Revista Científica
del Centro Universitario
de la Guardia Civil

Revista Lógoς Guardia Civil

Carlota Cuatrecasas Monforte
PhD in Law (International Major)
Judge at the Tarrasa 1st Criminal Court

ARTIFICIAL INTELLIGENCE AND CRIME INVESTIGATION

ARTIFICIAL INTELLIGENCE AND CRIME INVESTIGATION

1.- INTRODUCTION

1.1.- The use of Artificial Intelligence (AI) in the field of Justice: an approach

Bill Gates, the technologist par excellence of our era, uttered what is undoubtedly one of the best and most famous phrases reflecting the reality of the technological explosion we are experiencing today, which of course has its pros and cons: "*The first rule of any technology used in a business is that automation applied to an efficient operation will magnify the efficiency. The second is that automation applied to an inefficient operation will magnify the inefficiency*".

And such a statement, loaded with meaning and force, and applicable no doubt to the use of AI both in the private and public sector, should serve as a basis for any decision related to this technology. Because although many see it as a new potential saviour, the truth is that it has enormous potential to have an opposite effect – so it should be used with great caution.

AI is obviously a technological tool that is more than capable of enhancing the efficiency of human tasks; that is an indisputable fact. But it also has a huge potential to magnify their inefficiency and even perpetuate this, which is extremely dangerous, especially in the field of justice.

Historically, one of the greatest challenges facing societies has been to achieve true justice. This question, however, has not been and never will be resolved in a uniform and homogeneous way, given the great diversity of cultures and civilisations that have coexisted, do coexist and will coexist on our planet, and which have different values and different ways of conceiving and understanding what is just and what is unjust and, of course, how this should be managed.

Thus, there are undoubtably as many definitions of justice as there are people in the world, and this is precisely what, in my view, created a need for humanity to make social pacts in order to live together in peace; given how unsustainable and brutal it is for humans, as rational beings, to live without establishing common rules for coexistence that, shared by the majority, guarantee previously established individual and collective rights – as supported by the social contract theories developed by Hobbes, Locke and Rousseau.

And today such social pacts, existing to a greater or lesser extent and in different formats in practically all parts of the world, have a different substance in each territory. However, despite this great diversity, there is a specific common premise: the need for conflicts to be resolved within a reasonable timeframe and with reasonable speed. As the philosopher Seneca warned thousands of years ago, delayed conflict resolution can lead to enormous injustice, for "*nothing resembles injustice so much as delayed justice*".

In Spain, a democratic state governed by the rule of law and by virtue of the provisions of Article 117 of the Spanish Constitution, justice comes from the people and is administered in the name of the King by judges and magistrates who are members of the judiciary, one of the three branches of government.

And, specifically, all of the cogs that make the judicial machinery work in Spain, as is well known, are incorporated in the Administration of Justice, which is undoubtedly full of efficiencies that should be magnified; and also of course with inefficiencies that, on the contrary, should be reduced as far as possible or eliminated. So the use of AI must be tightly controlled in this area and, above all, by legally endorsed guarantees of success, as will be seen below.

Unfortunately, the service that the Administration of Justice has provided to citizens has not been as expected, meaning that the Administration of Justice has been and continues to be the worst rated by Spanish citizens[1], which is unacceptable. Above all, it is unacceptable because justice is the last and only option for thousands of people to resolve conflicts that influence the most important aspects of their lives and prevent them from living with the peace that every human being deserves.

Above all, Spanish justice is slow – sometimes extremely slow – and this is something that clearly needs to be improved. And I understand that progress on this issue means looking for possible solutions, and analysing what options exist today to evolve and improve the quality of the service provided by the Administration, which should be the main aspiration (and is, of course, the main responsibility) of every public servant.

And what better way to look for ways to progress than to explore all the possibilities offered by new technology available to us today and, in particular, AI, which is still a technology that has been little explored in the field of justice, above all because it is often unprofitable to invest in research for its use in public service.

Specifically, I would also like to point out that, although the quality of service offered by the courts should be the highest in any jurisdictional hierarchy and at any procedural stage, I understand that in the pre-trial phase of criminal proceedings the rights of citizens require special protection, since it is a very early stage in the investigation of alleged criminal acts, with information generally scarce and unclear; and with the risk that this entails both for the victims (and, eventually, for society in general) who need protection, and for the persons investigated as alleged perpetrators, since in all cases their right to the presumption of innocence must prevail. This conflict of rights is not usually easy to

[1] See, among others, the survey on Public Opinion and Tax Policy conducted by the Centro de Investigaciones Sociológicas (CIS) in July 2021. Last visited on 22 March 2023.
https://datos.cis.es/pdf/Es3332marMT_A.pdf

manage, especially with very limited material and human resources, both police and judicial.

1.2.- This implies the need for AI use to be regulated, particularly in the field of Justice.

Therefore, the benefits that AI could bring to improve quality and efficiency in the aforementioned criminal investigation process – as we will see – are evident, although they cannot be analysed in a responsible manner without weighing them against the possible risks, which could potentially detract from the benefits.

The good news is that these risks can be minimised and even eliminated through regulation; the law plays a fundamental role in the field of AI, as it is a key tool for setting limits and guaranteeing that AI is used to create value in society, by applying the provisions of article 18.4 of the Spanish Constitution.

Thus, it is mainly up to the legislative authorities to establish clear and comprehensive regulatory bases that ensure appropriate use and guarantee that AI is used to improve the quality of justice, specifically criminal justice, which falls within the scope of the obligations that Article 9.2 of the Magna Carta imposes on public authorities when it declares: "*It is incumbent upon public authorities to promote the conditions for real and effective freedom and equality of the individual and of the groups to which he belongs; to remove obstacles that prevent or hinder their full realisation and to facilitate the participation of all citizens in political, economic, cultural and social life*".

And, given that AI is a technology that knows no borders and must therefore be studied from a global perspective, this regulation should, in my view, come from two perspectives:

- On the one hand, a transnational perspective, through an international institution or organisation, which should set common minimum requirements for all countries in the world (or as many as possible) to guarantee the use of AI for the benefit of humanity, if possible in a legally binding way, in order to stop individuals from acting in their own interests and potentially harming the common good.

On the one hand, I understand that at the international level, the institution that would be best positioned to establish basic principles or minimum common standards on the use of AI worldwide would be the United Nations (UN), given that it has the membership of practically all countries in the world (a total of 193 member states), and has already adopted regulations of a similar nature in the past for the common good, namely the Universal Declaration of Human Rights, adopted and declared by the General Assembly in its resolution 217 A (III) of 10 December 1948, which undoubtedly marked a milestone in the history of humanity.

Thus, I believe that such a standard could be set out in a Universal Declaration of AI Principles, which, in my opinion, should contain at least the following principles:

- *Principle of respect for human dignity, with guaranteed human oversight and control of AI systems and priority given to social and environmental well-being;*
- *Principle of respect for the freedom and privacy of the human being, with a guarantee of individual management of personal data;*

- *Principle of transparency and comprehensibility of systems;*
- *Principle of fairness, equality, non-discrimination of human beings and inclusion;*
- *Principle of robustness, technical soundness and safety;*
- *Principle of accountability.*[2]

However, in the case of basic AI legislation, in my view, it is necessary to go a step further. This technology has the ability to escape human control, which could prove to be a major threat to our species. In this sense, as early as 1970, Brad Darrach, paraphrasing one of the fathers of AI, Marvin Minsky, said "*If humans are lucky,* they (machines)*may decide to keep them as pets. If they are unlucky, they will be treated as food*".[3]

Therefore – although this sounds somewhat catastrophic and should be taken with some caution – I believe that it would be absolutely essential that any basic or minimum regulations adopted at international level on AI should be legally binding, in order to avoid possible misconduct by rogue Member States with particular interests that could lead humanity to a point of no return; since it is clear that if AI were to leave human control, the impact would be global.

Furthermore, in order to guarantee compliance with such regulations, I believe that strict control mechanisms should be established, providing strong and punitive sanctions for those who breach them, since, as I have said, the damage caused by these could be irreparable; and, on the other hand, a national perspective, through the legislative powers of each State (or group of States, as in the case of the European Union), which should establish more specifically and in greater depth the methods of usage and limits of AI within their respective jurisdictions, while respecting the basic principles established at the international level.

With regard to AI, it is important that we all move in the same direction, despite the numerous interests (especially economic and political) that exist around it – because its potential for harm, as has been said, is real and potent, and it is extremely dangerous for everyone.

At present, however, unfortunately, we have neither basic regulation nor more specific regulation, at least in Spain and the EU, which is very disappointing.

Thus, on the one hand, with regard to basic regulations – although there have been numerous initiatives (both public and private) to establish the basic principles that should govern the use of AI – the truth is that none of them have the scope to which I have alluded.

However, appreciation should be shown for the recommendation adopted by UNESCO at its 41st General Conference held from 9 to 24 November 2021, consisting of the publication of a Recommendation on the Ethics of AI, due to the large number of

[2] Cuatrecasas Monforte, C. "*La Inteligencia Artificial como herramienta de investigación criminal (Artificial Intelligence as a criminal investigation tool)*". Editorial La Ley. 2022. Pág.50.

[3] Darrach, B. (1970). "Meet Shaky, the first electronic person: The fascinating and fearsome reality of a machine with a mind of its own". *Life.* Pág.66.

signatory states; although it lacks legally binding character, being a mere recommendation, which is clearly insufficient for the reasons already explained.

On the other hand, with regard to specific regulations, it is surprising that the European Union – which is usually extremely careful to respect the rights and freedoms of its citizens – in this case has arrived “late and badly”, despite the responsibility it bears as the largest single market in the world (and, in fact, it has not yet arrived, since the final text of the European AI Regulation has not yet been published). As a result of this lack of regulation, AI systems with the potential to violate fundamental rights on a massive scale have been used in some EU Member States in recent years, which many human *rights* organisations have denounced and is absolutely intolerable.

It is true that on 21 April 2021 the European Commission published the Proposal for Regulation of the European Parliament and the Council, establishing harmonised rules in the field of Artificial Intelligence (Artificial Intelligence Act) and amending certain legislative acts of the Union. And, subsequently, the Council of the European Union, in session nº3917/1 held on 5 and 6 December 2022, adopted its position (“general approach”) on the aforementioned proposal; although the publication in the coming months of the European Regulation on Artificial Intelligence, as I have already mentioned – which will undoubtedly mark a before and after in the use of AI in all areas, including criminal investigation and justice – is still pending.

Whilst awaiting the final text of the European Regulation on AI, I am particularly critical of the content of the European Commission's proposal, since in my opinion, the quality control mechanisms for AI systems that it provides for are clearly insufficient in terms of criminal investigation tools.

The aforementioned Proposal only imposes active public control in the case of systems that use biometric data for the identification of natural persons in real time or remotely, but this is not required for the rest of the AI tools whose use is permitted for criminal investigation purposes (articles 43 – ordinary procedure – and 47 – urgent procedure –), despite the fact that the vast majority of these are classified as high risk by the Proposal itself, which is not considered logical.

My suggestion in this respect is to go further, and I believe that the key to successful use of AI in the EU for criminal investigation purposes lies in the establishment of strong and comprehensive quality control mechanisms *ex ante* (prior to circulation) and *ex post* (to occasionally check that quality standards continue to be met, similar to motor vehicle checks like the Technical Inspection of Vehicles – TIV).

To this end, I believe that a good solution would be the creation of a European AI Agency, which would centralise the filtering and quality control of systems intended to be used within our borders (similar to the European Medicines Agency), with a central headquarters and – if required – delegations in the different Member States.

I believe that, whilst AI systems must always be subject to rigorous quality control, in the case of the public sector it must be completely exhaustive, given that in the private sector – if a company uses algorithms that later turn out to be not very transparent, for example – the user can decide to go to a competitor; but if the public sector uses AI systems to make decisions that may affect citizens, they cannot opt out or choose not to have the results imposed on them, so quality standards must be even higher if possible.

I understand that the European AI Agency to which I have referred would give a great deal of confidence not only to citizens – who need guarantees that the AI systems used, in this case by the authorities, are transparent, explainable, secure and contain quality data, in order to avoid infringements of fundamental rights – but also to the holders of the million-dollar patents that host them.

However, in the aforementioned Proposal of the European Commission, the most similar to the above is the creation of a European AI Committee (in its Title VI, Articles 56 to 58), which will ensure compliance with regulations established by the future Regulation on AI, and promote the adoption of new measures, which is a good solution but in my opinion insufficient.

Spain, in particular, has moved faster than the publication of the aforementioned European Regulation on AI and has already announced that it will have a Spanish Agency for the Supervision of Artificial Intelligence (AESIA), which will carry out supervision and risk minimisation tasks and seek to generate an AI research and business ecosystem[4], whose physical headquarters will be in La Coruña.

Although this is good news, given that it implies the creation of a body to control the use of AI in Spain, the fact is that it does not obey any joint action plan led by the EU by virtue of previously adopted joint legislation (which has not yet been approved). So there is a risk that, following the publication of the aforementioned European AI Regulation, Spain's plans will have to adhere to a different approach, given that what would be really effective would be for the EU to host the headquarters of a potential European AI Agency that would assume truly proactive powers (many more and with greater impact than those contemplated in the European Commission's Proposal) and, if necessary, decentralise some functions.

In view of the above however, it should be noted that even if there were basic or more specific regulations, I must warn the authorities that when using AI systems they should take the utmost caution and exercise the greatest possible diligence and responsibility, given the dangers involved (just as a doctor prescribing a medicine or a pharmacist dispensing it today, however approved its use may be by the European Medicines Agency).

In any case, it is important to underline that AI legislation should always go hand in hand with personal data protection legislation. Such systems rely on data – huge amounts of data – which have been described as the oil of the 21st century, and to a large extent the success or failure of AI systems depends on the quality of that data.

Thus, if poor quality data (abnormal, illegally obtained, processed non-compliantly, incorrect, etc.) is fed into the system, the results produced by the AI system that was

[4] See press release published on 13 September 2022 by the Ministry of Economic Affairs and Digital Transformation and the Ministry of Territorial Policy as part of the Digital Spain 2026 programme. Last visited on 18 March 2023.
https://portal.mineco.gob.es/RecursosNoticia/mineco/prensa/noticias/2022/20220913_ndp_sede_agencia_ia.pdf

trained with it will also be poor, which is known as the phenomenon: "*garbage in, garbage out*" and should obviously be avoided.

However, awareness is growing, and AI systems not only have more sophisticated and powerful *software* , but also use better quality data for training, both of which are essential for them to function optimally and correctly – the minimum expected, especially when used by the public sector.

Although it may seem simple, this is a complicated task, and the systems we have today (although they are advancing by leaps and bounds) still have a long way to go to reach acceptable levels of quality (which, in the case of use by the public sector, must be close to infallibility), which can sometimes lead to situations that could potentially violate rights and freedoms.

2.- AI FOR CRIME INVESTIGATION

That said, on the one hand, I would like to highlight the AI tools that could prove most useful for improving the efficiency of the Spanish criminal investigation process and, on the other, I would also like to highlight the most common potential risks that I have found in them and which, therefore, should be avoided.

In order to standardise and organise the information and facilitate the analysis of each of the aforementioned tools, I will classify them under three main headings:

1) Risk assessment and prediction tools

2) Crime investigation tools

3) Processing tools

a) Risk assessment and prediction tools

A.1. What are they?

Algorithms, as such, obviously cannot predict future events, but they can certainly estimate the probability of something happening based on existing data.[5]

Risk assessment and prediction tools are systems that use AI to analyse historical data and predict future behaviour and events, such as: where, when and by whom a crime is most likely to be committed; whether or not a defendant is a flight risk or a repeat offender; whether or not an inmate will be re-entering prison after a release; or whether or not a company will dispose of its assets after a lawsuit is filed against it.

And such systems can undoubtedly be of immense value in assisting law enforcement and judicial authorities in making important decisions, such as: allocating more or fewer police officers to a particular area and/or at a particular time; entering people into social reintegration programmes; ordering pre-trial detention measures;

[5] Waldman, A. (2019). "Power, Process, and Automated Decision-Making". *Fordham Law Review* (88). Page 5.

granting or denying prison sentences; or, where appropriate, imposing bail or ordering a seizure.

A.2. Areas of application

Today, these tools are used widely by the police, but there is more reticence in the judicial system despite the fact that, as we will see, they are already used in various judicial systems around the world, such as in the United States (USA).

However, the truth is that judges and magistrates often take into account the risk information provided by the police as another element to support their decisions and, in the event that such assessments have been carried out using AI systems for risk prediction and assessment, these would be applied, indirectly, by the aforementioned judicial authorities.

A.2.1. Police

The main priority of predictive policing systems is to optimise resources and thus improve the efficiency of police crime prevention tasks.

This is done by analysing, through AI, historical data held in police databases, and then creating risk scales that determine the likelihood of a certain behaviour or event occurring. This provides invaluable information that would allow, for example, increased surveillance in geographical areas and time slots classified as "hot", or of people classified as being at greater risk of crime or more vulnerable (i.e. more likely to be a victim of crime).

In this regard, and in order to highlight the differences between traditional policing (very reactive) and predictive policing, the European Union Agency for Fundamental Rights drew up the following (in my view, very illustrative) comparative table in 2019:

	TRADITIONAL POLICING	PREDICTIVE POLICING
Context	Offence committed or alert raised on a particular person	No crime committed and no alert raised on a particular person
Approach	Reactive	Proactive
Target	Arrest suspect(s)	Anticipating where and when crimes may be committed or by/against whom
Data used	Specific information related to the case	Generic information relating to several cases
Type of process	Data-driven processes and human processes are combined	It focuses mainly on processes based on large amounts of data.

Specifically, these predictive policing systems focus on three main areas:

- Crime prediction (or crime mapping), to predict where – geographic areas, more or less narrowly defined – and when – seasons, months, days, time slots or "rush hours", etc. – there is a higher risk of crime;

- Criminal identity prediction, identifying potential future offenders through criminal profiling, usually based on past circumstances and behaviour; and

- Vulnerable identity prediction, identifying potential individuals or groups of individuals who are more likely to become victims of crime in the future.

And all of this can undoubtedly be of enormous value in increasing the efficiency of policing and thus improving the quality of policing, which should undoubtedly lead to a reduction in crime – and that is what we should be aiming for.

Thus, it is clear that if resources are invested in optimising preventive police tasks using AI (capable of analysing and cross-checking huge amounts of data), and this is done in close collaboration with other agencies – i.e. social services, prison services, prosecution, courts, etc. – with the ability to promote and adopt the necessary and appropriate complementary measures to detect certain risks, the results obtained could represent a great advance towards social harmony, compared to reactive policing tasks (or even traditional preventive tasks but based only on – limited – data comparisons carried out by humans).

As I have already mentioned, predictive policing systems are already used by police forces in many countries around the world: in Europe (including Germany, Italy and the United Kingdom); in North America (Canada and the United States, including the states of California, Illinois and New York); in Latin America (including Chile); and in Asia (including Singapore).

In Spain, there is currently no official record of the use of AI tools for predictive policing, although the VioGén risk assessment system used by the State Secretariat for Security of the Ministry of the Interior, in compliance with the provisions of Organic Law 1/2004, of 28 December, and the Comprehensive Protection Measures against Gender Violence, which was put into operation on 26 July 2007, should be highlighted[6].

This tool – which in theory is understood to not use AI technology, but is merely an actuarial system – establishes a scale that determines the level of risk (negligible, low, medium, high and extreme) that a victim of gender-based violence and her children have of suffering new attacks by her aggressor.

To this end, although the Ministry of the Interior states that recidivism of male violence has decreased by 25% since the implementation of this tool,[7] the fact that this system is not publicly audited and or sufficiently transparent has led the Eticas

[6] http://www.interior.gob.es/web/servicios-al-ciudadano/violencia-contra-la-mujer/sistema-viogen

[7] See news published in La Vanguardia on 19 May 2019. *Algorithms against male violence.* Last visited on 24 March 2023. https://www.lavanguardia.com/tecnologia/20190519/462147339117/algoritmos-violencia-machista.html

Foundation, in collaboration with the Ana Bella Foundation, to publish an external audit report on 8 March 2022 with several controversial points to be examined.[8]

A.2.2. Judicial

The main priority of predictive justice systems is to optimise resources and thus improve the efficiency of legal proceedings in order to assess future risks, which is always a delicate task.

This is done by analysing historical data held in judicial databases using AI, and then creating risk scales that determine the likelihood of a certain behaviour or event occurring, which provides invaluable information allowing, for example, precautionary measures to be granted.

In the judicial sphere, this type of tool is used both in the *pre-trial* phase (to assist the judge in adopting precautionary measures), and *pre* and *post-trial* (to assist the judge in determining the dangerousness of the accused/prosecuted individuals, and analysing recidivism risk in those already convicted, for the purpose of possibly suspending the execution of sentences, granting of prison leave, etc.).

In contrast to the widespread use of risk prediction and assessment tools in the police field, as previously mentioned, this is not the case in the judicial field.

However, there are certainly jurisdictions where such tools are already widely used; for example in the US, where they are used in several states and counties, as reported by the Media Mobilizing Project in Philadelphia (Pennsylvania, USA) and MediaJustice in Oakland (California, USA), which developed a national database[9] that determines in which jurisdictions AI risk assessment tools are used in the US and what they use (e.g. COMPAS and PSA).

In this regard, it is interesting to highlight the aforementioned COMPAS system, an AI tool that uses *Machine Learning* technology in its operation, used by judges in certain US states (including New York, Wisconsin and California), which predicts the risk of an individual committing new crimes in the future.

However, an investigation by the ProPublica agency back in 2016 found that the algorithm used by the tool (which, by the way, is not public and remains hidden) contained bias against certain groups (including black people)[10], which is, of course, absolutely unacceptable.

In any case, the use of tools such as these, with the capacity to analyse and cross-check huge amounts of data as shown above, could be very useful to assist (not replace) judges in their task of assessing future risks; although, as will be seen below, this would only be legally viable if the fundamental rights of the persons concerned are respected

[8] See full content at https://eticasfoundation.org/wp-content/uploads/2022/03/ETICAS-FND-The-External-Audit-of-the-VioGen-System.pdf. Last visited on 28 March 2022.

[9] See https://pretrialrisk.com/national-landscape/where-are-prai-being-used/ Last visited on 12 March 2023.

[10] See Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). *Machine Bias*. ProPublica. Last visited on 17 March 2023.
https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

and absolutely guaranteed.

b) Crime investigation tools

B.1. What are they?

AI crime investigation tools are AI-powered systems that can be used by law enforcement, prosecutorial and judicial authorities to investigate and record crimes committed, including all circumstances that may influence their classification, as well as the guilt of offenders, in accordance with Article 299 of the Criminal Procedure Act.

However, there is no specific scientific category or classification in the field of AI that brings together all the aforementioned tools, although I consider it appropriate to carry out a joint study to group together all those systems that have similar characteristics and practical functions that are suitable for criminal investigation, in order to facilitate their analysis.

B.2. Classes

B.2.1. Tools that use biometric data

Biometric data are defined, inter alia, at the European level, in Article 4.14 of the General Data Protection Regulation (GDPR) which states that they are "personal data resulting from specific technical processing relating to the physical, physiological or behavioural characteristics of a natural person, which allow or confirm the unique identification of that natural person, such as facial images or dactyloscopic data"; and at the national level, in Article 5. l of Organic Law 7/2021 of 26 May on the protection of personal data processed for the purposes of the prevention, detection, investigation and prosecution of criminal offences and the execution of criminal penalties, which provides that they are "personal data obtained from specific technical processing, relating to the physical, physiological or behavioural characteristics of a natural person which enable or confirm the unique identification of that person, such as facial images or dactyloscopic data".

Thus, biometric data, by definition unique and non-transferable for each individual, are of enormous value for identification purposes, in particular in the field of criminal investigation.

And AI systems already exist today that use such data to perform the following functions:

- On the one hand, identifying and answering the question: who is this individual?; and

- On the other hand, checking or verifying an identity, i.e. to answer the question of: Is this individual really who they claim to be or who they are suspected of being?

In order to carry out such tasks, AI systems compare images of suspects, which are entered by the authorities, and verified images, which are found in vast police or judicial databases, and if there is a "*match*" between them, they trigger an alert that yields positive and very useful results.

B.2.1.1. Facial recognition

In terms of concept, facial recognition is a technology that allows the identification of people or verification of their identity through the shapes, proportions and features of their face, using AI.

In terms of potential uses, a good facial recognition AI system could help to identify or verify the presence of certain criminals in certain places, and could even replace (or complement), for example, the current line-up process.

Moreover, there are already applications that allow the profile of a suspect to be drawn up from a simple photograph in a matter of minutes (such as Clearview, although there are many controversies about its use[11]); and in China, for example, the use by police officers of so-called "smart glasses" – which have built-in facial recognition systems capable of identifying citizens with outstanding police warrants in real time while on patrol[12] – is already widespread, which could be highly effective in locating persons with arrest warrants, amongst others.

On the other hand, by incorporating facial recognition systems in video surveillance cameras, it would be possible to detect the entry into certain public places (i.e. the metro, a certain town, etc.) of those persons subject to precautionary measures or sentences prohibiting them from approaching/entering – and a direct warning could even be sent to the security forces – and this, of course, would not only help in preventing crimes involving breaches of precautionary measures or sentences (and other associated crimes), but would also greatly facilitate the investigation of such crimes in the event that they were to be committed.

Finally, among many other uses, facial recognition technology could also be used by forensic doctors to carry out identification of corpses using craniofacial superimposition techniques, which could be of great help in identifying victims, for example, in more complex cases.

B.2.1.2. Voice recognition

As a concept, voice recognition is a technology that enables people to be identified or their identify checked/verified through their speech by means of AI.

In terms of potential uses, utilising a good voice recognition AI system could help to identify unknown individuals engaging in conversations during a wiretapped phone conversation, and also to verify whether or not the voice in a recording corresponds to a certain person relevant to the investigation (e.g. the alleged perpetrator of threats), to give

[11] Amongst others, see the story published in La Vanguardia on 25 May 2022. "*Clearview AI: you must not sell portraits of British citizens in your database*". Last visited 25 March 2023.
https://www.lavanguardia.com/tecnologia/actualidad/20220525/8288861/golpe-realidad-clearview-ai-podra-extraer-vender-rostros-britanicos-pmv.html
And also the story published in The New York Times on 20 January 2020. "*The mystery company that could end privacy as we know it*."
https://www.nytimes.com/es/2020/01/20/espanol/negocios/clearview-reconocimiento-facial.html

[12] See the story published in The Wall Street Journal on 7 February 2018. "*Chinese Police Add Facial-Recognition Glasses to Surveillance Arsenal*". Last visited on 22 March 2023.
https://www.wsj.com/articles/chinese-police-go-robocop-with-facial-recognition-glasses-1518004353

a few examples.

B.2.1.3. Emotion recognition

As a concept, emotion recognition can be described as technology that allows certain feelings, intentions or moods to be detected using AI.

In order to be more effective, these types of systems generally combine facial and voice analysis of individuals and are able to detect micro-expressions and vocal tones or nuances that are virtually imperceptible to humans.

In terms of potential uses, utilising a good emotion recognition AI system could help to detect with high accuracy whether a person is telling the truth or not when giving a statement, to cite just one example.

There are also systems known as "aggression detectors", which, based on the fact that 90% of physical aggressions are preceded by increased facial stress and verbal aggression, identify, mainly through image and voice, those cases in which a physical attack is about to take place and immediately alert the police. Some of these systems have been installed in US schools, for example, to detect assaults and attacks mainly through voice analysis[13], although they do not have good references and evaluations for the moment, given that they still have a long way to go to reach acceptable levels of accuracy (and, moreover, they raise the question of privacy, as the agency ProPublica revealed in an investigation it carried out in 2019).[14]

B.2.1.4. Fingerprint and DNA recognition

Nowadays, fingerprint and DNA analysis techniques are undoubtedly the best and most reliable techniques to identify persons or to check/verify their identity. This is already carried out using programmes that allow automated data analysis, which is of great help to human technicians, who see their task reduced to a final moment, when there are only a few candidates left to analyse, after very useful preliminary filtering by "the machine".

As a concept, fingerprint recognition can be said, on the one hand, to be a technology that enables people to be identified or their identity to be checked/verified using AI by analysing the shapes of the papillary ridges of their fingers and their proportions; and, on the other hand, DNA recognition is a technology that enables people to be identified or their identity to be checked/verified using AI, in this case, by analysing human genomes.

In terms of potential uses, utilising a good AI system for fingerprint and DNA recognition could mainly help to access and automate the analysis of huge amounts of

[13] See the news report published in El Español on 29 June 2019. "*Aggression detectors in US schools - the solution to violence?*" Last visited on 15 March 2023. https://www.elespanol.com/omicrono/tecnologia/20190629/detectores-agresiones-escuelas-eeuu-solucion-violencia/409959757_0.html

[14] See the article published by ProPublica's Jack Gillum and Jeff Kao on 25 June 2019. *Aggression Detectors: The Unproven, Invasive Surveillance Technology Schools Are Using to Monitor Students* https://features.propublica.org/aggression-detector/the-unproven-invasive-surveillance-technology-schools-are-using-to-monitor-students/ Last visit on 16 March 2023.

data, with the possibility of even making final decisions (which, in my opinion, should always be supervised by a human) and thus increase the efficiency of searches and identifications of suspected criminals, victims, witnesses, corpses and missing persons in criminal cases, and facilitate the expert work in the most complex cases.

B.2.1.5. Signature and handwriting recognition

As a concept, it can be said that signature and handwriting recognition is technology that makes it possible to identify people or check/verify their identity using AI by analysing symbols and/or handwritten signs on a physical or digital medium.

In terms of potential uses, utilising a good AI system for signature and handwriting recognition could undoubtedly serve to automate and thus increase the efficiency of current handwriting and signature analysis that is of interest in criminal cases, which today relies on time-consuming expert evidence (yet such systems produce results in a matter of minutes) and can also lead to a considerable reduction in the associated costs.

B.2.2. Tools using Natural Language Processing (NLP) techniques

NLP techniques aim to translate human language (both spoken and written) into a language that "the machine" (or, more specifically, the algorithm) can understand.

B.2.2.1. Chatbots

As a concept, a "*chatbot*" is a tool that aims to engage in an *online*human-machine conversation orally, in writing, or with a combination of both communication forms (depending on the different input and output channels), using NLP techniques.

Within this class of systems, so-called cognitive *chatbots* or "*Smart chatbots*", which use AI (specifically "*Machine Learning*") to understand natural language using the aforementioned NLP techniques, with the potential to even execute spoken or written commands, are of particular interest for criminal investigation purposes.

In terms of potential uses, utilising a good *chatbot*could undoubtedly help potential victims in danger and witnesses of flagrant crimes, for example, to establish direct and immediate contact with the police, simply by pressing a button or uttering a keyword that would give the *chatbot* the order to contact the security forces and issue certain messages (with the respective geolocation), thus reducing to a minimum the possibility of being discovered by the possible perpetrators of the crime. And it could even collect information in real time through an automatic question and answer system with such interlocutors, which could provide the police with all the data they need to make their actions as efficient as possible.

It is also interesting to refer to Sweetie[15], a *chatbot* in the guise of a Filipino girl – a minor – created by the organisation Terre des Hommes in 2013, in order to target potential paedophiles who would follow her *online* conversations of a sexual nature.[16]

[15] A Sweetie 2.0 version is now available.

[16] See the news item published on the BBC web portal on 22 December 2017 "*Sweetie: 'Girl' chatbot targets thousands of paedophiles*". Last visited on 24 March 2023. https://www.bbc.com/news/av/technology-42461065

However, I understand that this would not be legally feasible in Spain, since prospective research is prohibited.

B.2.2.2. Text/document analysis systems

As a concept, textual/documentary analysis systems are those tools that enable information contained in text formats to be analysed and processed using AI.

In this sense, on the one hand, it is interesting to make reference to the enormous utility that this type of tool could have for filtering, ordering and classifying documents relating to police and/or judicial cases, according to the needs of the authority that uses them, allowing the latter to carry out searches and filters for specific and/or related information in a matter of seconds, which would be of great value, especially in so-called macro-cases.

Particularly paradigmatic in the judicial field is the so-called "Rolls Royce case", in which an AI system was used for the first time in the European Union (as it was at that time) to analyse huge amounts of documents in a court case. In this particular case, the Axcelerate system (from OpenText) was used, which with the application of AI and NLP techniques, through "*Machine Learning*", managed to analyse around 30 million documents, 2,000 times faster than a human being and with a cost reduction of 80%.[17]

As a result of this success, the system has since been used by the Serious Fraud Office (SFO) to support their investigations.[18]

On the other hand, it is also interesting to note the potential usefulness of such systems for analysing specific texts and, for example, for possible detecting false allegations.

Particularly interesting in this regard is, as an example, the VeriPol tool created by the Spanish National Police to determine whether a report of robbery with violence or intimidation is real or false, having reached an accuracy level of over 90% (compared to 75% achieved by expert human agents, according to the Ministry of the Interior),[19] which can undoubtedly be extremely useful, not only to detect possible false crime-reporting offences, but also to avoid the unnecessary use of police and, subsequently, judicial resources.

B.2.2.3. Systems for detecting and, where appropriate, moderating onlinecontent

As a concept, *online* content analysis systems are tools that make it possible to analyse

[17] See the news article published on the UK Government's Serious Fraud Office (SFO) website on 10 April 2018. "*AI powered "Robo-Lawyer" helps step up the SFO's fight against economic crime.*"https://www.sfo.gov.uk/2018/04/10/ai-powered-robo-lawyer-helps-step-up-the-sfos-fightagainst-economic-crime/

[18] See the news article published on the BBC web portal on 4 September 2018 "*A digital game or a powerful weapon against boardroom crime?*". Last visited on 20 March 2023. https://www.bbc.com/news/uk-45399995

[19] See the news article published in El Mundo on 27 October 2018. "*VeriPol, así sabe la Policía si tu denuncia es falsa*" ("VeriPol, this is how the police know if your report is false"). Last visited on 23 March 2023.
https://www.elmundo.es/espana/2018/10/27/5bd42db7e2704e27608b466e.html

and process information found on the Internet using AI and, where appropriate, to filter and restrict its content – especially in cases where it may be criminal.

In this regard, it should be noted that, for example, the perpetrators of the shootings at the mosque in Christchurch (New Zealand) on 15 March 2019, and at the synagogue in Poway (California, USA) on 27 April 2019, posted on the internet before committing their terrorist attacks, and this often occurs prior to certain crimes being committed.

Thus, using a powerful NLP system, it would certainly be possible, for example, to identify the increased threat levels posed by a particular individual or a particular group and take pre-emptive decisions to try to prevent future criminal acts.

B.2.3. Image analysis

Image recognition is an AI application that enables certain static or dynamic figures and/or symbols to be analysed and the information contained in them to be recognised.

In terms of its potential usefulness for criminal investigation purposes, it is interesting to note the existence of companies that are already developing programmes for investigating war crimes using AI systems that analyse images contained in videos and photographs, for the purpose of pre-constituting evidence to be presented before the International Criminal Court.[20]

It is also interesting to know that large technology companies (Facebook, YouTube, etc.) already have AI systems capable of detecting and recognising sensitive images that may contain criminal content, with the immediate consequence of blocking the accounts that upload and/or share them and reporting them to law enforcement agencies.

Finally, it is important to highlight AI systems for number plate identification ("*Automated Number Plate Recognition*" – ANPR) using Optical Character Recognition (OCR) techniques.

This technology, used by many police forces around the world, including the Barcelona Guardia Urbana (Urban Police), makes it possible, through the use of cameras installed in patrol cars, to detect number plates that are in their databases, because they belong to a stolen vehicle, are associated with an administrative incident, have been involved in a criminal offence, etc.[21]

c) Processing tools – a brief note

In order to achieve greater efficiency during the investigation process and, therefore, improve the quality of the service provided to citizens by the justice system, it is clear that it is not enough just to introduce changes in the way investigative measures are carried out and to their content, since if such advances do not go hand in hand, in parallel,

[20] These organisations include the US NGO Benetech and the Syrian Archive.

[21] See the news article published on the Barcelona City's Council's website on 7 February 2019. "*El nuevo sistema de reconocimiento automático de placas de matrícula de la Guardia Urbana*" ("The new automatic number plate recognition system used by the Guardia Urbana (Urban Police)". Last visited on 12 March 2023.
https://ajuntament.barcelona.cat/imi/es/noticia/el-nuevo-sistema

with a transformation in the way cases are handled to speed up case management, their effects will not have the intended (and expected) impact.

In this regard, it is interesting to note that today, without a doubt, AI is already sophisticated enough to enable programmes or systems to be created that analyse claims/allegations/complaints, determine jurisdiction and territorial, objective and functional competence, verify whether or not these are duly filed, detect the possible request for precautionary measures, check for the existence of counterclaims, etc., and, subsequently, if necessary, require them to be corrected, inform the Court Clerk or the judge, or directly proceed to admitting or dismissing these and subsequent processing, to the execution of spoken or written orders, etc.

And this would undoubtedly be extremely useful as a tool to assist Spanish courts and tribunals, where the processing of cases is often paralysed by structural and temporary problems affecting civil servants, who, despite their professionalism, are sometimes overwhelmed by the excessive workload and lack of personnel (as is the case with the Court Clerks and judges). This not only leads to an increase in the anguish of citizens, who have to go through long judicial processes, with the emotional (and often financial) burden that this implies, but also leads to other undesirable consequences, such as the subsequent reduction of the sentences imposed on the perpetrators of criminal acts (sometimes serious), due to the need to appreciate the mitigating circumstance of undue delay, and this is absolutely unacceptable.

In this regard, the PROMETEA tool is particularly interesting, which was introduced in 2017 at the Public Prosecutor's Office for the Autonomous City of Buenos Aires (and today is already operating in other Argentine provinces, and its implementation is underway in the Inter-American Court of Human Rights and the Constitutional Court of Colombia), and enables, through *Machine Learning*, files relating to the Prosecutor's Office to be processed and managed and, among other tasks, deadline checks to be carried out, establishing which solutions should be provided according to the problems of each by analysing the history of similar cases, automating data and documents, ordering and systematising information, etc.

Potential risks

Having revealed the possible benefits that the aforementioned AI tools could bring to criminal investigation, it is necessary, as part of a responsible exercise, to highlight some of the main potential risks most common to all of them and, in particular, to four of them.

First of all, the possible lack of accuracy and discriminatory potential of AI systems should be mentioned.

In this respect, there is still a long way to go, especially with regard to the computing power of the systems and, above all, the quality of the data used to train them.

Regarding this second aspect, it should be noted that, as already mentioned in previous pages, AI systems mainly rely on data, huge amounts of data, which are necessary to train the algorithms so that they can learn and improve over time.

Thus, for example, in the case of risk assessment and prediction tools used by the police, there is a risk of using information contained in police databases created during an era of so-called “dirty” or poorly controlled methods, known as the “*dirty police*” era, and thus is often of poor/low quality, which could lead to poor quality results and, worse, could lead to the perpetuation of undesirable patterns. However, awareness is now being raised and measures are starting to be taken to avoid this danger (for example, in the HART system used by the UK police, postcodes have been removed from the risk assessment form to avoid potential discrimination on the basis of an individual's geographical area of residence).

Secondly, reference should be made to the possible infringement of the right to privacy and the protection of personal data.

Believe it or not, these rights are often disregarded when it comes to the use of AI systems, which is highly dangerous, as the consequences of such an infringement could have absolutely irreparable consequences for the persons concerned.

In this regard, it is interesting to broach the definition of processing of personal data set out in the legislation, which should certainly serve as a guide for addressing this issue. Thus, Article 4 of the GDPR, in line with Article 5 of Organic Law 7/2021 of 26 May on the protection of personal data processed for the purposes of the prevention, detection, investigation and prosecution of criminal offences or the execution of criminal penalties, provides that the processing of personal data involves any operation or set of operations performed on personal data, whether or not this is carried out by means of automated procedures, such as “*collection, recording, organisation, structuring, storage, adaptation or modification, extraction, retrieval, consultation, use, communication by transmission, dissemination or any other form of enabling access, collation or interconnection, restriction, erasure or destruction*”*,* and this must also be considered in relation to the provisions of Article 18 of the Spanish Constitution and Organic Law 1/1982, of 5 May, on the civil protection of the right to honour, personal and family privacy and one's own image.

Thirdly, reference should be made to the possible existence of security gaps in the systems.

Due to the fact that the information processed in criminal investigations is highly sensitive and AI is also a technology that is very vulnerable to cyber-attacks and other unwanted information leaks, it is extremely important to focus on security, as the dire consequences of security breaches are also irreversible (especially when biometric data is used, as if a password or credit card number is “stolen”, for example, it can be changed, but if biometric data is “stolen”, the solution is not so obvious).

Finally, it is worth mentioning the possible lack of transparency.

When it comes to AI, I believe that transparency is the pivot around which everything revolves, given that it is absolutely impossible to verify the quality of the systems (whether or not they violate fundamental rights, what is the nature of the data used, what levels of accuracy are handled, etc.), if they are not transparent.

Thus, it is essential to demand transparency and explicability of AI systems, especially because, in order for their results to have evidential value in judicial proceedings, they should be subject to (real) contradiction, both during the investigation phase and, above all, during the plenary phase, with all the guarantees, and this could only be achieved if they were fully transparent and accessible.

In this regard, I would like to make a brief reflection that I think is important. Nowadays, there is no greater *black box* than that of the judge's brain, since it is currently impossible to know the real motivations that have led him or her to make a particular decision, however well argued it may be in legal terms. However, this fact, which is invariable and obvious, cannot serve as a pretext to justify the lack of transparency and explainability of AI systems, and the existence of black boxes within them, mainly because the impact and repercussion that a judge's decision can have on citizens is absolutely limited, compared to that which, multiplied exponentially, the result produced by an algorithm used, for example, by all the judges in Spain, can have.

3.- CONCLUSIONS

In February 2020, when presenting the European Data Strategy and Strategic Choices to ensure a human-centred development of AI, Margarethe Vestager, Executive Vice President of the European Commission (2019-2024), said: "*AI is neither good nor bad in itself, it all depends on why and how it is used*".

And this, of course, could be the perfect summary of all that has been said in these pages, given that, as has been reiterated on several occasions, AI is a technology that can bring enormous benefits, especially in the field of criminal justice, but which can also cause excessive dangers that we should try to avoid in order to guarantee its use for the benefit of human beings, which can only be achieved through legislation.

Of course, I understand that legislating on a phenomenon as complex as AI is no easy task, and even less so when it comes to authorising its use for criminal investigation purposes, and I therefore believe that the legislative powers should be advised on such a task in a cross-cutting manner by the most prestigious and brilliant experts on AI, personal data protection, cybersecurity, criminal law and criminal procedure, in this case, since there are many fundamental rights at stake, and it is clear that not everything that is technically possible is (or should be) legally feasible.

However, in any case, I believe that the focus should be on avoiding what, in my view, is the worst danger we face in authorising the use of AI systems in justice: that of their dehumanisation.

The fact is that most of the people who go to a police station or a court are in extremely vulnerable situations and come with stories full of fears, confidences and nuances that are impossible for a machine to handle with the empathy and warmth they require. This is true even if we have the most sophisticated and powerful AI systems at our disposal, because nothing can ever replace human contact, and that is what I believe should not be lost sight of, because it is what defines us as a species and what will allow us to preserve our dignity in any scenario.

Thus, without a doubt, we legal experts have an enormous challenge ahead of us and, of course, an immense responsibility, as we must rise to the needs emerging at this great historical moment, in order to try to lay the foundations of our evolution as a species while coexisting with a phenomenon as extraordinary but at the same time as dangerous as AI.