sup> WWF. Algorithms and artificial intelligence in Latin America. Op. cit., p. 29.

## **b. Administration of justice and criminal justice**

### **i. Evaluation of evidence and sentences. Mexico**

There is no record of artificial intelligence being used in the criminal justice system in Mexico. However, this possibility has been discussed. For example, Juliana Vivar published article<sup>1</sup> in the journal of the National Institute of Criminal Sciences (INACIPE), in which she analysed theoretically the possibility of using predictive models in criminal sentencing based on the Mexican legal framework.

In the article, Juliana Vivar analyses the characteristics that distinguish human beings from machines, such as freedom and autonomy, as well as the characteristics that are common to them, such as error and bias.

### **ii. Precautionary measures**

- *PRiSMA. Colombia*

In 2016, the Attorney General's Office in Colombia initiated a program operated through artificial intelligence called PRiSMA (Perfil de Riesgo de Reincidencia para Solicitud de Medidas de Aseguramiento). The objective is to determine the risk of recidivism of persons accused of crimes to support the decision of the Prosecutor General's Office on whether or not to request preventive detention.

PRiSMA is based on an algorithm that uses machine learning to process information from a database. The database was provided with information by the National Police and the Attorney General's Office.<sup>2</sup> From this, PRiSMA helped to predict the probability that a person would reoffend. If the system predicts that the person represents a high level of risk, the Prosecutor General's Office requests a more restrictive precautionary measure for liberty in order to prevent the person from committing a crime. If the system predicts a low risk, non-restrictive measures of liberty are requested.

The database that feeds PRiSMA has information on six million individuals with criminal records.<sup>3</sup> Using machine learning, patterns of behaviour associated with current and previous criminal events, judicial records, contraventions, and prison records are determined, and a level of risk related to the commission of crimes against property, violent crimes, and other crimes is established within a period of two years after the indictment.

<sup>1</sup> Juliana Vivar. Sentencing and penal algorithmic prediction. Herramienta o suplencia humana in INACIPE, Ciencias Penales e inteligencia artificial, 2021, Available in <https://revistaciencias.inacipe.gob.mx/index.php/02/article/view/437/342>

<sup>2</sup> Carmen Eloísa Ruiz López, Valentina del Sol Salazar and Humberto J. Sierra. Artificial Intelligence (AI) Operated Systems and Criminal Due Process. Perspectiva de aplicación en Colombia en Disrupción tecnológica, transformación digital y sociedad, Tomo III, Derecho, Innovación y tecnología: Fundamentos para el mundo digital. Diego Acosta González (and others): Juan Carlos Henao, Daniel Castaño (eds) - Bogotá, Universidad Externado de Colombia, 202, p. 841. Available at <https://bdigital.uexternado.edu.co/server/api/core/bitstreams/97830f4d-62c2-4d3c-b3cc-fea2145e11c8/content>.

<sup>3</sup> V. SALAZAR. Estrategias para la racionalización de la detención preventiva en Colombia: un análisis del instrumento PRiSMA, Colección Tesis de grado No. 107, Bogotá, Universidad Externado de Colombia, 2020, Title 4.1; crf. Also, Centro de Estudios sobre Seguridad y Drogas (CESED). "Prisma: el programa de la Fiscalía para predecir la reincidencia criminal," Faculty of Economics, Universidad de los Andes, available at [<https://cesed.uniandes.edu.co/prisma-el-programa-de-la-fiscalia-para-predecir-la-reincidencia-criminal/>] cited in Disrupción tecnológica, transformación digital y sociedad.

The Attorney General's Office<sup>1</sup> hoped to reduce criminal recidivism but also to reduce the use of pretrial detention. However, there may also be concerns that due process guarantees such as the right to defence, presumption of innocence, and judicial impartiality may be affected as part of the risk of the use of artificial intelligence.<sup>2</sup>

Codin Rights considered that PRiSMA may incur some biases such as stigmatization, discrimination, and criminalization of the most vulnerable people.<sup>3</sup>

There is no information that PRiSMA continues to be used.

### iii. Evaluation of evidence and sentencing

- *Projeto Sócrates Criminal procedure. Brazil*

In Brazil, in 2019, the "Projeto Sócrates" was announced, which, through artificial intelligence seeks to reduce by 25% the time to issue appellate judgments.<sup>4</sup> The system analyses the appeals received by the Brazilian Superior Court of Justice from 300,000 resolved cases and groups cases that are similar to decide them together.<sup>5</sup>

According to the Minister of Justice, Ricardo Villas, the aim is to implement this system to produce automated draft decisions based on the analysis of previous opinions.<sup>6</sup>

However, according to Codin Rights, there is concern about the possible risk of automating the systematic injustices historically produced by the Brazilian judicial system.<sup>7</sup>

### iv. Judicial management

- *PROMETEA. Argentina*

In 2019, the Deputy General Prosecutor's Office for Administrative and Tax Litigation of the Public Prosecutor's Office of Buenos Aires, Argentina adopted PROMETEA. This system incorporates artificial intelligence to automate the preparation of legal opinions based on analogous cases for which there are reiterated judicial precedents.<sup>8</sup>

According to its developers, PROMETEA implied reducing the resolution of a procurement document from 90 minutes to 1 minute.<sup>9</sup> In other words, it is a system that represents extraordinary efficiency. The gain in case resolution time allowed staff to devote more time to more complex cases.

A publication of the Inter-American Development Bank, PROMETEA has meant "a significant improvement in the overall efficiency of the institution [the Attorney General's

<sup>1</sup> Office of the Attorney General of the Nation; PRISMA Tool, Recidivism Risk Profile for the Request for Assurance Measures, Public Policy and Strategy Directorate.

<sup>2</sup> Disrupción tecnológica, transformación digital y sociedad, Tomo III, Derecho, Innovación y tecnología: Fundamentos para el mundo digital. Diego Acosta González (and others); Juan Carlos Henao, Daniel Castaño (eds) - Bogotá, Universidad Externado de Colombia, 2021.

<sup>3</sup> Codin Rights, Recidivism Risk Profile for the Request of Security Measures-Prisma (Colombia). Available at <https://notmy.ai/project-item/prisma-en/>

<sup>4</sup> Pilot project of Socrates, STJ's artificial intelligence program, is expected in August. <https://www.migalhas.com.br/quentes/299820/projeto-piloto-do-socrates--programa-de-inteligencia-artificial-do-stj--e-esperado-para-agosto>

<sup>5</sup> Leonie Wittenstein and Felipe Franco. Use of artificial intelligence and algorithmic systems in the judicial system KAS, Pág. 8, Disponible en <https://www.kas.de/documents/271408/16552318/Usode+de+la+Inteligencia+Artificial+y+los+sistemas+algoritmicos+en+el+sistema+judicial+%281%29.pdf/7cd2c325-6b7f-475e-2a8b-59b5544823cb?version=1.0&t=1680036938780>

<sup>6</sup> Codin Rights, Prójetto Sócrates (Brazil). Available at <https://notmy.ai/es/project-item/projeto-socrates/>

<sup>7</sup> Id.

<sup>8</sup> Elsa Estevez, Sebastián Linares and Pablo Fillottrani. PROMETEA. Transforming the administration of justice with artificial intelligence tools, IDB, 2020. Available at <https://publications.iadb.org/es/publications/spanish/viewer/PROMETEA-Transformando-la-administracion-de-justicia-con-herramientas-de-inteligencia-artificial.pdf>

<sup>9</sup> Ibid., p. 22.

Office]."<sup>1</sup>

The same publication acknowledges that "the greatest risk of applying AI in decision-making processes is that it amplifies existing biases rather than helping to remedy them".<sup>2</sup> However, in the case of PROMETEA, such risks were addressed through human rights-based controls.<sup>3</sup>

- *PretorIA. Colombia*

The Constitutional Court of Colombia is developing PretorIA, an artificial intelligence tool that seeks to streamline the selection of judicial tutela cases.<sup>4</sup> In Colombia, the Constitutional Court sets its jurisprudence based on the decision of tutelas.<sup>5</sup> PretorIA will basically help the Constitutional Court to pre-select the tutelas to be reviewed. The Constitutional Court receives more than 600,000 tutelas per year.<sup>6</sup>

PretorIA was inspired by PROMETEA, the tool used by the Buenos Aires Attorney General's Office mentioned in this report.

Saavedra and Upegui do not foresee tangible risks to the rights of individuals from the use of PretorIA because it has been announced that it will automate part of the activities in the preprocessing of information, but the decision maker is a person.<sup>7</sup>

## 7. Artificial intelligence and emerging challenges to due process in Latin America

Research conducted by human rights organisations in different parts of the world has agreed that the use of artificial intelligence-based systems in public security and justice systems represents a risk to human rights. AI in the criminal justice system can deepen discrimination and affect the right to privacy; it also has ethical implications.

However, as well it is necessary to review the possible effects of the use of tools based on artificial intelligence in the criminal justice system on the guarantees of due process, such as the presumption of innocence, the right to defence, the right to appeal the judgment, and judicial independence and impartiality. The right to personal liberty should also be taken into account.<sup>8</sup>

For example, facial recognition tools are used to identify persons for criminal investigation purposes. When such tools make mistakes, they affect the presumption of innocence and due process, and given the difficulty of challenging the decision made by an automated system, the right to defence becomes illusory. For this reason, it is necessary to identify the risks to human rights that technologies based on artificial intelligence generate. This identification will make it possible to regulate their use when appropriate or to prohibit them.

<sup>1</sup> Page 88

<sup>2</sup> *Ibid.*, p. 89

<sup>3</sup> Juan G. Corvalán, Enzo Maria Le Fevre Cervini, CERIDAP, 2020. Available at <https://ceridap.eu/prometea-experience-using-ai-to-optimize-public-institutions/?lng=en#:~:text=Prometea%20is%20a%20predictive%20artificial,exponential%20optimizer%20of%20bureaucratic%20processes>.

<sup>4</sup> Victor Saavedra and Juan Carlos Upegui. PretorIA and the automation of human rights case processing. *Derechos Digitales América Latina and Dejusticia*. P. 18. Available at [https://www.dejusticia.org/wp-content/uploads/2021/04/CPC\\_informe\\_Colombia.pdf](https://www.dejusticia.org/wp-content/uploads/2021/04/CPC_informe_Colombia.pdf)

<sup>5</sup> *Id.*

<sup>6</sup> *Ibid.* p. 22.

<sup>7</sup> *Ibid.*, p. 53.

<sup>8</sup> Fair Trials pointed out that using artificial intelligence-based systems in the criminal justice system violates the right to a fair trial and the presumption of innocence. Fair Trials, *Automated Injustice*, Op. cit. p. 4.

However, there are mechanical tasks related to the justice system, such as the search for judicial precedents or the generation of statistics, which do not affect the determination of rights. In these cases, artificial intelligence can contribute to noble purposes such as improving investigative capacities, efficiency, the reduction of waiting times, and greater predictability of decisions, among others.

In this understanding, the approach to AI in the justice system must be balanced so as not to condemn its use a priori but also not to incorporate it without considering its risks. When AI is condemned without a diagnosis, its benefits may be wasted. When AI is accepted without reviewing its risks, benefits, or capabilities, it may not be attributed to it.

#### **a. Personal liberty**

In legal systems, any infringement of the right to liberty by a police officer on a person must be objectively justified. For example, when the police observe someone committing a crime. For that reason, Article 9 of the International Covenant on Civil and Political Rights (ICCPR or the Covenant) requires that there must be justified grounds and a procedure established by law to detain a person. The American Convention on Human Rights (ACHR) concurs with the Covenant. It states that it is possible to restrict liberty only when there are causes and conditions previously established by the Constitutions or by laws enacted in accordance with them and in accordance with objectively defined procedures.<sup>1</sup>

Recognizing that it is legitimate for the police to perform crime prevention functions, these international standards are designed to protect individuals from arbitrary or unlawful interference by the state that endangers their liberty or even their lives. These standards are also applicable when the authorities rely on technology to perform security functions.

Thus, tracking people and random checks based on artificial intelligence tools affect, among other rights, personal freedom. First, because the systems reproduce discriminatory biases that especially affect disadvantaged social groups. This entails an arbitrary restriction of freedom (art. 7.3 of the ACHR). Secondly, because the systems commit errors in such a high percentage that they allow the police to follow and search people without a justified cause, which also represents an arbitrary restriction of liberty (Art. 7.3 of the ACHR). Thirdly, given the lack of information related to the way in which the tools based on artificial intelligence operate, the person does not know the reasons that led the police to search him or her (Art. 7.4 of the ACHR). The latter, in turn, violates the right of every person to judicial control of the detention and to challenge its legality (Art. 7.5 and 7.6 of the ACHR).

#### **b. Due process**

The use of artificial intelligence in the criminal justice system requires reflection on the application of due process guarantees. This report will not make an exhaustive study of all the due process guarantees affected, but it will provide an initial approximation to establish that when artificial intelligence-based tools are used in the criminal process, human rights must be at the centre.<sup>2</sup>

Respect for due process or fair trial is crucial for constructing a judicial decision that affects individuals' rights. Guaranteeing a fair trial gives validity and legitimacy to the

<sup>1</sup> I/A Court H.R., Case of Fernández Prieto and Tumbeiro v. Argentina. Case of Fernández Prieto and Tumbeiro v. Argentina. Merits and Reparations. Judgment of September 1, 2020. Series C No. 411, para. 66.

<sup>2</sup> Aleš Završnik in the article Criminal system, artificial intelligence systems and human rights analyses the impact on due process of artificial intelligence systems from the perspective of the European Convention on Human Rights.

process of judicial determination of rights.<sup>1</sup> At the end of the day, subjecting a person to criminal proceedings implies depriving him of such valuable rights as his freedom or property. Thus, the incorporation of artificial intelligence into the criminal process does not exempt the authorities from respecting this first-order right.

For example, when a prosecutor's office or a court relies totally or partially on an artificial intelligence system to decide which precautionary measure to request or order, evaluate evidence, or rule on a request for early release, it affects due process. In the following, we will review some due process guarantees that may be affected by the use of artificial intelligence.

### **c. Presumption of innocence**

The right to the presumption of innocence is a cornerstone of due process. It implies that a person accused of a crime should not be presumed guilty until the accusation has been proven beyond a reasonable doubt. For that reason, the accused person should enjoy the benefit of the doubt and be treated as innocent.<sup>2</sup>

This right can be violated from different perspectives by artificial intelligence systems. For example, when a system predicts that a person will commit a crime and, based on this, important decisions are made, such as denying him or her the possibility of being tried for freedom, the presumption of innocence is affected. In other words, the person is considered to have a certain level of risk because he/she coincides with a previously designed profile.<sup>3</sup>

It is not necessary for a person to be detained in order to affect his or her right to the presumption of innocence. It is sufficient for the police to carry out supposedly random searches based on decisions that rely on technologies that incorporate artificial intelligence to cause an affectation.

### **d. Right to legal representation and equality**

Everyone charged with a criminal offense has the right to be informed, in a language they understand and in detail, of the charge against them (art. 14.2.a of the ICCPR). They also have the right to have adequate time and facilities to prepare their defence (art. 14.2.b of the ICCPR) and to cross-examine the witnesses against them (art. 14.2.e of the ICCPR). In compliance with Article 14.2 of the ICCPR, these guarantees must be applied equally.

For that reason, people accused of crimes must understand how artificial intelligence systems work, including the information that feeds the databases that allow decisions to be automated. They must also be informed of how the decision was made and whether it was fully adopted by the AI tool or served as partial support to guide the judicial work. In case the decision was entirely taken by the AI, the accused persons should be in a position to present evidence challenging the decision.

The due process problems generated by the use of AI-based systems are similar to those encountered when using protected or anonymous witnesses in criminal prosecutions.<sup>4</sup>

<sup>1</sup> Carmen Eloísa Ruiz López, et. al. Systems operated by Artificial Intelligence (AI) and criminal due process. Perspective of application in Colombia, Op. cit. p. 844.

<sup>2</sup> Human Rights Committee, Case Sergey Khmelevsky v. Belarus, CCPR/C/139/D/2792/2016. 2024, para. 8.2.

<sup>3</sup> Jordi Nieva Fenoll. Inteligencia artificial y proceso judicial, 2019, p. 153.

<sup>4</sup> Aleš Završnik, Criminal justice, artificial intelligence systems and human rights, 2020, p. 577.

This is asserted because the way in which the systems are programmed is opaque,<sup>1</sup> their funding is unknown, and who designs them<sup>2</sup>. Not knowing how the algorithms work would be equivalent to forcing a person to be tried in a process for which he/she knows the applicable laws or in a language he/she does not know.<sup>3</sup>

When all these conditions are not guaranteed, the equality of procedural means is violated to the detriment of the accused person<sup>4</sup>, and the right to defence. This is so because the lawyer is prevented from preparing a technical defence due to the lack of knowledge of the functioning of the systems based on AI.<sup>5</sup>

#### **e. Right to appeal the judgment**

International instruments establish the right of the accused to appeal the judgment. (Article 14.5 of the ICCPR and Article 8.2.h) of the ACHR enshrine this right. According to General Comment No. 32 of the Human Rights Committee, the right to appeal the judgment encompasses the full review of the conviction. Such a review goes beyond the formal or legal aspects of the sentence. It must include a new analysis of the factual, evidentiary, and legal issues on which the challenged judgment is based.<sup>6</sup> It must also be carried out by a different and higher court judge.<sup>7</sup> The purpose of this right is to prevent a decision adopted in a flawed proceeding and with errors that cause prejudice to the defendant from becoming final.<sup>8</sup>

To respect this right, it must be ensured that persons accused of crimes can appeal in full against decisions affecting them that were made with artificial intelligence. The review needs to be able to analyse the factual and evidentiary issues on which the decision was based in order to identify possible errors. For example, if an AI-based tool decides that a person represents a high level of risk and, based on this, is ordered to be remanded in custody during the proceedings, the person should be able to challenge the decision. To do so, he or she has the right to know at least what technology was used, what information fed the database, what calculations were made, how the algorithm was trained, and what percentage of the decision corresponded to the system and what percentage to the judge. This data should be available for a different, higher authority to review the decision.

Without ensuring that the right to appeal the ruling can operate in the same way as when artificial intelligence is not used to make a decision in criminal proceedings, the use of such technology should not be allowed.<sup>9</sup>

#### **f. Independence and impartiality**

Independence and impartiality must also be ensured when AI-based tools are used in criminal proceedings. These guarantees are traditionally analysed in light of the performance of the adjudicators, but when an AI-based system is involved in a decision in criminal proceedings that determines rights, independence, and impartiality must be ensured from other coordinates, i.e., from the machines.<sup>10</sup>

<sup>1</sup> Carmen Eloísa Ruiz López, et. al. Systems Operated by Artificial Intelligence (AI) and criminal due process. Perspective of application in Colombia, Op. cit., p. 846.

<sup>2</sup> Ibid., p. 848.

<sup>3</sup> Jordi Nieva Fenoll. Inteligencia artificial y proceso judicial, 2019, p. 128.

<sup>4</sup> Human Rights Committee. Case Amir Abdiev v. Kazakhstan. CCPR/C/137/D/2618/2015. 2023, para. 7.8

<sup>5</sup> Carmen Eloísa Ruiz López, et. al. Systems Operated by Artificial Intelligence (AI) and criminal due process. Perspective of application in Colombia, Op. cit., p. 849.

<sup>6</sup> Mohamed vs. Argentina

<sup>7</sup> Case of Herrera Ulloa v. Costa Rica.

<sup>8</sup> Barreto Leiva vs. Venezuela

<sup>9</sup> Jordi Nieva predicted that the use of AI in the process will reconfigure challenges or even diminish them. See Artificial Intelligence and the Judicial Process, Op. cit., p. 144.

<sup>10</sup> Jordi Nieva Fenoll. Inteligencia artificial y proceso judicial, 2019, p. 128.

According to the Bangalore principles, judicial independence requires that the judge be free from connections and inappropriate influences with the executive and legislative branches.<sup>1</sup> These external connections and influences are also unwelcome when dealing with private actors such as corporations.

Also, following General Comment No. 32 of the Human Rights Committee, impartiality refers to the court's lack of prejudice or bias in making an appropriate decision in a case. The Bangalore principles add that impartiality refers to the judicial decision itself and to the decision-making process.

Applying such criteria to judicial decisions made with artificial intelligence should be analysed, at least where the information that feeds the algorithm comes from, who designs it, and the source of funding of the system. For example, judicial independence could be compromised by an AI-based system that automates judicial decisions based on data provided by the prosecution or the police. The same would be true if an AI tool used by the judiciary is funded by the prosecution or an external actor. Moreover, if the data automated by an AI system reproduces biases introduced by the prosecution and serves to reach a judicial decision, impartiality could be compromised. In other words, the executive branch should be prevented from influencing judicial decisions where AI is involved, just as it is when AI is not involved.<sup>2</sup>

The above is compounded by the fact that the incorporation of AI into the judicial decision may cloak it under the veil of technicality.<sup>3</sup>

## 8. Conclusions

Several governments in Europe are incorporating artificial intelligence-based technology for public security and for the operation of the criminal justice system. However, research shows that artificial intelligence is coming at a high cost to human rights. For example, crime prediction tools deepen discrimination against vulnerable populations on the basis of race or religious beliefs. This has been the case with the RADAR-iTe tool in Germany and NDAS in the UK.

Moreover, their effectiveness is in question as demonstrated by the mistakes made by HART in the UK and ProKid in the Netherlands.

In Latin America, the implementation of artificial intelligence in security work and the criminal justice system has been more gradual and with reservations. While crime prediction tools are not common, facial recognition is proliferating, as shown by the cases of Brazil, Mexico, El Salvador, and Chile. This has raised the alarm of several specialised human rights organisations, as there is a risk that these tools may reproduce biases and deepen discrimination against vulnerable populations such as people of African descent.

The use of artificial intelligence in the criminal justice system has been on a smaller scale. The cases analysed were Projeto Sócrates in Brazil and Prisma in Colombia.

<sup>1</sup> The Human Rights Committee noted that judicial independence requires that the judge and prosecutor be able to "interpret and apply the law, and evaluate facts and evidence freely without being subject to intimidation, obstruction or interference in the exercise of their function." Human Rights Committee. Case of Baltasar Garzón v. Spain. CCPR/C/132/D/2844/2016. 2021, para. 5.5.

<sup>2</sup> Jordi Nieva Fenoll. Artificial Intelligence and the Judicial Process, Op. cit.

<sup>3</sup> Carmen Eloísa Ruiz López, et. al. Systems Operated by Artificial Intelligence (AI) and criminal due process. Perspective of application in Colombia, Op. cit.

Regarding Socrates, which processes appeals, there are concerns that it automates systemic injustices inherent in the Brazilian judicial system. Regarding Prisma, which assessed the risk of recidivism, there were concerns about its impact on due process guarantees such as defence, presumption of innocence, and impartiality. However, there is no information that it continues to be used.

On the other hand, exercises in Uruguay, Chile and Colombia with crime prediction systems suggest that governments and universities are exercising caution in several Latin American countries. Instead of implementing the technology, they have conducted pilot tests to identify errors and possible biases.

Likewise, PROMETEA in Argentina and PretorIA in Colombia are good examples of the use of artificial intelligence for judicial management. In these countries, automation is helping to address common problems in justice systems, such as case processing delays. Also, in both cases, there is a growing awareness of the risks that artificial intelligence poses to human rights.

Artificial intelligence is a technological tool that is being used in Latin American justice systems. However, we must be alert to the use of automated decision-making systems that, far from achieving the results that their developers promise, may deepen existing inequalities against the most vulnerable populations.

The punitive power of the State, insofar as it implies depriving a person of such a precious good as freedom, should not be sacrificed under the false promise that systems based on artificial intelligence are objective and make decisions free of bias.

For this reason, it is proposed to rethink the classic guarantees of due process, such as the presumption of innocence, the right of defence, the right to appeal the ruling, and judicial independence, in the light of artificial intelligence systems. This report presents a first reflection that should be deepened so that no system based on artificial intelligence is implemented that allows violations of the right to due process.

Finally, governments should make transparent plans to incorporate artificial intelligence into criminal justice systems. The general population has a right to access such information in a language they understand. In addition, people who are subject to police review, detained, prosecuted, or convicted have a right to know how AI-based systems operate.

## 9. Recommendations

- Develop minimum elements for national and regional regulation of the use of artificial intelligence. For this purpose, good practices from Latin American countries as well as from Europe can be taken up, as is the case of the Law on Artificial Intelligence adopted by the European Parliament. Among them:
  - Prohibition of the use of mass surveillance technologies such as facial recognition;
  - Prohibition of the use of crime prediction systems;
  - Transparency and accountability regarding the acquisition and use of artificial intelligence systems;
  - Development of due process guarantees applicable to artificial intelligence

systems during all stages of the criminal process.

- Recover exercises such as those carried out in Colombia and Chile in which pilot tests were carried out without implementing systems based on artificial intelligence to evaluate their operation, the errors, and the risks they entail for human rights
- Explore opportunities to use artificial intelligence systems in the judiciary's work with due safeguards, such as implementing external audits and transparency, as in Argentina and Colombia.
- Conduct external audits of any technology that incorporates artificial intelligence, and that is implemented by public security forces or criminal justice systems
- Continue generating spaces for dialogue and interaction between authorities of public security systems, judicial branches, academia, and civil society to discuss the risks and opportunities presented by artificial intelligence for public security and criminal justice system tasks.