AIDP RESOLUTIONS

Section 3 – AI and the administration of criminal justice: 'Predictive policing,' 'predictive justice,' and evidence

Buenos Aires, 28–31 March 2023

Preamble

Aware that artificial intelligence (AI) is developing rapidly in contemporary society in various parts of the world. Already ubiquitous in peoples' lives in some countries, it may become part of daily life for a large part of the world's population in the future.

Noting that as a technological innovation, AI pushes consumers to buy new products, thus helping the global economy grow. Therefore, AI plays a non-negligible role in sustaining and even expanding the liberal market economy and the capitalist economic system.

Noting that the companies that create and market AI are frequently based in developed countries in the global North, and they often try to open markets all over the world.

Noting that 'digital divide' widens social inequalities among people. 'AI divide' may be the next phenomenon on the horizon.

***

Considering that AI can be defined as a set of theories and techniques used to create machines capable of simulating human intelligence.[1] As a scientific discipline, it is a blend of statistical and algorithmic mathematics, computer science, and the cognitive sciences. Symbolic AI is based on the rules of logic, whereas connectionist AI uses artificial neural networks.

Considering that machine learning is an example of connectionist AI, as is deep learning, which is a subset of machine learning that uses multiple layers of interconnected artificial neurons. As the number of neuronal layers enabling autonomous learning increase, the system's technological complexity increases, making the system more efficient and its calculations less explainable and traceable (deep learning).

Considering that machine learning can make technology extremely powerful, but its decision-making process can be so complex that it resembles a 'black box'.

[1] https://www.larousse.fr/encyclopedie/divers/intelligence_artificielle/187257

Noting that many AI systems used in the context of preventing, investigating, detecting, and punishing crime are machine-learning systems. Using self-learning algorithms, they carry out complex probability calculations in nanoseconds. To achieve their assigned goals, they process huge amounts of data and consume a lot of energy. Some of them, such as facial recognition systems, rely on deep learning.

***

Finding that in criminal justice administration, AI systems are used to prevent or detect criminal offenses based on risk assessment ('predictive policing').

Finding that AI systems are also used to help prosecutors and judges make decisions. More specifically, the term 'predictive justice' refers to (i) anticipating someone's behavior, e.g., to assessing the risk of fleeing in the pre-trial procedure or (re-) committing a crime so that decisions concerning them, such as pre-trial detention, sentencing, parole, and probation (actuarial justice, that nowadays may be supported by AI) may be made; and (ii) using AI to perform an ultraquick statistical analysis of prior decisions issued in similar cases and of relevant legal and regulatory provisions (quantitative legal analysis or LegalTech).

Considering that quantitative legal analysis is revolutionary in the sense that a mathematical calculation is meant to support or even to supplant legal reasoning.

Observing that the word 'predictive' used in the phrases 'predictive policing' and 'predictive justice' is confusing because AI systems calculate probabilities, but do not predict the future; these probabilities are based on correlations, not on causations. These calculations nevertheless have a performative effect on people, that is, might induce them to decide in line with their results. General speaking, AI's scientific roots encourage its users to trust and follow the probabilities calculated by the AI system, since 'automation bias' is higher when the system embodies a degree of scientific aura.

Finding that AI systems contribute to innovation in the search for evidence. They can quickly analyze big data and extract information that can be useful to investigators. AI systems can establish correlations between pieces of information that are invisible to the human eye. The crime analysis diagrams they produce can be highly valuable, elaborate information for investigators.

Finding that AI systems can produce information proffered as evidence for use in criminal trials. In particular, AI systems can provide forensic information by comparing biometric traits (e.g., facial images in facial recognition), the sound frequencies of different voices (vocal recognition), and DNA fragments (probabilistic genotyping).

Finding, lastly, that AI-assisted robots and 'smart' objects in various occupational areas and daily life may incidentally produce clues or evidence that may turn out to be useful in establishing facts in a criminal case.

***

Finding that despite significant progress in the past few years, AI systems are not completely reliable. Errors may be due to the poor quality of the data used or to how the algorithm is programmed or to the existence of false positives/negatives in correlations. The probabilities produced by an AI system may therefore be inaccurate.

Finding that the results produced by AI systems are not always entirely neutral. The accuracy of the probabilities calculated by AI systems depends not only on the quality of the collected and processed data, which may reflect bias, but also depends on how the systems have learned (unsupervised vs. supervised learning). Because they reproduce human decisions, self-learning algorithms are influenced by human foibles. One result is xenophobic,[2] racist, misogynist, etc. algorithms.

Finding that AI systems pose transparency problems. So-called 'black box AI' is so opaque that even specialists cannot determine how it arrives at its results. Even scientific experts cannot fully explain a system's reasoning to a court.

Considering that AI systems used in the field of criminal justice may be developed by the private sector. Such systems are products to be sold and must be profitable. The companies that develop them generally invoke trade-secret protection to refuse to reveal their algorithm's source code, without which the system's functioning cannot be properly analyzed.

***

Observing that not everything that is technologically possible is socially desirable. In a democracy, the political choices affecting the prevention, detection, investigation, and punishment of criminal offenses must be reflected in a law or a norm of equivalent binding force.

Reiterating that human rights must be fully protected when preventing, detecting, investigating, and punishing offenses, including when technological innovations are used in that context. Whereas AI often raises issues of privacy and personal data-protection law as well as the law of non-discrimination, all laws protecting human beings, in particular their freedom and dignity, as well as all the guarantees of a fair trial, including the presumption of innocence, are potentially threatened by the use of technologies that simulate human intelligence.

Considering that national laws and international and/or regional legal norms can set out the terms on which AI-related technological innovations may be allowed to contribute to the administration of criminal justice.

Reiterating that the ethical standards often referred to by the private sector do not have the same binding force as law.

***

[2] https://www.amnesty.org/en/documents/eur35/4686/2021/en/

Aware of:

- the *Recommendation of the council on artificial intelligence*, Organization for Economic Co-operation and Development, 22 May 2019, C/MIN(2019)3/FINAL;
- the *Recommendation on the Ethics of Artificial Intelligence*, United Nations Educational, Scientific and Cultural Organization, 23 November 2021, SHS/BIO/PI/2021/1;
- the *European ethical Charter on the use of Artificial intelligence in judicial systems and their environment*, European Commission for the Efficiency of Justice, Strasbourg, 3-4 December 2018;
- the *European Parliament Resolution of 6 October 2021 on artificial intelligence and criminal law and its use by the police and judicial authorities in criminal matters*, document 2020, 2016 (INI);
- the resolutions of the XIXth International Congress of Penal Law: *Information society and penal law*, Rio de Janeiro, 2014.

Resolutions

1. Use of AI systems by public authorities for assistance when preventing, detecting, or investigating criminal offences must be authorized in advance by a law or a norm of equivalent binding force.[3]

2. States must ensure that the decisions taken by authorities to focus on preventing, detecting, or investigating a particular type of crime is based on politically and democratically determined criteria rather than on the assumption that using AI technology will make it easier to prevent, detecting or investigating this type of crime.

3. To protect the legitimacy of the public authorities' activities preventing, detecting, and investigating criminal offences, states that wish to use AI systems must choose systems the functioning of which is fully transparent, explainable, and traceable (white box AI). They must ensure that intellectual property objections cannot be raised when seeking transparency, and they should prefer publicly available, open-source systems.

4. Laws or equivalent norms related to using AI systems in the prevention, detection, and investigation of criminal offenses must require that such systems have a high degree of technological reliability. A sufficiently precise regulation requiring appropriate verifications and evaluations, both external to and independent of the AI system's developer and provider, must limit to the greatest possible extent the risk of bias or any form of discrimination in machine learning, coding errors, and other technological malfunctions.

[3] Below we will shorten "norm of equivalent binding force" to "equivalent norm."

5. Laws or equivalent norms must require that AI systems used to assist in the prevention, detection, and investigation of criminal offenses be fully accessible, verifiable, and auditable by authorities that use them and by authorities that are in charge of verifications and evaluations.

6. Laws or equivalent norms authorizing the use of AI systems to assist in the prevention, detection, and investigation of criminal offenses must require that the training data be of high quality and representativeness.

Concerning data from police or judicial files, laws or equivalent norms must institute a system that ensures that such data are correct and up-to-date and that their use does not infringe the presumption of innocence. The presumption of innocence strictly prohibits the retention and use of data collected in response to the outcome of a predictive assessment when there is no subsequent finding of guilt, except if the data have relevance concerning another suspect.

As regards other data, in particular data accessible on social media, laws or equivalent norms must require compliance with the right to privacy and with personal data protection law when using such data. Appropriate verifications, independent of the police and judicial institutions, must be undertaken.

In general, laws or equivalent norms must be highly demanding with respect to the verification of the reliability of all data used by AI systems in connection with detecting, preventing, and investigating criminal offenses.

7. Laws or equivalent norms must require that before an AI system based on self-learning algorithms may be used in preventing, detecting, or investigating criminal offenses, the algorithms must be developed, trained, tested, and deployed under human supervision (human-in-the-loop machine learning).

These laws and equivalent norms must require a human evaluation before any action is taken to prevent, detect, or investigate criminal offenses based on the probabilities calculated by an AI system.

8. States and law enforcement authorities must ensure that their personnel who use AI to prevent, detect, or investigate criminal offenses receive hands-on training in the proper use of the relevant AI system, as well as training with respect to the risk of error and bias. They must ensure that such personnel have a thorough knowledge of the dangers AI may pose to human rights.

9. International, regional, national, or local authorities must establish independent bodies certifying the quality of AI systems intended to be used in preventing, detecting, or investigating criminal offenses. AI technology that cannot be operated and supervised in a transparent way, due to, inter alia, intellectual property rights, must not be certified.

The private sector should organize or unite to create AI-system quality labels with the goal of creating a virtuous circle for these products so that the authorities working to prevent, detect, or investigate criminal offenses are better able to determine which AI systems meet their needs.

10. All human rights must be protected when AI systems are used in preventing, detecting, or investigating criminal offenses. States and regional and international bodies must ensure that effective, proportionate, and dissuasive sanctions are imposed when such rights are violated.

Laws or equivalent norms must explicitly provide that where the cause of the violation of human rights is the technological malfunction of an AI system, the company that created the system will incur liability for fault or negligence or based on strict liability for defective products. They must also provide that investigations must be carried out to determine the cause of the violation.

11. All present resolutions are also applicable to preventing, detecting, investigating, and sanctioning administrative offenses by the competent authorities.

Resolutions specific to 'predictive policing'

12. States and regional and international human rights bodies must ensure that the use of AI systems in preventing and detecting criminal offenses does not lead to mass surveillance, which would result in a disproportionate reduction of individual freedoms (freedom of movement, freedom of expression, freedom of assembly, freedom of association, and freedom of religion).

In particular, states and local authorities must prohibit the use of AI systems to remotely identify individuals in publicly accessible spaces on the basis of their biometric data, as well as any other uses of AI systems that enable mass surveillance.

States are urged to be more transparent about their use of automated number plate recognition systems in publicly accessible space. When these systems include not only the taking of a picture of the licence plate, but also the taking of a picture of any individual in the vehicle, this option must be explicitly authorized by law. Applying facial recognition technology to the data collected through these pictures must be prohibited for the purposes of 'predictive policing'. It can only happen in the context of a specific investigation if there is a legal framework for it.

13. States must determine or have independent research bodies determine whether using AI systems in preventing criminal offenses helps decrease the number of offenses committed and, if so, in what proportion.

14. States must ensure that the financial cost of AI systems and their maintenance does not deprive the public crime-prevention services working on the *causes* of crime of funds (for psychological support, social support, training, and employment support).

15. Laws and equivalent norms must strictly prohibit the use of data as inculpatory evidence in criminal proceedings where those data were collected by an AI system in connection with crime prevention, that is, where there was no concrete suspicion that an offense had been committed and therefore the data were collected outside the scope of the legal framework governing criminal investigations.

If data collected by an AI system in the context of crime prevention are used as the basis for investigation ('starting information), in criminal investigation as starting information, the competent judicial authority must be informed of it. The data must be marked as such and the use of AI systems must be documented on the case file.

Resolutions specific to 'predictive justice'

16. Laws and equivalent norms must strictly prohibit the use of AI systems for actuarial justice purposes in sentencing.

Punishing or aggravating the punishment of someone based on the probability that they will commit a criminal offense in the future amounts to applying punishment based in part on a criminal act that has not occurred. That is contrary to human dignity, personal freedom, and fundamental principles of criminal justice.

The use of AI risk-assessment tools must be prohibited when severe security measures, such as detention, come into consideration. When states allow the use of such tools for less severe measures, the law must expressly authorize it, with sufficient procedural safeguard. However, AI probabilities cannot constitute the only basis for a decision.

17. States that wish to use AI to assist prosecutors or/and judges with quantitative legal analysis before taking decisions in criminal cases must limit use of this technology to minor offenses that represent a high volume of cases.

18. Before deciding to use AI to facilitate management of a high volume of cases involving minor offenses, states must assess whether it would be appropriate, in light of the *ultima ratio* principle, to decriminalize the conduct generating such cases.

19. Laws and equivalent norms must prohibit the use of quantitative legal analysis for assisting judges when ruling on guilt.

20. Laws and equivalent norms must prohibit the use of quantitative legal analysis for assisting judges with sentencing. The decision to punish a person and the type of sentence must be made by humans. Otherwise, justice may be dehumanized and people's human dignity may be threatened.

21. Laws and equivalent norms must prohibit the use of quantitative legal analysis for assisting judges with decisions in criminal matters that are issued before judgment and that involve coercive measures.

22. States must ensure that decisions taken with the assistance of quantitative legal analysis do not infringe the right of access to a human judge.

23. Laws and equivalent norms must prohibit the assistance of quantitative legal analysis unless the decision can be appealed by the person concerned. The decision at appeal level shall not be based solely on the quantitative legal analysis.

Resolutions specific to evidence gathered and/or produced by AI

24. Laws and equivalent norms on extracting data for analysis by an AI system must require that before asking a person for the access code of her/his software or hardware from which data may be extracted, the seizing authority must inform the person concerned of their right not to incriminate themselves.

25. Laws and equivalent norms on crime analysis must specify that the crime analysis diagrams produced by AI systems do not have probative value, but may serve as a guide for conducting investigation.

26. Laws and equivalent norms on using AI systems to gather evidence or produce information for criminal justice purposes must clearly indicate that the output of AI systems are only probabilities. They must require that all probability-based judgments indicate not only the probability calculated by the AI system that was used, but also the error rate of that system, as calculated by the certification body that evaluated it.

27. States and judicial authorities must ensure that the use of AI-calculated probabilities does not lower the existing standard of proof in criminal proceedings.

28. Laws and equivalent norms on using AI systems to gather evidence or produce information for criminal justice purposes must prohibit the use, as evidence, of probabilities calculated by AI systems that are not fully explainable (black box AI).

29. Laws and equivalent norms on using AI systems to gather evidence or produce information for criminal justice purposes must require, pursuant to the right to adversarial hearings, that, if data collected or produced by an AI system are used, all parties must be informed of it. The data must be marked as such and the use of AI systems must be documented on the case file.

Laws and equivalent norms must require that a party's production of an AI-calculated probability may be challenged by the other party.

30. Laws and equivalent norms must set forth the principle that the party producing the probability in court must systematically include complete information on how the AI system works and which data it uses.

31. Laws and equivalent norms on using AI systems to gather evidence or produce information for criminal justice purposes must, consistent with defense rights, provide that anyone accused of an offense based on a probability proffered as evidence be able to obtain the AI system's source code and training data so that these may be analyzed by an expert. Trade secret must not be allowed to impinge on defense rights.

32. Due to the high cost of obtaining an expert analysis of an AI system, states must ensure that anyone accused of an offense based on a probability calculated by an AI system have access not only to effective legal aid but also to financial aid for such specific expertise.