



Human Rights in the Era of Automated Decision Making and Predictive Technologies

Desara Dushi*

Abstract: Advances in Natural Language Processing and machine learning have made it possible to design predictive models that can be used to assist judicial proceedings. Such technologies transform the legal profession; guaranteeing that this does not disrupt the rule of law, access to justice, fair trial and contestability is a substantial challenge.

In what follows, I analyse the impact of AI systems on two fundamental rights: the right to a fair trial and due process, and the right to an effective remedy, concluding with potential solutions on human rights-based approaches to AI systems.

The right to a fair trial and due process

The use of algorithms in legal decision-making and legal prediction in courts raises concerns with regard to the right to a fair trial as stipulated in Article 6 of the [European Convention on Human Rights](#) (ECHR) and the principle of equality of arms. This is especially concerning in the criminal justice system where we witness a trend towards using automated processing techniques and algorithms, such as facial

*Desara Dushi, an [EMA](#) alumna, is a senior and postdoctoral researcher at the Law, Science, Technology and Society Research Group (LSTS) of Vrije University Brussels. She is involved in the [CoHuBiCol](#) and [ALTEP-DP](#) projects working in the cusps of law and technology.

recognition technology, or automated processing techniques for the determination of the length of a prison sentence. The use of such techniques has direct implications for the presumption of innocence, the right to be informed promptly of the cause and nature of an accusation, the right to a fair hearing and the right to defend oneself in person. In the field of crime prevention, where the main concerns relate to predictive policing (the algorithm drawing conclusions for possible future patterns of crime, impacting not only recidivists but also [individuals who have never been involved in criminal activity](#)), the approach may also affect the right to protection against arbitrary deprivation of liberty, stipulated in Article 5 of the ECHR, and the right not to be punished without a law, stipulated in Article 7 of the ECHR. In addition, there is concern that such systems may standardise pre-existing bias, which would be less likely to be identified due to the '[black box](#)' phenomenon behind the operation of such technologies. The lack of explicability behind the algorithm, their opacity and unpredictability, makes interpretation of their output impossible. Such lack of interpretation leads to inequality of arms, preventing individuals from accessing their right to an effective remedy. Algorithms are also being developed to [replace judges in courtrooms](#). While it might be too early to talk about robot judges who can decide cases autonomously, existing systems can assist judges and lawyers by quickly assessing relevant resources and unveiling patterns hinting to a potential outcome of the case (so-called 'predictive justice'). However, these technologies come with certain risks. They may be inappropriately used by judges and lawyers, who might over-trust the suggested predictions, or over rely on the suggested relevant resources. Increased reliance on predictive justice also jeopardises continuous learning on the part of judges and lawyers who might not look beyond the algorithms to prepare their cases.

The right to an effective remedy

The right to an effective remedy, stipulated in Article 13 of the European Convention on Human Rights, implies the right to a *reasoned* and individual decision that allows for contestability. This is closely related to the principle of accountability—basically that those who do wrong should be held accountable before the law and pay damages to the affected individual(s). But what if the damage is caused by a machine?

[The European Union General Data Protection Regulation](#) gives individuals the right to contest and request a review of automated decision-making that significantly affects their rights or legitimate interests. However, as [Hildebrandt](#) emphasises, the opacity of the AI systems makes contestability of their decisions non-viable. Without touching upon the never-ending debate over the liability of machines or their providers, let us focus on their explicability. One of the main requirements of a

trustworthy AI system, based on [the ethical guidelines of the European Commission](#), is the requirement of accountability. In other words, AI systems should be able to explain the 'how' and 'why' behind their outputs. This would [require](#) transparency over their data collection, algorithm training, data selection for modelling or profiling, and the management of individual consent, effectiveness and error rates of the algorithm among others. The opaqueness of the decision of algorithmic processes and its rationale create particular challenges to individuals' ability to apply their right to redress and to actually obtain an effective remedy in practice. This puts individuals in a position of inequality of arms, leaving them unable to defend themselves by applying their right to redress against any decision rendered by the machine that negatively affects them.

A way forward

In April 2021 the European Commission proposed a legal framework on artificial intelligence (the [AI Act](#)), which follows a risk-based approach in promoting the development of AI. While a good start, the safeguards proposed to regulate AI systems are mostly subject to a system of self-certification, rather inadequate to guarantee an effective control since such analyses risk being subjective rather than objectively analysing human rights risks.

There has also been criticism of the ambiguity of the criteria in the AI Act proposal. However, very clear criteria would risk limiting the scope of the proposal in light of the unpredictability of future AI developments.. A better approach would be a rights-based approach. This would entail the obligation of providing a human rights impact assessment for every new AI system and each time these systems are significantly modified, proving their compatibility with fundamental rights, and not only for the ones classified as high-risk.

An approach to preserve legal protection against technologies that have an impact on fundamental rights has been proposed by Hildebrandt. She argues that [legal protection by design](#) makes it possible for technology to be subject to democratic scrutiny and allows the contestability of its outcome by those affected by it. 'By design' means that procedural checks and balances are incorporated into the default settings of the technology. This will allow preservation of the core rule of law safeguards. In sum, while a complete ban on these technologies—which bring several benefits to our life would not be the ideal solution, when it comes to automated decisions that have a significant effect on fundamental rights, the central tenets of the rule of law, such as interpretability and contestability, should be guaranteed by design.