

**Judge. Mohamed Gamil Zakaria**

Prosecutor, Former Visiting Researcher at Florida International University

# AI Applications in the Criminal Justice System: The Next Logical Step or Violation of Human Rights

## ■ Correspondence:

Judge. Mohamed Gamil Zakaria, Prosecutor, Egypt. Former Visiting Researcher at Florida International University.

■ DOI: <https://doi.org/10.54873/jolets.v3i2.124>

■ E-mail: [gamil.mg.mz@gmail.com](mailto:gamil.mg.mz@gmail.com)

## ■ Citation:

Mohamed Gamil Zakaria, AI Applications in the Criminal Justice System: The Next Logical Step or Violation of Human Rights, Conference Research Papers The Third International Conference on Legal Aspects of Digital Transformation: Opportunities and Challenges, Faculty of Law at British University in Egypt, Cairo 17-18 June 2023, Journal of Law and Emerging Technologies, Volume 3, Issue 2, October 2023, p. 233-257



## **AI Applications in the Criminal Justice System: The Next Logical Step or Violation of Human Rights**

**Mr. Mohamed Gamil Zakaria**

### **Abstract**

Artificial intelligence (AI) is impacting almost every aspect of our lives, and automated decisions have begun to replace human decisions. Recently, criminal justice has also turned to AI to reduce crime rates and achieve effective action on the ground. Artificial intelligence and machine learning algorithms (ML) could predict and track crimes and criminals. In addition, they could be used in criminal courts. However, there is ample evidence that the application of this technology has affected basic human rights. In several countries, law enforcement agencies (LEAs), i.e., government agencies responsible for law enforcement, have begun using these technologies despite concerns about the presence of bias in the results and invasion of citizens' privacy. This research paper provides an overview of some AI applications used in criminal justice to predict crimes and offenders, detect and investigate crimes, and assist judges in criminal courts. It also examines the human rights implications of using these applications in this system.

**Keywords:** Artificial Intelligence (AI), Machine learning (ML), Face recognition, Predictive policing

## تطبيقات الذكاء الاصطناعي في نظام العدالة الجنائية: الخطوة المنطقية التالية أم انتهاك لحقوق الإنسان

محمد جميل زكريا

وكيل النائب العام بمصر، باحث بكلية القانون بجامعة فلوريدا الدولية

### الملخص

يؤثر الذكاء الاصطناعي على كل جانب من جوانب حياتنا تقريبا، وبدأت القرارات الآلية تحل محل القرارات البشرية. وفي الآونة الأخيرة، اتجهت أيضا العدالة الجنائية إلى استخدام الذكاء الاصطناعي للحد من معدلات الجريمة وتحقيق إجراءات فعالة على أرض الواقع. يمكن للذكاء الاصطناعي وخوارزميات التعلم الآلي (ML) التنبؤ بالجرائم والمجرمين وتتبعها. بالإضافة إلى ذلك، يمكن استخدامها في المحاكم الجنائية. ولكن على الجانب الآخر، هناك أدلة كثيرة على أن تطبيق هذه التكنولوجيا قد أثر على حقوق الإنسان الأساسية.

في العديد من البلدان، بدأت وكالات إنفاذ القانون (LEAs)، أي الوكالات الحكومية المسؤولة عن إنفاذ القانون، في استخدام هذه التقنيات على الرغم من المخاوف بشأن وجود تحيز في النتائج وانتهاك خصوصية المواطنين.

تقدم هذه الورقة البحثية نظرة عامة على بعض تطبيقات الذكاء الاصطناعي المستخدمة في العدالة الجنائية للتنبؤ بالجرائم والجناة، والكشف عن الجرائم والتحقيق فيها، ومساعدة القضاة في المحاكم الجنائية. بالإضافة إلى أنها تناقش الآثار المترتبة على حقوق الإنسان لاستخدام هذه التطبيقات في هذا النظام.

**الكلمات الرئيسية:** الذكاء الاصطناعي (الذكاء الاصطناعي)، التعلم الآلي (ML)، التعرف على الوجوه.

## AI Applications on Crime Prediction

LEAs perform their tasks in a data-driven, intelligent world. They have had to evolve their working methods to keep pace with modern crimes such as cybercrime, fraud, and human trafficking. Novel technologies have impacted crime prediction, and one of those technologies is AI. The Council of Europe's Commissioner for Human Rights Recommendation defines AI as 'a machine-based system that can make recommendations, predictions, or decisions for a given set of objectives. It does so by utilizing machine and/or human-based inputs to: (i) perceive real and/or virtual environments; (ii) abstract such perceptions into models manually or automatically; and (iii) use model interpretations to formulate options for outcomes.»<sup>(1)</sup> Nowadays, AI is used to gain early insights about future crimes and criminals and helps to make some predictions such as what kind of criminal act, who is the victim, who are the perpetrators, and where will they take place. In addition, AI applications for facial recognition are improving and are being used to make assumptions about possible perpetrators based on facial expressions. This research section will focus on two main areas where AI is being used in predictive policing and facial recognition. It will also discuss some of the AI applications used, their input data or variables, and their goals.

### 1.1. AI Predictive Policing

A report sponsored by the National Institute of Justice in the United States of America 'USA' defines predictive policing as 'the application of analytic techniques-especially quantitative techniques-to identify likely targets for police intervention and prevent crime or solve past crimes through statistical prediction.'<sup>(2)</sup> According to Perry et al. (2013), existing predictive policing

(1) 'Unboxing Artificial Intelligence: 10 Steps to Protect Human Rights' (Commissioner for Human Rights) <[https://www.coe.int/en/web/commissioner/view/-/asset\\_publisher/ugj3i6qSEkhZ/content/unboxing-artificial-intelligence-10-steps-to-protect-human-rights](https://www.coe.int/en/web/commissioner/view/-/asset_publisher/ugj3i6qSEkhZ/content/unboxing-artificial-intelligence-10-steps-to-protect-human-rights)> accessed 28 April 2022

(2) Perry WL and others, 'Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations' (RAND Corporation 2013) <[https://www.rand.org/pubs/research\\_reports/RR233.html](https://www.rand.org/pubs/research_reports/RR233.html)> accessed 28 April 2022

approaches include methods for predicting locations and times of crimes, methods for predicting offenders, offender identities, and predicting crime victims. <sup>(1)</sup> LEAS have used AI models to evaluate past crime data structures and graphs and apply early crime prevention measures. AI predictive policing could be defined as the techniques used to collect, refine, and analyze past crime data and then apply AI models to make assumptions and predictions. In the U.S., some police agencies have purchased predictive policing algorithms from private technology companies. Some of these predictive programs predict which types of people are more likely to commit crimes and who the potential victims might be. They can also estimate which locations are more prone to criminal activity than others. One example of AI algorithms that use AI to predict potential criminals is the Heat List. The list, used by the Chicago Police Department, contains 400 people the department considers particularly prone to violence, either as offenders or victims.<sup>(2)</sup> The algorithm they rely on uses several variables for prediction, including the number of prior gun charges, the number of prior arrests for violent offenses, gang affiliation, and propensity for criminal activity. <sup>(3)</sup> The Chicago Police Department ‘CPD’ claims that the algorithm could be a practical approach to reducing violence. However, it faces many criticisms, from prejudging and discriminating against certain individuals to risking privacy violations.<sup>(4)</sup> In addition, there are the consequences of being on the list, such as being subject to constant surveillance by police, even though the person may be innocent or the result of the algorithm may not be correct. An example of this is McDaniel, a person who was advised by police that he was at risk of either committing a shooting

(1) Perry WL and others, ‘Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations’ (RAND Corporation 2013) <[https://www.rand.org/pubs/research\\_reports/RR233.html](https://www.rand.org/pubs/research_reports/RR233.html)> accessed 28 April 2022

(2) ‘Chicago Police Use «heat List» to Prevent Violence’ (Police1) <<https://www.police1.com/chiefs-sheriffs/articles/chicago-police-use-heat-list-to-prevent-violence-aYQ4dZmhaqIvEB2g/>> accessed 28 April 2022

(3) ‘Chicago’s Strategic Subject List, a.k.a. «Heat List» · Predictive Policing’ <<https://teampturn.gitbooks.io/predictive-policing/content/systems/chicago.html>> accessed 12 April 2022

(4) ‘Chicago Police «Heat List» Renews Old Fears About Government Flagging and Tagging’ (American Civil Liberties Union) <<https://www.aclu.org/blog/privacy-technology/surveillance-technologies/chicago-police-heat-list-renews-old-fears-about>> accessed 12 April 2022

or being shot, even though there was no indication of a discrepancy in his criminal record.<sup>(1)</sup> Another example of an AI algorithm that predicts risky areas is PredPol. The American private company PredPol claims to use a machine learning algorithm to make predictions after training this algorithm on past crime records (2 to 5 years). The inputs to PredPol's system include incident records, crime type, time, and location.<sup>(2)</sup> The algorithm predicts the highest risk areas or locations where criminal activity is likely to occur at a given time. Police officers on patrol are then provided with a highlighted map of hotspots so they can check them regularly.<sup>(3)</sup> For example, if a restaurant is predicted to be a theft hotspot based on the frequency of thefts in that area and at certain times, police officers checking that location regularly will deter potential thieves and prevent crime. Based on an evaluation of 117 days, PredPol predicted 4.7% of crimes in the city of Los Angeles.<sup>(4)</sup> Another predictive policing program is Hunch Lab. According to Cheetham (2019), a patrol management system predicts areas of risk, suggests tactics to address those risks, and suggests patrol assignments. The algorithm uses not only past crime data but also other data such as time of year, day of the week, proximity to bars, lighting, weather, and so on. It might uncover some correlations, such as fewer crimes on cold days or that cars are stolen more often when parked near schools.<sup>(5)</sup> The last example is Beware. The program, developed by Intrado and used by the police department in Fresno - California, works automatically when officers respond to calls. It uses public datasets such as (criminal convictions, public arrest records, social media data, and information

(1) Stroud M, 'An Automated Policing Program Got This Man Shot Twice' (The Verge, 24 May 2021) <<https://www.theverge.com/22444020/heat-listed-csk-entry>> accessed 12 April 2022

(2) (PredPol) <<https://predpol.com/>>

(3) Ferguson AG, 'Policing Predictive Policing' (2016) 94 Washington University Law Review 1109 <<https://heinonline.org/HOL/P?h=hein.journals/walq94&i=1163>> accessed 8 February 2022

(4) Ferguson AG, 'Policing Predictive Policing' (2016) 94 Washington University Law Review 1109 <<https://heinonline.org/HOL/P?h=hein.journals/walq94&i=1163>> accessed 8 February 2022

(5) 'Why We Sold HunchLab' (Azavea, 23 January 2019) <<https://www.azavea.com/blog/2019/01/23/why-we-sold-hunchlab/>> accessed 8 February 2022

compiled by commercial data brokers)<sup>(1)</sup> to consider a person's relative risk of committing a crime by creating a personal threat score. Police claim that the information provided by these tools helps them track violent criminals and ensure public safety. However, civil libertarians view these tools as a violation of public privacy because they process some sensitive personal information without people's consent.<sup>(2)</sup>

## 1.2. AI Facial recognition

Face recognition could be defined as “the way software determines the similarity between two facial images to evaluate a claim.” The technology is used for a variety of purposes, from logging a user into their phone to searching for a specific person in a photo database.”<sup>(3)</sup> Recently, AI technology in the field of facial recognition has proliferated rapidly. Recently, it has become possible to predict the age, race, and gender of individuals based on their facial expressions.”<sup>(4)</sup> States, private companies, and researchers are making every effort to use AI facial recognition to predict criminal behavior even before crimes are committed by analyzing people's faces to make assumptions about whether or not the person is likely to be a criminal. However, there are many calls for this research to be stopped because it violates individual privacy.<sup>(5)</sup> An example of the claims demonstrating the success of facial recognition algorithms is the face recognition model developed by three researchers at Harrisburg College in 2020. They proclaimed that the algorithm could predict

(1) ‘Reasonably Suspicious Algorithms: Predictive Policing at the United States Border’ (N.Y.U. Review of Law & Social Change, 27 September 2017) <<https://socialchangenyu.com/review/reasonably-suspicious-algorithms-predictive-policing-at-the-united-states-border/>> accessed 12 April 2022

(2) Jovenal J, ‘The New Way Police Are Surveilling You: Calculating Your Threat «Score»’ Washington Post (10 January 2016) <[https://www.washingtonpost.com/local/public-safety/the-new-way-police-are-surveilling-you-calculating-your-threat-score/2016/01/10/e42bccac-8e15-11e5-baf4-bdf37355da0c\\_story.html](https://www.washingtonpost.com/local/public-safety/the-new-way-police-are-surveilling-you-calculating-your-threat-score/2016/01/10/e42bccac-8e15-11e5-baf4-bdf37355da0c_story.html)> accessed 8 February 2022

(3) ‘How Does Facial Recognition Work?’ <<https://www.csis.org/analysis/how-does-facial-recognition-work>> accessed 28 April 2022

(4) Agbo-Ajala O and Viriri S, ‘Deeply Learned Classifiers for Age and Gender Predictions of Unfiltered Faces’ (2020) 2020 The Scientific World Journal e1289408 <<https://www.hindawi.com/journals/tswj/2020/1289408/>> accessed 12 April 2022

(5) News ABC, ‘States Push Back against Use of Facial Recognition by Police’ (ABC News) <<https://abcnews.go.com/Politics/wireStory/states-push-back-facial-recognition-police-77510175>> accessed 13 April 2022



crime from a person's face. They claimed it had 80% accuracy and «no racial bias.» However, the model has been heavily criticized for its biased results <sup>(1)</sup>, as its input data already comes from human-based datasets, which are also biased. Another study by Wu and Zhang (2016) <sup>(2)</sup> showed, after experiments, that data-driven face classifiers can infer criminality through supervised machine learning. Moreover, they discovered that the law-abiding public differed from criminals in their facial features. Similarly, in 2018, the National Institute of Standards and Technology (NIST) evaluated 127 facial recognition software algorithms from 39 developers and found massive progress in face-based biometric matching algorithms. <sup>(3)</sup> In addition, the Chinese government studied the activities of Chongqing residents and was able to predict suspicious individuals, incorporating facial expressions into the analysis. <sup>(4)</sup> On the other hand, an experiment conducted by Bowyer et al. (2020) shows that the concept (crime from the face) is an illusion, and it is dangerous to believe in it. The authors concluded that it is better to think about more valuable solutions for humanity than to waste valuable resources and time on this type of research due to the unintended biases in the data sets. <sup>(5)</sup> Also in the United Kingdom UK, the Metropolitan Police's facial recognition technology has been monitoring crowds with Live Facial Recognition 'LFR' since August 2016. This technology was evaluated by two researchers from the College of Essex. They found that four out of five people identified as

(1) Holmes A, 'Researchers Said Their «unbiased» Facial Recognition Could Identify Potential Future Criminals — Then Deleted the Announcement after Backlash' (Business Insider) <<https://www.businessinsider.com/harrisburg-university-lab-facial-recognition-identify-future-criminals-2020-5>> accessed 9 February 2022

(2) Wu X and Zhang X, 'Responses to Critiques on Machine Learning of Criminality Perceptions (Addendum of ArXiv:1611.04135)' [2016] arXiv:1611.04135 [cs] <<http://arxiv.org/abs/1611.04135>> accessed 9 February 2022

(3) robin.materese@nist.gov, 'NIST Evaluation Shows Advance in Face Recognition Software's Capabilities' (NIST, 30 November 2018) <<https://www.nist.gov/news-events/news/2018/11/nist-evaluation-shows-advance-face-recognition-softwares-capabilities>> accessed 9 February 2022

(4) Simon Denyer, 'China Bets on Facial Recognition in Big Drive for Total Surveillance' (Washington Post, 2018) <<https://www.washingtonpost.com/news/world/wp/2018/01/07/feature/in-china-facial-recognition-is-sharp-end-of-a-drive-for-total-surveillance/>> accessed 24 February 2022

(5) Bowyer KW and others, 'The Criminality From Face Illusion' [2020] arXiv:2006.03895 [cs] <<http://arxiv.org/abs/2006.03895>> accessed 9 February 2022

possible suspects by the MET'S technology were innocent. <sup>(1)</sup> In addition, Axon, a major provider of body cameras and software in the U.S., announced it would ban the use of facial recognition systems on its devices, saying the technology was not trustworthy enough to ethically justify its use due to its potential for racial impact and inaccuracy. <sup>(2)</sup> California also considered banning automatic facial recognition on police cameras<sup>(3)</sup> because it considers identity verification of individuals without their consent to be an invasion of public privacy and California law.

## 2. AI in Crime Investigations and Prosecution

Criminal investigation methods have evolved rapidly in recent years. Nowadays, prosecutors rely on digital evidence and support their work with modern technologies. Criminal investigators need to be able to access and analyze vast amounts of data from multiple sources: Cloud data, mobile devices, crime scene data, criminal history databases and social media platforms. They should have the latest tools, rely on the most capable experts, and undergo high-quality training to keep up with the changing technological environment. AI can highlight patterns in the data collected, analyze it efficiently, and provide prosecutors, law enforcement officials, and other agencies with some information they need, which usually takes a lot of time and effort. This section focuses on how AI is being used to facilitate crime detection and investigations.

### 2.1 Multimedia and Social Media Data Analysis

Big Data refers to large amounts of data that are very complex and cannot

- 
- (1) '81% of «suspects» Flagged by Met's Police Facial Recognition Technology Innocent, Independent Report Says' (Sky News) <<https://news.sky.com/story/met-polices-facial-recognition-tech-has-81-error-rate-independent-report-says-11755941>> accessed 9 February 2022
  - (2) 'Opinion | A Major Police Body Cam Company Just Banned Facial Recognition - The New York Times' <<https://www.nytimes.com/2019/06/27/opinion/police-cam-facial-recognition.html>> accessed 9 February 2022
  - (3) Sam Dean, 'California Considers Ban on Facial Recognition's New Frontier: Police Body Cameras' (Los Angeles Times, 7 June 2019) <<https://www.latimes.com/business/technology/la-fi-tn-face-recognition-ban-california-police-body-camera-20190607-story.html>> accessed 9 February 2022

be processed using traditional methods. This has complicated multimedia analysis in post-event investigations. Investigators have to review a huge number of data samples to find evidence. In addition, analyzing many images and videos and trying to extrapolate valuable information is very tiring and leads to errors. AI can master all multimedia data from CCTVS, low-quality security cameras, recordings, and mobile devices. It analyzes records, learns, and performs complex tasks, such as recognizing and matching faces with databases, recognizing and matching voices, recognizing criminal acts, and identifying weapons. <sup>(1)</sup>

In addition, enormous amounts of information are shared across different platforms. Users interact with each other and share their opinions and thoughts about everything. And some countries, such as China and Iran, monitor their citizens through their social media platforms. <sup>(2)</sup> In the U.S., the Federal Bureau of Investigations ‘FBI’ scans online media to support criminal investigations they conduct. Their agents could initiate an investigation when crimes are prevented; during this investigation, they search all available online information. <sup>(3)</sup> AI could help in detecting criminal behaviors and crimes after being fed with these datasets. For example, a platform was proposed by Shoeibi et al. (2021) to analyze the connections between Twitter users and their shared content. In the platform, AI techniques were used to detect crimes and evaluate users’ agreement with criminal topics. <sup>(4)</sup>

---

(1) Rigano C, ‘Using Artificial Intelligence to Address Criminal Justice Needs’ (National Institute of Justice) <<https://nij.ojp.gov/topics/articles/using-artificial-intelligence-address-criminal-justice-needs>> accessed 28 February 2022

(2) Shahbaz A and Funk A, ‘Social Media Surveillance’ (Freedom House) <<https://freedomhouse.org/report/freedom-on-the-net/2019/the-crisis-of-social-media/social-media-surveillance>> accessed 1 March 2022

(3) Panduranga H and Pablo M, ‘Federal Government Social Media Surveillance, Explained | Brennan Center for Justice’ (2022) <<https://www.brennancenter.org/our-work/research-reports/federal-government-social-media-surveillance-explained>> accessed 13 April 2022

(4) Shoeibi N and others, ‘AI-Crime Hunter: An AI Mixture of Experts for Crime Discovery on Twitter’ (2021) 10 Electronics 3081 <<https://www.mdpi.com/2079-9292/10/24/3081>> accessed 28 February 2022

## 2.2 Gunshot detection

Many crimes are recorded as audio on personal devices such as cell phones and smartwatches. AI could help analyze these recordings and provide useful information to investigators. One example is Cadre Labs' research aimed at detecting gunshots in audio recordings. They used real audio records extracted from YouTube and recorded from unknown devices. They also used audio recordings of test shots from various firearms recorded by other devices. The algorithm was fed the data and was able to successfully identify gunshots, determine the number of different firearms displayed, and was moderately successful in recognizing the (class, caliber, and make/model) of the firearms used. <sup>(1)</sup> Another firearm detection technology used in many major cities in the US is ShotSpotter. This involves multiple acoustic sensors with four microphones per device. The technology detects all blast sounds, with the AI algorithm filtering out the noise, creating an audio recording and sending it to ShotSpotter's Incident Response Center "IRC." Experts at the IRC review the recording, and if they classify it as gunshots, they immediately notify the police.<sup>(2)</sup> However, there are many concerns about this technology. One of the examples of ShotSpotter's ineffectiveness is the profiling of Michael Williams. He spent a year in jail based on evidence found by ShotSpotter before his murder charge was dismissed because the prosecution deemed this evidence insufficient. <sup>(3)</sup> The problem is that this technology is not transparent enough and cannot be verified, which is not accepted in criminal court proceedings. Apart from the lack of transparency, placing live microphones in public areas may also raise privacy concerns.

(1) Lilien R, 'Development of Computational Methods for the Audio Analysis of Gunshots | Office of Justice Programs' (2018) <<https://www.ojp.gov/ncjrs/virtual-library/abstracts/development-computational-methods-audio-analysis-gunshots>> accessed 7 March 2022

(2) Goodman B, 'Shotspotter - The New Tool to Degrade What Is Left of the Fourth Amendment Comments' (2021) 54 UIC John Marshall Law Review [i] <<https://heinonline.org/HOL/P?h=hein.journals/jmlr54&i=810>> accessed 10 March 2022

(3) stanley jay, 'Four Problems with the ShotSpotter Gunshot Detection System' (American Civil Liberties Union, 2021) <<https://www.aclu.org/news/privacy-technology/four-problems-with-the-shotspotter-gunshot-detection-system/>> accessed 25 February 2022

### 2.3 AI Interrogators.

Recently, there have been several calls for the use of AI in interrogation rooms. A 2014 study showed that AI in the form of a computer-generated agent performed well in getting some individuals—who participated in the experiment—to admit their past involvement in criminal activity or to convict of crimes.<sup>(1)</sup> ‘AI Interrogator was able to achieve more accurate results than human interrogator through AI facial emotion recognition, AI verbal recognition, and deception detection software.’<sup>(2)</sup> Another study found that AI algorithms are able to detect changes in human emotions with 90% accuracy versus 75% accuracy of humans.<sup>(3)</sup> In addition, a study conducted at the Nirma Institute of Technology in India found that the AI model can analyze human emotions through voice and speech recognition.<sup>(4)</sup> In addition, AI lie detectors could replace traditional lie detectors. The lie detector is used to identify liars<sup>(5)</sup>. This is a digital video camera connected to a computer that relies on artificial neural networks to learn and recognize patterns in data. The inventors of the device say the AI identifies nonverbal micro gestures on people’s faces and picks up on unconscious reactions such as signs of stress, mental tension, and feigned happiness.<sup>(6)</sup>

### 2.4 DNA analysis.

DNA traces found at crime scenes that may carry the DNA of the perpetrator

- 
- (1) Pollina DA and Barretta A, ‘The Effectiveness of a National Security Screening Interview Conducted by a Computer-Generated Agent’ (2014) 39 *Computers in Human Behavior* 39 <<https://www.sciencedirect.com/science/article/pii/S0747563214003458>> accessed 28 February 2022
  - (2) Noriega M, ‘The Application of Artificial Intelligence in Police Interrogations: An Analysis Addressing the Proposed Effect AI Has on Racial and Gender Bias, Cooperation, and False Confessions’ (2020) 117 *Futures* 102510 <<https://www.sciencedirect.com/science/article/pii/S0016328719303726>> accessed 28 February 2022
  - (3) ‘At First Blush, You Look Happy—or Sad, or Angry’ (At first blush, you look happy—or sad, or angry, 2018) <<https://news.osu.edu/at-first-blush-you-look-happy--or-sad-or-angry/>> accessed 28 February 2022
  - (4) Dasgupta PB, ‘Detection and Analysis of Human Emotions through Voice and Speech Pattern Processing’ (2017) 52 *International Journal of Computer Trends and Technology* 1 <<http://arxiv.org/abs/1710.10198>> accessed 28 February 2022
  - (5) Leonetti C, ‘AbraCadabra, Hocus Pocus, Same Song, Different Chorus: The Newest Iteration of the Science of Lie Detection’ (2017) 24 *Richmond Journal of Law & Technology* 1 <<https://heinonline.org/HOL/P?h=hein.journals/jolt24&i=67>> accessed 28 April 2022
  - (6) Kennedy P, ‘Artificial Intelligence Lie Detector Developed by Imperial Alumnus | Imperial News | Imperial College London’ (Imperial News, 2014) <<https://www.imperial.ac.uk/news/144486/artificial-intelligence-detector-developed-imperial-alumnus/>> accessed 28 February 2022

should be treated with caution. They are often mixed with the traces of other innocent people, leading to wrongful convictions.<sup>(1)</sup> Scientists are trying to develop AI techniques to extract DNA profiles from the DNA soup and determine whether a DNA sample came from someone who was at the crime scene or was innocently transferred.<sup>(2)</sup>

An example of a novel technological solution is The Pace System, a machine learning-based tool offered to determine the number of contributors in a DNA mixture. Moreover, it does not involve any computational effort, and the estimation could be performed by any standard computer<sup>(3)</sup>

### 3. AI in Criminal Courts.

AI systems are not yet replacing human judges in criminal courts, but they may soon do so. (4) Right now, AI algorithms are being used as tools for judges to make some probation and sentencing decisions. Kleinberg et al.-2017 found in their research that AI applications can suggest better decisions than humans by comparing the performance of both algorithms and human judges. They found that the use of algorithms could reduce the incarceration rate by 41.8% without changing the crime reduction rate. (5) An example of tools to assist judges is HART, which stands for Harm Assessment Risk Tool. It is a machine learning algorithm used in the UK. It analyzes a person's data, including their criminal history and sociodemographic background,

---

(1) Worth K, 'Framed for Murder by His Own DNA' (The Marshall Project, 19 April 2018) <<https://www.themarshallproject.org/2018/04/19/framed-for-murder-by-his-own-dna>> accessed 3 March 2022

(2) Richmond K, 'AI Could Revolutionise DNA Evidence – but Right Now We Can't Trust the Machines' (The Conversation, 2020) <<http://theconversation.com/ai-could-revolutionise-dna-evidence-but-right-now-we-cant-trust-the-machines-129927>> accessed 3 March 2022

(3) Marciano MA and Adelman JD, 'PACE: Probabilistic Assessment for Contributor Estimation— A Machine Learning-Based Assessment of the Number of Contributors in DNA Mixtures' (2017) 27 *Forensic Science International: Genetics* 82 <<https://www.sciencedirect.com/science/article/pii/S1872497316302162>> accessed 3 March 2022

(4) Niler E, 'Can AI Be a Fair Judge in Court? Estonia Thinks So' [2019] *Wired* <<https://www.wired.com/story/can-ai-be-fair-judge-court-estonia-thinks-so/>> accessed 9 April 2022

(5) Sunstein CR, 'Algorithms, Correcting Biases' (Social Science Research Network 2018) SSRN Scholarly Paper 3300171 <<https://papers.ssrn.com/abstract=3300171>> accessed 9 April 2022

and assesses their risk of reoffending.<sup>(1)</sup> Another risk assessment tool used in the United States is COMPAS, which stands for Correctional Offender Management for Alternative Sanctions. It assesses the likelihood that a defendant will be a recidivist and decides whether he/she has to be released on bail. <sup>(2)</sup> COMPAS is not the only application used in the United States. There are many applications, and one of them is the prediction of bond amounts in Fort Lauderdale, Florida.<sup>(3)</sup>

In addition, in the PRC, over 70% of cases are simple, and criminal laws can be easily applied. Court procedures can be automated to save judicial resources for complex cases<sup>(4)</sup>. For example, in 2016, the Supreme People's Court (SPC) issued several opinions on further developing the distinction between simple and sophisticated cases to optimize the allocation of judicial resources <sup>(5)</sup>. An example of AI algorithms used by Chinese courts is the Traffic Accidents Dispute Resolution System "TADRS". It examines and analyzes data such as weather, traffic signs, accidents, and relevant legal regulations. Then, using Deep Learning, it scans previous court rulings to make suggestions to judges in the courts.<sup>(6)</sup>

AI is also being used to predict future court decisions. Machine learning

- 
- (1) Oswald M and others, 'Algorithmic Risk Assessment Policing Models: Lessons from the Durham HART Model and «Experimental» Proportionality' (2018) 27 Information & Communications Technology Law 223 <<https://doi.org/10.1080/13600834.2018.1458455>> accessed 9 April 2022
  - (2) 'COMPAS (Software)', , Wikipedia (2022) <[https://en.wikipedia.org/w/index.php?title=COMPAS\\_\(software\)&oldid=1078441070](https://en.wikipedia.org/w/index.php?title=COMPAS_(software)&oldid=1078441070)> accessed 9 April 2022
  - (3) Eckhouse L, 'Opinion | Big Data May Be Reinforcing Racial Bias in the Criminal Justice System' Washington Post (10 February 2017) <[https://www.washingtonpost.com/opinions/big-data-may-be-reinforcing-racial-bias-in-the-criminal-justice-system/2017/02/10/d63de518-ee3a-11e6-9973-c5efb7ccfb0d\\_story.html](https://www.washingtonpost.com/opinions/big-data-may-be-reinforcing-racial-bias-in-the-criminal-justice-system/2017/02/10/d63de518-ee3a-11e6-9973-c5efb7ccfb0d_story.html)> accessed 9 April 2022
  - (4) Zheng GG, 'China's Grand Design of People's Smart Courts' (2020) 7 Asian Journal of Law and Society 561 <<https://www.cambridge.org/core/journals/asian-journal-of-law-and-society/article/chinas-grand-design-of-peoples-smart-courts/476879522161B47A5BE10DBC4BDE8215#r4>> accessed 9 April 2022
  - (5) «Several Opinions of the Supreme People's Court on Further Promoting the Separation of Complex and Simple Cases and Optimizing the Allocation of Judicial Resources-Beijing Court Network.» (2016). Retrieved March 23, 2022, from <https://bjgy.chinacourt.gov.cn/article/detail/2018/04/id/3281212.shtml>.
  - (6) Zhu W and Xueyuan Z, 'Exploring the Element-Based Trial of Road Traffic Disputes——From the Reform Practice of Sichuan High Court' (2018) <<http://zlib.cglhub.com/auto/db/detail.aspx?db=950015&rid=52758211&agfi=0&cls=0&uni=True&cid=0&showgp=False&prec=False&md=265&pd=3&msd=265&psd=3&md=265&pdd=3&count=10&reds=%D2%AA%EF%BF%BD%EF%BF%BD%CA%BD;%EF%BF%BD%EF%BF%BD%EF%BF%BD;%CA%BD%EF%BF%BD%EF%BF%BD>> accessed 9 April 2022

algorithms could be fed with case law and previous judgments of judges to predict the next court decisions. Researchers were able to predict the decisions of the European Court of Human Rights (ECTHR), and they used natural language processing tools<sup>(1)</sup> [49] to analyze court case data. They were able to predict the violation of 9 articles of the ECHR with an average hit rate of 75%.<sup>(2)</sup>

#### 4. AI and human rights

Some believe that the application of AI technology in criminal justice can help overcome human bias, fix its flaws, and improve the system. They prefer to develop these technologies in law enforcement and justice, according to what has been shown recently by experiments and studies on human bias and error. For example, a study by Danziger et al. found that judges' decisions are influenced by some factors and daily routines, such as whether the decision was made at the beginning or end of the session and whether or not it was made after a meal break. The experiment has shown that judges are more inclined to grant probation after having a meal.<sup>(3)</sup> Another example is the analysis of Amos Tversky and Daniel Kahneman, which points to biases in judges' decisions.<sup>(4)</sup> On the other hand, critics see AI as violating human rights, discriminating against people, and disregarding their privacy. This research section identifies three human rights that are affected by the application of AI in criminal justice, namely non-discrimination and equality, protection of privacy and personal data, and procedural justice.

(1) 'Natural Language Processing', , Wikipedia (2022) <[https://en.wikipedia.org/w/index.php?title=Natural\\_language\\_processing&oldid=1081150798](https://en.wikipedia.org/w/index.php?title=Natural_language_processing&oldid=1081150798)> accessed 9 April 2022

(2) Medvedeva M, Vols M and Wieling M, 'Using Machine Learning to Predict Decisions of the European Court of Human Rights' (2020) 28 Artificial Intelligence & Law 237 <<https://search.ebscohost.com/login.aspx?direct=true&AuthType=shib&db=aci&AN=143543764&authtype=shib&site=ehost-live&scope=site&custid=s8862125>> accessed 9 April 2022

(3) Danziger S, Levav J and Avnaim-Pesso L, 'Extraneous Factors in Judicial Decisions' (2011) 108 Proceedings of the National Academy of Sciences of the United States of America 6889

(4) Tversky A and Kahneman D, 'Judgment under Uncertainty: Heuristics and Biases' (1974) 185 Science (New York, N.Y.) 1124



#### 4.1 Non-discrimination and equality

Regardless of their origin, race, ethnicity, color, or gender, all people should be treated equally before the law and should not be discriminated against at any stage of the criminal justice process. All human rights treaties prohibit discrimination; first, Article 7 of the (Universal Declaration of Human Rights) states: ‘All human beings are equal before the law and are entitled without distinction to equal protection of the law. All are entitled to equal protection against any discrimination contrary to this Declaration and against any incitement to such discrimination.’<sup>(1)</sup>[53] Similarly, Article 14 of the European Convention on Human Rights states, ‘‘The enjoyment of the rights and freedoms set forth in this Convention shall be secured without discrimination on any ground such as sex, race, color, language, religion, political or other opinion, national or social origin, association with a national minority, property, birth or other status.’’<sup>(2)</sup> Even in countries that are not parties to the above treaties, such as the United States, there are many laws prohibiting all forms of discrimination, such as the Americans with Disability Act, the Civil Rights Act, and the Pregnancy Discrimination Act.<sup>(3)</sup> There is considerable evidence of the presence of bias in the results of AI algorithms in general, e.g., the biased results in commercial activities such as those of Amazon, which appears to exclude minority neighborhoods from some services,<sup>(4)</sup> and Google, which linked the names of blacks to arrest-related ads.<sup>(5)</sup> And in evaluating AI applications for crime prediction, the

(1) Nations U, ‘Universal Declaration of Human Rights, Article 10’ (United Nations) <<https://www.un.org/en/about-us/universal-declaration-of-human-rights>> accessed 30 March 2022

(2) ‘Article 14 of the European Convention on Human Rights’, , Wikipedia (2021) <[https://en.wikipedia.org/w/index.php?title=Article\\_14\\_of\\_the\\_European\\_Convention\\_on\\_Human\\_Rights&oldid=1061490489](https://en.wikipedia.org/w/index.php?title=Article_14_of_the_European_Convention_on_Human_Rights&oldid=1061490489)> accessed 17 March 2022

(3) ‘Federal Antidiscrimination Laws’ (www.nolo.com) <<https://www.nolo.com/legal-encyclopedia/federal-anti-discrimination-laws-29451.html>> accessed 17 March 2022

(4) Ingold D and Soper S, ‘Amazon Doesn’t Consider the Race of Its Customers. Should It?’ (Bloomberg.com, 2016) <<http://www.bloomberg.com/graphics/2016-amazon-same-day/>> accessed 29 March 2022

(5) ‘Racism Is Poisoning Online Ad Delivery, Says Harvard Professor’ (MIT Technology Review, 2013) <<https://www.technologyreview.com/2013/02/04/253879/racism-is-poisoning-online-ad-delivery-says-harvard-professor/>> accessed 29 March 2022

fact is that these algorithms are fed input data that contain human biases. Consequently, they will provide discriminated results and predictions. To illustrate this point, AI hotspots may increase geographic discrimination for two reasons: first, the historical data processed by the algorithms are not clean and are predominantly focused on places where minors and poor people live, <sup>(1)</sup>[58] as these areas are most heavily policed. Second, the input data contain only police conceptualization and may include illegal practices and over-policing of some areas. <sup>(2)</sup>This will result in these areas being classified as hotspots, and thus more arrests will be made on a one-sided basis. To further illustrate, the so-called heat lists that classify certain individuals as potential criminals can stigmatize them based on partial data. Predictive policing is based on existing, incomplete data in which minorities are unfairly overrepresented. The use of predictive policing would only exacerbate racial disparities in crime data. Predictive AI policing methods would lead to a reciprocal loop in which biased data lead to discriminatory practices, which in turn lead to even more biased data, which in turn lead to even more discriminatory practices.<sup>(3)</sup> Furthermore, due to the lack of transparency of these AI models, LEAS may assume that the algorithm's results are neutral and accurate, but in reality they are not absolutely objective. Also, the possibility of detecting discrimination will be difficult, if not impossible. Moreover, discrimination affects the society as a whole. According to (Schlehahn et al., 2015), depending on the results of the algorithm, which are mainly derived from their criminal records, the continued treatment of ex-offenders as those most likely to commit crimes will undoubtedly affect their

(1) O'Neil C, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (First edition, Crown 2016)

(2) Babuta A, Oswald M and Rinik C, *Machine Learning Algorithms and Police Decision-Making: Legal, Ethical and Regulatory Challenges* (Royal United Services Institute 2018)

(3) O' Donnell Renata M, 'Challenging Racist Predictive Policing Algorithms Under the Equal Protection Clause' (NYU Law Review, 1 June 2019) <<https://www.nyulawreview.org/issues/volume-94-number-3/challenging-racist-predictive-policing-algorithms-under-the-equal-protection-clause/>> accessed 18 March 2022

rehabilitation and drive most of them to criminal behavior again. <sup>(1)</sup>Moreover, algorithms known to predict whether or not a defendant will reoffend can also lead to discrimination. For example, COMPAS, which helps judges in the U.S. decide whether someone should be released on parole,<sup>(2)</sup> has been found to discriminate against blacks. One study that examined cases against both whites and blacks found that the results for white defendants were skewed toward the lower-risk categories, while this was not the case for black defendants.<sup>(3)</sup>

## 4.2 Personal Data Protection

The protection of personal data is a human right. Everyone has the right to access their personal data, to keep it private, to object to its use, and to know when and where it is being used or disclosed. Privacy is a fundamental human right recognized in Article 12 of the Universal Declaration of Human Rights, which states, «No one shall be subjected to arbitrary interference with his or her private life»<sup>(4)</sup>. Currently, we can see AI algorithms being applied to every aspect of our lives. The datasets processed by AI usually contain personal data such as personal information, criminal records, and personal images. Accordingly, it is imperative that international organizations and states enact laws and regulations to protect personal data. In Europe, the General Data Protection Regulation (GDPR), which came into effect in 2018, gives Europeans the right to opt out of any automated decision because it carries legal consequences.<sup>(5)</sup> In the U.S., there is no federal law that regulates data

(1) Schlehahn E and others, 'Benefits and Pitfalls of Predictive Policing', 2015 European Intelligence and Security Informatics Conference (2015)

(2) Zuiderveen Borgesius FJ, 'Strengthening Legal Protection against Discrimination by Algorithms and Artificial Intelligence' (2020) 24 The International Journal of Human Rights 1572 <<https://doi.org/10.1080/13642987.2020.1743976>> accessed 18 March 2022

(3) Mattu JA Jeff Larson, Lauren Kirchner, Surya, 'Machine Bias' (ProPublica, 2016) <[https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing?token=HM\\_vaeiiqqmnZ9h19EC5CdRyKi-ACk6M](https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing?token=HM_vaeiiqqmnZ9h19EC5CdRyKi-ACk6M)> accessed 18 March 2022

(4) Nations U, 'Universal Declaration of Human Rights' (United Nations) <<https://www.un.org/en/about-us/universal-declaration-of-human-rights>> accessed 17 March 2022

(5) 'Art. 22 GDPR – Automated Individual Decision-Making, Including Profiling' (General Data Protection Regulation (GDPR)) <<https://gdpr-info.eu/art-22-gdpr/>> accessed 30 March 2022

privacy. However, some states have their own laws, such as the California Privacy Rights Act, which gives citizens the right to be informed about the data collected about them, to request the deletion of their collected personal data, and to know whether or not the personal data will be shared. <sup>(1)</sup> Another example is the Illinois Biometric Information Privacy Act ‘BIPA.’ <sup>(2)</sup> [67] AI can violate people’s privacy when used in the criminal justice system. One of the controversial AI applications that can impact privacy is facial recognition. For example, Clearview, an American facial recognition company, has built a huge facial recognition database that contains more than three billion images, most of which come from Facebook, YouTube, and millions of websites. It identifies each person by their face. All it takes is to upload a person’s image and all of that person’s public images are displayed, along with links to the websites where those images appear. <sup>(3)</sup> More than 2,400 police agencies use Clearview’s application to identify suspects and victims. <sup>(4)</sup> It is also used to identify violent protesters. <sup>(5)</sup> There are numerous concerns about Clearview’s application in terms of violating individuals’ privacy by using their data and photos without their consent. In addition, it is too risky for this sensitive data to be uploaded to servers that are not vetted for their ability to protect the data. Many companies such as Google, Facebook, Microsoft, and Twitter have asked Clearview to stop collecting data from their websites and users. <sup>(6)</sup> Many complaints and lawsuits have been filed against Clearview; in Germany, for

(1) ‘The California Privacy Rights Act of 2020’ <<https://iapp.org/resources/article/the-california-privacy-rights-act-of-2020/>> accessed 30 March 2022

(2) ‘740 ILCS 14/ Biometric Information Privacy Act.’ <<https://www.ilga.gov/legislation/ilcs/ilcs3.asp?ActID=3004&ChapterID=57>> accessed 4 April 2022

(3) 2022 Hill K, ‘The Secretive Company That Might End Privacy as We Know It’ The New York Times (18 January 2020) <<https://www.nytimes.com/2020/01/18/technology/clearview-privacy-facial-recognition.html>> accessed 3 April 2022

(4) This Week in Startups, E1100: Clearview AI CEO Hoan Ton-That on Balancing Privacy & Security, Engaging with Controversy (2020) <[https://www.youtube.com/watch?v=wNLK\\_f6m4e0](https://www.youtube.com/watch?v=wNLK_f6m4e0)> accessed 4 April 2022

(5) Prazan P and Fossi C, ‘Miami Police Used Facial Recognition Technology in Protester’s Arrest’ (NBC 6 South Florida, 2020) <<https://www.nbcmiami.com/investigations/miami-police-used-facial-recognition-technology-in-protesters-arrest/2278848/>> accessed 4 April 2022

(6) Cox K, ‘Facebook, YouTube Order Clearview to Stop Scraping Them for Faces to Match’ (Ars Technica, 7 February 2020) <<https://arstechnica.com/tech-policy/2020/02/facebook-youtube-order-clearview-to-stop-scraping-them-for-faces-to-match/>> accessed 4 April 2022

example, the «Hamburg Data Protection Authority (HDPa)» upheld the right of every European to not be represented in Clearview’s database. The HDPa also required the company to delete the complainant’s biometric data and claimed that Clearview AI must comply with the General Data Protection Regulation because its activities affect European residents. <sup>(1)</sup>In the U.S., specifically in Illinois, the American Civil Liberties Union “ACLU” filed a lawsuit against the company for violating the Illinois Biometric Privacy Act and collecting biometric data on individuals without their consent. <sup>(2)</sup> The latest example from China is Guo Bing, who did not accept Wildlife World Park’s decision to use facial recognition technology and filed suit against the park. The Chinese court ordered the park to delete his facial data, holding that the improper use of this sensitive data would jeopardize people’s privacy. <sup>(3)</sup>[74] In addition, the collection and processing of data from social media platforms by AI algorithms can violate people’s privacy. This data usually contains sensitive personal information, and LEAS can process it without people’s consent, as was the case with Trump’s election campaign. <sup>(4)</sup>

### 4.3 Procedural Fairness

The right to a fair trial is one of the fundamental human rights guaranteed in any democratic society and is an essential guarantor of the rule of law. Judges should uphold and monitor fairness in judicial proceedings. They should not be influenced by any party outside the judiciary in the performance of their duties. Criminal courts must be independent and impartial. One of the international

(1) ‘Clearview AI Deemed Illegal in the EU’ (2021) <<https://noyb.eu/en/clearview-ai-deemed-illegal-eu>> accessed 4 April 2022

(2) Campbell IC, ‘Clearview AI Hit with Sweeping Legal Complaints over Controversial Face Scraping in Europe’ (The Verge, 27 May 2021) <<https://www.theverge.com/2021/5/27/22455446/clearview-ai-legal-privacy-complaint-privacy-international-facial-recognition-eu>> accessed 4 April 2022

(3) Qiao R-S, Wang C and Clark D, ‘Zoo Loses the «First Face Recognition Case» at Appeal’ (Lexology, 22 April 2021) <<https://www.lexology.com/library/detail.aspx?g=882322e8-fcb4-4f19-83e9-a4ff71d50e1d>> accessed 30 March 2022

(4) Confessore N, ‘Cambridge Analytica and Facebook: The Scandal and the Fallout So Far’ The New York Times (4 April 2018) <<https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html>> accessed 13 April 2022

protective instruments for this right is the Universal Declaration of Human Rights. It states, 'Everyone is entitled in full equality to a fair and public hearing by an independent and impartial tribunal of his rights and obligations and of any criminal charge against him (Article 10).'<sup>(1)</sup> Another instrument is the European Convention on Human Rights. It protects the right to a fair trial, the presumption of innocence, and the right to examine or have examined adverse witnesses (Article 6).<sup>(2)</sup> Moreover, in the United States, the Sixth Amendment guarantees the rights of criminal defendants, including the right to an impartial jury and the right to know the nature of the charges and evidence against them.<sup>(3)</sup> The application of AI to the criminal justice system may violate the right to a fair trial. First, the algorithm that predicts the outcome of a case or assesses the potential risk of recidivism is usually not transparent about how it works and arrives at its result. Because of the black box problem,<sup>(4)</sup> the algorithm's decision will not be understandable.<sup>(5)</sup> Consequently, judges and parties, both plaintiffs and defendants, would not be able to review and comment on the algorithm's decisions because of this opaque mechanism. Second, judges do not know much about the capabilities and limitations of AI algorithms and are not well informed about their development and new legal technologies. Moreover, people often tend to follow the results of AI systems without verifying their accuracy.<sup>(6)</sup> An experiment by Salem et al. (2015) showed that people are willing to place too much trust in automated decision-making

(1) Nations U, 'Universal Declaration of Human Rights' (United Nations) <<https://www.un.org/en/about-us/universal-declaration-of-human-rights>> accessed 17 March 2022

(2) 'Article 6 of the European Convention on Human Rights', , Wikipedia (2022) <[https://en.wikipedia.org/w/index.php?title=Article\\_6\\_of\\_the\\_European\\_Convention\\_on\\_Human\\_Rights&oldid=1079331684](https://en.wikipedia.org/w/index.php?title=Article_6_of_the_European_Convention_on_Human_Rights&oldid=1079331684)> accessed 5 April 2022

(3) 'Sixth Amendment' (LII / Legal Information Institute) <[https://www.law.cornell.edu/constitution/sixth\\_amendment](https://www.law.cornell.edu/constitution/sixth_amendment)> accessed 6 April 2022

(4) kenton will, 'What Is a Black Box Model?' (Investopedia, 2022) <<https://www.investopedia.com/terms/b/black-box.asp>> accessed 6 April 2022

(5) Steponaite VK and Valcke P, 'Judicial Analytics on Trial: An Assessment of Legal Analytics in Judicial Systems in Light of the Right to a Fair Trial' (2020) 27 Maastricht Journal of European and Comparative Law 759 <<https://doi.org/10.1177/1023263X20981472>> accessed 6 April 2022

(6) Dymitruk M, 'The Right to a Fair Trial in Automated Civil Proceedings' (2019) 13 Masaryk University Journal of Law and Technology 27 <<https://journals.muni.cz/mujlt/article/view/11624>> accessed 7 April 2022

systems.<sup>(1)</sup> Finally, targeted surveillance by AI algorithms for predictive policing may negatively impact the right to presumption of innocence. Such tools select individuals as possible criminals, and these individuals may be overrepresented in the data because they use public services more than others, making their data available. Innocent people are considered suspicious and can be arrested without cause, even though there is no evidence that they have done anything illegal.

### **Safeguards and Recommendations**

- It is necessary to ensure that the AI system used is transparent. Judges, prosecutors, and defendants are owed an explanation for the automated decisions to justify them and be sure that no human rights are violated. Even if there is no explainable AI algorithm, there are still some short-term solutions that could help judges and prosecutors address the transparency problem, such as switching to AI algorithms that use publicly available data and moving away from opaque algorithms.
- Judges must not rely solely on the evidence provided by AI algorithms, but must seek to support it with other, more reasoned evidence.
- Law enforcement officials should be sufficiently qualified to use these algorithms before they are deployed in the system.
- The use of these algorithms should be regulated by laws. These laws should address the privacy of citizens and regulate how algorithms process people's data, when it is processed, and how privacy is guaranteed.
- The law must ensure that the public has the right to an explanation when they are affected by automated decisions, and it must also ensure that

---

(1) Salem M and others, 'Would You Trust a (Faulty) Robot? Effects of Error, Task Type and Personality on Human-Robot Cooperation and Trust', Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction (Association for Computing Machinery 2015) <<https://doi.org/10.1145/2696454.2696497>> accessed 6 April 2022

public interests outweigh the trade secret claims of private companies.

- Criminal justice experts should work along with data scientists to identify the algorithms used in the system and the factors the algorithms use in determining their results. This will help ensure that the algorithm complies with laws and human rights.

## References

1. Perry, W. L., McInnis, B., Price, C. C., Smith, S., & Hollywood, J. S. (2013). Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations.
2. Stroud, M. (2021, May 24). An automated policing program got this man shot twice. The Verge. Retrieved February 4, 2022, from <https://www.theverge.com/22444020/heat-listed-csk-entry>
3. Ferguson, A. G. (2016). Policing Predictive Policing.
4. Agbo-Ajala, O., & Viriri, S. (2020). Deeply Learned Classifiers for Age and Gender Predictions of Unfiltered Faces.
5. Wu, X., & Zhang, X. (2016). Responses to Critiques on Machine Learning of Criminality Perceptions
6. Bowyer, K. W., King, M., Scheirer, W., & Vangara, K. (2020). The Criminality from Face Illusion.
7. Oswald, M., Grace, J., Urwin, S., & Barnes, G. C. (2018). Algorithmic risk assessment policing models: lessons from the Durham HART model and ‘Experimental’ proportionality. *Information & Communications Technology Law*
8. Danziger, S., Levav, J., & Avnaim-Pesso, L. (2011). Extraneous factors in judicial decisions.
9. Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases.



10. Dymitruk, M. (2019). The Right to a Fair Trial in Automated Civil Proceedings.
11. Salem, M., Lakatos, G., Amirabdollahian, F., & Dautenhahn, K. (2015). Would You Trust a (Faulty) Robot? Effects of Error, Task Type and Personality on Human-Robot Cooperation and Trust.

