Vol. 92 issue 1, 2021

RIDP

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

MAKLU

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.

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

Summary

THE USE OF AI TOOLS IN CRIMINAL COURTS: JUSTICE DONE AND SEEN TO BE DONE?

By Vanessa Franssen[*] and Alyson Berrendorf[**]

Abstract

Artificial intelligence (AI) is impacting all sectors of society these days, including the criminal justice area. AI has indeed become an important tool in this area, whether for citizens seeking justice, legal practitioners or police and judicial authorities. While there is already a large body of literature on the prediction and detection of crime, this article focuses on the current and future role of AI in the adjudication of criminal cases. A distinction is made between AI systems that facilitate adjudication and those that could, in part or wholly, replace human judges. At each step, concrete examples are given, and it is evaluated what are, or could be, the advantages and disadvantages of such systems when used in criminal courts.

1 Introduction

Artificial intelligence (AI) has become a subject that is difficult to ignore these days. All sectors of society seem to be impacted: the financial sector, health care, national security, public administration, (social) media, and even the legal area. In this area in particular, AI seems to have become a major and perhaps unavoidable tool, whether for citizens seeking justice, legal practitioners or police and judicial authorities.

Without a doubt, this new form of intelligence makes it possible to perform more quickly and effectively certain basic and tedious tasks that, prior to the arrival of AI systems, used to consume a lot of time. Data analysis and cross-referencing have achieved results in terms of performance and accuracy that would have been hard to imagine some years ago. For instance, predictive policing tools such as PredPol (for Predictive Policing),[1] Pal-

[*] Professor of criminal law and criminal procedure, University of Liège; Affiliated Senior Researcher, KU Leuven. For correspondence: <vanessa.franssen@uliege.be>.

[**] PhD Candidate (Research fellow F.R.S.-FNRS) in the field of criminal law and criminal procedure, University of Liège. For correspondence: <alyson.berrendorf@uliege.be>.

[1] For some basic information about the functioning of this system, see PredPol, 'What. Where. When' <https://www.predpol.com/> accessed 14 July 2021.

antir,[2] CAS (for Criminality Anticipation System)[3] or CloudWalk,[4] are already widely used by police authorities around the globe to predict where future crime will take place.[5] Supporters of such technologies highlight the effectiveness and efficiency of predictive policing, enabling police authorities to better target their interventions and even anticipate certain criminal phenomena, whereas critics argue that these systems are less effective than they appear to be, resulting in a self-fulfilling prophesy and targeting mainly lower social classes and ethnic minorities,[6] and that they contain biases and have disempowering effects.[7] Moreover, a human decision, based on digitized and potentially reductive information, remains necessary.[8] Regardless of who is right or wrong in this particular debate, the development of AI tools creates high expectations, also in the field of criminal justice, as they can support and perhaps even replace human actors in preparing and trying criminal cases.

This raises the question whether AI tools will enable us to achieve, if not 'ultimate' justice, at least greater justice at a less high cost and more equity. The development of AI indeed looks promising in terms of speed, efficiency and effectiveness of justice, but it may well prove to be Pandora's box. What is the potential of AI tools for adjudication by criminal courts and what are the pitfalls? These are the questions that we will address in this short article, thereby explicitly focusing on the *adjudication* of criminal cases, not on the prediction and detection of crime as there is already a large body of literature on this

[2] See eg Pascal Hérard, 'Technologies de prédiction du crime: Palantir a scruté les citoyens de la Nouvelle-Orléans en secret pendant 6 ans' (*TV5Monde*, 3 March 2018) <https://information.tv5monde.com/info/technologies-de-prediction-du-crime-palantir-scrute-les-citoyens-de-la-nouvelle-orleans-en> accessed 14 July 2021; Ali Winston, 'New Orleans ends its Palantir predictive policing program' (*TheVerge*, 15 March 2018) <https://www.theverge.com/2018/3/15/17126174/new-orleans-palantir-predictive-policing-program-end> accessed 14 July 2021.

[3] See eg Serena Oosterloo and Gerwin van Schie, 'The Politics and Biases of the "Crime Anticipation System" of the Dutch Police' in Jo Bates, Paul D. Clough, Robert Jäschke and Jahna Otterbacher (eds), *Proceedings of the International Workshop on Bias in Information, Algorithms, and Systems* (CEUR Workshop Proceedings 2018) 30, 41.

[4] See eg Yuan Yang, Yingzhi Yang and Sherry Fei Ju, 'China seeks glimpse of citizens' future with crime predicting AI' (*Financial Times*, 23 July 2017) <https://www.ft.com/content/5ec7093c-6e06-11e7-b9c7-15af748b60d0> accessed 14 July 2021.

[5] See eg Walter L Perry, Brian McInnis, Carter C Price, Susan Smith and John S. Hollywood (2013), 'Predictive Policing. Forecasting Crime for Law Enforcement' (*Rand Research Brief*, 2018) <https://www.rand.org/ pubs/research_briefs/RB9735.html> accessed 15 July 2021.

[6] See eg Cathy O'Neil, *Algorithmes – La Bombe à retardement* (Les Arènes 2018) 135-137; Mehdi Harmi, 'Algorithmes – Peuvent-ils prédire l'avenir ?' (2021) Science et Vie 76; Interview with Angèle Christin: Hubert Guillaud, 'La justice prédictive (2/3) : prédictions et régulations' (*Le Monde*, 13 September 2017) < https://www.lemonde.fr/blog/internetactu/2017/09/13/la-justice-predictive-23-predictions-et-regulations /> accessed 5 July 2021.

[7] See eg Rosamunde Elise van Brakel, 'Pre-emptive Big Data Surveillance and its (Dis)empowering Consequences: The Case of Predictive Policing' in Bart van der Sloot, Dennis Broeders and Erik Schrijvers (eds), *Exploring the Boundaries of Big Data* (Amsterdam University Press 2016) 117, 141.

[8] Emre Bayamlioglu and Ronald Leenes, 'The "rule of law" implications of data-driven decision-making: a techno-regulatory perspective' (2018) Law, Innov Technol 295, 313.

subject.[9] To this end, we will make a distinction between AI systems that facilitate adjudication (Part 3) and those that could, in part or wholly, replace human judges (Part 4). At each step, we will give some concrete examples and evaluate what are, or could be, the advantages and disadvantages of such systems in the area of criminal justice. However, before doing so, it may be good to clarify some elementary concepts related to AI, necessary for a good understanding of the problems inherent to software developed for and used in criminal justice (Part 2).

2 Some Basic Notions

Despite the growing interest in AI, at all levels, there still exists quite some confusion about certain basic concepts.[10] In fact, this is not entirely surprising, considering the lack of international consensus on what exactly constitutes an AI system. At EU level, recent attempts have been made to define more clearly this notion. According to the High-Level Expert Group on AI set up by the European Commission, AI systems are:

> software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal (…).[11]

In its proposal for an AI Act,[12] made public in February 2021, the European Commission put forward a more precise and simpler definition of AI systems. While one may welcome the Commission's attempt to define AI in a narrower way, the proposed definition has also provoked quite some critical reactions.[13]

---

[9] See eg Neil Shah, Nandish Bhagat and Manan Shah, 'Crime forecasting: a machine learning and computer vision approach to crime prediction prevention' (2021) 4 VCIBA 9; Rohit Patil, Muzamil Kacchi, Pranali Gavali, Komal Pimparia, 'Crime Pattern Detection, Analysis & Prediction using Machine Learning' (2020) 7 IRJET 119; Shruti Gosavi and Shraddha Kavathekar, A Survey on Crime Occurrence Detection and prediction Techniques' (2018) 8 Int. j. eng. technol. manag. appl. sci. 1405.

[10] Rembrandt Deville, Nico Sergeyssels and Catherine Middag, 'Basic concepts of AI for legal scholars' in Jan De Bruyne and Cedric Vanleenhove (eds), *Law & Artificial Intelligence* (Intersentia 2021) 1.

[11] High-Level Expert Group on Artificial Intelligence, 'Ethics Guidelines for Trustworthy AI' (*European Commission*, 8 April 2019) <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai> accessed 5 July 2021.

[12] European Commission, 'Proposal for a regulation of the European parliament and of the council laying down harmonized rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts', COM(2021), Brussels.

[13] See eg Michael Vaele and Frederik Zuiderveen Borgesius, 'Demystifying the Draft EU Artificial Intelligence, Analysing the good, the bad, and the unclear elements of the proposed approach' (2021) 22 Computer L. Rev. Int. 97; Nathalie Smuha, Emma Ahmed-Rengers, Adam Harkens, Wenlong Li, James MacLaren, Riccardo Piselli and Karen Yeung, 'How the EU can achieve Legally Trustworthy AI: A Response to the European Commission's Proposal for an Artificial Intelligence Act' (*SSRN*, 5 August 2021) <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3899991&download=yes> accessed 19 October 2021.

While the debate on the definition of AI is ongoing, one can clearly observe two approaches to AI in the scientific literature, namely 'knowledge-based learning' and 'data-based learning'.

The first approach is based on explicit knowledge in the form of a model:

> Typically, an expert in the field trie[s] to pour his knowledge into a model (eg a set of rules, patterns or logical statements). This model [i]s subsequently implemented as a series of instructions – and thus as an algorithm – in the machine to obtain its goal.[14]

Under this approach, the input of the developer or programmer of the AI system is thus detrimental for its output.

In contrast, under the data-based approach, the system itself recognizes patterns, based on the numerous examples of inputs received. This approach has become predominant today and is obviously at the heart of machine learning.[15] Machine learning becomes 'deep learning'[16] when tasks become more complex and the system more autonomous and therefore, inevitably, also opaquer.[17] When such self-directed learning processes are used in a legal context, they are likely to raise significant questions. For example, risk assessment tools that predict future reoffending (*infra*) are based on machine learning techniques, involving a statistical analysis of large datasets on (past) criminal conduct.[18] However, this complexity cannot be an excuse for the opacity of the legal process, as will be discussed below.

3 AI as a Tool to Facilitate Adjudication

After these short terminological explanations, we can now focus on the first use of AI in the field of criminal justice: AI as a support tool for the adjudication of criminal cases. Three types of tools can be distinguished. In order of growing complexity, these are: 1) tools that make existing legal information more easily accessible and searchable, 2) tools that make predictions about the outcome of legal cases, and 3) tools that aim to predict human behaviour relevant for sentencing purposes.

---

[14] Rembrandt Deville, Nico Sergeyssels and Catherine Middag, 'Basic concepts of AI for legal scholars' in Jan De Bruyne and Cedric Vanleenhove (eds), *Law & Artificial Intelligence* (Intersentia 2021) 1, 4.

[15] ibid 1.

[16] The main difference between machine learning and deep learning is that machine learning algorithms will process quantitative and structured data whereas the algorithms used in deep learning will be based on unstructured data.

[17] Emmanuel Barthe, 'L'intelligence artificielle et le droit' (2017) 54 Information, données & documents 23, 24; Morgane Hubert, 'Les algorithmes prédictifs au service du juge : vers une déshumanisation de la justice pénale ? Regards critiques de juges d'instruction' (Master thesis, Catholic University of Leuven 2020) 29.

[18] ibid 28.

In the following sections, we will briefly present each type of tool on the basis of some concrete examples (Section 3.1), before examining the added value (Section 3.2) and potential risks (Section 3.3) these tools entail for the criminal justice system.

3.1 Some examples of AI support tools

3.1.1 Legal search engines

In the legal area, one of the first applications of algorithms consisted in the creation of legal databases, containing legislation, case law and/or legal literature. While some of these databases are quite simple, offering only limited search functions (eg on the basis of the date of a judgment, the case number or the parties' names), others are more sophisticated and powerful, combining several parameters. Clearly, these tools make legal information accessible in a more efficient manner. By now, they have become part of legal practitioners and judges every-day toolset to prepare their cases. Some of these databases have been developed by or for public authorities and are publicly available, free of charge; others are owned by private companies (eg legal publishers) and require prior subscription, sometimes at a relatively high price. In recent years, there seems to be a trend among public authorities to outsource the development and maintenance of certain databases to the private sector for cost-efficiency and/or technical reasons.[19] As we will discuss later, this trend raises a number of concerns.

3.1.2 Tools predicting legal outcomes

In a next step, AI tools based on assisted machine learning[20] have been developed for the purpose of predicting the most likely solution to a dispute. Thanks to various mathematical and statistical tools, it would be possible to assess, with varying degrees of accuracy, the probability of success of certain proceedings.

Examples of such prediction tools in the area of civil and administrative law are Predictice, Supra Legem or Case Law Analytics. In the case of Predictice, an algorithm will calculate the probability of resolution of a dispute as well as the range of compensation

[19] For instance, in the United States, the publication of judgments and their electronic access is largely enabled by legal publishers. Similarly, but much more recently, in Belgium, the rulings of the Court of Cassation are no longer published by the Court itself but by Larcier and the electronic access to older judgments has been made possible thanks to the efforts of the law library of the KU Leuven (see <https://justice.belgium.be/fr/ordre_judiciaire/cours_et_tribunaux/cour_de_cassation/jurisprudence> and <https:// bib.kuleuven.be/rbib/collectie/archieven/arrcass/arresten-van-het-hof-van-cassatie>). Furthermore, in order to make all judicial decisions publicly available, as the Belgian legislator ambitiously set forward in 2019, the Ministry of Justice is exploring public-private partnerships. For a critical analysis, see Jean De Codt, 'La parole du juge sous le boisseau de sa quantification numérique. À propos de la publicité des jugements à l'ère 2.0' (2021) Journal des Tribunaux 22, 24.

[20] Emmanuel Barthe, 'L'intelligence artificielle et le droit' (2017) 54 Information, données & documents 23, 24.

amounts, and all this will be exported and made comprehensible in the form of a customizable report.[21] The predictive algorithm of Supra Legem analyzes hundreds of thousands of administrative decisions (for instance, in the field of migration law), applying different criteria.[22] This analysis would allow to identify trends in case law that are otherwise invisible.[23] Furthermore, the Case Law Analytics software examines the risks regarding a contract or a litigation. According to the programme description, the combination of mathematics and law renders it possible to measure in a precise way the influence of a specific element of a case, in the decision of the judge in charge of the case, or to know how to adjust in the best way a clause in the contract.[24]

Case Crunch, yet another AI software, makes predictions about, for instance, insurance claims, based on the analysis of the outcome of past claims. In a 2017 competition, the system proved to be right about the expected outcome in 87%, compared to 62% of correct assessments by 100 top lawyers of London law firms.[25]

Similarly, in the field of labour law, Legal Insights, an AI tool developed by Wolters Kluwer for the Belgian legal market, makes predictions about the chances of winning a dismissal case.[26]

Closer to the criminal law field are studies using AI tools that involve machine learning and natural language processing to predict the outcome of a case before certain courts, such as the United States Supreme Court[27] or the European Court of Human Rights.[28] The accuracy of these predictions ranges between 70% and 79%.

Clearly, such prediction tools enable legal practitioners (and their clients) to assess, *ex ante*, the success rate of future litigation and to prepare their cases more effectively. Yet for judges too, these tools may prove to be useful as they allow them to position themselves against previous decisions taken by other courts in their legal system and thus to

[21] Predictice, 'Au Cœur de la justice' <https://predictice.com/> accessed 5 July 2021.
[22] For example, the subject matter of the decision, the characterics of the claimant as well as the defendant, and the meaning of the legal provision.
[23] Michaël Benesty, 'Supra Legem' <https://www.data.gouv.fr/fr/reuses/supra-legem/> accessed 5 July 2021.
[24] Case Law Analytics, 'Analysez votre risque juridique grâce à l'IA' <https://www.caselawanalytics.com/> accessed 5 July 2021.
[25] Case Crunch, 'Lawyer Challenge' <https://www.case-crunch.com> accessed 14 July 2021.
[26] Wolters Kluwer, 'Legal Insights' <https://www.wolterskluwer.com/fr-be/solutions/legal-insights/> accessed 14 July 2021.
[27] Daniel Martin Katz, Michael J Bommarito II and Josh Blackman, 'A General Approach for Predicting the Behavior of the Supreme Court of the United States' (*PLoS ONE*, 12 April 2017) <https://doi.org/10.1371/journal.pone.0174698> accessed 14 July 2021.
[28] Nikaloas Aletras, Dimitrios Tsarapatsanis, Daniel Preoţiuc-Pietro and Vasileios Lampos, 'Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective' (*PeerJ Computer Science*, 24 October 2016) <https://doi.org/10.7717/peerj-cs.93> accessed 14 July 2021.

ensure more consistency in case law, even in legal systems that do not adhere to the doctrine of binding precedent.

3.1.3 Tools predicting future behaviour of suspects or convicts

Still less frequently applied are AI tools that make *ex ante* individual[29] predictions about future behaviour, in particular criminal behaviour. These predictive algorithms have been used 'to predict (…) who is likely to fail to appear at their court hearing, and who is likely to reoffend at some point in the future.'[30] Such predictions may indeed be relevant for the execution of judicial decisions (eg parole decisions), but potentially also at the sentencing stage as they give the judge an evidence-based indication about the suspect's likelihood to reoffend. Based on this prediction, the judge could pronounce a more adequate sentence.

Among the best-known programmes is COMPAS (acronym for Correctional Offender Management Profiling for Alternative Sanctions), used in the United States. It is a decisional support tool[31] developed by a private company[32] that makes a prediction on the basis of the defendant's criminal file and a questionnaire-based interview. More precisely, 'this software predicts a defendant's risk of committing a misdemeanor or felony within 2 years of assessment from 137 features about an individual and the individual's past criminal record.'[33] These features include, for instance, age, sex and criminal history, but not race. While originally designed for pre-trial release decisions and post-sentencing decisions (eg parole), the tool's use has been gradually expanded to sentencing decisions, as the *Loomis* case (*infra*) illustrates.

COMPAS is a classic example of supervised machine learning. The algorithm is trained with past data that are analyzed on the basis of a decision tree and develops model relationships between independent and dependent variables. In a next step, this model is tested on new cases to improve its performance. Once sufficiently trained, the algorithm is applied in individual cases to determine the defendant's recidivism score on a scale of

---

[29] Unlike predictive policing tools (*supra*), which focus on geographical areas. As indicated in the introduction, this article will not focus on predictive policing tools, but rather on tools that could help judges assess the risk of future criminal behaviour to the extent that this is relevant for their decision-making in a particular case.

[30] Julia Dressel and Hany Farid, 'The accuracy, fairness, and limits of predicting recidivism' (*Science Advances*, 17 January 2018) <https://advances.sciencemag.org/content/4/1/eaao5580> accessed 5 July 2021, quoting Walter L Perry, Brian McInnis, Carter C Price, Susan Smith and John S Hollywood, *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations* (*Rand Corp,* 2013) <https://www.rand.org/pubs/research_reports/RR233.html> accessed 5 July 2021.

[31] Interview with Angèle Christin: Hubert Guillaud, 'La justice prédictive (2/3) : prédictions et régulations' (*Le Monde*, 13 September 2017) <https://www.lemonde.fr/blog/internetactu/2017/09/13/la-justice-predictive-23-predictions-et-regulations/> accessed 5 July 2021.

[32] Originally named Northpointe, rebranded to Equivant in 2017, shortly after the critical ruling of the Wisconsin Supreme Court in the notorious *Loomis* case (*infra*).

[33] Julia Dressel and Hany Farid, 'The accuracy, fairness, and limits of predicting recidivism' (*Science Advances*, 17 January 2018) <https://advances.sciencemag.org/content/4/1/eaao5580> accessed 5 July 2021.

ten. In the *Loomis* case, the result of this assessment was added to the case file as part of the presentencing investigation report submitted to the sentencing court, which informs the latter in view of imposing a sentence.

It is important to emphasize that, even though the defendant's criminal record and his/her answers to the questionnaire are taken into account, the risk assessment is *not* specific to the defendant personal situation but based on a comparison of his/her data (at a certain point in time) with a similar data group collected in the past. Nor does the tool allow to account for future changes in the individual's personal situation or at policy-making level (eg increased efforts to reintegrate sentenced persons in society, improvements in the educational system, better economic conditions).

COMPAS is not entirely unique, but definitely the most discussed (and, as we will see, the most criticized) AI tool in its kind. Other more or less comparable tools are HART (for Hart Assessment Risk tool), used by the Durham police,[34] and CAIS (for Correctional Assessment and Intervention System), developed by the National Council on Crime and Delinquency in the United States for post-sentencing treatment by social services agencies.[35] The earlier mentioned CloudWalk technology would also enable individual risk assessments thanks to facial recognition, thereby sophisticating traditional predictive policing.[36]

All these tools are instruments that, in an ideal scenario, provide evidence-based information to judges (and other authorities) and thus enable them to make better informed, allegedly more objective and less biased decisions. Yet is this really true? This question will be answered in the next two sections.

3.2 Advantages

The application of AI as a support tool for adjudication in criminal cases definitely entails a number of advantages, as private companies developing such tools are quick to point out. Generally speaking, the appeal of algorithmic processing lies in its speed, efficiency and (apparent) objectivity.[37]

---

[34] See eg Chris Baraniuk, 'Durham Police AI to help with custody decisions' (*BBC*, 10 May 2017) <https://www.bbc.com/news/technology-39857645> accessed 14 July 2021.

[35] See eg Elgin Karls, Eric La Nguyen, Dan Spika and Kendall Vega, 'A Demonstration Analysis of the Correctional Assessment and Intervention Analysis (CAIS)', Report prepared for the National Council on Crime and Delinquency (University of Wisconsin Madison 2018) <https://lafollette.wisc.edu/images/publications/workshops/2018-NCCD_final.pdf> accessed 14 July 2021.

[36] See eg Daniel Faggella, 'AI for Crime Prevention and Detection – 5 Current Applications' (*Emerj*, 2 February 2019) <https://emerj.com/ai-sector-overviews/ai-crime-prevention-5-current-applications/> accessed 14 July 2021.

[37] Sonia Desmoulin-Canselier, 'L'emprise des algorithmes – A propos de Frank Pasquale, The Black Box Society. The Secret Algorithms That Control Money and Information' (*La vie des idées*, 20 June 2018) <https://laviedesidees.fr/L-emprise-des-algorithmes.html> accessed 14 July 2021.

The first advantage of AI as a tool in criminal justice is the underlying computing power and speed of execution. Without a doubt, the evolution of technologies, in particular machine learning, has allowed the automated processing of a large amount of data, consisting of judicial decisions, legal rules, but also examples of cases.[38] Legal search engines are a straight-forward example thereof, but only a first step. The ultimate goal and advantage would be to automate all repetitive tasks and eradicate time-consuming tasks.

Moreover, a second advantage of these intelligent decision-making tools would be the speed of the proceedings, which would have the correlative advantage of relieving the courts and which could tackle the (systemic) problem of judicial delays.

Equally importantly, the use of these tools by judges in their decision-making would also prove – and this constitutes a third advantage – to be a factor of consistency for judicial practices. This is particularly true for AI tools predicting legal outcomes in fairly simple legal proceedings that require limited human judgment. Typically, this concerns highly regulated fields of law with quite precise legal requirements such as traffic law, labour law or immigration law.

Finally, a fourth argument put forward in favour of the introduction of AI tools in the criminal justice area would be the neutrality and accuracy of these systems. For example, according to the companies developing these new software programmes, they would be more objective and accurate than human beings in analyzing a criminal's chances of reoffending. This assessment has, however, been questioned and invalidated by independent actors, as we will discuss below.

3.3 Disadvantages

Notwithstanding the above advantages, the use of AI in courts also presents various and non-negligible drawbacks, which arguably are more important in the field of criminal law than in the area of civil or administrative law. In this Section, we will identify three main disadvantages. The first one is probably intrinsic to the operation of AI systems, especially those involving machine learning, and concerns the data used for their operation. Systems that are data-based require the collection and exploitation of a large set of data. Problems may arise in the selection of data and/or in their subsequent exploitation. Second, we will address the opacity of AI systems – a problem better known under the concept of 'black box'. On the one hand, this opacity is due to the way in which AI systems are designed and function, hampering the interpretability and explicability of decisions based on algorithms. On the other hand, this 'black box' phenomenon is further increased by the fact that many AI tools are created by private companies, invoking their business secrecy to shield off the design and functioning of those systems. Third, the effectiveness and accuracy of at least some tools seem to be overrated too.

[38] Serge Abiteboul and Florence G'Sell, 'Les algorithmes pourraient-ils remplacer les juges?' *Le Big data et le droit* (Dalloz 2019) 12 <https://hal.inria.fr/hal-02304016v2/document> accessed 25 October 2021.

3.3.1 The issue of data quality and risks of errors and biases

AI systems essentially require the input and processing of data. At the level of the collection and exploitation of data, several issues may arise.

As pointed out by Arthur Holland Michel, these issues 'can be categorized as incomplete data, low quality data, incorrect or false data, and discrepant data (data that differ from the data the system was designed for).'[39] The output of the system inherently depends on the quality of the input. It is a common mistake to believe that the more data you have, the higher your success rate will be. Of course, a large amount of data is needed to train the algorithms, but if the data are of poor quality, incomplete, incorrect and/or badly transcribed, this will impact the quality of the result too.[40]

In addition, there may be errors and/or biases built into the system. Errors are generally due to the data that are used, the functioning of the algorithm, or security flaws. Biases, whether explicit or implicit, may result from the (past) data used to train the system or the factors and their relative weight in the decision tree.[41] Whereas machine learning techniques appear to be value-neutral,[42] the same cannot be said of the people who design or programme them. Consequently, 'a predictive algorithm's recommendation actually masks an underlying series of subjective judgments on the part of the system designers about what data to use, include or exclude, how to weight the data, and what information to emphasize or deemphasize.'[43] In the same way, 'being aware that these "implicit biases" exist, and that everyone possesses them – even scientists – is an important step toward drawing fair and unbiased conclusions.'[44] Therefore, once the data have been collected and integrated, it is still necessary to check whether the process is free of errors and biases, in order to achieve the expected result.

The aforementioned COMPAS software provides an apt illustration of these issues. As explained, this software was designed without taking into account the ethnicity of individuals and thus one could legitimately believe that this type of data would not play a

---

[39] Arthur Holland Michel, 'Known Unknowns: Data Issues and Military Autonomous Systems' (2021) United Nations Institute for disarmament research 1, 3.

[40] Serge Abiteboul and Florence G'Sell, 'Les algorithmes pourraient-ils remplacer les juges?' *Le Big data et le droit* (Dalloz 2019) 12 <https://hal.inria.fr/hal-02304016v2/document> accessed 25 October 2021.

[41] For other sources of bias, see eg Dana Pessach and Erez Shmueli, 'Algorithmic Fairness' (*Cornell University*, 21 January 2020) <https://arxiv.org/abs/2001.09784> accessed 21 October 2021.

[42] Yannick Meneceur, 'Les systèmes judiciaires européens à l'épreuve du développement de l'intelligence artificielle' (2018) 2 Revue pratique de la prospective de l'innovation 13.

[43] Robert Brauneis and Ellen Goodman, 'Algorithmic Transparency for the Smart City' (2018) 20 Yale J. L. & Tech. 103, 119; Harry Surden, 'Values Embedded in Legal Artificial Intelligence' (2017) Univ. Colo. Law Legal Studies 5.

[44] Hanna Wallach, 'Big Data, Machine Learning, and the Social Sciences: Fairness, Accountability, and Transparency' (*Hanna Wallach*, 19 December 2014) <https://hannawallach.medium.com/big-data-machine-learning-and-the-social-sciences-927a8e20460d> and <https://www.microsoft.com/en-us/research/publication/big-data-machine-learning-and-the-social-sciences-fairness-accountability-and-transparency/> accessed 15 July 2021.

role in the subsequent recidivism risk assessment. Nevertheless, a study by ProPublica, an independent non-profit newsroom, revealed that, indirectly, ethnicity did matter and even outweighed other explicitly included factors due to the cross-referencing of different data such as place of residence or profession, but also because certain ethnicities were overrepresented in the data used to train the system.[45] As a result, black suspects were more likely to obtain a high-risk score than white ones. This shows the risk of biases is significant and not easy to address.[46]

The use of COMPAS in the notorious *Loomis* case,[47] however, also shows that the quality (and accuracy) of the risk assessment is determined by the quality of the collected data. In particular, the Wisconsin Supreme Court pointed out that the tool had been trained with group data collected nation-wide, without cross-validation for the Wisconsin population.[48] Consequently, certain (demographic, social, economic, legal, etc.) specificities of the State in which the tool was implemented might not be sufficiently reflected. Therefore, the Wisconsin Supreme Court required that the presentencing investigation report including a risk assessment based on the COMPAS tool entails a notification of the limitations of the system.

Moving beyond the COMPAS example, in some legal systems, there are simply not enough and/or not sufficiently precise data available to feed a recidivism prediction tool.[49] What is more, AI tools trained with past data inevitably create models that replicate what happened in the past. They are unable to account for new policies or circumstances. Therefore, the data used to train the AI system need to be updated regularly to account for changing legislation and policies, even if they are not directly linked to sen-

---

[45] Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, 'Machine Biais – There's software used across the country to predict future criminals. And it's biased against blacks' (*ProPublic*a, 23 May 2016) <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> accessed 5 July 2021.

[46] As Angèle Christin points out, 'many of these tools today are tainted by a suspicion of unconstitutionality and removing the postcode or job will not remove their bias, particularly because the history of data and judicial systems is steeped in discrimination.' Interview with Angèle Christin: Hubert Guillaud, 'La justice prédictive (2/3): prédictions et régulations' (*Le Monde*, 13 September 2017) <https://www.lemonde.fr/blog/internetactu/2017/09/13/la-justice-predictive-23-predictions-et-regulations/> accessed 5 July 2021.

[47] For a summary of the case, see Adrien van den Branden, *Les robots à l'assaut de la justice* (1st ed., Bruylant 2019) 4-5. For a critical analysis, see Katherine Freeman, 'Algorithmic Injustice: How the Wisconsin Supreme Court Failed to Protect Due Process Rights in State v. Loomis' (2016) 18 N.C. J.L. & Tech. On. 75.

[48] Wisconsin Supreme Court, State v. Loomis, 881 N.W.2d 749, 770–71 (Wis. 2016).

[49] In Belgium for instance, there is (still) no nation-wide database on criminal convictions and subsequent reoffending. Available empirical data are based on criminal records of individual convicts (which contain limited information and, until recently, were not always up-to-date or fully correct) and on dedicated academic studies (which usually focus on specific populations and/or types of offences, and which often use different concepts of repeat offending). For a (fairly) recent study, see Benjamin Mine and Luc Robert, 'Recidive na een rechterlijke beslissing. Nationale cijfers op basis van het Centraal Strafregister'/'La récidive après une décision judiciaire. Des chiffres nationaux sur la base du Casier judiciaire central' (2015) Final report, Institut National de Criminalistique et de Criminologie 1, 12-27 <https://nicc.fgov.be/upload/publicaties/rapport_38_1.pdf> accessed 25 October 2021.

tencing and the execution of sentences, such as better education or an improved economic situation. Yet, even then, AI may 'blin[d] us to everything that is not quantifiable and digitizable.'[50]

Therefore, all in all, such prediction tools should be used with particular precaution and judges should be duly informed of the origin and quality of the data, and the way in which they are processed.

For less sophisticated AI support tools, such as legal search engines, the problem regarding the quality of the data is likely to be more easily detectable and can be more easily remedied. Imagine a legal database that only contains judgments of one single court or a limited number of journals. Then, it is quite obvious that the search results will be less rich and will only offer a partial understanding of the matter than with a legal database covering all case law and a large body of legal literature.

3.3.2 Opacity of AI systems

Machine learning systems as inherent black boxes?

Originally referring to the on-board recording system in transportation means, the term 'black box' also refers to an opaque and closed device, inaccessible to the eye.[51] When transposed to AI systems, black box means that 'the processes happening inside of them are difficult – and sometimes impossible – to fully understand.'[52] The problem of opacity is particularly pressing in case of machine learning, especially systems based on neural networks.[53]

The black box problem is intrinsically linked to the requirements of explanation and justification, which are highly important in a legal context and even more so in criminal cases where the defendant's liberty and other fundamental rights are at risk. Considering that justice must not only be done, but also seen to be done, it is crucial that the technical functioning of the tool, the data selected by the algorithm, the decision-making factors and process can be explained in a language that all parties are able to understand.[54] Put

[50]Free translation from the original quote « qui nous rend aveugles à tout ce qui n'est pas quantifiable ou numérisiable ». See Sonia Desmoulin-Canselier, 'L'emprise des algorithmes – A propos de Frank Pasquale, 'The Black Box Society. The Secret Algorithms That Control Money and Information' (*La vie des idées*, 20 June 2018) <https://laviedesidees.fr/L-emprise-des-algorithmes.html> accessed 27 July 2021.

[51] ibid.

[52] Mira Ortegon, 'Dismantling the Black Box: Why Governments Should Demand Algorithmic Accountability' (*Brown Political Review*, 30 March 2019) <https://brownpoliticalreview.org/2019/03/dismantling-black-box-governments-demand-algorithmic-accountability/> accessed 25 July 2021.

[53] Moustafa Zouinar, 'Evolutions de l'Intelligence Artificielle : quels enjeux pour l'activité humaine et la relation Humain-Machine au travail ?' (*Open Edition Journals*, 15 April 2020) <https://journals.openedition.org/activites/4941#quotation> accessed 21 June 2021. See also Jenna Burrell, 'How the Machine Thinks: Understanding Opacity in Machine Learning Algorithms' (2016) Big Data & Society 1.

[54] Lêmy Godefroy, 'L'office du juge à l'épreuve de l'algorithme' in Jean-Pierre Clavier (eds), *L'algorithmisation de la justice* (Larcier 2020) 111.

differently, it must be possible to 'reconstruct the relevance of the pathways'[55] operated by the algorithm. Yet, most importantly, judges need to understand the functioning of the AI system that supports them in order to be able to assess its added value. Without such understanding, the prediction made by the system, whether regarding the probable outcome of a case or the risk of reoffending, could lead to undesirable situations and ill-founded or even arbitrary decision-making, heavily impacting the future of the convicted person.[56]

Again, the *Loomis* case provides a good illustration of this problem.[57] Mr Loomis argued that his due process (or, transposed to the European legal context, fair trial) rights had been violated because the sentencing court based its decision on the risk assessment made by COMPAS, notwithstanding it was impossible, for him and the court, to review how factors are weighed by the system and how risk scores are produced. Since he could not review the functioning of the AI system, it was also impossible for him to challenge the risk score the system had put forward. The Wisconsin Supreme Court dismissed this argument, but in a concurring opinion, Justice Abraham pointed out that the 'lack of understanding of COMPAS was a significant problem in the instant case' and argued that 'in considering COMPAS (or other risk assessment tools) in sentencing, a circuit court must set forth on the record a *meaningful process of reasoning* addressing the relevance, strengths, and weaknesses of the risk assessment tool.'[58]

Other positive examples, imposing higher transparency requirements upon the use of AI systems, can be found outside the ('hard core'[59]) criminal law field. Certain public administrations that use AI tools to detect administrative infringements (eg illegal parking or illegal renting of holiday housing) have started to publish so-called 'algorithm registers' providing more information on the systems that are used and their functioning.[60] These examples prove that opacity of AI systems is not a kind of fatality, but an issue that can be addressed if there is a will to do so.[61]

---

[55] Emmanuel Poinas, *Le tribunal des algorithme. Juger à l'ère des nouvelles technologies* (Berger Levrault 2019) 240.

[56] Yannick Meneceur, 'Les systèmes judiciaires européens à l'épreuve du développement de l'intelligence artificielle' (2018) 2 Revue pratique de la prospective de l'innovation 7, 14.

[57] See also Alyssa M. Carlson, 'The Need for Transparency in the Age of Predictive Sentencing Algorithms' (2017) 103 Iowa L. Rev. 303.

[58] Emphasis added.

[59] *Jussila v Finland* App no 73053/01 (ECtHR, 23 November 2006), para 43. For a comprehensive analysis of the difference between criminal law and quasi-criminal law, see Vanessa Franssen and Christopher Harding (eds), *Criminal and Quasi-criminal Enforcement Mechanisms in Europe. Origins, Concepts, Future* (Hart Publishing *forthcoming* 2022).

[60] See eg the algorithm registers published by the cities of Amsterdam and Helsinki <https://algoritme-register.amsterdam.nl/en/ai-register/> ; <https://ai.hel.fi/en/ai-register/> accessed 15 July 2021.

[61] Nazrin Huseinzade, 'Algorithm transparency: How to Eat the Cake and Have It Too' (*European Law Blog*, 27 January 2021) <https://europeanlawblog.eu/2021/01/27/algorithm-transparency-how-to-eat-the-cake-and-have-it-too/#comments> accessed 21 July 2021.

An additional complicating factor in this black box discussion is, however, that AI systems are often developed by private companies and that these actors are reluctant to communicate about the functioning of their systems, hiding behind business secrecy to protect their proprietary code. As a result, 'transparency is hard to come by.'[62] Once again, the *Loomis* case aptly illustrates this issue. The proprietary nature of COMPAS prevented the sentencing court and Mr Loomis to understand how the system functions, to evaluate its scientific validity (or accuracy) and, subsequently, its contestability. Although the Wisconsin Supreme Court recognised the software's limitations, it found that Northpointe had a clear financial interest in not disclosing the algorithm that it had itself developed.[63] In addition, the company also pointed out that with knowledge of the algorithm, criminals could potentially distort the risk assessment and exploit the model to their advantage, which would make the algorithm ineffective.[64]

Moreover, AI systems provided by the private sector also create other impediments to information. As indicated above, legal search engines have become an indispensable tool for legal practitioners and judges to prepare their cases. The more powerful ones are, however, expensive and thus not accessible to all. Suspects who are able to afford a lawyer from a top law firm are thus likely to get a better defence. Considering the criminal justice system is often underfunded, judges too will probably have more limited access to useful legal information. If one adds to this the trend to outsource certain public services, including databases providing access to case law, one will understand why the argument that AI renders justice more cost-efficient should be somewhat mitigated.

In conclusion, it seems fundamental that the algorithmic tools underlying legal decisions are accessible, or at least, explicable, at the end of the process. The complexity of the decision-making process should not be used as an excuse for the opacity of the system, and private companies' business interests should be adequately balanced against suspects' fundamental rights.[65]

3.3.3 Overstated effectiveness

Finally, a third concern is the accuracy and effectiveness of AI tools predicting legal outcomes and future risks. While the developers of AI tools like to emphasize how accurate

[62] Mira Ortegon, 'Dismantling the Black Box: Why Governments Should Demand Algorithmic Accountability' (*Brown Political Review*, 30 March 2019) <https://brownpoliticalreview.org/2019/03/dismantling-black-box-governments-demand-algorithmic-accountability/> accessed 21 July 2021.

[63] Adrien van den Branden, *Les robots à l'assaut de la justice* (1st ed., Bruylant 2019) 8.

[64] ibid.

[65] For some interesting examples outside the criminal law field on how data protection rules can be used to enhance AI transparency, see Nazrin Huseinzade, 'Algorithm transparency: How to Eat the Cake and Have It Too' (*European Law Blog*, 27 January 2021) <https://europeanlawblog.eu/2021/01/27/algorithm-transparency-how-to-eat-the-cake-and-have-it-too/#comments> accessed 21 July 2021.

and reliable the predictions made by these tools are, independent studies show a different reality.

For instance, with respect to the COMPAS tool, an academic study published in 2018 analyzed the reliability of the software by comparing the results obtained by the COMPAS to those of different individuals with no or very little legal background, estimating the likelihood of reoffending by the same convicted persons. To this end, randomly selected individuals were asked one simple question: 'Do you think this person will commit another crime within two years?' In contrast, COMPAS analyzed 137 factors to arrive at a conclusive result. It turned out that the accuracy of both assessments, made by AI and human beings, was very similar: While COMPAS obtained a 65,2% efficiency score, the human beings were right in 67% of the cases. Thus, even people without any criminal justice expertise would have achieved the same result as a computer programme designed and trained to make such risk assessments.[66]

As regards AI tools predicting legal outcomes, it is important to highlight that they are usually designed for very specific legal disputes, typically in fields of law that are highly regulated and technical, thus requiring limited human judgment (eg traffic law), or disputes where the number of legal factors to account for is relatively limited (eg insurance claims, labour law disputes, return decisions in the field of migration law). The more complex the nature of the dispute and the more open-ended legal norms are, the less likely such tools will lead to a satisfactory prediction. This point will be further elaborated in the next part.

4 Judge-Made vs AI-Made Criminal Justice

In this last part, we will take the analysis a step further and focus on AI tools that, rather than supporting human judges, could replace them. While examples of such 'robot judges' are still fairly limited[67] and, to our knowledge, inexistent in the field of criminal justice, we can nevertheless already examine, on a theoretical level, some advantages these tools are likely to bring compared to a human judge (Section 4.1). Next, we will present a number of disadvantages that can result from automated adjudication (Section 4.2).

4.1 Advantages

A robot judge potentially has several advantages over the human judge, in particular in terms of consistency in decision-making, reliability, cost and speed.

[66] Julia Dressel and Hany Farid, 'The accuracy, fairness, and limits of predicting recidivism' (*Science Advances*, 17 January 2018) <https://advances.sciencemag.org/content/4/1/eaao5580> accessed 5 July 2021.

[67] So far, some proposals have been made for disputes with small values and to extend some online dispute resolution mechanisms. See eg Michael Grupp 'How to Build a Robot Lawyer' in Markus Hartung, Micha-Manuel Bues and Gernot Halbleib (eds), *Legal Tech. How Technology is Changing the Legal World. A Practioner's Guide* (Beck-Hart-Nomos 2018) 249, 257.

First, in terms of consistency and reliability, it is well-known that human judges are not in a constant mood and thus that their decision may vary depending on the day and the hour of the hearing or decision, or the events that may occur in their lives.[68] In contrast, the algorithm will work every day, at every time of day, in the same way. This coherence is based on the proper functioning of the AI system; it is indeed a safe bet that systems implementing algorithmic justice will base their analysis on the same data recorded in their digital library, even the least known. Consistent decision-making contributes to the coherence and stability of the legal system, and thereby enhances legal certainty.

Next, robot judges may be more cost-efficient than human judges. For sure, the initial investments may be high, but once those are made, the robot judge has a lower operational cost than a human one, for a similar number of cases handled.[69] Moreover, thanks to the robot judge, physical trials and the costs related to their organisation could perhaps be abandoned in a number of cases (to the extent, of course, that this is compatible with the right to a fair trial),[70] which would lead to a substantial reduction in the costs of legal proceedings.[71]

Finally, we could mention the speed and effectiveness of robot judges. The AI system would outperform the human judge for certain tasks, in particular for dealing with repetitive demands and the calculation of damages or interests.

4.2 Disadvantages

Despite the aforementioned advantages, the disadvantages of AI decision-making should neither be neglected, nor underestimated.

4.2.1 Law is not code

First of all, our law contains a lot of open-ended norms, requiring a certain degree of human appreciation. To take just one example, the application of fundamental rights

[68] Shai Danziger, Jonathan Levav and Liora Avnaim-Pesso, 'Extraneous factors in judicial decision' (*PNAS*, 26 April 2011) <http://www.pnas.org/content/108/17/6889> accessed 5 July 2021; Serge Abiteboul and Florence G'Sell, 'Les algorithmes pourraient-ils remplacer les juges?' *Le Big data et le droit* (Dalloz 2019) 9 <https://hal.inria.fr/hal-02304016v2/document> accessed 25 October 2021.

[69] Adrien van den Branden, *Les robots à l'assaut de la justice* (1st ed., Bruylant 2019) 48.

[70] The European Court of Human Rights has ruled on a number of occasions that for some criminal cases an oral hearing may not be indispensable, in particular 'where there are no issues of credibility or contested facts which necessitate a hearing,' where an oral hearing would be 'an obstacle to the particular diligence required' in certain cases and where the legal questions raised are not particularly complicated. In such cases, 'the courts may fairly and reasonably decide the case on the basis of the parties' submissions and other written materials.' See eg *Jussila v Finland* App no 73053/01 (ECtHR, 23 November 2006), paras 40-48; *Nusret Kaya and Others v Turkey* App nos 43750/06, 43752/06, 32054/08, 37753/08 and 60915/08 (ECtHR, 22 April 2014), para 84. This case law may offer useful guidance when considering to introduce robot judges in certain criminal cases.

[71] Loïck Gérard and Dominique Mougenot, 'Titre 1 – Justice robotisée & droits fondamentaux' in Jean-Benoit Hubin, Hervé Jacquemin and Benoît Michaux (eds), *Le juge et l'algorithme : juges augmentés ou justice diminuée ?* (1st ed., CRIDS 2019) 39.

such as the right to a fair trial or the right to privacy implies a balancing of interests and almost inevitably entails a certain level of uncertainty as to the outcome of a particular case. Such norms cannot easily be coded in algorithmic language.[72]

What is more, legal norms are based on values and social norms. When applied by human judges, 'they identify which values are at stake in a given decisional environment and ask, where necessary, if those values have been properly balanced.'[73] Yet when designers and programmers of AI systems are led to arbitrate on values, the importance of legal rules and the interpretation of these rules, this may be highly problematic, especially since they have not been legally trained and are thus unfamiliar with the technique of law and its purpose.[74]

4.2.2 Static decision-making

As explained above, data-driven learning systems are, for the most part, backward-looking: 'algorithms can only learn from existing datasets, which are grounded in past experiences and past trends.'[75] This has the advantage of reliability and predictability, but also entails the risk of static decision-making. Indeed, an AI system will struggle when faced with new situations or evolving legal norms. One may thus wonder whether AI is capable of producing new and innovative decisions. For instance, the European Convention on Human Rights constitutes a living instrument and the European Court of Human Rights regularly adjusts its case law to keep pace with societal evolutions, as well as international and national law. Recent examples concern the principle of *non bis in idem*[76] and the protection of the right to privacy in the context of mass surveillance practices.[77]

---

[72] Kareb Hao, 'AI still doesn't have the common sense to understand human language' (*MIT Technology Review*, 31 January 2020) <https://www.technologyreview.com/2020/01/31/304844/ai-common-sense-reads-human-language-ai2/> accessed 14 July 2021.

[73] Kiel Brennan-Marquez, 'Plausible Cause: Explanatory Standards in the Age of Powerful Machines' (2017) 70 Vand. L. Rev. 1249.

[74] Engerrand Marique and Alain Strowel, 'Gouverner par la loi ou les algorithmes: de la norme générale de comportement au guidage rapproché des conduites' (2017) 10 Dalloz IP/IT 517, 521.

[75] Mira Ortegon, 'Dismantling the Black Box: Why Governments Should Demand Algorithmic Accountability' (*Brown Political Review*, 30 March 2019) <https://brownpoliticalreview.org/2019/03/dismantling-black-box-governments-demand-algorithmic-accountability/> accessed 25 July 2021.

[76] *A and B v Norway* App nos 24130/11 and 29758/11 (ECtHR, 15 November 2016). For a critical analysis of the Court's new approach, see Michiel Luchtman, 'Ne bis in idem at the Interface of Administrative and Criminal Law Enforcement – Sufficiently Connected in Substance, Time and Space?' (2019) 90 Revue internationale de droit pénal 335, 343-347.

[77] *Big Brother Watch and Others v The United Kingdom* App nos 58170/13, 62322/14 and 24960/15 (ECtHR, 25 May 2021). For a critical analysis of the Court's new approach, see Nóra Ni Loideain, 'Not So Grand: The Big Brother Watch ECtHR Grand Chamber Judgment' (*Information Law and Policy Centre*, 28 May 2021) <https://infolawcentre.blogs.sas.ac.uk/2021/05/28/not-so-grand-the-big-brother-watch-ecthr-grand-chamber-judgment/> accessed 25 October 2021.

Slightly older examples regard the acceptability of irreducible life sentences without parole possibility in terrorism and other serious crime cases in the light of Article 3 of the European Convention on Human Rights.[78]

4.2.3 Independence and impartiality

In the context of algorithmic justice, the questions of independence and impartiality of the robot judge cannot be ignored. Impartiality refers to the judge's own qualities and denotes the absence of prejudice or bias. It has a *subjective* (ie with regard to the judge's personal behaviour) and an *objective* dimension (ie the appearance created toward the parties or the general public).[79] Independence is more about the status or position of the judge, especially in relation to other branches of power, but also in relation to the influence of third parties.

When transposed to robot judges, one could believe that the digital process is free of any reproach. The robot judge would be incorruptible, emotionless, uninfluenced and neutral. But is this really true?

On the one hand, the impartiality of the AI system can be problematic, particularly because of the opacity of the system and its operation (*supra*). From the outside, this process seems inaccessible or impenetrable, and may thus create the impression of arbitrary or unfair decision-making.[80] This impression is likely to be exacerbated by the aforementioned risk of systemic biases.[81] On the other hand, with respect to independence, how can we check and control the private actors who finance and design the AI decision-making tools to make sure they do not influence judicial decisions? Is there not a risk that justice will lose its independence to a private justice system, intrinsically linked to the companies that create the algorithms?[82]

As the Consultative Committee of European Judges pointed out, these technologies 'must not prevent judges from applying the law independently and impartially (...). An

---

[78] See eg *Babar Ahmad and Others v The United Kingdom* App nos 24027/07, 11949/08, 36742/08, 66911/09 and 67354/09 (ECtHR, 10 April 2012); *Vinter and Others v The United Kingdom* App nos 66069/09, 130/10 and 3896/10 (ECtHR, 9 July 2013).

[79] For a fairly recent criminal case where the Grand Chamber found a violation of the impartiality requirement, see *Morice v France* App no 29369/10 (ECtHR, 23 April 2015).

[80] See eg Alekh Agarwal, Alina Beygelzimer, Miroslav Dudík, John Langford and Hanna Wallach, 'A Reductions Approach to Fair Classification' (*PMLR*, 2018) 1 <http://proceedings.mlr.press/v80/agarwal18a.html> accessed 21 October 2021.

[81] That said, interpretability and understandability of the AI system alone will not solve the problem of biases. See eg Connor O'Sullivan, 'Interpretability in Machine Learning' (*Towards data science*, 21 October 2020) <https://towardsdatascience.com/interpretability-in-machine-learning-ab0cf2e66e1> accessed 25 October 2021; Connor O'Sullivan, 'What is Algorithm Fairness? An introduction to the field that aims at understanding and preventing bias in machine learning models' (*Towards data science*, 5 March 2021) <https://towardsdatascience.com/what-is-algorithm-fairness-3182e161cf9f> accessed 25 October 2021.

[82] Morgane Hubert, 'Les algorithmes prédictifs au service du juge : vers une déshumanisation de la justice pénale ? Regards critiques de juges d'instruction' (Master thesis, Catholic University of Leuven 2020) 48, 49.

excessive dependence on technology and those who control it is a risk for justice. (...) Judges should not be subject, for reasons of efficiency alone, to the imperatives of technology and those who control technology.'[83]

4.2.4 Transparency and contestability

When a suspect is convicted or acquitted, it is important for him/her to understand that decision, either to accept it or, on the contrary, to identify the means that would allow him/her to contest it. The court's reasoning is essential to respect the right to a fair trial, including the equality of arms, the adversarial principle, and the right to a legal remedy. For this reason, the General Data Protection Regulation and the so-called Law Enforcement Directive are very strict on applying automated decision-making to individuals.[84]

To the extent AI tools are used in the decision-making process, it is therefore crucial that the defendant has access to all the stages of the algorithmic reasoning that led to the decision. As explained earlier, this is far from obvious due to the black box phenomenon and the private interests of the company that developed the AI tools. Therefore, it is important to minimize the risks generated by the multiplication of black boxes that algorithmically determine individual trajectories. A balance must be struck between business secrecy protecting the creation of software by private companies and the transparency requirements intended to protect the individual.[85] Finally, it is also necessary to rethink and establish control and surveillance mechanisms to prevent justice from moving from the public to the private domain.[86]

5 Conclusion

As we have seen, AI tools can have substantial advantages, but also considerable disadvantages, both for justice in general and for individual defendants, that should not be minimized. However, it would be a mistake to categorically reject these AI tools in the area of criminal justice. It is indeed necessary to strike the right balance, which would

[83] Consultative Committee of European Judges, 'Justice et technologies de l'information', Opinion No 14, 2011, paras 8 and 34; Commission européenne pour l'efficacité de la justice, 'Lignes directrices sur la conduite du changement vers la Cyberjustice' (*CEPEJ-GT-QUAL*, 7 December 2016) <https://rm.coe.int/1680748154#_Toc461547117> accessed 23 June 2021.

[84] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), recital 71 and art 22; Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA, recital 38 and art 11.

[85] Georges De Leval and Frédérique Georges, *Droit judiciaire, t. 1, Institutions judiciaires et éléments de compétence* (Larcier 2014) 101.

[86] Morgane Hubert, 'Les algorithmes prédictifs au service du juge : vers une déshumanisation de la justice pénale ? Regards critiques de juges d'instruction' (Master thesis, Catholic University of Leuven 2020) 96.

combine the positive elements of the one and the other and limit their negative features. If these tools can provide valuable support to human judges on the basis of reliable, transparent and verifiable information, one should not deprive criminal justice actors of these instruments. These tools could constitute an additional guarantee, not a danger for parties, provided they are properly mastered.[87] Ultimately, 'if we can preserve a human-centric approach to leveraging this powerful technology, predictive analysis can be one of the building blocks of a more transparent and more efficient legal system.'[88]

References

Abiteboul S and G'Sell F, 'Les algorithmes pourraient-ils remplacer les juges?' *Le Big data et le droit* (Dalloz 2019) <https://hal.inria.fr/hal-02304016v2/document> accessed 25 October 2021

Agarwal A, Beygelzimer A, Dudík M, Langford J and Wallach H, 'A Reductions Approach to Fair Classification' (*PMLR*, 2018) 1 <http://proceedings.mlr.press/v80/agarwal18a.html> accessed 21 October 2021

Aletras N, Tsarapatsanis D, Preoţiuc-Pietro D and Lampos V, 'Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective' (*PeerJ Computer Science*, 24 October 2016) <https://doi.org/10.7717/peerj-cs.93> accessed 14 July 2021

Angwin J, Larson J, Mattu S and Kirchner L, 'Machine Biais – There's software used across the country to predict future criminals. And it's biased against blacks' (*ProPublic*a, 23 May 2016) <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> accessed 5 July 2021

Baraniuk C, 'Durham Police AI to help with custody decisions' (*BBC*, 10 May 2017) <https://www.bbc.com/news/technology-39857645> accessed 14 July 2021

Barthe E, 'L'intelligence artificielle et le droit' (2017) 54 Information, données & documents 23

Bayamlioglu E and Leenes R, The "rule of law" implications of data-driven decision-making: a techno-regulatory perspective' (2018) Law, Innov Technol 295

Benesty M, 'Supra Legem' <https://www.data.gouv.fr/fr/reuses/supra-legem/> accessed 5 July 2021

[87] Serge Abiteboul and Florence G'Sell, 'Les algorithmes pourraient-ils remplacer les juges?' *Le Big data et le droit* (Dalloz 2019) 14-15 <https://hal.inria.fr/hal-02304016v2/document> accessed 25 October 2021.

[88] Roland Vogl, 'Legal Tech in the USA' in Markus Hartung, Micha-Manuel Bues and Gernot Halbleib (eds), *Legal Tech. How Technology is Changing the Legal World. A Practioner's Guide* (Beck-Hart-Nomos 2018) 380, 392.

Brauneis R and Goodman E, 'Algorithmic Transparency for the Smart City' (2018) 20 Yale J. L. & Tech. 103

Brennan-Marquez K, 'Plausible Cause: Explanatory Standards in the Age of Powerful Machines' (2017) 70 Vand. L. Rev. 1249

Burrell J, 'How the Machine Thinks: Understanding Opacity in Machine Learning Algorithms' (2016) Big Data & Society 1

Carlson A, 'The Need for Transparency in the Age of Predictive Sentencing Algorithms' (2017) 103 Iowa L. Rev. 303

Case Crunch, 'Lawyer Challenge' <https://www.case-crunch.com> accessed 14 July 2021

Case Law Analytics, 'Analysez votre risque juridique grâce à l'IA' <https://www.caselawanalytics.com/> accessed 5 July 2021

Commission européenne pour l'efficacité de la justice, 'Lignes directrices sur la conduite du changement vers la Cyberjustice' (*CEPEJ-GT-QUAL*, 7 December 2016) <https://rm.coe.int/1680748154#_Toc461547117> accessed 23 June 2021

Consultative Committee of European Judges, 'Justice et technologies de l'information', Opinion No 14, 2011

Danziger S, Levav J and Avnaim-Pesso L, 'Extraneous factors in judicial decision' (*PNAS*, 26 April 2011) <http://www.pnas.org/content/108/17/6889> accessed 5 July 2021

De Leval G and Georges F, *Droit judiciaire, t. 1, Institutions judiciaires et éléments de compétence* (Larcier 2014)

Desmoulin Canselier S, 'L'emprise des algorithmes – A propos de Frank Pasquale, The Black Box Society. The Secret Algorithms That Control Money and Information' (*La vie des idées*, 20 June 2018) <https://laviedesidees.fr/L-emprise-des-algorithmes.html> accessed 14 July 2021

Deville R, Sergeyssels N and Middag C, 'Basic concepts of AI for legal scholars' in De Bruyne J and Vanleenhove C (eds), *Law & Artificial Intelligence* (Intersentia 2021)

Dressel J and Farid H, 'The accuracy, fairness, and limits of predicting recidivism' (*Science Advances*, 17 January 2018) <https://advances.sciencemag.org/content/4/1/eaao5580> accessed 5 July 2021

European Commission, 'Proposal for a regulation of the European parliament and of the council laying down harmonized rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts', COM(2021), Brussels

Faggella D, 'AI for Crime Prevention and Detection – 5 Current Applications' (*Emerj*, 2 February 2019) <https://emerj.com/ai-sector-overviews/ai-crime-prevention-5-current-applications/> accessed 14 July 2021

Freeman K, 'Algorithmic Injustice: How the Wisconsin Supreme Court Failed to Protect Due Process Rights in State v. Loomis', 2016 18 N.C. J.L. & Tech. On. 75

Gérard L and Mougenot D, 'Titre 1 – Justice robotisée & droits fondamentaux' in Hubin J-B, Jacquemin H and Michaux B (eds), *Le juge et l'algorithme : juges augmentés ou justice diminuée ?* (1st ed., CRIDS 2019)

Godefroy L, 'L'office du juge à l'épreuve de l'algorithme' in Clavier J-P (ed.), *L'algorithmisation de la justice* (Larcier 2020)

Gosavi S and Kavathekar S, A Survey on Crime Occurrence Detection and prediction Techniques' (2018) 8 Int. j. eng. technol. manag. appl. sci. 1405

Grupp M, 'How to Build a Robot Lawyer' in Hartung M, Bues M-M and Halbleib G (eds), *Legal Tech. How Technology is Changing the Legal World. A Practioner's Guide* (Beck-Hart-Nomos 2018)

Guillaud H, 'La justice prédictive (2/3) : prédictions et régulations' (*Le Monde*, 13 September 2017) <https://www.lemonde.fr/blog/internetactu/2017/09/13/la-justice-predictive-23-predictions-et-regulations/> accessed 5 July 2021

Hao K, 'AI still doesn't have the common sense to understand human language' (*MIT Technology Review*, 31 January 2020) <https://www.technologyreview.com/2020/01/31/304844/ai-common-sense-reads-human-language-ai2/> accessed 14 July 2021

Hérard F, 'Technologies de prédiction du crime: Palantir a scruté les citoyens de la Nouvelle-Orléans en secret pendant 6 ans' (*TV5Monde*, 3 March 2018) <https://information.tv5monde.com/info/technologies-de-prediction-du-crime-palantir-scrute-les-citoyens-de-la-nouvelle-orleans-en> accessed 14 July 2021

Holland Michel A, 'Known Unknowns: Data Issues and Military Autonomous Systems' (2021) United Nations Institute for disarmament research 1

Hubert M, 'Les algorithmes prédictifs au service du juge : vers une déshumanisation de la justice pénale ? Regards critiques de juges d'instruction' (Master thesis, Catholic University of Leuven 2020)

Huseinzade N, 'Algorithm transparency: How to Eat the Cake and Have It Too' (*European Law Blog*, 27 January 2021) <https://europeanlawblog.eu/2021/01/27/algorithm-transparency-how-to-eat-the-cake-and-have-it-too/#comments> accessed 21 July 2021

Karls E, La Nguyen E, Spika D and Vega K, 'A Demonstration Analysis of the Correctional Assessment and Intervention Analysis (CAIS)', Report prepared for the National Council on Crime and Delinquency (*University of Wisconsin Madison*, 2018) <https://lafollette.wisc.edu/images/publications/workshops/2018-NCCD_final.pdf> accessed 14 July 2021

Katz D M, Bommarito II M J and Blackman J, 'A General Approach for Predicting the Behavior of the Supreme Court of the United States' (*PLoS ONE*, 12 April 2017) <https://doi.org/10.1371/journal.pone.0174698> accessed 14 July 2021

Luchtman L, 'Ne bis in idem at the Interface of Administrative and Criminal Law Enforcement – Sufficiently Connected in Substance, Time and Space?' (2019) 90 Revue internationale de droit pénal 335

Marique E and Strowel A, 'Gouverner par la loi ou les algorithmes: de la norme générale de comportement au guidage rapproché des conduites' (2017) 10 Dalloz IP/IT 517

Meneceur Y, 'Les systèmes judiciaires européens à l'épreuve du développement de l'intelligence artificielle' (2018) 2 Revue pratique de la prospective de l'innovation 13

Ni Loideain N, 'Not So Grand: The Big Brother Watch ECtHR Grand Chamber Judgment' (*Information Law and Policy Centre*, 28 May 2021) <https://infolawcentre.blogs.sas.ac.uk/2021/05/28/not-so-grand-the-big-brother-watch-ecthr-grand-chamber-judgment/> accessed 25 October 2021

Mine B and Robert L, 'Recidive na een rechterlijke beslissing. Nationale cijfers op basis van het Centraal Strafregister'/'La récidive après une décision judiciaire. Des chiffres nationaux sur la base du Casier judiciaire central'(2015) Final report, Institut National de Criminalistique et de Criminologie 1 <https://nicc.fgov.be/upload/publicaties/rapport_38_1.pdf> accessed 25 October 2021

O'Neil C, *Algorithmes – La Bombe à retardement* (Les Arènes 2018)

O'Sullivan C, 'Interpretability in Machine Learning' (*Towards data science*, 21 October 2020) <https://towardsdatascience.com/interpretability-in-machine-learning-ab0cf2e66e1> accessed 25 October 2021

—— 'What is Algorithm Fairness? An introduction to the field that aims at understanding and preventing bias in machine learning models' (*Towards data science*, 5 March 2021) <https://towardsdatascience.com/what-is-algorithm-fairness-3182e161cf9f> accessed 25 October 2021

Oosterloo S and Van Schie G, 'The Politics and Biases of the "Crime Anticipation System" of the Dutch Police' in Bates J, Clough P D, Jäschke R and Otterbacher J (eds), *Proceedings of the International Workshop on Bias in Information, Algorithms, and Systems* (CEUR Workshop Proceedings 2018) 30

Ortegon M, 'Dismantling the Black Box: Why Governments Should Demand Algorithmic Accountability' (*Brown Political Review*, 30 March 2019) <https://brownpoliticalreview.org/2019/03/dismantling-black-box-governments-demand-algorithmic-accountability/> accessed 25 July 2021

Patil R, Kacchi M, Gavali P and Pimparia K, 'Crime Pattern Detection, Analysis & Prediction using Machine Learning' (2020) 7 IRJET 119

Perry W L, McInnis B, Price C C, Smith S and Hollywood J S, *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations* (*Rand Corp*, 2013) <https://www.rand.org/pubs/research_reports/RR233.html> accessed 5 July 2021

Pessach D and Shmueli E, 'Algorithmic Fairness' (*Cornell University*, 21 January 2020) <https://arxiv.org/abs/2001.09784> accessed 21 October 2021

Poinas E, *Le tribunal des algorithme. Juger à l'ère des nouvelles technologies* (Berger Levrault 2019)

Predictice, 'Au Cœur de la justice' <https://predictice.com/> accessed 5 July 2021

PredPol, 'What. Where. When' <https://www.predpol.com/> accessed 14 July 2021

Shah N, Bhagat N and Shah M, 'Crime forecasting: a machine learning and computer vision approach to crime prediction prevention' (2021) 4 VCIBA 9

Smuha N, Ahmed-Rengers E, Harkens A, Li W, MacLaren J, Piselli R and Yeung K, 'How the EU can achieve Legally Trustworthy AI: A Response to the European Commission's Proposal for an Artificial Intelligence Act' (*SSRN*, 5 August 2021) <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3899991&download=yes> accessed 19 October 2021

Surden H, 'Values Embedded in Legal Artificial Intelligence' (2017) Univ. Colo. Law Legal Studies 5

Vaele M and Zuiderveen Borgesius F, 'Demystifying the Draft EU Artificial Intelligence, Analysing the good, the bad, and the unclear elements of the proposed approach', (2021) 22 Computer L. Rev. Int. 97

Van Brakel R. E, 'Pre-emptive Big Data Surveillance and its (Dis)empowering Consequences: The Case of Predictive Policing' in van der Sloot B, Broeders D and Schrijvers E (eds), *Exploring the Boundaries of Big Data* (Amsterdam University Press 2016)

Van den Branden A, *Les robots à l'assaut de la justice* (1st ed., Bruylant 2019)

Vogl R, 'Legal Tech in the USA' in Hartung M, Bues M-M and Halbleib G (eds), *Legal Tech. How Technology is Changing the Legal World. A Practioner's Guide* (Beck-Hart-Nomos 2018)

Wallach H, 'Big Data, Machine Learning, and the Social Sciences: Fairness, Accountability, and Transparency' (*Hanna Wallach*, 19 December 2014) <https://hannawallach.medium.com/big-data-machine-learning-and-the-social-sciences-927a8e20460d> and <https://www.microsoft.com/en-us/research/publication/big-data-machine-learning-and-the-social-sciences-fairness-accountability-and-transparency/> accessed 15 July 2021

Winston A, 'New Orleans ends its Palantir predictive policing program' (*TheVerge*, March 2018) <https://www.theverge.com/2018/3/15/17126174/new-orleans-palantir-predictive-policing-program-end> accessed 14 July 2021

Wolters Kluwer, 'Legal Insights' <https://www.wolterskluwer.com/fr-be/solutions/legal-insights/> accessed 14 July 2021

Yang Y, Yang Y and Fei Ju S, 'China seeks glimpse of citizens' future with crime predicting AI' (*Financial Times*, 23 July 2017) <https://www.ft.com/content/5ec7093c-6e06-11e7-b9c7-15af748b60d0> accessed 14 July 2021

Zouinar M, 'Evolutions de l'Intelligence Artificielle : quels enjeux pour l'activité humaine et la relation Humain-Machine au travail ?' (*Open Edition Journals*, 15 April 2020) <https://journals.openedition.org/activites/4941#quotation> accessed 21 June 2021

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