

RIDDP

**Gert Vermeulen, Nina Peršak
& Nicola Recchia** (Eds.)

Artificial Intelligence, Big Data and Automated Decision-Making in Criminal Justice

Revue Internationale de Droit Pénal
International Review of Penal Law
Revista internacional de Derecho Penal
Международное обозрение уголовного права
刑事法律国际评论
المجلة الدولية للقانون الجنائي
Revista Internacional de Direito Penal
Rivista internazionale di diritto penale
Internationale Revue für Strafrecht



AIDP – Association Internationale de Droit Pénal | The International Association of Penal Law is the oldest association of specialists in penal law in the world. Since 1924, it is dedicated to the scientific study of criminal law and covers: (1) criminal policy and codification of penal law, (2) comparative criminal law, (3) international criminal law (incl. specialization in international criminal justice) and (4) human rights in the administration of criminal justice. The Association's website provides further information (<http://www.penal.org>).

RIDP – Revue Internationale de Droit Pénal | The International Review of Penal Law is the primary publication medium and core scientific output of the Association. It seeks to contribute to the development of ideas, knowledge, and practices in the field of penal sciences. Combining international and comparative perspectives, the RIDP covers criminal law theory and philosophy, general principles of criminal law, special criminal law, criminal procedure, and international criminal law. The RIDP is published twice a year. Typically, issues are linked to the Association's core scientific activities, ie the AIDP conferences, Young Penalist conferences, world conferences or, every five years, the International Congress of Penal Law. Occasionally, issues will be dedicated to a single, topical scientific theme, validated by the Scientific Committee of the Association, comprising high-quality papers which have been either presented and discussed in small-scale expert colloquia or selected following an open call for papers. The RIDP is published in English only.

Peer review: All contributions are subject to double-layered peer review. The primary scientific and peer review responsibility for all issues lies with the designated Scientific Editor(s). The additional scientific quality control is carried out by the Executive Committee of the Editorial Board, which may turn to the Committee of Reviewers for supplementary peer review.

Disclaimer: The statements and opinions made in the RIDP contributions are solely those of the respective authors and not of the Association or MAKLU Publishers. Neither of them accepts legal responsibility or liability for any errors or omissions in the contributions nor makes any representation, express or implied, with respect to the accuracy of the material.

© 2021 Gert Vermeulen, Nina Peršak & Nicola Recchia (Editors) and authors for the entirety of the edited issue and the authored contribution, respectively. All rights reserved: contributions to the RIDP may not be reproduced in any form, by print, photo print or any other means, without prior written permission from the author of that contribution. For the reproduction of the entire publication, a written permission of the Editors must be obtained.

ISSN – 0223-5404
ISBN 978-90-466-1130-2
D/2021/1997/46
NUR 824
BISAC LAW026000

Maklu- Publishers

Somersstraat 13/15, 2018 Antwerpen, Belgium, info@maklu.be
Koninginnelaan 96, 7315 EB Apeldoorn, The Netherlands, info@maklu.nl
www.maklu.eu

USA & Canada

International Specialized Book Services
920 NE 58th Ave., Suite 300, Portland, OR 97213-3786, orders@isbs.com, www.isbs.com

Editorial Board

Executive Committee

General Director of Publications & Editor-in-Chief | Gert VERMEULEN, Ghent University and Institute for International Research on Criminal Policy, BE

Co-Editor-in-Chief | Nina PERŠAK, University of Ljubljana, SI
Editorial Secretary | Hannah VERBEKE, Ghent University, BE
Editors | Gleb BOGUSH, Moscow State University, RU | Dominik BRODOWSKI, Saarland University, DE | Juliette TRICOT, Paris Nanterre University, FR | Michele PAPA, University of Florence, IT | Eduardo SAAD-DINIZ, University of São Paulo, BR | Beatriz GARCÍA MORENO, CEU-ICADE, ES
AIDP President | John VERVAELE, Utrecht University, NL
Vice-President in charge of Scientific Coordination | Katalin LIGETI, University of Luxembourg, LU

Committee of Reviewers – Members | Isidoro BLANCO CORDERO, University of Alicante, ES | Steve BECKER, Assistant Appellate Defender, USA | Peter CSONKA, European Commission, BE | José Luis DE LA CUESTA, Universidad del País Vasco, ES | José Luis DíEZ RIPOLLÉS, Universidad de Málaga, ES | Antonio GULLO, Luiss University, IT | LU Jianping, Beijing Normal University, CN | Sérgio Salomão SHECAIRA, University of São Paulo and Instituto Brasileiro de Ciências Criminais, BR | Eileen SERVIDIO-DELABRE, American Graduate School of International Relations & Diplomacy, FR | Françoise TULKENS, Université de Louvain, BE | Emilio VIANO, American University, USA | Roberto M CARLES, Universidad de Buenos Aires, AR | Manuel ESPINOZA DE LOS MONTEROS, WSG and Wharton Zicklin Center for Business Ethics, DE – **Young Penalists** | BAI Luyuan, Max Planck Institute for foreign and international criminal law, DE | Nicola RECCHIA, Goethe-University Frankfurt am Main, DE

Scientific Committee (names omitted if already featuring above) – Executive Vice-President | Jean-François THONY, Procureur général près la Cour d'Appel de Rennes, FR – **Vice-Presidents** | Carlos Eduardo JAPIASSU, Universidade Estacio de Sa, BR | Ulrika SUNDBERG, Ambassador, SE | Xiumei WANG, Center of Criminal Law Science, Beijing Normal University, CN – **Secretary General** | Stanislaw TOSZA, Utrecht University, NL – **Secretary of Scientific Committee** | Miren ODRIOZOLA, University of the Basque Country, ES – **Members** | Maria FILATOVA, HSE University, RU | Sabine GLESS, University of Basel, CH | André KLIP, Maastricht University, NL | Nasrin MEHRA, Shahid Beheshti University, IR | Adán NIETO, Instituto de Derecho Penal Europeo e Internacional, University of Castilla-La Mancha, ES | Lorenzo PICOTTI, University of Verona, IT | Vlad Alexandru VOICESCU, Romanian Association of Penal Sciences, RO | Bettina WEISSER, University of Cologne, DE | Liane WÖRNER, University of Konstanz, DE | Chenguang ZHAO, Beijing Normal University, CN – **Associated Centers (unless already featuring above)** | Filippo MUSCA, Istituto Superiore Internazionale di Scienze Criminali, Siracusa, IT | Anne WEYENBERGH, European Criminal Law Academic Network, Brussels, BE – **Young Penalists** | Francisco FIGUEROA, Buenos Aires University, AR

Honorary Editorial Board - Honorary Director | Reynald OTTENHOF, University of Nantes, FR – **Members** | Mireille DELMAS-MARTY Collège de France, FR | Alfonso STILE, Sapienza University of Rome, IT | Christine VAN DEN WYNGAERT, Kosovo Specialist Chambers, NL | Eugenio Raúl ZAFFARONI, Corte Interamericana de Derechos Humanos, CR

Summary

Preface: Capabilities and Limitations of AI in Criminal Justice <i>by Gert Vermeulen, Nina Peršak and Nicola Recchia</i>	7
Setting the Scene	
Algorithmic Decisions within the Criminal Justice Ecosystem and their Problem Matrix, <i>by Krisztina Karsai</i>	13
AI and Big Data in Predictive Detection and Policing	
Applying the Presumption of Innocence to Policing with AI, <i>by Kelly Blount</i>	33
Click, Collect and Calculate: The Growing Importance of Big Data in Predicting Future Criminal Behaviour, <i>by Julia Heilemann</i>	49
Augmented Reality in Law Enforcement from an EU Data Protection Law Perspective: The DARLENE Project as a Case Study, <i>by Katherine Quezada-Tavárez</i>	69
On the Potentialities and Limitations of Autonomous Systems in Money Laundering Control, <i>by Leonardo Simões Agapito, Matheus de Alencar e Miranda and Túlio Felipe Xavier Januário</i>	87
Crimes Involving AI: Liability Issues and Jurisdictional Challenges	
AI Crimes and Misdemeanors: Debating the Boundaries of Criminal Liability and Imputation, <i>by Anna Moraiti</i>	109
AI and Criminal Law: The Myth of 'Control' in a Data-Driven Society <i>by Beatrice Panattoni</i>	125
The Impact of AI on Corporate Criminal Liability: Algorithmic Misconduct in the Prism of Derivative and Holistic Theories, <i>by Federico Mazzacova</i>	143
The Challenges of AI for Transnational Criminal Law: Jurisdiction and Cooperation <i>by Miguel João Costa and António Manuel Abrantes</i>	159
AI-Assisted and Automated Actuarial Justice or Adjudication of Criminal Cases	
Lombroso 2.0: On AI and Predictions of Dangerousness in Criminal Justice <i>by Alice Giannini</i>	179

The Use of AI Tools in Criminal Courts: Justice Done and Seen to Be Done?
by Vanessa Franssen and Alyson Berrendorf..... 199

Automated Justice and Its Limits: Irreplaceable Human(e) Dimensions of Criminal
Justice, *by Nina Peršak* 225

ALGORITHMIC DECISIONS WITHIN THE CRIMINAL JUSTICE ECOSYSTEM AND THEIR PROBLEM MATRIX

By Krisztina Karsai*

Abstract

This paper highlights the social-legal environment of criminal justice through identifying and defining the different needs and possibilities of deploying algorithmic decision-making solutions in the distinct stages of the criminal procedure. A peculiar paradox prevails in this area: although no comprehensive policy on the use of algorithms and algorithmic decision-making exists in the justice process, the application of tools using such technology is almost universal. The objective of this paper is to introduce the main challenges in this regard and to present arguments as to why the applications of algorithms within criminal justice is not evidential simply because technology enables it. The paper follows theoretical criminology methods and addresses issues and principles both from criminology and from criminal law perspectives. Six main criteria are identified, which support explaining both the lack of necessity and the lack of compliance with system-relevant values and characters of criminal justice regarding the application of algorithms. The following criteria are discussed: adaptation traps (the interplay perceived between algorithmizing data and information relevant for criminal justice); the myth of objective truth and of convincing the judge (identifying the main goal of the criminal procedure and describing the goals are to be achieved if algorithms are to play any role in the procedure); the very theoretic paradigms of criminal law and criminology (how these system-shaping paradigms will be eroded – or revolutionized – by algorithmic thinking); the immanent non-mathematisable values of criminal justice (how non-coded values can or cannot play a role in algorithmic solutions); the 'bad' subjectivity (whether or not subjectivity of the judge should be excluded), and the purity of the data (why specific data related to criminal justice are simply unusable as training datasets for algorithmic solutions).

1 Introduction and Objectives

This paper highlights the social-legal environment of criminal justice by identifying and defining the different needs and possibilities of deploying algorithmic decision-making solutions (ADM)¹ in the distinct stages of criminal procedures. The objectives of this paper are to introduce the main conceptual challenges in this regard and to present arguments as to why the application of algorithms is far from evidential within the criminal justice system, even though technology at hand would provide means for us to do so. The paper follows theoretical criminology methods and addresses issues and principles

* Full Professor of Law, University of Szeged, Faculty of Law and Political Sciences, Institute of Criminal Law and Criminal Science. For correspondence: <karsai.krisztina@juris.u-szeged.hu>.

¹ 'Algorithms need not be software: in the broadest sense, they are encoded procedures for transforming input data into a desired output, based on specified calculations. The procedures name both a problem and the steps by which it should be solved.' Tarleton Gillespie, 'The Relevance of Algorithms' in Tarleton Gillespie and others (eds), *Media Technologies Essays on Communication, Materiality, and Society* (MIT Press Scholarship Online 2014).

from both criminology and criminal law. The theoretical framework of this study is shaped by the (Central) European continental legal system and by both static and dynamic characteristics of criminal justice; therefore, application of my conclusions to other legal systems would first require wise adaptation and further research. Finally, I elaborate on some of the issues – or purported ‘traps’ that push most criminal law professionals into the so-called ‘uncanny valley’² about the rise of algorithmic or machine decision making, so that identification can facilitate further research and discussion of the issues involved.

In 1963, Lawlor stated that ‘given a chance, computers can help the legal profession in at least three very important ways. Computers can help find the law, they can help analyse the law and they can help lawyers and lower court judges to predict or anticipate decisions.’³ He believed that the trustworthy prediction of judicial decisions is dependent upon a scientific understanding of the functioning of the law and how facts and legal norms impact judges and judicial decisions. Indeed, almost sixty years later we have made immense advances in this field, but the latent infiltration of algorithmic solutions without clear scientific reasoning has guided development in a direction different from Lawlor’s prediction. Moreover, the Collingridge dilemma⁴ is undeniably evident – the dilemma of control over new technologies versus innovation has yet to be; rather, the development was simply allowed to flow. The ‘latent infiltration’ of such technologies is obvious.⁵ As expected, the technological and legal advent of ‘big data’ as a cultural, technological and scholarly phenomenon can be identified as key factors in the changed landscape. As Boyd and Crawford stated, ‘big data’ is the interplay of ‘[t]echnology: maximizing computation power and algorithmic accuracy to gather, analyse, link, and compare large data sets. Analysis: drawing on large data sets to identify patterns in order to make economic, social, technical, and legal claims. Mythology: the widespread belief that large data sets offer a higher form of intelligence and knowledge that can generate insights that were previously impossible, with the *aura of truth, objectivity, and accuracy*’.⁶

² ‘Uncanny valley’ is a specific psychological phenomenon in connection with robots: humans develop unsettling feelings if robots are too human-like, introduced in 1970. Bibi van den Berg B, ‘The Uncanny Valley Everywhere?’ in Simone Fischer-Hübner and others (eds), *Privacy and Identity Management for Life* (Springer 2010).

³ Reed C. Lawlor, ‘What computers can do: analysis and prediction of judicial decisions’ [1963] ABAJ 49.

⁴ ‘[A]ttempting to control a technology is difficult...because during its early stages, when it can be controlled, not enough can be known about its harmful social consequences to warrant controlling its development; but by the time these consequences are apparent, control has become costly and slow.’ The dilemma described by David Collingridge is cited and analysed by Audley Genus, Andy Stirling, ‘Collingridge and the dilemma of control: Towards responsible and accountable innovation’ [2017] Research Policy <<http://dx.doi.org/10.1016/j.respol.2017.09.012>>.

⁵ Rebecca Wexler, ‘Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System’ [2018] SLRV 70.

⁶ Danah Boyd, Kate Crawford, ‘Critical questions for Big Data. Provocations for a cultural, technological, and scholarly phenomenon’ [2012] *Information, Communication & Society* 662.

2 Algorithms in the Criminal Justice Ecosystem

The global conceptualisation of the broad infiltration and deep intrusion of algorithms into the everyday life of our societies speaks about ‘governing algorithms’⁷ and about ‘algorithmic governmentality’⁸ and calls for ethical and legal opportunities of applying ADM solutions within the justice sectors of societies. Moreover, ‘we are witnessing a gradual movement away from the traditional, retrospective, individualized model of criminal justice, which prioritises a deliberated and personalised approach to pursuing justice and truth, towards a prospective, aggregated model, which involves a more ostensibly efficient, yet impersonal and distanced, approach. “Actuarial justice” is based on a “risk management” or “actuarial” approach to the regulation of crime and the administration of justice.’⁹

The utilization of tools applying such technologies is almost universal, but their applications are fragmented. In many situations, the impact of the decision on people can be significant, especially as regards access to credit, employment,¹⁰ medical treatment¹¹ and judicial sentences,¹² among other things.¹³ Although the emergence of these devices is considered a technological advancement, we have failed to consider systemic differences and specifics, and have perhaps even neglected to realise them. The novelty of the problems combined with the exclusion of analogue solutions (not algorithmic decision-making) have led to the present situation of running into confrontation due to their presence and worldwide massive penetration. The application of ADM solutions can be fundamentally different depending on the stage of the criminal justice ecosystem (or ‘pipeline’) at which they are implemented. ADM solutions can be deployed for prevention, detection, investigation, prosecution of crimes, in courts’ proceedings and during the penal execution.

⁷ Solon Barocas and others, ‘Governing Algorithms: A Provocation Piece’ [2013] SSRN <<http://dx.doi.org/10.2139/ssrn.2245322>> accessed 15 October 2021.

⁸ Antoinette Rouvroy, Thomas Berns, ‘Gouvernementalité algorithmique et perspectives d’émancipation’ *Réseaux* [2013] 177 <<http://doi.org/10.3917/res.177.0163>> accessed 15 October 2021.

⁹ Amber Marks and others, ‘Automatic Justice? Technology, Crime, and Social Control’ in Roger Brownsword and others (eds) *The Oxford Handbook of the Law and Regulation of Technology* (OUP 2017).

¹⁰ Sifting through personal emails for personality profiling in Finland. See more in Brigitte Alfter and others, *Automating Society: Taking Stock of Automated Decision-Making in the EU* (AlgorithmWatch 2019)

¹¹ Allocating treatment for patients in the public health system in Italy. See more in Brigitte Alfter et al. (2019) 88.

¹² Police officers apply facial recognition algorithms to identify suspects (or victims) appearing in recordings from a crime scene, and judges use risk assessment ADM solutions for bail, sentencing and parole decisions based on an individual’s demographic characteristics and criminal history (and in order to predict recidivism).

¹³ Some further examples include the ADM solutions used by airport security for assessing risks posed by airline passengers (no-flight lists); the automated processing of traffic offences in France or algorithmic identification of children possibly vulnerable to be neglected (Denmark). Konrad Lischka, Anita Klinger, *Wenn Maschinen Menschen Bewerten* (Bertelsmann 2017).

3 Problem Matrix of the Criminal Justice Application of Algorithms

The purpose of criminal justice is to punish the perpetrator of a crime, ie one who violates the coexistence rules of society (retaliation), and to prevent that person or anyone else from committing another (new) crime (prevention goal and deterrence objective). At the system level, we believe this punishment ensures the functioning of society and the fight against crime and, where appropriate, the reduction of crime. It is undeniable that ADM solutions are to some extent present in the full spectrum of criminal justice, in the most general sense, helping human decision-making with a purpose appropriate to their use. In light of the Collingridge dilemma, however, we are here today because the latent infiltration of these devices has left us incapable of answering the original questions. Summarising these questions is appropriate, even if we are yet unable to reassuringly answer them. Of course, a narrative exists in which meaningful answers can be presented, but that does not fall within the frame of modern, rule of law criminal justice that humanity has spent the past two centuries developing and fighting for – instead, this is something completely different – and I will return to this at the end of this paper. In my opinion, the original questions are based along the lines of six main criteria; these problems stem directly from the purpose, function and internal structure of criminal justice as a social subsystem (six criteria of the problem matrix).

Algorithmic solutions can be used for description (recognition of patterns), for the exploration of correlations and for prediction. The resulting new information and facts and their application in the criminal justice chain can be identified along specific sub-objectives. The descriptive algorithm may be suitable for:

- identifying the perpetrator, the victim, the witness (facial recognition, linking personal datasets);
- establishing a pattern for certain offences committed;
- establishing a pattern for scenes;
- establishing probability of who the perpetrator of a committed crime was;
- determining how the court ‘usually’ decides.

By all means, the results of the descriptive analysis can be based on a variety of theoretical constructions, ie what data are used and why; thus, for example, past similar characteristics or identical conditions; factors to which the outcome is ‘presumably’ related, etc. The resulting outcome can serve as a basis for actions of the authorities, ie official decision-making, thus supporting a decision to order police action and law enforcement operations. The last two possible outcomes, although this has not yet appeared in criminal cases, would present the possibility for the results of the algorithm to support the court’s decision or possibly even replace it.

As such, based on established patterns, a predictive algorithm would be capable of foreseeing (for example):

- the likelihood and location of certain criminal offences;

- the likelihood of recidivism;
- the possible location of a sought person or object;
- the likelihood of becoming a victim;¹⁴
- the likelihood of becoming an offender;
- how the court is likely to decide;
- whether an incarcerated person is at risk for attempted suicide.¹⁵

While the results obtained in this way can be used in a manner similar to prescriptive systems, this method allows for the imposition of protective measures (specific crime prevention) and recidivism quality to be taken into consideration. On the other hand, classification can be distinct based on whether we are discussing big data-based algorithms or ‘traditional’ risk assessment-based, statistical algorithms, where the latter are characterised by the inclusion of variables and data verified by the criminological methodology in the algorithm, whereas in the former case, the algorithm works with data outside of criminal justice, and affirmations are often not delivered by criminology or another type of social research.

3.1 Adaptation traps

The basic premise of criminology is that crime as a social phenomenon and as the individual phenomenon (the commission of a crime by an individual) can depend on factors outside of criminal justice, ie it is based not only on what constitutes a crime, how thorough the police are or how well-functioning crime prevention is, but also on the individual’s own circumstances, which is why the use of big data (almost always personal or person-related data) seems revolutionary and promising. That is not the problem. However, the appearance of big data and the essentially unlimited possibilities for its analysis means that a wide variety of data can be examined collectively by algorithms, namely the exploration of data that has not been researched in relation to its own characteristics or to crime, due to the absence of a basic theoretical model. If by connecting large amounts of data, an algorithm is able to detect correlations, *we tend to think that this pattern indicates a causal link*. In many cases, however, accepted scientific methodology cannot be used to establish the correctness and explanatory power of the correlation, yet we still have a tendency to assume that if there is an abundance of data, there must also be a

¹⁴ According to Perry and his research team, predictive policing, as part of the criminal justice ecosystem, can be divided into four broad categories: 1. Methods for predicting crimes: These are approaches used to predict places where and times at which there is an increased risk of crime. 2. Methods for predicting offenders: These approaches identify individuals likely to become offenders in the future. 3. Methods for predicting perpetrators’ identities: These techniques are used to create profiles that accurately match likely offenders with specific past crimes. 4. Methods for predicting victims of crimes: Like those methods that focus on offenders, crime locations and times of heightened risk, these approaches are used to identify groups or, in some cases, individuals who are likely to become victims of crime. Walter R. Perry, “Hollywood: Predictive policing: The Role of Crime Forecasting in Law Enforcement Operations” [Rand Corporations 2013].

¹⁵ A historical overview provided by Kevin Ashley, ‘A Brief History of Changing Roles of Case Prediction in AI and Law’ [2019] *Law in Context* 36.

pattern (correlation). Moreover, the mentioned so-called 'aura'¹⁶ surrounds the myth of big data; however, this so-called aura is not scientifically proven.

The most significant and most common purpose of utilising ADM systems has been and continues to be the exploitation of algorithms to *improve human capabilities*, as such systems are capable of processing *much more* information in *much less time* than humans. The efficiency factor is thus the most paramount theoretical justification for the application of these systems. Coupled with this is the aforementioned – scientifically unproven – fallacy that more data will lead to the discovery of new layers of reality that have hitherto been hidden from the human mind. Drawing upon all of these factors, we may falsely arrive at the conclusion that the application of such systems will provide us with due assistance within criminal justice, so as to say that deconstructing the past (the criminal act committed) in an attempt to reconstruct the future (the application of punishment) will result in novel findings.

As I mentioned, ADM tools that are utilized in criminal justice have two functions. Even if covariance causes are unknown, identifying the correlations can be useful in building an understanding of the functioning of a given crime phenomenon. Then, coupled with the results of predictions based on the former, we may be able to influence the overall social patterns of crime (obviously in the direction of decline), and accordingly, from a crime control perspective, the lack of a scientific foundation could then be acceptable.

3.2 The myth of objective truth of the past and of the conviction of the judge

The conviction of the perpetrator and the coercive execution of penalties are based on the assumption that under criminal proceedings, the decision made is in accordance with the truth, that the crime was committed by the accused and was committed as stated in the judgement. Judicial certainty establishes truth, that is, it describes what happened. Unveiling the past in its entirety is not actually possible, nor can criminal justice rely on complete certainty, so the approach used is approximate: it requires judicial conviction (European systems) or certainty beyond reasonable doubt (Anglo-Saxon systems). And although professional regulations try to minimise the risk of its occurrence, the possibility of error is an inherent part of the system; on the one hand, in terms of the limits of perception about the past,¹⁷ and on the other, in terms of its (human) evaluation process. This, in fact, also means that the truth accepted by a judicial decision (the exploration of a past act) can also be perceived on a probabilistic basis in the sense of how close it is to the real events. However, we have no means for measuring this; the judge or jury making the decision must assume 100% certainty. Everyone else – depending on their procedural position or on the lack of it – would have a different estimate if asked. This also means that from an external objective point of view, a probabilistic decision is made in any case. If algorithms were developed to calculate the probability of commission by the accused

¹⁶ Boyd and Crawford, Critical questions for Big Data (n 6).

¹⁷ Fenyvesi Csaba, 'World Tendencies of Forensic Sciences in XXI Century Criminalist' [2014] Journal of Yaroslav the Wise National Law University 9.

based on all available data (data pertaining to the act, the perpetrator, investigative actions, etc.), then, depending on the scaling, the resulting outcomes would be, for example: “It’s more than 75% likely to have been committed by XY” or “it’s more than 50% likely to have been...” or possibly “it’s not more than 60% likely to have been...”. This is quite incongruous to the current paradigm of thinking, even if we see that a 100% conviction, as mentioned, actually refers to the subjective probability of the decision-making judge (or members of the body). In other words, from a different perspective, we can pose the following question: on an overall societal level, which probability would we rather accept to be the decisive one – the approximate truth established by the court or the probability offered by the algorithm?

For the sake of completeness, the criterion according to which we should choose among the two mentioned options of decisive probability should be accompanied by two supplementing observations. First, probabilistic decision-making appears prominently in criminal justice in the course of expert activity: as the relative probability of occurrence of the alleged fact or the probability of the plausibility of the claim¹⁸ (for example, in determining from which weapon a bullet was fired, whose fingerprint, whose DNA, who the father is, who signed it, whose voice can be heard on a recording, could the driver have stopped, had he been travelling at the permitted speed, what is the active substance content of the confiscated drug based, etc..). Moreover, on account of advances in modern technology, we are witnessing a ‘rise in the level of probability’¹⁹ in criminal justice. The development of this field has been facilitated by several notorious cases in which the probability established by the expert had filtered into the judge’s decision through the ‘the prosecutor’s fallacy or error’.²⁰

The second observation supplementing this critical criterion regards the fact that *in other branches of justice*, such as in property matters, certainty may not be required, in the light of general life experience, to rule out conflicting alternatives and the reasonable doubt they create. In civil litigation, depending on the subject matter of the litigation, the expected level of probability required for the formation of a judicial conviction may even vary from case to case.²¹ The role of probability in relation to causation, culpability and the amount of compensation is particularly characteristic for civil litigation.²²

¹⁸ Orbán József, ‘Comparison of Applicability of Bayesian and Frequentist Statistics in Criminal Law’ [2013] *Internal Security* 1; Michael J Saks, ‘History of the Law’s Reception of Forensic Science’ in Jay A Siegel, Pekka J Sauko (eds) *Encyclopaedia of Forensic Sciences* [2013].

¹⁹ Fenyvesi, ‘World tendencies’ (n 17).

²⁰ See more William C. Thompson, Edward L. Schumann, ‘The Prosecutor’s fallacy and the defence attorney’s fallacy’ [1987] *Law and Human Behavior* 3.

²¹ Mark Schweizer, ‘The civil standard of proof – what is it actually?’ [2016] *The International Journal of Evidence & Proof* Volume 3; Marco Di Bello, ‘Plausibility and probability in juridical proof’ [2019] *The International Journal of Evidence & Proof* 1; Alex Biedermann, Joelle Vuille, ‘The decisional nature of probability and plausibility assessment in juridical evidence and proof’ [2018] *International Commentary on Evidence* 1.

²² Julia Mortera, Philip Dawid, ‘Probability and Evidence’ and Basil C. Bitas, ‘Probability in the Courtroom’ in Tamás Rudas (ed.) *Handbook of Probability* [SAGE 2008].

To sum up, in relation to the theory of past truth and its role in criminal proceedings, it must be decided whether we seek to capture the probability offered by the algorithm, guided by the idea that it will be somehow more advantageous than the current paradigm in which the judicial decision (and any mistakes) is accepted as an approximate framework of past truth.

3.3 The twin towers of criminal science

Criminal law and criminology provide basic theoretical constructs that work towards the exclusion of algorithmic solutions, namely due to the existence of a general consensus and to a lack of refutation otherwise contradicting their inclusion. Here, I discuss four major paradigms.

Crime is a *complex social phenomenon*; this phenomenon represents the unity of two mutually inseparable yet conceptually distinct components. One element is the violation of criminal law norms, ie the human behaviour; the other is the person violating the criminal law norm. Quantitatively, the two components are also distinct. The reasons for this can be found both in the design of the legal systems and in the activities of law enforcement authorities. The difference between the two sides is strengthened by the fact that the number of persons and acts are not the same, because a person can commit several acts to be prosecuted, or an act that physically appears to be a single occurrence of crime can lead to more people being prosecuted.²³ It is a criminologically well-known fact that the whole crime, the crime as a whole, cannot be known (eg due to underreporting etc.) and therefore cannot be statistically captured to its full extent.²⁴ The known part of the crime, the crimes that have become known, and the range of offenders detected are the ones to whom statistical methods can be applied. The difference between total and known crime is *latent crime*, the extent, structure, temporal and spatial changes of which are unknown. The branch of criminology concerned with latent crime seeks to explore an unknown set, but even at a societal level, this is only suitable for an approximate description of the phenomenon. It is also clear that this is not a matter of technology – at least to the best of our knowledge today, as accuracy or certainty would only be provided by real-time recording and subsequent ‘traceability’ of all past events. The dark figure of crime (latent crime) therefore remains a black box²⁵, as algorithms, information and data cannot be extracted from this segment. This also means that future estimates based on revealed crime data are not based on reality, but only on a part of it, and are therefore

²³ See more John MacDonald, ‘*Measuring Crime and Criminality*’ (Routledge 2017).

²⁴ The opposite opinion will be represented by Perry and his team by mentioning that criminals and victims follow common life patterns; overlaps in those patterns indicate an increased likelihood of crime. Geographic and temporal features influence the where and when of those patterns; as they move within those patterns, criminals make ‘rational’ decisions about whether to commit crimes, taking into account such factors as the area, the target’s suitability and the risk of getting caught. Perry and others, *Hollywood...* (n 14).

²⁵ See more Ales Zavrsnik, ‘Algorithmic justice: Algorithms and big data in criminal justice settings’ EJC [2019] 1.

necessarily biased, so hardly any argument can be made in favour of accepting their ‘truth’.

Central to criminal justice is the crime, the commission of which triggers the machinery, which is then directly aimed at proving the commission of a crime and establishing responsibility. *Crime is a normative category*, having different content through time and space. What constitutes a crime in a given country is a moral social issue embedded in the ‘technical’ framework of criminal law, which sometimes has a political connotation. It follows that, although a significant proportion of behaviours treated as crimes here and now are and will always be crimes (murder, theft, sexual violence, robbery etc.); their legal classification may nonetheless change, not to mention other ‘non-classical’ offences (tax evasion, computer fraud, abortion, etc.). This means that even the data available on the crimes committed cannot form a future estimate precisely, because of the capacity of the legislator to bring about changes.²⁶

Modern (rule of law-based) criminal justice systems focus on *the perpetrator’s act*; criminal justice or the application of criminal law must focus on the act committed – the perpetrator and his or her characteristics cannot be decisive factors in assessment. Meanwhile, the dominance and acceptance of criminal law does not limit the use of ADM tools that do not predict an individual’s future actions. For example, the latest ‘star weapons’²⁷ in crime prevention are those that basically predict individual future crime based either on individual or community data and undermine the criminal law provisions on the criminal act. However, this development can logically be understood as meaning that if prediction were to be viewed as a means for crime prevention rather than law enforcement, then criminal procedural guarantees could be otherwise construed – but in reality, such reclassification could not be a means of circumventing human rights.

If we want an algorithm that works with factual data related to specific crimes and that it thus ‘identifies’ the perpetrator or ‘predicts’ the commission of a crime, we also face the limitations of perceiving the reality, as *each criminal case is unique*, and the cases (the acts of crime) are shaped by many unknown causes or variables. Individual cases with their unique sets of factual circumstances are less processable algorithmically. Here, too,

²⁶ Data are further distorted by internal theoretical and practical rules of criminal law. For example, formerly, the crime of child pornography (essentially any conduct related to recordings of under-18s) had for years been recorded in statistics in the order of tens of thousands of commissions in Hungary, because if the perpetrator had owned and possessed multiple recordings, then the case was recorded as counts based on the number of recordings that were seized, each recording being equivalent to one count. This practice was modified to tallying the number of minors involved – based on a statistical approach; from a statistical perspective, we could say that this type of crime has drastically decreased in Hungary – but in reality, it has not.

²⁷ Leo Kelion, ‘Crime Prediction Software “Adopted by 14 UK Police Forces”’ *BBC News* (UK, 4 February 2019) <<https://www.bbc.com/news/technology-47118229>> accessed 15 October 2021.

See Charles Raab and others, *Ethics Advisory Report for West Midlands Police* (The Alan Turing Institute 2017) <<https://www.turing.ac.uk/research/publications/ethics-advisory-report-west-midlands-police>> accessed 15 October 2021.

the question that could be raised is whether or not we would accept the tendency-based decisions of algorithms.

3.4 Non-mathematisable system-identical values

The operation of the rule of law, the peaceful coexistence of people and, of course, the changing world all hold many values that cannot be directly expressed in the specific legal norm in the “legal algorithm”.²⁸ At best – in terms of algorithmisation – the written constitution of a country or binding international legal instruments underpin these values. At worst, the historical constitution or customary international law would have to be added to the sources of interpretation in order for the ‘legal formula’ to properly function. Justice, fairness, fundamental and human rights are values that should be treated as *constant ‘variables’ in the ‘legal algorithm’*. Unfortunately, the content outlines of these values are not constant; at best, their core can be provided with an interpretation that is clear and thus can be subject to algorithmisation. If this path is followed, the recorded content can, of course, be coded into any ‘legal algorithm’, but this is likely to narrow the scope (content) of the principle.

Moreover, these values affect not only the meaning of the ‘legal algorithm’ but also the application of the ‘legal algorithm’ within criminal justice. In particular, this affects the right to individual liberty or human dignity; therefore, the algorithmisation of the requirement of due process seems to be an impossible undertaking.

It should also be pointed out that it was the software used in the US to predict recidivism (old-fashioned systems based on psychological risk analysis) that shed light on the fact that social coexistence values have developed in the 21st century that make the emerging correlations and patterns unacceptable, or at least render the use of the resulting correlations in decision-making (or in decision support) impossible. This is because these relationships are associated with protected characteristics such as gender, race, religion, and so on. And while correlations, and possibly causal relationships may be statistically true, in a social – and possibly political – context, we do not want this to be the case. Meanwhile, from a statistical perspective, algorithms are precisely designed to ‘discriminate’ on the basis of social consensus on certain values; in particular, the human rights requirement of the prohibition of discrimination makes certain forms of discrimination unacceptable.²⁹

In principle, of course, it is also conceivable that our algorithm can also recognise patterns that are output by other variables that are not directly coded into the ‘legal algorithm’ and thus have an impact on fairness, due process, and fundamental rights requirements in a particular case. The problem, however, is that compiling enough of such cases

²⁸ A legal norm as a tool of algorithmic problem solving is: if in case A (hypothesis), XY behaves in a manner of B (disposition), then C will be the consequence (sanction, compulsory measure which is a legal effect or legal disadvantage).

²⁹ Lilian Edwards, Michael Veale, ‘Slave to the Algorithm’ DLTR [2017] 12.

is highly unlikely, even in legal systems serving the peaceful coexistence of a larger population, so their use as learner data would not lead to adequate results.

Regarding this criterion, it is worth returning once again to the question of objective truth and the legal consequences associated with it. The 'litmus test'³⁰ of criminal justice is the achievement of the aforementioned judicial certainty, so the role of doubt is therefore crucial: if there is no certainty, if doubt remains, it can only lead to acquittal. However, this key feature of criminal justice cannot be used if the result of the calculation of the algorithm is scalable, ie if it shows the result as a percentage or on any scale instead of a clear 'yes' or a 'no' (eg the probability of guilt or the probability of the crime having been committed, or perhaps whether the person perpetrated the act or not). Under the current principle (if doubt exists, conviction will not be made) would certainly not be applicable in such a scalable paradigm, but this could be circumvented by still accepting some 'degree' of quantified doubt within 'judgement certainty'. This would mean, for example, an 80% or 90% probability of a conviction.³¹

3.5 The 'bad' subjectivity

Exclusion of the subjective component of the human factor in relation to automatic data processing and consequent decision-making may arise as an advantage. Throughout the entire ecosystem of the criminal justice, achieving the set goals requires a series of human decisions.

More specifically, a judicial decision is a human decision, so it is evident that the subject of the judge influences the content of his or her decision. An important system-shaping element is that the legal education, the professional conditions, and the socialisation of becoming a judge guarantee non-subjective professionalism; therefore, it is assumed that undesirable subjectivity does not appear in decisions. It is critical to clearly define what we consider to be a subjective component that we would prefer to banish from algorithmic decision-making. The subjectivity of the judge — in modern criminal justice — is key to humanity, values that are central to democracies based on the rule of law. Emphasis should be placed on the importance of life experience for the judiciary, which comprises the totality of the actions or other manifestations of different people observed in different life situations and includes their comparison and the ability to build upon and draw conclusions from these. The Subject-Judge helps to recognise the problem and then translate the decision into real life. The judge mediates community content and community values and integrates the procedure and the decision into the social coexistence. Based on all this, the subjectivity of the judge – the human decision-maker – actually serves as a kind of control or a factor that allows for the decision to remain within a lucid and logical framework and enables the decision to be made at all. On the other hand, clearly other features of subjectivity also influence the decision, but not at the professional level.³² The

³⁰ Zavrsnik, Algorithmic justice (n 25).

³¹ Further details see Marco Di Bello, 'Trial by statistics: is a high probability of guilt enough to convict?' *Mind* [2019] 512.

³² See more, for example Tania Sourdin, 'Judge v. robot?' [2018] *UNSW Law Journal* 4.

basic premise of modern criminal procedure is the exclusivity and omnipotence of the intellect, along with the promise that there will be no room for emotion alongside reason.³³

Thus, the idea that automatic decision-making could eliminate the 'bad' subjective component is understandable. However, this is neither practical nor possible in our current context of criminal justice, especially in the phases of the criminal procedure that follow the detection stage. For if this were the goal, we would then merely replace the subject-relevant psychological risk of judicial decision-making with the risk logic of algorithms. This, of course, could be pure objectivity and thus an acceptable new paradigm, but if algorithms in criminal justice cannot work with pure data (see the next criterion – under section 3.6.), no added value would result from swapping the two types of risk. Furthermore, the possibility of judicial discretion capable of dealing with the uniqueness of cases would be lost, and the consideration of the essential values already mentioned in the fourth criterion would be undermined.

Professional competence is treated as an axiom of judicial (human) decision-making, as one of the most important safeguards and, at the same time, as a key element of a fair, non-arbitrary and humane justice system. One might also think that if this component were taken away from the formula, incompetent and unacceptable decisions would be made. However, the combination of probabilistic decision-making and criteria that focus on the human decision-maker yields surprising results that may override this paradigm, where appropriate. Scientific research also deals with whether it is possible to foresee the expected decision of a court, even without a detailed legal examination of the cases, ie whether the professional component in decision-making could as such be waived.

In a 2016 study, decisions of the European Court of Human Rights (ECtHR) were examined using the *natural language processing method* and, based on these, the researchers made predictive conclusions about the decisions, the existence of the alleged violation or its exclusion.³⁴ Research was carried out on Articles 3, 6 and 8 of the ECHR in approximately 600 cases: relevant information and textual relationships were filtered out of the text file and transformed into learning data, and then the output became the binary code of the decision, ie whether there was an infringement in that case. This was then compared to the actual decision, based on which it was concluded that the algorithm worked with an average accuracy of 79% (in an average of 79 cases out of 100 cases, the decision of the algorithm was in line with the decision made by the human judge), which is a fairly good result. Interesting results could be acquired through examination of the cases that received a 'wrong' prediction, and whether the conditions that set them apart from

³³ See more Russel Cropanzano and others, '*Social Justice and the Experience of Justice*' [Routledge 2011]. Moreover, in a research study published in 2017, the researchers used the US Supreme Court as an example and demonstrated that the judges implicitly reveal their leanings during oral arguments, even before arguments and deliberations had been concluded. Bryce J. Dietrich and others, 'Emotional Arousal Predicts Voting on the U.S. Supreme Court' [2018] *Political Analysis* 2.

³⁴ Nikolaos Aletras and others, 'Predicting judicial decisions of the European Court of Human Rights' *PeerJ Computer Science* [2016] <<https://doi.org/10.7717/peerj-cs.93>> accessed 1 October 2021.

the rest are identifiable. Unusual constellations may result depending on the outcome of the study, for example, cases in which a pattern could not be determined by the algorithm, and with this, whether some circumstances had existed that 'diverted' the decision (eg, political or 'harm-reducing' is a less legitimate argument).

A surprisingly similar study, published in 2017, is also worth mentioning. The study concerned the decisions of the US Supreme Court, and did not involve examining the legal reasoning of the underlying cases.³⁵ The researchers examined the votes of judges in 240,000 decisions (28,000 cases between 1816 and 2015), judicial factual data (subject matter, fact of cited legislation, date of submission, the deciding court, procedural acts in the main proceedings, etc.), judges' appointment data (identifiable political ideology), and decision-making characteristics (such as the likelihood of a dissenting opinion, etc.). The applied algorithm estimated the votes of each judge with an accuracy of almost 70–72%.

Professionalism and professional competence, as key components of human judicial decision-making, were not included in the variables in a targeted and meaningful way in any of the mentioned research. And yet, in most cases, the algorithm resulted in predictions in line with the judicial decisions. This can be a good basis for further research and a possible paradigm shift in machine decision-making and leaves us to decide whether – combining these results with the second criterion – the 72% or 79% accuracy can be a rate sufficient enough to replace decision-making. If based on the studies, we were to accept those rates and that decisions made by the algorithms are to a vast degree the same as those made by the human court, would we forgo court decision-making (efficiency, human resources, cost factors, time factor)?

These research studies were conducted with non-criminal case analysis; moreover, the algorithms delivered predictions of courts, whereas the examination of these was limited to the questions of law – as both the ECtHR and the US Supreme Court cases were aimed at reviewing the legal compliance of an earlier decision. However, the discovery and identification of relevant facts is also necessary for the decisions of the ECtHR and the US Supreme Court, and thus is still comparable to judicial activity in criminal matters. Considering these, the degree of accuracy of the research results becomes surprising – these are the results that *could be used to resolve cases* if the second criterion of the matrix discussed here is resolved and accepted. The question of what would happen if we were to compare the results of algorithms with different logic working with the same data set will continue to remain open for some time to come. It is also questionable whether algorithms for predicting the decisions of courts dealing with facts and law and those of courts dealing with questions of law alone should choose different paths, or whether the difference between the two types of judicial activity necessarily disappears during the process of algorithmisation.

³⁵ Daniel Martin Katz and others, 'A General Approach for Predicting the Behaviour of the Supreme Court of the United States' [2017] *Plos ONE* 4.

3.6 Purity of algorithms – ‘you are what you eat’

The responses that result from the analysis of data and information (to support or replace human decisions) depend on the input data. For both algorithms following traditional methodologies and big data-based ADM solutions, the key issue is what data the algorithm should be allowed to work with (ie learn from). It is a basic requirement that both the training and further datasets used be clear and unbiased; otherwise, the output, either the pattern or the prediction, will certainly be objectionable. Within the justice field, this requirement raises several issues that are rather difficult to resolve.

If the data used to ‘feed’ the algorithm are the previous court decisions themselves (meaning human decisions), then we assume and accept that all previous judicial (or possibly other official) decisions were legally correct, as we allow the algorithm to be based on drawn patterns. This starting point is obviously correct in legal terms, but a retrospective *possibility to remedy* errors exists in all legal systems, typically in the form of extraordinary legal remedies, therefore, changes as subsequent rewritten input modifies the database and changes the output. And this can have an impact on the decisions based on it, so such solutions should hardly be avoided.

If past judicial decisions serve as the basis for the use of the ADM tool in criminal justice, the patterning potential of the *subjective* factors that may have appeared in past judicial decisions (discrimination, racism, etc.) may be filtered through the ‘uncovered’ contexts of the algorithm, ie we would exclude the infiltration of the human subject from the individual case, but allow it in its cumulative effect.³⁶ A ‘shining’ example of this is the software COMPAS, which is used in many US jurisdictions for predicting reoffending patterns and to support judicial sanctioning. COMPAS was accused of *being biased* against black defendants because it classified a greater share of black defendants as high-risk reoffenders than white defendants. Meanwhile, the algorithm assigned defendants scores from 1 to 10 that indicated how likely they were to reoffend based on more than 100 factors, including age, sex, and criminal record, while race was not used (!) as an indicator.³⁷ And yet, the differences incurred may have been the result of an interplay between the indicators, which represented (perhaps biased) former human decisions as criminal records or prior arrests (eg heavier policing in predominantly black neighbourhoods in certain areas).

The question of *data distillation* as a fundamental activity in any database construction may arise. However, if the data entered as input are judicial decisions made based on the appropriate rules, representing ‘truth’ and legal correctness, it would be challenging to find a legitimate basis for cleaning the data.

³⁶ Zavrsnik, Algorithmic justice (n 25).

³⁷ Anthony W. Flores and others, ‘False Positives, False Negatives, and False Analyses’ *Community Resources for Justice (US)* 2017 <https://www.crj.org/assets/2017/07/9_Machine_bias_rejoinder.pdf> accessed 15 October 2021.

Changing the reality of social coexistence is accompanied by a change in the law; it seems easy to incorporate legislative changes into the algorithms, but legal practice will not change, as the algorithms will not look for a new direction or new interpretation based on the 'old' pattern. This type of algorithmisation *freezes the practice of law* because even if human factors remain in the decision-making process, they do not receive a new or different impulse (the ADM solution provides results shaped by the 'old' pattern). And if human factors were to be fully excluded, the 'algorithmic jurisprudence' necessarily remains unchanged; it will no longer be possible to change it in an organic way as adapted by judges to the changing world.

4 Further Discussion

The present problem matrix contains *inherently interrelated factors* – the professional and scientific paradigms of the rule of law in criminal justice fundamentally lead the way, ie whether and to what extent there is room for algorithmic decision-making throughout the chain. It is also conceivable that, if we accept the correctness of previous decisions, the algorithms can ponder the chances of litigants winning the lawsuit, and the law will allow the parties to accept the higher probability and give legally binding status to the decisions so calculated. This, of course, does not override the specifics of criminal justice, but it can bring significant efficiencies and better resource management in other areas of law. The disengaged capacity can then be used in the human-intensive decision-making processes of criminal justice. Of course, it is also possible that the criteria problems discussed should be bridged through other solutions with a different basic paradigm and *allow the twin towers of criminal science to collapse*, in order to replace them with a new philosophy and a new type of criminal justice. It is possible that in the new paradigm, our knowledge of crime will be provided by databases, the relationships between information or facts will be delivered by algorithms, crime action will shift towards prediction and crime prevention will be the main activity. If this is not successful, then automated justice will make the judgment.³⁸ However, we have yet to arrive there. Until we can deliver proper answers to the questions raised by the stealth infiltration of algorithms, my position is the non-application of algorithmic systems within the criminal justice ecosystem (pipeline), and I propose for academia and legislative bodies to re-examine (in their own jurisdiction) the application of any of ADM solutions in the light of these criteria and formulate doubts and new limitations they may bear.

References

Aletras N, Tsarapatsanis D, Preoțiu-Pietro D and Lampos V, 'Predicting judicial decisions of the European Court of Human Rights' (2016) PeerJ Computer Science' <<https://doi.org/10.7717/peerj-cs.93>>

³⁸ See Završnik, Algorithmic justice (n 25).

Alfter B, Müller-Eiselt R and Spielkamp M, 'Automating Society: Taking Stock of Automated Decision-Making in the EU' (AlgorithmWatch 2019) <https://algorithmwatch.org/de/wp-content/uploads/2019/02/Automating_Society_Report_2019.pdf>

Ashley KD, 'A Brief History of Changing Roles of Case Prediction in AI and Law' [2019] *Law in Context*

Barocas S, Hood S and Malte Ziewitz A, 'Governing Algorithms: A Provocation Piece' [2013] <<https://www.semanticscholar.org/paper/Governing-Algorithms%3A-A-Provocation-Piece-Barocas-Hood/5a518d93366180456130e7d003b4aaf3a0a2bae7>>

Biedermann A and Vuille J, 'The Decisional Nature of Probability and Plausibility Assessment in Juridical Evidence and Proof' (2018) 16 *International Commentary on Evidence* 1.

Bitas BC, 'Probability in the Courtroom' in Rudas T (ed.), *Handbook of Probability: Theory and Applications* (SAGE 2008)

Boyd D and Crawford K, 'Critical questions for Big Data. Provocations for a cultural, technological, and scholarly phenomenon' (2011) 15 *Information, Communication & Society* 662

Cropanzano R, Stein JH and Nadisic T, *Social Justice and the Experience of Justice* (Routledge 2011)

Di Bello M, 'Plausibility and Probability in Juridical Proof' (2019) 12 *The International Journal of Evidence & Proof* 161

— — 'Trial by Statistics: Is a High Probability of Guilt Enough to Convict?' (2019) 128 *Mind* 1045

Dietrich BJ, Enos RD and Sen M, 'Emotional Arousal Predicts Voting on the U.S. Supreme Court' (2018) 27 *Political Analysis*

Edwards L and Veale M, 'Slave to the Algorithm: Why a "right to an explanation" is probably not the remedy you are looking for' (2017) 16 *Duke Law & Technology Review* 18

Fenyvesi C, 'World Tendencies of Forensic Sciences in XXI Century' (2014) *Criminalist. Journal of Yaroslav the Wise National Law University*, Apostille Publishing House LLC, Ukraine 9/2014, 10-21

Flores AW, Lowenkamp CT and Bechtel K, 'False Positives, False Negatives, and False Analyses: A Rejoinder to "Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And it's Biased Against Blacks"' *Community Resources for Justice* (US) 2017 <<http://www.crj.org/>>

Genus A and Stirling A, 'Collingridge and the Dilemma of Control: Towards Responsible and Accountable Innovation' [2017] *Research Policy* <<http://dx.doi.org/10.1016/j.respol.2017.09.012>>

Gillespie T, 'The Relevance of Algorithms' in Gillespie T, Boczkowski PJ and Foot KA (eds), *Media Technologies Essays on Communication, Materiality, and Society* (MIT Press Scholarship Online 2014), 167-94 <<https://www.researchgate.net/publication/281562384>>

Katz DM, Bommarito MJ, Blackman J, 'A General Approach for Predicting the Behaviour of the Supreme Court of the United States' (2017) 12 *Plos ONE* 1

Lawlor RC, 'What Computers Can Do: Analysis and Prediction of Judicial Decisions' (1963) 49 *American Bar Association Journal* 337

Lischka K and Klingel A, 'Wenn Maschinen Menschen bewerten'. Internationale Fallbeispiele für Prozesse algorithmischer Entscheidungsfindung. Arbeitspapier, May 2017, Bertelsmann Stiftung. DOI 10.11586/2017025

Marks A, Bowling B and Keenan C, 'Automatic Justice? Technology, Crime, and Social Control' in Brownsword R, Scotford E and Yeung K (eds), *The Oxford Handbook of the Law and Regulation of Technology* (OUP 2017)

MacDonald J, *Measuring Crime and Criminality* (Routledge 2017)

Mortera J and Dawid P, 'Probability and Evidence' in Rudas T (ed.), *Handbook of Probability: Theory and Applications* (SAGE 2008)

Orbán J, 'Comparison of Applicability of Bayesian and Frequentist Statistics in Criminal Law' (2013) 5 *Internal Security* 197

Perry WL, McInnis B, Price CC, Smith SC and Hollywood JS, *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations* (Rand Corporation 2013)

Rouvroy A and Berns T, 'Gouvernementalité algorithmique et perspectives d'émancipation' (2013) 1 *Réseaux* (No 177), 163 (DOI 10.3917/res.177.0163) Translated by Elizabeth Libbrecht [Algorithmic Governmentality and Prospects of Emancipation]

Saks MJ, 'History of the Law's Reception of Forensic Science' in Siegel JA and Sauko PJ (eds), *Encyclopaedia of Forensic Sciences* (Academic Press 2013)

Schweizer M, 'The Civil Standard of Proof – What is It Actually?' (2016) 20 *The International Journal of Evidence & Proof* 217

Sourdin T, 'Judge v. robot? Artificial intelligence and judicial decision making' (2018) 41 *UNSW Law Journal* 1114

The Alan Turing Institute, 'Report on Ethics Advisory Report for West Midlands Police' (2017) <<https://www.turing.ac.uk/research/publications/ethics-advisory-report-west-midlands-police>>

Thompson WC and Schumann EL, 'The Prosecutor's Fallacy and the Defence Attorney's Fallacy' (1987) 11 *Law and Human Behavior* 167

Van den Berg B, 'The Uncanny Valley Everywhere?' in Fischer-Hübner S, Duquenoy P, Hansen M, Leenes R and Zhang G (eds), *Privacy and Identity Management for Life* (Springer 2010)

Wexler R, 'Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System' (2018) 70 *Stanford Law Review* 1343

Zavrsnik A, 'Algorithmic Justice: Algorithms and Big Data in Criminal Justice Settings' (2019) *European Journal of Criminology* <<https://doi.org/10.1177/1477370819876762>>

Artificial intelligence (AI) is impacting our everyday lives in a myriad of ways. The use of algorithms, AI agents and big data techniques also creates unprecedented opportunities for the prevention, investigation, detection or prosecution of criminal offences and the efficiency of the criminal justice system. Equally, however, the rapid increase of AI and big data in criminal justice raises a plethora of criminological, ethical, legal and technological questions and concerns, eg about enhanced surveillance and control in a pre-crime society and the risk of bias or even manipulation in (automated) decision-making. In view of the stakes involved, the need for regulation of AI and its alignment with human rights, democracy and the rule of law standards has been amply recognised, both globally and regionally. The lawfulness, social acceptance and overall legitimacy of AI, big data and automated decision-making in criminal justice will depend on a range of factors, including (algorithmic) transparency, trustworthiness, non-discrimination, accountability, responsibility, effective over-sight, data protection, due process, fair trial, access to justice, effective redress and remedy. Addressing these issues and raising awareness on AI systems' capabilities and limitations within criminal justice is needed to be better prepared for the future that is now upon us.

This special issue on 'Artificial intelligence, big data and automated decision-making in criminal justice' comprises topical and innovative papers on the above issues, centred around AI and big data in predictive detection and policing, liability issues and jurisdictional challenges prompted by crimes involving AI, and AI-assisted and automated actuarial justice or adjudication of criminal cases.

Gert Vermeulen is Senior Full Professor of European and international Criminal Law and Data Protection Law, Director of the Institute for International Research on Criminal Policy (IRCP), Di-rector of the Knowledge and Research Platform on Privacy, Information Exchange, Law Enforcement and Surveillance (PIXLES) and Director of the Smart Solutions for Secure Societies (i4S) business development center, all at Ghent University, Belgium. He is also General Director Publications of the AIDP and Editor-in-Chief of the RIDP.

Nina Peršak is Scientific Director and Senior Research Fellow, Institute for Criminal-Law Ethics and Criminology (Ljubljana), Advanced Academia Fellow (CAS Sofia), Member of the European Commission's Expert Group on EU Criminal Policy, Independent Ethics Adviser, and Co-Editor-in-Chief of the RIDP.

Nicola Recchia is Postdoc Researcher in Criminal Law at the Goethe-University Frankfurt, Germany. He is also member of the Young Penalists Committee and of the Scientific Committee of the AIDP.

www.maklu.be
ISBN 978-90-466-1130-2

