9. ULUSLARARASI SUÇ VE CEZA FİLM FESTİVALİ
"SANAL DÜNYADA ADALET"
TEBLİĞLER

9th INTERNATIONAL CRIME AND PUNISHMENT FILM FESTIVAL
"JUSTICE IN THE VIRTUAL WORLD"
ACADEMIC PAPERS

Editör/Editor:

Prof. Dr. Adem SÖZÜER

Yayın Kurulu/ Editorial Board

Araş. Gör. Dr. Büşra DEMİRAL BAKIRMAN

Araş. Gör. Büşra ÖZÇELİK RENÇBER

Araş. Gör. Eren SÖZÜER

Araş. Gör. Mehmet Utku SAVARİ

Yazarlar/ Authors

Prof. Dr. Bernd HOLZNAGEL	*Prof. Dr. Kemal İNAN*
Prof. Dr. Krisztina KARSAI	*Prof. Dr. Liane WÖRNER*
Prof. Dr. Murat Volkan DÜLGER	*Prof. Dr. Olgun DEĞİRMENCİ*
Prof. Dr. Pınar MEMİŞ KARTAL	*Prof. Dr. Stephen C. THAMAN*
Prof. Dr. Zeljko NIKAC	*Assoc. Prof. Dr. Cristian Dumitru MIHEŞ*
Doç. Dr. Güneş OKUYUCU ERGÜN	*Doç. Dr. Reşit KARAASLAN*
Dr. Öğr. Üyesi Ali Osman KARAOĞLU	*Dr. Öğr. Üyesi Buket ABANOZ ÖZTÜRK*
Dr. Öğr. Üyesi Cüneyt PEKMEZ	*Dr. Öğr. Üyesi Haluk TOROSLU*
Dr. Öğr. Üyesi Kerim ÇAKIR	*Dr. Öğr. Üyesi Tuba KELEP PEKMEZ*
Dr. Öğr. Üyesi Zafer İÇER	*Dr. Andor GÁL*
Araş. Gör. Büşra ÖZÇELİK RENÇBER	*Res. Asst. Nicolai PREETZ*
Araş. Gör. Ömer Metehan AYNURAL	*Branko LESTANIN*
Öğr. Gör. Yüksek Mühendis Başak BULUZ	*Av. Ceren ÖZBEK*

onikilevha

İÇİNDEKİLER

Algorithmic decisions within the criminal justice pipeline and human rights – an introduction

Krisztina Karsai[1]

1. State of Art - General Challenges

While already in 1963 REED C. LAWLOR indicated 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."[2] LAWLOR predicted that "in a few years' lawyers will rely more and more on computers to perform many tasks for them, they will not rely on computers simply to do their bookkeeping, filing or other clerical tasks. They also them in their research and in the analysis and prediction of judicial decisions. In the latter tasks they make use of modern logic and the mathematical theory of probability, at least indirectly. (…) In the future, lawyers who consider the interest of their client's paramount will not hesitate to employ computers to aid them in solving the problems of their clients to the extent that computer techniques are applicable at the time. And judges who are called upon to decide important may raise their eyebrows at attorneys who do not use of computer facilities in their review and analysis of the law. As scientific methods for analysing the law are further developed. justice may become blind and we may approach Pound's standards of full, equal and exact justice and Holmes's dream of a scientific evaluation of the values of legal decisions."[3] He believed that trustworthy prediction judges' decisions depends ***on a scientific understanding*** of the functioning of the law and how facts and legal norms impact the judge and the judicial decisions. Indeed, almost sixty years later we have made long strides in this field, but the latent infiltration of algorithmic solutions without clear scientific reasoning has led the development in a direction different from what LAWLOR had predicted. Moreover, the COLLINGRIDGE dilemma[4] is more than obvious – the dilemma of control over the new technologies or innovation as such was not solved, but instead just let the development flow. Let's invite REBECA WEXLER to summarise the result of this 'latent infiltration' (for the US at least): "at every stage—policing and investigations, pretrial incarceration, assessing evidence of guilt at trial, sentencing, and parole— machine learning systems and other software programs increasingly guide criminal justice outcomes. Predictive policing technologies identify "hot spot" neighbourhoods. Social media analytics flag at-risk individuals. Forensic scientists use software programs to analyse crime scene evidence, including DNA, fingerprints, ballistics, and face matches. And judges and parole boards rely on risk assessment instruments, which purport to predict an individual's future behaviour, to decide who will make bail or parole and even what sentence to impose."[5]

All the mentioned stages of the criminal justice pipeline are strongly linked to human rights as mirrored in national legal regulatory frameworks and in international protection regimes. This paper aims to highlight the most problematic issues of ***existing ADM*** systems in relation to

---

[1] Full professor of criminal law, at the Faculty of Law, University of Szeged, Hungary

[2] REED C. LAWLOR: What computers can do: analysis and prediction of judicial decisions. American Bar Association Journal 1963, 49 337–344, 337

[3] LAWLOR 1963, 344

[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." DAVID COLLINGRIDGE: The social control of technology. Frances Pinter 1980, 19

[5] REBECCA WEXLER: Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System. Stanford Law Review Vol 70 May 2018, 1343-1429, 1347

human rights and to provide a more in-depth explanation of the fair trial requirement in connection with criminal cases. Current legal and social research waves and a multitude of exciting research papers[6] focus on the application of algorithmic decision-making solutions in every sector of society, and special attention is dedicated to human rights implications – with more than solid justification. I dedicated the foregoing papers to the theoretical and practical principles (paradigm criteria) that shape general thinking about the role and application of algorithms within criminal justice. Overcoming all theoretical burdens seems like a mission impossible (as today), so instead, it is more practical to convey understanding of the traps and struggles of algorithms and automated data processing within criminal justice from the point of view of human rights' protection, because the latter approach has specific and clear ruling on international and domestic levels and falls under jurisdictions of international courts – at least in Europe (European Court of Human Rights and Court of Justice of the European Union). In this paper, I apply the regulatory framework and jurisprudence of the European Court of Human Rights on the European Convention on Human Rights – other regional regimes of human rights protection will be not included for now.

2. Aims of the Paper

In this paper, I intend to continue my research on the connections between the use of algorithmic decision-making systems and the criminal justice system. Part One[7] already touched upon the fundamental definition issues[8] and identified the key terms and risks of any regulation and provided a cross-cultural approach in order to elaborate upon the recent – diverging – legal narratives of the relevant geographic actors. Part Two[9] highlighted the social-legal environment

---

[6] Some examples: EDWARDS, LILIAN AND VEALE, MICHAEL: Slave to the Algorithm? Why a 'Right to an Explanation' Is Probably Not the Remedy You Are Looking For', Duke Law & Technology, 2017, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2972855; HENDRICKX, FRANK AND ALINE VAN BEVER: Article 8 ECHR: Judicial Patterns of Employment Privacy Protection. In: The European Convention on Human Rights and the Employment Relation (eds) F. Dorssemont, K. Lörcher, and I. Schömann. Oxford: Hart Publishing, 2013, 183-208; O'NEIL, CATHY: Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. New York: Crown, 2016; PASQUALE, FRANK A.: Platform Neutrality: Enhancing Freedom of Expression in Spheres of Private Power. Rochester, NY: Social Science Research Network. Retrieved 7 June 2017 (https://papers.ssrn.com/abstract=2779270); VOORHOOF, DIRK AND P. HUMBLET: The Right to Freedom of Expression in the Workplace under Article 10 ECHR. In: The European Convention on Human Rights and the Employment Relation. Oxford: Hart Publishing 2013, 183-208; STEPHANIE BORNSTEIN: Anti-discriminatory Algorithms. https://ssrn.com/abstract=3307893 Alabama Law Review Vol 70/2, 2018, 519-572

[7] KARSAI, KRISZTINA, Algorithmic Decision Making and Issues of Criminal Justice - A General Approach (March 12, 2020). Descrierea CIP a Bibliotecii Naţionale a României OMAGIU. MIRIȘAN, VALENTIN In honorem Valentin Mirişan: Gânduri, Studii şi Instituţii / ed.: conf. univ. dr. Cristian Dumitru Miheş - Bucureşti : Universul Juridic, 2020 Conţine bibliografie ISBN 978-606-39-0615-2, Available at SSRN: https://ssrn.com/abstract=3612106 146-161

[8] „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: Media Technologies Essays on Communication, Materiality, and Society (eds.: Tarleton Gillespie, Pablo J. Boczkowski, and Kirsten A. Foot), 2014 MIT Press Scholarship Online. 167-194 (https://www.researchgate.net/publication/281562384). However, in this paper I take algorithms as simple as it is possible without losing the essence and catch their core components relevant to the present discussion.

[9] Still under review for publication. KRISZTINA KARSAI: Algorithmic decisions within the criminal justice pipeline and their problem matrix. Manuscript, 2020. The paper identifies six main criteria that explain both the lack of necessity and the lack of compliance with system relevant values and characters of criminal justice for the case of application of algorithms. The following criteria will be introduced here: the adaptation traps (how data and information relevant for criminal justice and algorithmicising interplay); the myth of objective truth and of convincing the judge (what is the main goal of the criminal procedure and how the goals are to be achieved if algorithms 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-

of criminal justice as identifying and defining the different needs and possibilities of deploying algorithmic decision-making (ADM[10]) solutions in the distinct stages of the criminal procedure. Part Three – this paper – is dedicated to the challenges emanated by the human right protection regimes regarding the application of algorithms within criminal justice.

This paper aims to highlight the specific requirements of applying ADM solutions within criminal justice from the point of view of the protection of human rights (discrimination, privacy, freedom of expression, fair trial and due process).[11]

3. Human Rights Standards and Algorithmic Decision-Making Systems

ADM solutions are each designed for a specific stage of the criminal justice pipeline; therefore, each uses different datasets and machine learning concepts and obviously they must comply with different legal requirements depending on the given procedural stage. For instance, the rights of the concerned individual to be informed are fundamentally different in the stages of prevention and investigation or at the court. Where public authorities use ADM solutions (affecting the given person), the content of the information obligation depends on the legal status of the person concerned (and the law applicable to him / her), so the information obligation may vary in different phases. Moreover, human rights requirements are inherently different depending on the affected group of persons: the algorithms work with data of potential offenders, of victims (possibly witness), or of members of the criminal justice system. Therefore, the main question here is ***whether*** the result and the quality of the algorithmic decisions regarding human rights ***diverge from those*** taken by human decision makers. The "algorithmic accountability" for human right flaws (or violations) is a core issue to be addressed in the design and use – and evaluation – of ADM solutions.

The emergence of human rights standards and their specification in relation to artificial intelligence and the use of ADM systems has been an important topic in the last decade in academia, in legal research, in the civil sector and the political arena.[12] In this study, I do not

---

mathematizable values of criminal justice (how non-coded values can or cannot play a role in algorithmic solutions); the "bad" subjectivity (whether the subjectivity of the judge shall be excluded from the judiciary or better not), and the purity of the data (why specific data related to criminal justice can be taken hardly for algorithmic solutions as training datasets).

[10] In this paper, the term of artificial intelligence will used and understood as a type of ADM solution.

[11] The very specific requirements of the protection of the right to life and the right of personal freedom play an important role in the context of smart weapons and algorithmically controlled drones (which are simply robots, e.g. algorithms with physical body), which is not addressed in this paper.

[12] Council of Europe - Recommendation CM/Rec(2012)3 of the Committee of Ministers to member States on the protection of human rights with regard to search engines, 2012; Council of Europe – Committee of the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data: Guidelines on the protection of individuals with regard to the processing of personal data in a world of big data, 2017 [Council of Europe study, Big Data, 2017]; Council of Europe – Committee of experts on internet intermediaries: Algorithms and Human Rights. Study on the human rights dimensions of automated data processing techniques and possible regulatory implications, 2017 [Council of Europe study, DGI (2017) 12]; Council of Europe – the CEPEJ (European Commission for the Efficiency of Justice): European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment, 2018; European Union – Ethics Guidelines for Trustworthy AI, High Level Expert Group on AI to the European Commission, 2019; Council of Europe – Ad hoc Committee on Artificial Intelligence: The Impact of Artificial Intelligence on Human Rights, Democracy and the Rule of Law (Catelijne Muller) 2020 [Council of Europe, AI&HR]; European Union Agency for Fundamental Rights: Getting the Future Right. Artificial Intelligence and Fundamental Rights, 2020; European Commission - Whitepaper on artificial intelligence, COM(2020) 65 final.

undertake to present these results in detail, but it is important to emphasize their role: they will have an extraordinary impact on national and European legislation, as I believe they have set the boundaries within which it is advisable to legislate this technological development. The ECtHR (European Court of Human Rights) was not asked to decide in cases when ADM systems were applied, the closest connections can be observed in relation with video surveillance systems or wiretapping etc. – which are far from being automated or semi-automated decision systems. However, it is necessary to start or maintain the discussion to provide a solid basis for necessary legislative steps, as well as for judicial decisions.

ADM solutions are in many cases ***developed by private companies***, which is a crucial factor in complying with human rights requirements. Public actors are undoubtedly bound by human rights regulatory systems (e.g. in criminal justice) in Europe, so human rights standards can be incorporated into the use of bespoke algorithms. The question, however, is to what extent have human rights requirements been considered in the design of technological solutions, especially in the absence of full human rights acknowledgement in relations between individual actors (private entities): the so-called horizontal effects of human rights are highly discussed in academic scholarship[13] and only sporadically applied by private entities. Neither business operators on large scale (multinational companies) nor small tech factories are obliged directly to comply the ECHR (European Convention on Human Rights) however, they must follow the national constitutional regime and the technical norms of law incorporating human rights standards. But still, the "traditional asymmetry of power and information between state structures and human beings is shifting towards an asymmetry of power and information between operators of algorithms (who may be public or private) and those who are acted upon and governed."[14] On a global scale, the introduction of the "Ruggie Principles"[15] made clear that corporations must respect human rights as well, and the duty of any states is to take appropriate steps (policy, legislation, regulation and to make companies to respect.[16]

A further challenge thereto is the application of ***big data*** feed algorithms for decisions targeting individuals, whilst such algorithms are used to identify (or analyse) group behaviour and its tendencies or patterns. In criminal justice, "these tools make determinations about the likelihood of a particular individual reoffending on the basis of others who share similarities to them. It is

[13] LOTTIE LANE: The Horizontal Effect of International Human Rights Law in Practice. European Journal of Comparative Law and Governance. 22 Mar 2018, Volume 5: Issue 1, 5–88; KNOX, JOHN H. "Horizontal Human Rights Law." The American Journal of International Law, vol. 102, no. 1, 2008, pp. 1–47. JSTOR, www.jstor.org/stable/40007767. Accessed 17 Dec. 2020.; IBRAHIM KANALAN: Horizontal Effect of Human Rights in the Era of Transnational Constellations: On the Accountability of Private Actors for Human Rights Violations. July 2016 DOI: 10.1007/978-3-319-29215-1_17; European Yearbook of International Economic Law 2016 432-460.

[14] Council of Europe study, DGI (2017) 12; 33

[15] Guiding Principles on Business and Human Rights: Implementing the United Nations 'Protect, Respect and Remedy' Framework. 2011. The principles were developed by the Special Representative of the Secretary-General on the issue of human rights and transnational corporations and other business enterprises. The Human Rights Council of the UN endorsed the Guiding Principles in its resolution 17/4 of 16 June 2011.

[16] Also, the efforts of the EU on a governance level to establish more effective human right protection within private parties' relationships cannot be underestimated, but such tools or regulatory regimes are still under negotiation, we will see the real impacts after the legislative process has ended. The new EU sanction regime against human rights violations can affect individuals directly in the framework of the EU's Common Foreign and Security Policy. See more https://ec.europa.eu/commission/presscorner/detail/en/ip_20_1939 and https://www.europeanpapers.eu/en/europeanforum/new-eu-sanctions-regime-against-human-rights-violations

foreseeable that these tools could be applied not only to inform, but to actually make decisions in areas such as sentencing, parole or entry into rehabilitation or diversion programmes."[17]

The risk of human rights violations arises beyond the context of applying algorithmic solutions by authorities and private entities in a very specific situation generated by private actors: when an ADM solution will be used ***for illegal purposes by criminals*** and how during this act, human rights will be affected. Criminals can facilitate and improve their attacks on digital systems by maximizing opportunities for profit in a shorter time, exploiting new victims, and creating more innovative criminal business models, while reducing their chances of being caught. Discernibly, this situation is a special case of the horizontal effect of human rights. Generally, the state has the obligation to design the regulatory regimes (constitutional law, police law, criminal law) in a way that human rights violation can be prevented and sanctioned. Criminal activities as such often violate human rights (defamation, murder, kidnapping etc.), but the criminalisation and blameworthiness of any statute offence does not foot on the mentioned horizontal effect theoretically but on the individual violation of social values and of the rules of peaceful coexistence in the society. It is therefore necessary for regulatory systems (national legal systems, international instruments) to respond to the challenge, i.e., that while committing without the use of ADM is a criminal offense, when ADM is deployed, the act could fall outside the scope of criminal law.[18]

4. Human Rights Implications – Selected Issues

4.1. Prohibition of Discrimination (Article 14 ECHR)

General concerns can be raised connection with bias (prohibition of discrimination), with big data (data protection of individuals) and with loss of autonomy (human dignity).[19] However, especially in the criminal justice pipeline, special and particular human rights are also affected and should be highlighted and discussed. Since algorithmic decision-making systems may be based on correlation between data sets and efficiency considerations, there is a danger that such systems perpetuate or exacerbate indirect discrimination through stereotyping. 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 draw patterns. This starting point is obviously correct in legal terms, but in every legal system there is some kind of ***retrospective possibility to remedy*** possible errors (typically in the form of extraordinary legal remedies), the possible rewriting of the subsequent change to the database changes the output. And this can have an impact on the decisions based on it, so such solutions should hardly be avoided.

---

[17] LORNA MCGREGOR – DARAGH MURRAY – VIVIAN NG: International human rights law as a framework for algorithmic accountability. International & Comparative Law Quarterly, Volume 68 Issue 2 pp. 309-343, Published online by Cambridge University Press: 17 April 2019

[18] Relevant criminal statutes shall also cover behaviours in which the criminals are using ADM solutions – with further technological developments, such cases will appear in average criminality. Some possible examples include: defamation will be committed by pattern-recognizing, self-learning algorithms; other person-related defamatory content (pictures, videos) will be search, detected, and published by self-learning algorithms; generating and using fake information about persons through algorithms – in official databases (falsification, fraud etc.); in public online spaces (basis for false accusation); algorithmic stock market manipulations etc. See more Malicious Uses and Abuses of Artificial Intelligence. Trend Micro Research - United Nations Interregional Crime and Justice Research Institute (UNICRI) - Europol's European Cybercrime Centre (EC3), 2020.

[19] British Law Society (2019)

A further paradigm shift is urged by the fact that if past official / judicial decisions are 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 "uncovered" contexts of the algorithm, i.e., we would exclude the infiltration of the human subject from the individual case, but allow it in its cumulative effect.[20] A 'shining' example of this is the software COMPAS, which is used in many jurisdictions within the US to predict reoffending and supporting judges in sanctioning.[21] COMPAS was accused of ***being biased*** against black defendants, because it classifies a greater share of black defendants as high-risk re-offenders than white defendants. Meanwhile, the algorithm assigns defendants scores from 1 to 10 that indicate how likely they are to reoffend based on more than 100 factors, including age, sex and criminal record, while race is not used (!) as indicator.[22] And still. The differences may be results of an interplay between the indicators, which represents (perhaps biased) former human decisions as criminal records or prior arrests (e.g. heavier policing in predominantly black neighbourhoods in certain areas).

It is obvious, "if algorithmic decision-making systems are based on previous human decisions, it is likely that the same biases which potentially undermine the human decision-making are replicated and multiplied in the algorithmic decision-making systems, only that they are then more difficult to identify and correct.[23]

Of course, the question of ***data distillation*** may arise, which is, by definition, a fundamental activity in any database construction. However, if judicial decisions that are made based on the appropriate rules and representing "truth" and legal correctness are entered into the database, it is difficult to find legitimation for the basis on which the data be cleaned.

4.2. Right to Privacy (Article 8 ECHR)

Among human rights, in connection with big data processing and algorithmic processing of personal data, one of the most affected is the right to privacy. Criminal justice is 'the' procedure where privacy is extremely affected due to the nature and aim of the state's actions within this field. However, the numerous existing examples affect privacy to a different degree and with different intensity, depending on the stage at which the ADM solution was used within the criminal justice pipeline and on the purpose of its application. In the early stage of the criminal justice pipeline if ADM solutions are applied for predictive purposes in a public security context (prediction of behaviours, prediction of infection of areas by criminality etc.), it is difficult to separate information protected by privacy and private data connected to any risks relevant for further interventions by law enforcement. Mass surveillance for identifying ***possible*** perpetrators and facial recognition for law enforcement authorities, encountering non-

[20] ALES ZAVRSNIK: Algorithmic justice: Algorithms and big data in criminal justice settings. European Journal of Criminology 2019 (DOI: 10.1177/1477370819876762) 15

[21] See ANTHONY W. FLORES - CHRISTOPHER T. LOWENKAMP - KRISTIN BECHTEL: 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/)

[22] Within each risk category, the proportion of defendants who reoffend is approximately the same regardless of race; this is Northpointe's definition of fairness. The overall recidivism rate for black defendants is higher than for white defendants (52 percent vs. 39 percent). Black defendants are more likely to be classified as medium or high risk (58 percent vs. 33 percent). While the algorithm does not use race directly, many attributes that predict reoffending nonetheless vary by race. For example, black defendants are more likely to have prior arrests, and since prior arrests predict reoffending, the algorithm flags more black defendants as high risk even though it does not use race in the classification. Black defendants who do not reoffend are predicted to be riskier than white defendants who do not reoffend. See press references: Washington Post 17 October 2016

[23] Council of Europe study, DGI (2017) 12, 28

concerned persons by digital surveillance, interception of online communication, wiretapping etc. of ***suspects***, the eventual unlawful gathering of evidence using a variety of other inventions as ADM solutions could have mass consequence on the right to privacy. Also, the indefinite storage of fingerprints, cellular samples and DNA profiles for law enforcement purposes not only affects privacy, but also the presumption of innocence (as one components of the fair trial requirement).

As regards the horizontal effects of the right of the privacy, the acute and centred topic of personal data protection can be mentioned here. ADM solutions are used in online profiling of individuals whose surfing, clicking, and searching patterns are recorded by the systems (cookies, digital fingerprinting etc.). Behavioural data is also processed by other smart devices (e.g. cell phone, smart watch etc.). All these could have consequences on the right provided by Article 8 ECHR. The main concern of using data from profiles for different purposes through algorithms is that the ***data loses its original context***. Repurposing of data is likely to affect a person's informational self-determination. Also, the multiple use of data (beyond the original context) is hard to track and it is difficult to control the legal usage. The use of data from profiles, including those established based on data collected by search algorithms and search engines, directly affects the right to a person's informational self-determination.[24] On the other hand it also raises the question of accessing and processing data ***by unauthorized third parties***, such as corporations enhancing their marketing capabilities or security agencies engaging in unlawful surveillance.[25] The state regulatory regimes shall tackle this challenge in order to create a legal environment preventing, restricting and (if needed) sanctioning the unlawful use of private data.

The interrelation of these issues and the resulting extreme legislative and law enforcement difficulties are embodied in the heroic legal and law enforcement struggle ***against online child pornography***. The privacy rights of child victims and their criminally protected components, the protection of the right to sexual self-determination, and the fundamental rights aspects of the fight against sexual exploitation in general need to be reconciled with the privacy of potential perpetrators (consumers of child pornography and real paedophiles of perpetrating sexual activities with minors) and with the fair trial requirement (presumption of innocence) in case of applying preventive devices using ADM technologies (learning algorithms in chatrooms, pattern-recognizing ADM solutions for text analysis in forums or in chatrooms). The legal environment should be designed in a balanced way between the human rights of the eventual perpetrators and the protection of children against online exploitation and abuse.

4.3. Freedom of Expression (Article 10 ECHR)

Algorithms (search engines) are broadly applied for content-filtering and content-removal mainly in mass social media platforms. As the Committee of Ministers (Council of Europe) stated in 2012

> *"[a] prerequisite for the existence of effective search engines is the freedom to crawl and index the information available on the Web. The filtering and blocking of Internet content by search engine providers entail the risk of violation of freedom of expression guaranteed by Article 10 of the Convention in respect to the rights of providers and users to distribute and access information. Search engine providers should not be obliged to monitor their networks and services proactively in order to detect possibly illegal content, nor should they conduct any ex ante filtering or*

[24] Council of Europe study, Big Data, 2017, 22
[25] Council of Europe study, Big Data, 2017, 22

blocking activity, unless mandated by court order or by a competent authority. However, there may be legitimate requests to remove specific sources from their index, for example in cases where other rights outweigh the right to freedom of expression and information; the right to information cannot be understood as extending the access to content beyond the intention of the person who exercises her or his freedom of expression."[26]

The Ad Hoc Committee in 2020 stressed out that using facial recognition for security reasons in public areas may interfere with a person's freedom of opinion and expression, simply because the protection of 'group anonymity' no longer exists, if everyone in the group could potentially be recognized. This chilling effect could have the consequence that people change their behaviour because of the use of facial recognition, for example they no longer participate in peaceful demonstrations. The same goes for the situation in which all of our data is used for AI-enabled scoring, assessment and performance (e.g. to receive credit, a mortgage, a loan, a job, a promotion, etc.). People might become more hesitant to openly express a certain opinion, read certain books or newspapers online or watch certain online media.[27]

Content filtering (filtering of speech) can affect the freedom of expression as stated above, but furthermore algorithms today are not capable of detecting irony or critical analysis. The filtering of speech to eliminate harmful content through algorithms therefore faces a high risk of over-blocking and removing speech that is not only harmless but can contribute positively to public debate.[28] On the other hand, there is an opposite burden on the side of the developers: the algorithmic solutions have to protect the individual's rights, but also provide the very specific protection of community needs: algorithms are used to filter and remove communication contents that constitute criminal offences, e.g. slander, call for committing a terrorist offence or genocide, hate speech, holocaust denying or nowadays spreading fake news etc. Numerous human rights (privacy, right to freedom, fair trial) can be affected if authorities decide to intervene based on the ***algorithmic filtering*** and banning – allegedly, i.e., labelled by the algorithms as – harmful content.

The current state (research[29] and press[30]) is that algorithms today (e.g., Facebook's hate speech filtering system) are not advanced enough to operate without the need for human assistance to detect harmful content; manual classification is still necessary, but because this is particularly time consuming, and therefore the duly action (removal or flagging) is often delayed, hence human right's violations are – for now – encoded until technology is able to achieve a more sophisticated level. But still, in case of automatic filtering, curtailing of freedom of speech can easily become a usual operation, if the recognition algorithms are not finetuned properly and automated removal systems are processing the removal.

The need for effective protection of human rights is emerging in situations where private entities are able to violate or disrupt the freedom of expression or the ***free flow of opinions*** – suppose if algorithms-led (artificial intelligence) chatbots and fake news generators were producing tons of content with fictional opinions. The state's obligation for designing the regulatory framework for suffocating acts or behaviours that violate human rights is obvious here, and criminal law regulation can have an essential role in it.

[26] Council of Europe, CM/Rec(2012)3, Appendix 12-14
[27] Council of Europe, AI&HR, 2020, 9
[28] Council of Europe study 2017, 20
[29] For example: RAGINI GOKHALE - MARIA FASLI: Matrix factorization for co-training algorithm to classify human rights abuses. December 2018, (DOI: 10.1109/BigData.2018.8622397) Conference: 2018 IEEE International Conference on Big Data (Big Data) http://repository.essex.ac.uk/24234/1/fasli_big_data_03.pdf
[30] For example: https://futurism.com/facebook-human-algorithm-hate-speech (2018)

5. Fair Trial and Due Process – Article 6 ECHR

Fairness in the court procedure itself is more than a codified human right within the ECHR. The right to fair trial is part of the rule of law and an essential pillar of democratic society. According to the jurisdiction of the ECtHR the fair trial requirement has more components than covered herein, therefore, due to the limited space of this paper – I have selected four of them (see below). With regard to algorithms, such solutions can have two functions within the procedures. On the one hand algorithms (artificial intelligence) can support the judge in decision-making through data-analysis (recognition of patterns) and prediction. On the other hand, the algorithm, or the ADM solution (or the artificial intelligence led solution) itself can substitute or replace human decision-making, if we trigger an automatic decision-making system within the decision process.

However, it would be unrealistic to assume that the latter penetrate in the justice systems of Europe anytime in the near future, although it's worth noting that there are already countries where social and legal 'experiments' have been launched in connection with civil claims. If the need and interest is greater for disputes to ***be resolved somehow*** than in the decision being made by a human, a paradigm shift is undoubtedly inevitable. As in Brazil, for example, where the judicial system is burdened by a backlog of many millions of unclosed cases (not criminal cases indeed) and where the access to justice is impossible in more backward areas (no courts, no personnel, etc.), this has happened again with the introduction of several MI-driven algorithmic decision-making systems.[31]

In case of fulfilling a supporting role, the ADM solutions or AI based systems have a certain level of priority in the human psyche - experimental psychological research shows that despite people's knowledge and competences, they are often willing to follow the advice of the AI system without verifying its correctness. Humans trust ADM (AI) systems even if they 'act' or decide in an obviously inappropriate manner. A research involving lawyers demonstrated that people tend to use computer systems to reduce the burden of the decision-making process rather than to increase the quality of their own decisions. DYMITRUK concluded that it is therefore unlikely that that the use of decision support systems would improve civil proceedings.[32]

5.1. Fairness – Equality of Arms and Adversarial Procedure (disclosure of evidence)

Equality of arms is an inherent feature of a fair trial. It requires that each party be given a reasonable opportunity to present his case under conditions that do not place him at a disadvantage vis-à-vis his opponent. Equality of arms requires that a fair balance be struck between the parties and applies to both criminal and civil cases.[33]

As a rule, Article 6 § 1 requires that the prosecution authorities disclose to the defence all material evidence in their possession for or against the accused. In this context, the relevant

[31] https://www.globallegalinsights.com/practice-areas/ai-machine-learning-and-big-data-laws-and-regulations/brazil

[32] See Mari Dymitruk: The right to a fair trial in automated civil proceedings. Masaryk University Journal of Law and Technology, Volume 13 (1) 2019, 31

[33] Guide on Article 6 of the Convention – Right to a fair trial (criminal limb). 140. European Court of Human Rights 29/115 Last update: 31.12.2019 [Guide 2019]

considerations can also be drawn from Article 6 § 3 (b), which guarantees to the applicant "adequate time and facilities for the preparation of his defence" (Guide 2019, 157). However, the entitlement to disclosure of relevant evidence is not an absolute right. In criminal proceedings there may be competing interests, such as national security or the need to protect witnesses who are at risk of reprisals or to keep secret the methods used by the police to investigate crime, which must be weighed against the rights of the accused. In some cases, it may be necessary to withhold certain evidence from the defence to preserve the fundamental rights of another individual or to safeguard an important public interest. However, only measures that are strictly necessary are permissible under Article 6 § 1 regarding restriction of the rights of the defence. Moreover, in order to ensure that the accused receives a fair trial, any difficulties caused to the defence by a limitation on its rights must be sufficiently counterbalanced under the procedure followed by the judicial authorities (Guide 2019, 160).

If algorithmic solutions help law enforcement authorities or prosecutors at least in managing information about a case, then '***knowledge asymmetry'*** within the criminal procedure is deepened, which is difficult to balance through usual and traditional institutions of the criminal procedure with respect to the specialities of the technology (machine learning, big data, black box solutions, and sometimes the protection of intellectual property). Under these circumstances, the application of any algorithm by state authorities represents an overwhelming disproportion between the defence party and the state party (public prosecutor) in terms of possessing information.

QUATTROCOLO and her colleagues pointed out that *"the more evidence becomes technological, the less the parties, and especially the defence, are able to challenge it. Such impairment has, at least, two reasons. The prosecution is usually able to access to the newest technology, with an "indirect" financial exposure, relying on public money, while the defence seldom can afford it. Moreover, if on the one hand, the use of automated systems, per se, suggests neutrality of the method, discouraging any challenge, on the other hand, the defence is scarcely afforded the access to the technology that could allow to challenge the prosecutor's methods*."[34] Digital evidence (if the evidence is result of algorithmic process) can be challenged based on doubt of algorithmic logic or of accuracy of data, meanwhile these challenges seem to not be met by the traditional forensic methods because – if even the experts can only guess how the output was calculated – the trust towards 'experts' becomes non-existent. QUATTROCOLO risks to underline that this shift may even change the role of the parties and the judge in the criminal proceedings.[35]

In the context of disclosure of evidence, complex issues may arise concerning the disclosure of ***electronic data***, which may constitute a certain mass of information in the hands of the prosecution. In such a case, an important safeguard in the sifting process is to ensure that the defence is provided an opportunity to be involved in laying down the criteria for determining what might be relevant for disclosure. Moreover, as regards identified or tagged data, any refusal to allow the defence to have further searches of such data carried out in principle raises an issue concerning the provision of adequate facilities for the preparation of the defence (Guide 2019, 164). Equality of arms requires that data owners provide access to the used data in order to check and guarantee the accuracy of data. Furthermore, the question may be hypothetical, but would it be possible to feed the data into to a 'defence side' algorithm (or ADM solution)?

[34] SERENA QUATTROCOLO, COSIMO ANGLANO, MASSIMO CANONICO, MARCO GUAZZONE: Technical Solutions for Legal Challenges: Equality of Arms in Criminal Proceedings. Global Jurist, 20(1), 2020, 11

[35] SERENA QUATTROCOLO: Artificial intelligence, computational modelling and criminal proceedings. Springer, 2020, 93

If so, will any differences in the outputs have any consequences for the questions to be answered?

In many countries, software and ADM solutions are developed by private companies that have their own interests and rights over the algorithms. It should not be overlooked that "new technologies entering criminal proceedings are bringing intellectual property claims with them (…) and [therefore] future developers of data-driven systems are therefore likely to depend more heavily on trade secret protections."[36] The question here is that whether the logic, the functional model or even the source code of such a solution can be disclosed to the defendant if evidence provided by the algorithmic systems enjoys legal protection as a business (trade) secret. Such ***trade secret barriers*** have already begun to appear (in US) twenty years ago in connection with the first broadly used DNA tests (then later with breath alcohol test), their manufacturer refused to disclose information even in criminal cases, not when the assessment of facts was at stake.[37] Under the fair trial requirement of the European human right regime, similar questions need to be answered under the umbrella of equality of arms. Civil law countries in Europe provide more or less effective protection of business through their procedural codes, which comply with ECHR (no general claim no to disclaim).[38]

5.2. Fairness – Reasoning of Decisions

According to established case-law, reflecting a principle linked to the proper administration of justice, judgments of courts and tribunals should adequately state the reasons upon which they are based (Guide 2019, 166). However, the ECHR – according to the ECtHR – does not require jurors to give reasons for their decision and Article 6 does not preclude a defendant from being tried by a lay jury even where reasons are not given for the verdict. The controversial outcome of the Taxquet case[39] may apply for decisions of algorithmic decision making by analogy. Nevertheless, for the requirements of a fair trial to be satisfied, the accused, and indeed the public, must be able to understand the verdict that has been given; this is a vital safeguard against arbitrariness (Guide 2019, 171). We can apply these expectations for algorithmic decisions as well, but the biggest challenge that remains is the issue of understanding. If algorithms bring an output that will be used as evidence or as the basis of a decision, the affected person should be able to understand the logical and cognitive 'route' leading to the given decision. If no scientific evidence exists (see big data myth[40]) on the causality and if even experts are merely guessing as to what happens within the algorithms, we can hardly expect people – concerned persons and the broader community – to understand and accept decisions of such. Moreover, if the developer or owner of the ADM solution applied in the court

[36] REBECCA WEXLER: Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System. Stanford Law Review Volume 70 (5) 2018, 1350

[37] The first relevant verdict (Vermont State, 2000) excluded such an evidence from the procedure, because the manufacturer did not share the requested information about its product. But most courts eventually found the DNA evidence admissible despite the trade secret methodologies used to analyse it. Wexler describes and analyses the changed landscape after 2015 (Chubbs judgement) in the US and argues that a criminal trade secret privilege is ahistorical, harmful to defendants, and unnecessary to protect the interests of the secret holder. Meanwhile, compared to substantive trade secret law, the privilege overprotects intellectual property. Further, privileging trade secrets in criminal proceedings fails to serve the theoretical purposes behind either trade secret law or privilege law. The trade secret inquiry sheds new light on how evidence rules do, and should, function differently in civil and criminal cases. See more in WEXLER, 2018, 1392

[38] The main differences of the three major paradigms or regulatory systems (European countries, USA and China) have been introduced in KRISZTINA KARSAI: Algorithmic decision making and issues of criminal justice – a general approach In: MIHEŞ, CRISTIAN DUMITRU (ed.) In honorem Valentin Mirişan: Gânduri, Studii şi Instituţii Bucuresti, Universul Juridic SRL., (2020) pp. 146-161

[39] Case of Taxquet v. Belgium (Application no. 926/05)

[40] HANNAH-MOFFAT, KELLY: Algorithmic risk governance: Big data analytics, race and information activism in criminal justice debates. Theoretical Criminology 2019, Vol. 23(4); ZAVRSNIK, 2019

proceeding refers to trade or business secret (as mentioned before), the transparency and the explainability of court decision (based or supported by ADM solution) can be ensured by very clear legal norms for disclosure.[41]

5.3. Fairness – Entrapment

The ECtHR has recognised the need for the authorities to have recourse to special investigative methods, notably in organised crime and corruption cases. It has held, in this connection, that the use of special investigative methods – in particular, undercover techniques – does not in itself infringe the right to a fair trial (Guide 2019, 210). Nowadays it is obvious that algorithms could perform activities that are carried out by undercover agents – chatbots communicating illegal content (harassment, racist or hate speeches) or specific algorithms ordering illegal goods from darknet providers. Only creativity (and human rights) restricts law enforcement in using algorithms for entrapment. We witness the opening of new area of undercover law enforcement activities by operating systems like Sweetie 2.0. for detecting and investigating child pornography.

In this regard, it is worth to mention an extremely innovative but, in some respects, concern raising project from the Netherlands, the Sweetie (1.0 and 2.0). In this Dutch project, the 'electronic' agent, an MI run chatbot with a 3D avatar of a 10-year-old Philippine girl engages in a dialogue in open and private chatrooms to identify paedophile perpetrators.[42] All chat conversations are recorded by the system that assists analysts in classifying 'online predators' by analysing the data. The chat data is processed per chat and a profile is created for each chat partner to assist with the identification of recurring chat partners.[43]

The use of Sweetie for engaging suspects raises two concerns. On one hand "operating Sweetie in public chatrooms upon a general suspicion may constitute non-targeted entrapment that is considered random virtue-testing. (…) This raises concerns in terms of the fair trial rights of the suspects, as undercover powers are usually the exception to the rule, and generally aimed at suspects against whom there is a prior suspicion."[44] Furthermore the direct interaction of Sweetie with possible perpetrators may lead to unlawful incitement of crimes, namely "a direct interaction bears the risk of influencing the suspect and thus leading him or her into committing an offence he/she would have otherwise not committed. Consequently, Sweetie may lead to the facilitation of the crimes it actually intends to prevent."[45]

---

[41] Which is not the situation for example in the U.S. The famous Loomis case showed that business secrecy could be an obstacle for transparency even in cases of criminal courts. Wisconsin v. Loomis, 2015AP157-CR (July 16, 2016)

[42] Tracks Inspector, a Dutch software company developed the software for the Sweetie 2.0 project of the Terre des Hommes children's right organisation. See more https://www.terredeshommes.nl/en/programmes/sweetie-20-stop-webcam-child-sex

[43] "Some concerns raise from the point of view of privacy; when it comes to Sweetie, we can distinguish two situations from a privacy perspective: (1) Sweetie being present in a public chatroom, and (2) Sweetie directly interacting with a suspect one-on-one in a private (video)chat. About Sweetie merely being present in a public chatroom, the privacy infringement seems limited. But of course, Sweetie just being present does not yet serve a clear law enforcement purpose. This may change though if Sweetie records (logs) the conversations in the chat. In these cases, there might be a substantial infringement of privacy. However, as it stands this is still the subject of debate. (…) In any case the goal of Sweetie 2.0 is not to monitor public chatrooms and discussions. Rather, Sweetie is deployed as a lure in the public chatroom in order to engage with potential child sex offenders. As such, the possible privacy infringements that take place in the context of one-one-one conversations and interactions are likely to be more relevant." SIMONE VAN DER HOF - ILINA GEORGIEVA - BART SCHERMER - BERT-JAAP KOOPS (eds.): Sweetie 2.0. Using Artificial Intelligence to Fight Webcam Child Sex Tourism. Springer 2019. 51-52

[44] VAN DER HOF ET AL. 2019, 53

[45] VAN DER HOF ET AL. 2019, 52

In the assessment by ECtHR the substantive test of incitement will be carried out. On the basis of the available information, the ECtHR was able to determine with a sufficient degree of certainty that the authorities investigated the applicant's activities in an essentially passive manner and did not incite him or her to commit an offence, which will normally be sufficient for the ECtHR to conclude that the subsequent use in criminal proceedings against the applicant of the evidence obtained by the undercover measure does not raise an issue under Article 6 § 1 (Guide 2019, 230). The 'essentially passive manner' means that the law enforcement authorities cannot create the offence in order to be able to prosecute it.

However, the heroic combat against global child pornography requires finetuned answers, hence there are no truly effective methods for preventing this type of criminality.[46] Furthermore ECtHR has found in cases dealing with children and vulnerable groups, that positive state obligations can trump negative ones when it comes to securing the physical and moral welfare of children, and states are required to have effective criminal law provisions in place that would not only protect minors, but also effectively deter against grave acts committed against them.[47] Beyond the appreciation of interests and state obligations, the specific nature of the offence needs special treatment as well. Online webcam sexuality – even if it is a forbidden one – requires a certain level of mutual reactions, otherwise there is no possibility of being engaged in it (committing it). This means that the threshold of passivity on the side of state agent Sweetie should be raised for webcam cases, and there is a need for specific assessment in child pornography cases.

5.4. Presumption of Innocence

As mentioned, algorithms can be and are (and will be) used for the prediction of tendencies or concrete events, even if they are acts of human behaviour. If the prediction occurs in a context related to criminal law, for example, prediction on becoming an offender or on being a high-risk citizen from the point of view susceptibility to be violent, the ***key idea is not to presume*** the innocence (hot spot policing; stop and search or arrest based on algorithmic counting; new tools on predicting future crimes but also tools on content filtering could be taken as relevant thereto etc.). The key issue to be addressed here is whether prediction ***turns into prejudice*** and so maybe into racism or discrimination. "Often the systems are based on existing police databases that intentionally or unintentionally reflect systemic biases. Depending on how crimes are recorded, which crimes are selected to be included within the analysis and which analytical tools are used, predictive algorithms may thus contribute to prejudicial decision-making and discriminatory outcomes. (…) As a result, bias may become standardised and may then be less likely to be identified and questioned as such."[48] If Article 6 ECHR § 2 states that everyone charged with a criminal offence shall be presumed innocent until proven guilty according to law, a need for clarification arises concerning whether the entire criminal justice pipeline is affected. According to the jurisprudence of ECtHR

> *"Article 6 § 2 is aimed at preventing the undermining of a fair criminal trial by prejudicial statements made in close connection with those proceedings. Where no such proceedings are or have been in existence, statements*

[46] The threshold to engage in webcam sex tourism is low and the chances of getting caught are as of yet minimal. Perpetrators can further reduce the chances of being caught by using fake identities, but also various anonymisation services and hidden servers (to name just a few) to prevent detection. As such, the chances of identifying a suspect after the webcam sex stream has been concluded is likely low. The best chance of finding a suspect, is thus to catch them in the act. See in VAN DER HOF ET AL. 2019, 54

[47] See the case KU v. Finland (Application no. 2872/02 43, 46§)

[48] Council of Europe study, DGI (2017) 12,

attributing criminal or other reprehensible conduct are more relevant to considerations of protection against defamation and adequate access to court to determine civil rights, raising potential issues under Articles 8 and 6 of the Convention. Moreover, the prejudicial statements must concern the same criminal offence in respect of which the protection of the presumption of innocence in the context of the latter proceedings is claimed. (…) A fundamental distinction must be made between a statement that someone is merely suspected of having committed a crime and a clear declaration, in the absence of a final conviction, that an individual has committed the crime in question. The latter infringes the presumption of innocence, whereas the former has been regarded as unobjectionable in various situations examined by the Court. (Guide 2019, 328, 330), The presumption of innocence may be infringed not only by a judge or court but also by other public authorities The voicing of suspicions regarding an accused's innocence is conceivable as long as the conclusion of criminal proceedings has not resulted in a decision on the merits of the accusation (Guide 2019, 333).

To round up, the presumption of innocence applies formally only in the last stage of the criminal procedure (trial); however, limited applicability related to its moral value (reputation related aspect) has been acknowledged by the ECtHR in pre-trial (and post-trial) situations. As GALETTA pointed out "the legal presumption of innocence is recognised and regulated by law, whereas the moral presumption of innocence is not. These considerations urge the need to investigate whether the moral presumption of innocence might have any legal relevance in the pre-trial stage of a criminal proceeding."[49] This could be especially relevant for law enforcement authorities by applying algorithms for detecting offences, collecting evidence and for predictive policing.

The main conflicting issue to be highlighted here is (as usual in connection with policing) whether a low ***threshold for suspicion*** (level of perceiving the reality and its hypothetical connection to the given person) can impact the presumption of innocence. In other words, whether human rights are violated if authorities intervene after algorithmic solutions have identified or predicted someone's involvement in an already committed or likely-to-be-committed-in-the-future offence. More generally, whether the digital technologies will challenge presumption of innocence. MENDOLA argues that predictive policing using algorithms provides an important result "the possibility for improved accuracy on data about criminals, a reverse approach in policing with the possibility to exonerate an individual from all charges, a considerable reduction of costs for policing and a demonstrated crime reduction in urban areas constitute very precious supports." On the other hand, he also stated that "some endemic drawbacks such as false positives, bad data/lack of transparency and a reciprocal lack of trust between government and its citizens" can be also observed. Based on his research, it can be underline that "there is ***a general contrast or lack of coherence between the current ECtHR case law*** orientation on presumption of innocence and reasonable suspicion with the predictive policing model." Currently, the structure of Article 6 ECHR is founded on a reactive framework where law enforcement acts in consequence of the commission of a crime. Conversely, predictive policing is based on pre-crime models in which intelligence and preventive actions are employed before the criminal event. This new surveillance paradigm is not even integrated into Article 6 ECHR; thus, such right do not offer adequate answers to violation of presumption of innocence perpetrated by surveillance systems. (…) Overall, the

[49] ANTONELLA GALETTA: The changing nature of the presumption of innocence in today's surveillance societies: rewrite human rights or regulate the use of surveillance technologies. European Journal of Law and Technology, Vol. 4, No. 2, 2013, 4

influence of predictive policing on the presumption of innocence and reasonable suspicion principles may constitute clear evidence of malleability of such fundamental guarantees, in which predictive projections has to be seen as extra source in policing, evaluated and contextualized by public officers within a broader perspective."[50] GALETTA calls for expansion of the scope of the presumption of innocence or urges the implementations of new provisions on the use of surveillance technologies, which he claims can be extended to the entire terrain of using new proactive technologies.[51] I think that we are at the stage of development where we could overcome the COLLINGRIDGE dilemma through dedicated regulatory steps in these cases through establishing domestic legislation on using such technologies within the criminal justice pipeline, but changes should be made instantly.

In the jurisdiction of the ECtHR the creation of suspicion is one of the main themes when analysing the impact of surveillance technologies (on the presumption of innocence). The ECtHR underlines that suspicions stem from public perceptions that can be somehow warped by stigmatisation. The Court admits implicitly that public perceptions can play a crucial role in defining the culpability of an individual, especially in the post-trial stage. In addition, the Court seems to acknowledge that they have a certain relevance within legal reasoning, considering that individuals should be given protection against false perceptions. As a consequence, in S. and Marper the ECtHR ascribes public perceptions to legal significance. Nonetheless, it recognises that the presumption of innocence does not only give rise to a human right, but also to a moral value that should be safeguarded.[52]

***

The concept of human rights and their further development is necessary as – we know it well – "the Convention is a living instrument which … must be interpreted in the light of present-day conditions"[53]'. The challenge posed by ADM solutions and AI led systems within (criminal) justice calls for present-day new interpretations. And clear legislative steps both in national and international context.

Thereby, I argue for non-application of algorithmic systems within the criminal justice pipeline until we can deliver proper answers to the questions it raises. [54]

[50] MARCO MENDOLA: One Step Further in the 'Surveillance Society': The Case of Predictive Policing. Tech and Law Center, October 2016. www.techandlaw.net; 22
[51] GALETTA 2013, 8
[52] GALETTA 2013, 6
[53] Tyrer v. the United Kingdom (application No. 5856/72)
[54] This research was supported by the project nr. EFOP-3.6.2-16-2017-00007, entitled Aspects on the development of intelligent, sustainable and inclusive society: social, technological, innovation networks in employment and digital economy. The project has been supported by the European Union, co-financed by the European Social Fund and the budget of Hungary.