



EUROPEAN COURT OF HUMAN RIGHTS
COUR EUROPÉENNE DES DROITS DE L'HOMME

Opening of the Judicial Year 2025

Judicial Seminar

Artificial Intelligence and the Right to a Fair Trial

Speech by Vytautas Mizaras

31 January 2025

1. Introduction

Over time, as technology has advanced from early computers to the era of artificial intelligence (AI), perspectives on its role in the judiciary have evolved as well. It continues to trigger debates surrounding the integration of algorithmic decision-making in legal practice and, more specifically, in the judiciary. The absence of a human judge in such cases is likely to violate the right to a fair trial

These reflections highlight the broader tension between optimism about technology's ability to enhance efficiency and precision, and concern about its implications for human judgment, accountability and fairness in the rule of law.

1.1. The objectives of the presentation

This presentation aims to explore potential roles of AI systems in the judiciary – which I will refer to as “e-judges” – within the framework of the right to a fair trial as protected by national constitutional law and international law (*inter alia*, Article 6 of the European Convention on Human Rights – hereinafter “the Convention”).

The right to a fair trial lies at the heart of human rights protection. It represents one of the most fundamental principles of any democratic society governed by the rule of law.

Central to this principle is the indispensable role of a human judge in legal proceedings. Given that the justice system is inherently a social process, two critical questions arise:

- Would an e-judge be compatible with the values underpinning such a system?
- Would a fully autonomous AI system be capable of analysing claims and evidence, rendering decisions, and entirely replacing human judges?

2. Taxonomy of e-judges

Rapidly evolving AI technologies are likely to reshape various areas of the judicial process.

The idea of e-judges – artificial intelligence systems integrated into the judicial process – attracts both fascination and scepticism.

In the following section of my presentation, I explore three conceptual models of e-judges, each one of which offers distinct advantages and challenges.

I have separated the three categories of e-judges according to their involvement in judicial proceedings and the degree of human involvement.

2.1. “Argument development AI”

The first conceptual model is “argument development AI”, where an e-judge (or, to be more precise, an e-clerk) is designed to support judicial reasoning by analysing vast repositories of legal texts, precedents and case-law.

Imagine a digital assistant that combs through past judicial opinions, identifying the most relevant arguments and organising them into a coherent, structured form. Such a system would not replace the judge but would act as a guide, presenting a range of possibilities and ensuring that no critical angle is overlooked.

This model’s primary virtue lies in its ability to enhance thoroughness and efficiency.

Argument development AI can alleviate this pressure by performing exhaustive analyses in seconds, allowing judges to focus their attention on the most important aspects of a case.

The shortcomings of this model relate to algorithms that need to be carefully calibrated to avoid bias, and their outputs must remain transparent and easy to explain.

Importantly, the judge’s role as the ultimate arbiter of justice must remain intact; **AI should serve as a tool, not a substitute for human judgment.**

2.2. “Supervised e-judge”: a human judge-in-the-loop

The second model, the **supervised e-judge**, takes a more active role in the judicial process by assisting in the drafting of judgments.

This system goes beyond identifying relevant arguments.

It synthesises them into preliminary drafts based on the judge’s input and the case materials provided. The judge’s task then shifts from creating judgments from scratch to refining and finalising drafts generated by AI.

This approach has clear practical advantages.

Drafting judgments is a time-consuming process that demands precision and clarity. By automating the initial stages, the supervised e-judge can significantly reduce the workload of judges, enabling them to devote more time to complex legal reasoning and deliberation.

Additionally, this model can help standardise the structure and language of judgments, making them more accessible to the public and easier to analyse for legal professionals.

This model is both pragmatic and likely to be implemented in practice, because it is designed to assist judges rather than replace them.

The human judge retains a central role in delivering justice, as the draft opinion produced by the “supervised e-judge” is presented only as a preliminary draft, requiring careful review.

In this framework, the judge remains the ultimate arbiter, making decisions based on the facts of the case and a human sense of justice.

Put simply, the *human judge stays firmly in the loop* and at the heart of the justice-delivery process.

2.3. Autonomous e-judge

The most ambitious model is the autonomous e-judge, an AI system capable of independently deciding certain types of cases.

This model could be particularly suited to routine or low-stakes matters, such as traffic violations or small claims disputes.

By handling these cases autonomously, the system could free up human judges to focus on more complex and high-stakes issues.

Autonomous e-judges could dramatically reduce case backlogs, ensuring faster resolution and greater access to justice for under-served groups.

Moreover, by applying consistent standards across cases, these systems could enhance the perception of fairness and impartiality in the judicial process.

But the introduction of fully autonomous e-judges raises profound ethical and practical concerns.

Can a machine truly understand the human context of a legal dispute? How should errors or controversial decisions be addressed? And, most importantly, could we as a society entrust delivering justice – an innate concept that we humans feel – to an algorithmic agent?

Each of these conceptual models of an e-judge offers unique opportunities to enhance the judicial process. Yet their successful integration depends on careful design, rigorous oversight and an unwavering commitment to preserving the human values at the heart of justice.

In any discussion about judicial systems enhanced by AI, maintaining a “**judge-in-the-loop**” is paramount.

This concept ensures that while technology can assist, it does not supplant the essential human qualities of judgment, empathy and discretion. By keeping judges actively involved, we honour the foundational principles of justice and fairness that resonate deeply with the public. This balance between human oversight and technological aid is critical for a system that aims to be both efficient and just.

The integration of AI in judicial processes should always strive to reinforce, rather than replace, the human elements that are essential to justice. By prioritising a judge-in-the-loop approach, we preserve the trust and confidence that society places in its legal institutions, ensuring that our pursuit of justice remains both fair and humane.

The next section of the presentation will explore whether such a model of AI decision-making aligns with the essential elements of a fair trial.

The goal is to determine whether these fundamental guarantees can be preserved when justice is administered not by humans, but by machines.

3. The right to a fair trial and AI

3.1. The right to a fair trial: the scope of protection

Judges have the moral obligation to ensure, i.e. to monitor and preserve, fairness in every step of the procedure and throughout proceedings as a whole.

If these requirements are met, it can be concluded that the trial has been fair. Conversely, a breach of any of these elements can result in the violation of the right to a fair trial.

The following part of the presentation looks into different aspects of the right to a fair trial.

Specifically, I have divided the guarantees of that right in relation to the use of AI in the justice system **into two groups**:

- (a) those that can be to some extent preserved through the use of AI systems; and
- (b) those that may be compromised.

Among the guarantees that may to some extent be compatible with the use of AI systems, I would include the following: **independence and impartiality, the length of proceedings and legal certainty**.

In contrast, **the guarantees that may be at risk owing to the use of AI systems include: the adversarial principle and equality of arms, the reasoning of judicial decisions and the principle of a public hearing**.

I will now discuss how the use of AI in the justice system can, in my opinion, most likely be an obstacle to the implementation of the right to a fair trial.

3.2. Problematic aspects of ensuring the guarantees of the right to a fair trial

3.2.1. The adversarial principle and equality of arms

The European Court of Human Rights (hereinafter “the Court”) has consistently held that the right to adversarial proceedings and the principle of equality of arms are fundamental components of a fair trial.

These rights ensure that all parties have the opportunity to know and respond to the evidence or observations presented that may influence a court’s decision.

When integrating AI-based tools into the decision-making process, such principles take on renewed significance.

For the right to adversarial proceedings to be preserved, it is crucial to make both quantitative and qualitative information about the AI system being used publicly accessible – especially to the parties involved in the trial.

Such transparency allows parties to understand the construction of the system’s scales, assess their limitations and challenge them effectively before a judge (CEPEJ 2019, 47).

Transparency must extend to the workings of the AI algorithm itself, which must be communicated in clear, understandable human language.

If algorithms are shielded as trade secrets, the parties are deprived of the opportunity to fully engage with all the evidence influencing the decision. This undermines the right to adversarial proceedings, as it

denies parties the ability to test and debate the reliability of the AI system – a core requirement of procedural fairness.

The principle of equality of arms demands a “fair balance” between the parties in judicial proceedings (ECtHR 2019, 65). Fairness in this context is not merely procedural; it is substantive, ensuring that neither side is placed at a significant disadvantage. However, the increasing use of AI in courtrooms poses a potential challenge to this equilibrium.

AI-based tools may enhance efficiency and accessibility for certain parties – such as the State, well-resourced companies or individuals with high levels of technological literacy. They can simultaneously create barriers for others, particularly those less familiar with technology or without access to adequate resources (CEPEJ 2019, 47-48). For instance, as AI systems and self-learning algorithms grow more complex, judicial actors may increasingly depend on computer experts to navigate and explain these technologies. This reliance introduces a risk: financially dominant parties are more likely to secure access to such expertise, potentially skewing the balance of power in their favour. Without careful oversight, this disparity can lead to an erosion of the fairness guaranteed by an equality of arms.

To prevent such outcomes, courts must adopt measures to ensure that all parties, regardless of their financial or technological capacities, have equal opportunities to access the tools and expertise required to meaningfully participate in proceedings. Addressing these risks is not only a matter of fairness but also essential to maintaining public trust in the justice system.

To uphold the principle of equality of arms, it is essential that when a judge intends to rely on AI-based technology in decision-making, the affected party is granted access to the AI algorithm. This access should enable the party to challenge the algorithm’s scientific validity, scrutinise the weight assigned to various elements, and identify any potential errors in its conclusions (CEPEJ 2019, 55).

3.2.2. The reasoning of judicial decisions

Despite significant advancements, AI systems remain unable to provide meaningful explanations or justifications for their decisions – at least for now.

While AI-generated outcomes may often be accurate, statistical and mathematical correlations alone are insufficient to meet the rigorous standards of a reasoned decision. This shortfall is particularly critical in criminal matters, where individualised justification is not just expected but required.

The absence of a human judge in such cases is likely to violate the right to a fair trial, as there would be no judgment articulating the legal reasoning behind the decision. Without this foundation, the essential transparency and accountability of judicial decisions would be lost.

If AI-based technology is to be employed in the judiciary, it must assist rather than replace human judges.

The inclusion of human reasoning is not merely a procedural formality but a substantive requirement to ensure that judgments are comprehensible and legitimate and that they meet the high standards of justification demanded by the rule of law.

The inability of AI systems to generate reasoned decisions poses a significant challenge to the effective exercise of the right of appeal. Without a clear explanation of the reasoning behind a decision, it becomes practically impossible for a party to mount a meaningful appeal.

This issue is compounded when the algorithm or AI system in question is developed by a private company that shields its workings under the guise of trade secrets. Such opacity undermines the transparency and accountability necessary for justice, creating barriers that erode trust in the judicial process and jeopardise the fundamental right to a fair trial.

Legitimacy is foundational to the judiciary, and its importance becomes even more pronounced when AI is incorporated into decision-making. Machine-learning systems, often described as data-driven black boxes, lack explicit legal knowledge or expertise. Consequently, their predictions or recommendations may not be viewed as reasonable or persuasive by human participants in the judicial process.

A critical element of the judiciary's legitimacy – and the parties' respect for its decisions – is the perception that judgments are rendered by a fellow human being, capable of understanding and empathising with their circumstances.

It is questionable whether individuals will feel that they sufficiently understand an algorithm's conclusions to accept and respect a judgment based on them.

Thus, when an AI assistant supports a judge in the decision-making process, it is imperative for the human judge to explain why the AI system's outcome is convincing and to provide a legal justification that meets the standards of a reasoned decision. Without such a clarification, the legitimacy of the process – and the public's trust in it – may be undermined.

3.2.3. A public hearing

The requirement of a public hearing poses significant challenges for the concept of an autonomous e-judge.

AI systems, by their nature, lack transparency and are often described as opaque "black boxes".

This opacity creates a serious risk:

- Justice administered by an autonomous e-judge could occur in secrecy, shielded from public scrutiny.
- The risk is further compounded when these AI systems are developed by private companies that treat their algorithms as trade secrets, limiting oversight and accountability.
- Additionally, the simple cases that might, in theory, be delegated to an autonomous e-judge could often be heard by a court of first and only instance. Such cases are entitled to an oral hearing, which raises the question of how an AI system – lacking human presence and engagement – could fulfil this fundamental requirement of the right to a fair trial.

These concerns highlight the critical tension between technological efficiency and the principles of transparency, accountability and public confidence in the judicial process.

An autonomous e-judge, by its very nature, would fail to meet key procedural safeguards guaranteed under protection of the right to a fair trial. Without a public courtroom to administer justice or a human judge to publicly deliver the judgment, such a system would not comply with the rights to a public and oral hearing or the public delivery of judgments. These are foundational elements of transparency and accountability in the judicial process.

Conversely, when AI systems are employed as assistance tools, with human judges maintaining an active role in court proceedings, these requirements can still be fulfilled. The involvement of a human judge ensures the administration of justice remains public, transparent and accessible, thereby preserving the essential guarantees of the right to a fair trial.

3.2.4. Independence and impartiality

Judicial independence requires that judges critically evaluate AI-generated recommendations and prioritise values and considerations that algorithms cannot fully capture or preserve.

A judge should never abdicate his or her responsibility to exercise independent judgment and must avoid blindly deferring to an AI system's output.

When AI systems are used in the decision-making process, they must comply with the principles of the rule of law and judicial independence (CEPEJ 2019, 8). The critical question is whether a judge's instinct and judgment are strong enough to challenge an AI-generated result when justice demands it. Judicial independence cannot be safeguarded if judges are overly influenced by AI systems, undermining their own reasoning and discretion.

4. Conclusions

1. If AI systems are introduced into the courtroom, there is a risk that foundational principles such as the adversarial principle, equality of arms, the right to a public hearing and judicial independence may be undermined. In some respects, AI – particularly in the form of an autonomous e-judge – conflicts with the core values of the right to a fair trial. The most significant obstacle lies in its inability to provide reasoned legal justification for decisions. This shortfall leads to the conclusion that, with current technology, the use of an autonomous e-judge in the courtroom is not feasible.

2. Nevertheless, in the light of the Court's approach to interpreting the Convention as a "living instrument" that evolves alongside societal developments, we cannot entirely rule out the possibility of an autonomous e-judge in the future. The concept of an autonomous e-judge remains a work in progress. An autonomous e-judge could be tasked with resolving smaller, repetitive cases of a non-complex nature – cases characterised by clear outcomes and minimal evidentiary or interpretive challenges

For this to happen, society must approach AI with an open mind while ensuring that rigorous standards are upheld. Still, significant progress must be made before an autonomous e-judge can fully meet the demands of a fair trial.

The "human touch" remains irreplaceable, as the role of a human judge extends far beyond that of an information processor or problem solver. Judicial proceedings without human involvement would be undesirable from a societal perspective, given the unique capacity of human judges to deliver justice grounded in empathy, discretion and moral reasoning.

3. By contrast, objections to AI assisting human judges appear less significant, provided judges maintain a critical stance towards its use. AI assistants in courts might offer clear advantages: they can enhance legitimacy and fairness by identifying and reducing extraneous influences and individual biases, thereby minimising human error. Additionally, AI can streamline judicial processes, making court proceedings less time-consuming and reducing costs. Greater efficiency through AI has the potential to enhance legal certainty and improve access to justice.

4. The primary challenge lies not in AI itself but in its imperfection. Current AI systems cannot fully ensure the guarantees of the right to a fair trial while meeting the requirements of the rule of law. Law must not adapt to technology, nor should the principles of the rule of law – especially those enshrined in Article 6 of the Convention – be weakened to accommodate AI.

Instead, technology must be designed to operate in compliance with the rule of law. Lawyers and legal professionals must "dictate" the tasks and boundaries for technology developers, ensuring that innovations serve justice rather than undermine it.

While technology will undoubtedly continue to transform the world, it must not replace the rule of law. If it does, we risk becoming hostages to technological manipulation – **entering an era of “technologocracy”**, where technology, rather than law, dictates societal norms and decisions. Such a future would jeopardise the foundational values of justice, fairness and human dignity.