

**EUROPEJSKI PRZEGLĄD PRAWA I STOSUNKÓW MIĘDZYNARODOWYCH /  
EUROPEAN REVIEW OF LAW & INTERNATIONAL RELATIONS**

**Nr [No] 2/2026/78**

**Kwartalnik / Academic quarterly  
Ukazuje się od 2004 roku / Published since 2004**

**WYDAWCA / PUBLISHER**

**Europejska Wyższa Szkoła Prawa i Administracji w Warszawie, ul. Ogrodowa 58 /  
European University of Law and Administration in Warsaw, 58 Ogrodowa St.**

**ISSN 2081-0903 e-ISSN 2957-0980**

**KOLEGIUM REDAKCYJNE / EDITORIAL BOARD**

**Redaktor naczelny / Editor-in-chief**

Prof. Jerzy J. Wiatr

**Zastępca redaktora naczelnego / Deputy editor-in-chief**

Prof. Barbara Bajor

**Członkowie / Members**

Prof. Artur Kotowski

Dr Magdalena Dziedzic

**Członkowie zagraniczni / Foreign members**

Dr Tamar Gegelia (Tbilisi, GE)

Prof. Laura Miraut Martín (Las Palmas, ES)

Prof. Aleksander Waszkiewicz / Alexander Vashkevich (Wilno / Vilnius, LT)

**Sekretarz redakcji / Editorial secretary – Tadeusz T. Nowacki**

**Zastępca sekretarza redakcji / Deputy editorial secretary – Daria Pajek**

**Redaktor graficzny / Graphic editor – Małgorzata Głuszczyk-Draczuk**

**Projekt okładki / Cover design – Stefan Sobczak & Tomasz Czajka**

**RADA PROGRAMOWA / PROGRAM BOARD**

**Przewodniczący / Chairman**

Prof. Tomasz Szanciło (Warszawa / Warsaw, PL)

**Wiceprzewodniczący / Vice-chairman**

Prof. Roman Wieruszewski (Warszawa / Warsaw, PL)

**Członkowie / Members**

Prof. Lilija Aczłowa / Liliya Achilova

(Taszkent / Tashkent, UZ)

Prof. Richard Albert (Austin, US)

Prof. Ignacio Ara Pinilla (San Cristóbal de La Laguna, ES)

Prof. Anton Bebler (Lublana / Ljubljana, SI)

Prof. Trinidad Bergamasco (Córdoba, AR)

Prof. Dominik Bierecki (Słupsk, PL)

Prof. Sławka Dimitrowa / Slavka Dimitrova (Burgas, BG)

Prof. Giorgi / George Goradze (Tbilisi, GE)

Prof. Antonios Karaiskos (Kioto / Kyoto, JP)

Prof. Michał Kowalski (Warszawa / Warsaw, PL)

Prof. Roman Kuźniar (Warszawa / Warsaw, PL)

Dr Khushboo Malik (Delhi, IN)

Prof. Felipe de la Mata Pizaña (Meksyk / Mexico City, MX)

Prof. Rubén Miranda Gonçalves (Las Palmas, ES)

Prof. Wataru Miyasaka (Tsukuba, JP)

Prof. Giacomo Oberto (Bologna / Bologna, IT)

Prof. Adeola Olufunke Kehinde (Lagos, NG)

Dr Mohammad Owais Faruki / Farooqui

(Szardża / Sharjah, AE)

Prof. Seweryn / Sev Ozdowski (Sydney, AU)

Prof. Krzysztof Pałeczki (Kraków, PL)

Prof. Jeffrey Pojanowski (Notre Dame, US)

Dr Andra Nicoleta Puran

(Bukareszt / Bucharest, RO)

Prof. Carlos Jorge Quete (Luanda, AO)

Dr Fábio da Silva Veiga (Porto, PT)

Prof. Jacek Sobczak (Warszawa / Warsaw, PL)

Prof. Antonio Sorela Castillo (Logroño, ES)

Prof. Beata Stępień-Załučka (Rzeszów, PL)

Prof. Emőd Veress (Miskolc / Miskolc, HU)

Prof. Anna Zalcewicz (Warszawa / Warsaw, PL)

Prof. Igor Zgoliński (Bydgoszcz, PL)

Prof. Luigi Zingone (Trydent / Trento, IT)

**ADRES / ADDRESS**

00-876 Warszawa ul. Ogrodowa 58, budynek B, 2. piętro / 00-876 Warsaw 58 Ogrodowa St., Building B, 2nd floor, PL  
tel. +48 22 18 24 942

**epism.ewspa.edu.pl**

Drugi w tym roku numer naszego kwartalnika, podobnie jak to miało miejsce w roku ubiegłym, ma charakter tematyczny. Publikujemy w nim referaty wygłoszone na międzynarodowej konferencji naukowej zatytułowanej: „Artificial Intelligence in Legal Proceedings Worldwide – A Threat to Human Rights?” Prezentację i omówienie tych tekstów zawiera wstęp autorstwa redaktora naukowego numeru, prof. Tomasza Szanciły.

Uczelnia nasza stawia sobie m.in. za cel podejmowanie na konferencjach naukowych oraz w pracach badawczych ważnych problemów, z którymi mierzy się obecnie prawodawstwo oraz praktyka prawnicza. Jednym z takich problemów jest wpływ sztucznej inteligencji na procedury sądowe i odpowiedź na pytanie, czy jej zastosowanie zagraża prawu człowieka do sądu.

Międzynarodowy wymiar konferencji pozwolił na poszerzenie perspektywy badawczej i analitycznej. Referaty publikujemy w większości w angielskiej wersji językowej (w jakiej je wygłoszono).

Uzupełnieniem bieżącego numeru są dwa artykuły spoza konferencji, w tym (ostatni w kolejności) omawiający teorię salda w kontekście porównawczym.

prof. zw. dr hab. Roman Wieruszewski,  
rektor EWSPA

The second issue of our quarterly this year, as was the case last year, is a thematic issue. It features papers presented at an international academic conference entitled: ‘Artificial Intelligence in Legal Proceedings Worldwide – A Threat to Human Rights?’ An introduction by the issue’s academic editor, Prof. Tomasz Szanciło, provides an overview and discussion of these texts.

One of the aims of our university is to address, at academic conferences and in research, the significant issues currently facing legislation and legal practice. One such issue is the impact of artificial intelligence on court proceedings and the question of whether its use threatens the human right to a fair trial.

The international scope of the conference enabled us to broaden our research and analytical perspectives. We are publishing most of the papers in English (the language in which they were delivered).

This issue is supplemented by two articles not arising from the conference, including the final one, which discusses the theory of balance in a comparative context.

Prof. (titular) Roman Wieruszewski, DSc,  
Rector of EULA

## Spis treści / Table of contents

2/2026/78

Od redaktora naukowego numeru ( <b>Tomasz Szanciło</b> ) .....	5
From the scientific editor of this issue ( <b>Tomasz Szanciło</b> ) .....	7
<b>Mohammad Owais Farooqui</b> (ZEA), A Procedural Right to a Human Judge: Foundations and Framework for Safeguarding Fair Trials in the Age of AI [Prawo do bycia sądzonym przez człowieka: podstawy i ramy sprawiedliwego procesu sądowego w erze sztucznej inteligencji] .....	9
<b>Kamil Zaradkiewicz</b> , Robot Judge and the Right to a Fair Trial [Sędzia-robot a prawo do rzetelnego procesu sądowego] .....	38
<b>Inesa Stolper</b> , Judicial Impartiality and Artificial Intelligence from the Perspective of Article 6 of ECHR [Bezstronność sądu a sztuczna inteligencja z perspektywy art. 6 EKPC] .....	57
<b>Artur Kotowski</b> , The Legal Method in the Context of Futurology [Metoda prawnicza w kontekście futurologii] .....	79
<b>Krzysztof Grzesiowski</b> , The Algorithm of Justice: Objective Interpretation of Law as a Safeguard of Human Rights in Judicial Proceedings [Algorytm sprawiedliwości – obiektywna wykładnia prawa jako gwarancja praw człowieka w postępowaniu sądowym] .....	91
<b>Andreea Tabacu</b> , Who Is Right: Professional or AI? [Kto ma rację – profesjoniści czy sztuczna inteligencja?] .....	103
<b>Agnieszka Góra-Błaszczkowska, Ganesh V. Radhakrishnan</b> , Artificial Intelligence in Judicial Systems: Constitutional Frameworks and Democratic Safeguards in Poland and India. A Comparative Governance Analysis [Sztuczna inteligencja w systemach sądowych – ramy konstytucyjne i zabezpieczenia demokratyczne w Polsce i Indiach. Analiza porównawcza systemu zarządzania] .....	123

<b>Ganiyu Yahaya</b> , Artificial Intelligence and the Right to a Fair Hearing in Nigerian Judicial Proceedings: Navigating the Line Between Efficiency and Justice [Sztuczna inteligencja a prawo do rzetelnego rozpatrzenia sprawy w nigeryjskich postępowaniach sądowych – balansowanie między wydajnością a sprawiedliwością] .....	142
<b>Igor Zgoliński</b> , Wirtualna sala narad i inne nowe instrumenty postępowania karnego. Uwagi na tle zagrożeń i kosztów cyfryzacji [Virtual Courtrooms and Other New Tools in Criminal Proceedings: Notes on the Risks and Costs of Digitization] .....	172
<b>Mateusz Szenawa</b> , Zagrożenia gospodarcze związane z rozwojem sztucznej inteligencji – ze szczególnym uwzględnieniem technologii deepfake – na podstawie wybranych regulacji prawnych Unii Europejskiej [Economic Threats Associated with the Development of Artificial Intelligence, with Particular Reference to Deepfake Technology, Based on Selected European Union Legal Regulations] .....	185
<b>Dominik Bierecki</b> , Teoria salda w prawie niemieckim i gruzińskim a perspektywa polska [The Balance Theory in German and Georgian Law and the Polish Perspective] .....	209
<b>Informacje dla Autorów</b> .....	222
<b>Information for Authors</b> .....	225

## Od redaktora naukowego numeru

23 i 24 października 2025 roku odbyła się Międzynarodowa Konferencja Naukowa *Artificial Intelligence in Legal Proceedings Worldwide – A Threat to Human Rights?* (przeprowadzona w całości w języku angielskim), zorganizowana przez Europejską Wyższą Szkołę Prawa i Administracji w Warszawie we współpracy z POLITEHNICA București (Politechniką Bukareszteńską) – Narodowym Uniwersytetem Nauki i Technologii w Pitești (Rumunia).

Ten numer „Europejskiego Przeglądu Prawa i Stosunków Międzynarodowych” jest tematyczny (z jednym wyjątkiem, gdyż ostatni artykuł dołączono z uwagi na aktualność problematyki teorii salda – zwłaszcza wobec sporów na tle kredytów powiązanych z walutą obcą – w dyskusji prawniczej). Jego tematem jest zastosowanie sztucznej inteligencji w procedurach sądowych – nie tylko w Polsce, ale również za granicą – i stanowi zwieńczenie październikowej konferencji. Zawiera rozszerzone wersje części wygłoszonych referatów; mogą i powinny stać się one przyczynkiem do dalszych dyskusji.

Problematyka konferencji jest nader istotna, gdyż wykorzystanie sztucznej inteligencji w postępowaniu sądowym stanowi obecnie jedno z kluczowych zagadnień w debatach prawników. W dobie dynamicznego rozwoju technologii AI coraz częściej podnosi się bowiem kwestię możliwości wykorzystania tych rozwiązań w postępowaniu sądowym. Podstawowym celem jest rozwiązanie problemu przewlekłości postępowań sądowych – będącego bolączką nie tylko polskiego wymiaru sprawiedliwości. Kluczowe jest jednak, aby zachować zarazem konwencyjny i konstytucyjny standard prawa do sądu. W związku z tym wprowadzenie sztucznej inteligencji do wymiaru sprawiedliwości nie powinno być celem samym w sobie, ale winno prowadzić do przyspieszenia postępowań sądowych, z zachowaniem procesowych praw stron związanych z uprawnieniem do rzetelnego procesu.

Omówieniu problemów związanych z zastosowaniem sztucznej inteligencji w procedurach sądowych posłużyło niemal dwadzieścia referatów przedstawionych przez naukowców i praktyków z różnych krajów i części świata: Polski, Rumunii, Litwy, Indii, Zjednoczonych Emiratów Arabskich i Nigerii. Prelegenci podzielili się doświadczeniami, przedstawiając regulacje obowiązujące w wielu odmiennych systemach prawnych.

Każdego dnia konferencji odbywały się dwa panele. W pierwszym dniu panele te dotyczyły *Human Rights and Artificial Intelligence in Judicial Proceedings* oraz *The Use of Artificial Intelligence in Judicial Proceedings*. Jasne jest bowiem, że nie byłoby możliwe

rozpatrywanie zastosowania sztucznej inteligencji w jakichkolwiek procedurach sądowych bez uwzględnienia praw człowieka. W szczególności warto zwrócić uwagę na artykuł dotyczący robota jako sędziego i – w tym kontekście – prawa do rzetelnego procesu w ujęciu konstytucyjnym. Powstaje bowiem bardzo istotne i podstawowe pytanie: Czy robot może być sędzią – i jakie ewentualnie ograniczenia zachodzą w tej materii? O tym zaś, że sztuczna inteligencja nie zna granic, przekonują artykuły prezentujące rozwiązania z innych krajów – omówiono zwłaszcza te stosowane w procedurach sądowych.

Drugi dzień rozpoczął się panelem *Impact of Artificial Intelligence on Judicial Decisions*. Należy tutaj zadać sobie zwłaszcza pytanie: Kto ma rację – profesjonalista (sędzia) czy sztuczna inteligencja? I jak się ma do tego bezstronność i niezależność sędziowska. Są to kluczowe kwestie. Przede wszystkim warto się zaś się zastanowić: czy algorytm może być obiektywny, niezależny i bezstronny, czy też nabranie tych cech zależy od zakresu danych wprowadzonych przez człowieka?

Ostatni panel *Artificial Intelligence in Various Aspects of Law* dotyczył odniesienia sztucznej inteligencji do różnych kwestii prawnych związanych z procedurami sądowymi. Warto zwrócić uwagę, czy mówimy o futurologii, czy o rzeczywistości – i jak można wykorzystać w tym aspekcie metodę prawniczą. A powróciwszy do teraźniejszości, można się przekonać, jak wszystkie omawiane aspekty przekładają się na procedury np. w Nigerii, a więc państwie, którego rozwiązania prawne są nam praktycznie nieznanne, a jest to przecież jedno z największych państw Afryki.

Powyższe przekonuje, że tematy referatów związanych z zastosowaniem sztucznej inteligencji w postępowaniu sądowym były bardzo różnorodne. Można wysnuć ogólniejszy wniosek, że zastosowanie AI w sądownictwie – nie tylko w Polsce – napotyka (i zapewne jeszcze napotka) istotne bariery, szczególnie gdy chodzi o drugą instancję. O ile bowiem można sobie wyobrazić, że w pierwszej instancji wyrok wydaje sztuczna inteligencja, co zresztą dzieje się już w niektórych krajach, o tyle w drugiej instancji jej zastosowanie budzi jednak zdecydowanie więcej emocji, przede wszystkim społecznych. Wydaje się więc, że należy traktować zastosowanie AI w głównej mierze jako narzędzie pomocy sędziom przy wydawaniu sprawiedliwych i prawidłowych orzeczeń.

Konferencja stanowiła głos w ważnej dyskusji, która już się toczy i nadal będzie się toczyć, a jej efekt pozostaje niewiadomą.

dr hab. Tomasz Szanciło, profesor uczelni  
Europejska Wyższa Szkoła Prawa i Administracji w Warszawie  
sędzia Izby Cywilnej Sądu Najwyższego

## *From the scientific editor of this issue*

On October 23 and 24, 2025, the International Scientific Conference ‘Artificial Intelligence in Legal Proceedings Worldwide – A Threat to Human Rights?’ (conducted entirely in English) took place, organized by the European University of Law and Administration in Warsaw in cooperation with the POLITEHNICA București (Bucharest), National University of Science and Technology in Pitești (Romania).

This issue of the ‘European Review of Law and International Relations’ is a special, thematic edition (with one exception: the final article was included due to the topicality of the theory of balance—particularly in light of disputes involving foreign-currency-denominated loans—in legal discourse). Its theme is the application of artificial intelligence in court proceedings—not only in Poland but also abroad—and it serves as the conclusion to the October conference. It contains expanded versions of some of the papers presented; these can and should serve as a starting point for further discussion.

The subject of the conference is of great importance, as the use of artificial intelligence in court proceedings is currently one of the key issues in lawyers’ debates. In an era of rapid AI technological development, the question of whether these solutions can be used in court proceedings is being raised with increasing frequency. The primary goal is to address the problem of protracted court proceedings—a problem that plagues not only the Polish justice system. However, it is crucial to preserve both the conventional and constitutional standards of the right to a fair trial. Therefore, the introduction of artificial intelligence into the justice system should not be an end in itself, but should lead to the acceleration of court proceedings while preserving the procedural rights of the parties related to the right to a fair trial.

Nearly twenty papers presented by scholars and practitioners from various countries and regions of the world—including Poland, Romania, Lithuania, India, the United Arab Emirates, and Nigeria—served to discuss the issues related to the use of artificial intelligence in court proceedings. The speakers shared their experiences, presenting the regulations in force across many different legal systems.

Two panels were held each day of the conference. On the first day, these panels focused on ‘Human Rights and Artificial Intelligence in Judicial Proceedings’ and ‘The Use of Artificial Intelligence in Judicial Proceedings.’ It is clear, after all, that it would be impossible to consider the use of artificial intelligence in any judicial proceedings without

taking human rights into account. In particular, it is worth noting the article on robots as judges and—in this context—the constitutional right to a fair trial. This raises a very important and fundamental question: Can a robot be a judge—and what limitations, if any, apply in this regard? Articles presenting solutions from other countries—the ones used in court procedures were discussed in particular—demonstrate that artificial intelligence knows no bounds.

The second day began with a panel discussion titled ‘Impact of Artificial Intelligence on Judicial Decisions.’ Here, we must ask ourselves, in particular: Who is right—the professional (the judge) or artificial intelligence? and how does this relate to judicial impartiality and independence? These are key questions. Above all, it is worth considering: Can an algorithm be objective, independent, and impartial, or does acquiring these qualities depend on the scope of the data entered by a human?

The final panel, ‘Artificial Intelligence in Various Aspects of Law,’ focused on the application of artificial intelligence to various legal issues related to court proceedings. It is worth considering whether we are discussing futurology or reality—and how the legal method can be applied in this context. Returning to the present, one can see how all the discussed aspects translate into procedures, for example, in Nigeria—a country whose legal system is virtually unknown to us, even though it is one of the largest countries in Africa.

The above demonstrates that the topics of the presentations related to the application of artificial intelligence in court proceedings were very diverse. A more general conclusion can be drawn: the application of AI in the judiciary—not only in Poland—faces (and will likely continue to face) significant barriers, especially when it comes to the appellate level. While it is conceivable that a judgment could be issued by artificial intelligence in the first instance—which, incidentally, is already happening in some countries—its use in the appellate courts, however, provokes far more controversy, primarily of a social nature. It therefore seems that the use of AI should be viewed primarily as a tool to assist judges in rendering fair and correct rulings.

The conference served as a voice in an important discussion that is already underway and will continue, though its outcome remains uncertain.

*Tomasz Szanciło, associate professor, PhD in Law, post-doctoral degree  
European University of Law and Administration in Warsaw  
Justice of the Supreme Court of the Republic of Poland, Civil Chamber*



Mohammad Owais Farooqui\*

## A Procedural Right to a Human Judge: Foundations and Framework for Safeguarding Fair Trials in the Age of AI

[Prawo do bycia sądzonym przez człowieka – podstawy i ramy sprawiedliwego procesu sądowego]

### Abstract

The integration of artificial intelligence into judicial systems is accelerating worldwide, raising fundamental concerns about the right to a fair trial. This article argues that a procedural right to a human judge constitutes a necessary safeguard within existing international fair trial frameworks. It contends that Article 6 of the European Convention on Human Rights (ECHR) and Article 14 of the International Covenant on Civil and Political Rights (ICCPR) implicitly require human adjudication and that this requirement should be made explicit. A review of scholarship and comparative legal practice indicates that human judges provide attributes of reasoned judgment, empathy, and accountability that current AI systems cannot replicate. A comparative analysis of European, United States, and Chinese approaches reveals a convergent, if variously grounded, commitment to human oversight of judicial AI. Drawing on these foundations, the article proposes a six-element legal framework governing the conditions under which AI tools may be used in judicial proceedings without infringing fair trial rights. The framework centres on human final authority, transparency, contestability, legal accountability, non-discrimination, and proportionality. The article further examines pathways for implementing this right through treaty interpretation, constitutional provisions, and model legislation.<sup>1</sup>

**Keywords:** procedural rights, human judge, fair trial, Artificial Intelligence (AI), judicial oversight, comparative analysis.

\* **Mohammad Owais Farooqui** – associate professor, PhD in Law, Department of Public Law, College of Law, University of Sharjah (affiliation), UAE; <https://orcid.org/0000-0003-0154-802X>; mfarooqui@sharjah.ac.ae / profesor uczelni, dr nauk prawnych, Katedra Prawa Publicznego, Wydział Prawa, Uniwersytet w Szardży (afiliacja), ZEA.

<sup>1</sup> M. Górski, *Why a Human Court? On the Right to a Human Judge in the Context of the Fair Trial Principle*, 'European Criminal Law Associations' Forum' 2023, 1, pp. 83–88; <https://eucrim.eu/articles/why-a-human-court-on-the-right-to-a-human-judge-in-the-context-of-the-fair-trial-principle/> [accessed: 28.03.2025]; M. Zou, E. Lefley, *Generative Artificial Intelligence and Article 6 of the European Convention on Human Rights: The Right to a Human Judge?* [in:] M. Zou, M. Ebers, C. Poncibò and R. Calo (eds.), *The Cambridge Handbook of Generative AI and the Law*, Cambridge University Press 2025, pp. 1–20.

---

## Introduction

---

The settlement of disputes has traditionally been entrusted to human judges. From the earliest courts, the figure of a human judge has served as the embodiment of justice, impartiality, and the rule of law. Rapid advances in artificial intelligence now challenge this paradigm. AI algorithms are being deployed for legal research, outcome prediction, and, increasingly, to assist or guide judicial decision-making in courts around the world. This technological shift raises a fundamental question: Can—or should—machines assume the role of human judges in the administration of justice? This question implicates core tenets of the right to a fair trial and has generated a growing body of scholarship asking whether there is, or ought to be, a right to a human judge.

These developments raise fundamental questions about whether the guarantees of a fair trial—forged in an era when human adjudication was assumed—remain adequate in an age when consequential legal decisions may be shaped or even rendered by algorithmic systems.<sup>2</sup>

This paper proposes that a procedural right to a human judge be recognized as a necessary component of the right to a fair trial in the AI era. By procedural right to human judge, I am referring to a justiciability right that final judicial determinations impacting upon rights or duties of a person will be made or superintended by human judicial officers, as opposed to being totally devolved to AI systems. This is based on normative and practical issues. Human adjudication is normatively associated with human dignity and the rule of law: to be judged by his or her peers (or human authorities) is a part of treating the parties with respect and empathy.<sup>3</sup> In reality, thanks to the existing AI technology, there is no moral reasoning, empathy, and contextual judgment that human judges can offer in the courtroom.<sup>4</sup> It is possible that an algorithmic judge, however efficient he is, can deliver coldly formalistic justice unresponsive to individual situations, in other words, a trial that is not really fair.

The stakes are high. Providing that society is ready to accept the fully automated judicial decision this will be the seismic shift in the way justice is conceived and served. Proponents of AI in the judiciary argue that algorithms can increase efficiency, consistency, and perhaps even accuracy in adjudication.<sup>5</sup> Indeed, AI tools promise to alleviate case backlogs and assist judges in han-

---

<sup>2</sup> T. Vasdani, *Robot Justice: China's Use of Internet Courts*, Toronto 2020, Feb.; <https://www.lexisnexis.ca/en-ca/ihc/2020-02/robot-justice-chinas-use-of-internet-courts.page> [accessed: 28.03.2025].

<sup>3</sup> M. Zou, E. Lefley, *Generative...*, *ibid.*

<sup>4</sup> M. Górski, *Why a Human...*, *ibid.*; <https://eucrim.eu/articles/why-a-human-court-on-the-right-to-a-human-judge-in-the-context-of-the-fair-trial-principle/> [accessed: 28.03.2025].

<sup>5</sup> I. Banks, *Operationalising Human Oversight of AI-Supported Judicial Decision-Making: A Systemic Perspective*, 'Algorithms, Data and Society' 2023; <https://algosoc.org/research/the-rise-of-technocourts-and-governing-the-appification-of-justice> [accessed: 28.03.2025].

dling complex data.<sup>6</sup> But these gains must be weighed against fundamental rights and the public's confidence in the justice system. A court's legitimacy rests not only on correct outcomes but also on fair processes and public trust. I must ask whether a litigant can trust—and accept—a judgment if it is rendered by a machine. As this paper will explore, studies in procedural justice suggest that people value having their voice heard by an impartial human decision-maker, and that the perceived legitimacy of decisions may suffer if rendered by an impersonal algorithm.<sup>7</sup>

Today's debate on AI in the judiciary has only recently moved from speculative to serious. While early scholarship has begun to ask whether we have a right to a human judge, this paper breaks new ground by offering a concrete framework and comparative perspective that previous works have not. For example, Marcin Górski's 2023 essay argues that under modern fair trial standards "courts cannot operate without a human element"<sup>8</sup> and concludes that using AI in place of judges "does not seem to be reconcilable" with the fairness required by a fair trial. Similarly, Zou and Lefley (2025) focus on Article 6 ECHR and, through a human dignity lens, contend that the right to a human judge is implicit—emphasizing the intrinsic value of human judgment, empathy, and being heard by a human decision-maker.<sup>9</sup> Nonetheless, both of these sources do not give a step-by-step plan on how to execute such a right within various legal frameworks. The paper expands on their observations; however, it goes a step ahead: it provides the first practical system to instantiate a procedural right to human judge and discusses the way this right can be achieved in practice all over the world. By so doing, it addresses an important gap in the literature by bridging the gap between principles and the practice and across a variety of jurisdictions as compared to the Euro-centric focus of previous research.

Within the passages which follow, I establish a procedural right to a human judge. In the following Part, the contemporary situation: the emerging body of research on AI and fair trial rights, and the practice of various jurisdictions, will be reviewed. The following Part offers a comparative study of the strategies in Europe, the United States, and China, which elucidates the differences and common trends in the human regulation of AI in the judicial system. Then I presented our new offering: a systematic legal approach that defines the way in which AI can be incorporated into the judicial process without undermining the human essence of the process. This includes a proposed test

---

<sup>6</sup> *Ibid.*

<sup>7</sup> A. Završnik, M. M. Plesničar, *The Right to a Human Judge in Criminal Justice Settings*, paper presented at the American Society of Criminology 79th Annual Meeting, San Francisco, 15.11.2024; [https://convention2.allacademic.com/one/asc/asc24/index.php?program\\_focus=view\\_paper&selected\\_paper\\_id=2163748&cmd=online\\_program\\_direct\\_link&sub\\_action=online\\_program](https://convention2.allacademic.com/one/asc/asc24/index.php?program_focus=view_paper&selected_paper_id=2163748&cmd=online_program_direct_link&sub_action=online_program) [accessed: 28.03.2025].

<sup>8</sup> M. Górski, *Why a Human...*, *ibid.*

<sup>9</sup> M. Zou, E. Lefley, *Generative...*, *ibid.*

or set of principles to determine the permissible uses of AI in light of fair trial guarantees. Finally, I discuss how this proposed right can be implemented and address potential counterarguments. I conclude that recognizing a right to a human judge is both a principled and a practical step to safeguard fair trials amid rapidly evolving technology.

---

## **Literature Review**

---

**The Question of a Human Judge:** The notion of an explicit “right to a human judge” has only recently emerged from the realm of science fiction into legal scholarship. Traditional fair trial doctrine assumed judges are human without saying so. With AI now entering courtrooms, jurists and scholars have begun to articulate why the human element is indispensable to justice. This section surveys the burgeoning debate on AI adjudication, distills the key concerns about replacing human judges, and shows how a consensus is forming around preserving a human judge’s role—setting the stage for our novel framework.

**Human Judges vs. AI Adjudicators—Key Concerns:** A central theme in the literature is the comparative strengths of human versus algorithmic decision-makers. Many scholars suggest that a fully autonomous AI judge would inherently breach core fair trial criteria. An AI is not “independent” in the sense a human judge is it remains a tool designed and controlled by those who program or train it. In Górski’s analysis, an algorithmic court would always lack the *personal conscience and autonomy* that human judges possess, meaning such a system could never truly satisfy the requirement of an independent and impartial tribunal.<sup>10</sup> While it is theoretically possible to engineer safeguards for an AI’s impartiality, ensuring the strict level of judicial independence demanded by instruments like Article 6 ECHR is extraordinarily difficult in practice.<sup>11</sup> Górski (2023) posits that we should not “change the anthropocentric mindset” of adjudication and that the fair trial principle itself implies the necessity of human judges.<sup>12</sup> Indeed, he argues that under present conditions replacing judges with AI “does not seem to be reconcilable” with the fairness component of the right to a fair trial. In short, current AI technology cannot replicate the moral agency, intuition, and accountability of a human judge.

**Efficiency vs. Fairness—The Ongoing Debate:** Proponents of “robot judges” often emphasize gains in speed and consistency. AI could reduce back-

---

<sup>10</sup> M. Górski, *Why a Human...*, *ibid.*

<sup>11</sup> Convention for the Protection of Human Rights and Fundamental Freedoms (European Convention on Human Rights, 1950), Article 6(1).

<sup>12</sup> *Ibid.*

logs and eliminate some randomness in human decisions. Tech optimists like Richard Susskind envision online courts and AI-driven processes that resolve disputes faster and at lower cost.<sup>13</sup> In China, officials tout AI's ability to process millions of routine cases, relieving overburdened courts. These efficiency arguments carry weight—justice delayed is justice denied, after all—but critics counter that efficiency means little if the process is perceived as unfair or inhumane. As legal ethicist Tania Sourdin cautions, unchecked automation may “erod[e] the quality of justice” in pursuit of speed.<sup>14</sup> The danger lies in the fact that the litigants are turned into cogs in a machine instead of being subject to a rational debate. Empirical evidence on procedural justice supports this concern: multiple studies have shown that individuals are highly interested in the manner in which decisions are made—that they are heard, are treated with respect and are judged by a person who listens and understands. Even an AI judge that is correct on a score would not score high on such metrics of procedural justice since it cannot actually listen and empathize. Završnik and Plesničar (2024) cautioned in a recent paper at a conference that automating judicial rulings poses the risk of undermining its empirical legitimacy through loss of transparency, consistency, and effective communication.<sup>15</sup> People cannot believe that a complicated algorithm was able to listen to their own story. In particular in criminal justice, it has been observed that the absence of human-like characteristics in an AI such as compassion and moral judgment may compromise the rehabilitative goals and societal acceptability of sentencing. Overall, although AI may have the potential to enhance efficiency, it may affect the perceived fairness and legitimacy, which may be viewed by many people as a trade-off that is simply unacceptable.

**Emerging Calls for a “Right to a Human Judge:”** In light of these concerns, calls to safeguard human adjudication have grown explicit. The Council of Bars and Law Societies of Europe (CCBE), representing European lawyers, took a strong stance in 2021: it declared that “a judge should not be allowed to delegate all or part of his/her decision-making power to an AI tool,” urging that laws ensure “the entire decision-making process remains a human-driven activity” and that “a right to a human judge should be guaranteed at all stages of the proceedings.”<sup>16</sup> In a 2023 statement, the CCBE reaffirmed that effective human oversight of AI is a “precondition” for a justice system under

<sup>13</sup> R. Susskind, *Online Courts and the Future of Justice*, Oxford University Press: Oxford 2019.

<sup>14</sup> T. Sourdin, *Judge v. Robot? Artificial Intelligence and Judicial Decision-Making*, 'University of New South Wales Law Journal' 2018, 41(4), pp. 1114–1134; See also: T. Sourdin, *Judges, Technology and Artificial Intelligence: The Artificial Judge*, Edward Elgar Publishing: Cheltenham 2021.

<sup>15</sup> A. Završnik, M. M. Plesničar, *The Right...*, *ibid*.

<sup>16</sup> Council of Bars and Law Societies of Europe (CCBE), CCBE Position Paper on the Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act), 8.10.2021; available at: [https://www.ccbe.eu/fileadmin/speciality\\_distribution/public/documents/IT\\_LAW/ITL\\_Position\\_papers/EN\\_ITL\\_20211008\\_CCBE-position-paper-on-the-AIA.pdf](https://www.ccbe.eu/fileadmin/speciality_distribution/public/documents/IT_LAW/ITL_Position_papers/EN_ITL_20211008_CCBE-position-paper-on-the-AIA.pdf) [accessed: 28.03.2025].

the rule of law, stressing that human judges must retain full responsibility for decisions.<sup>17</sup> In the CCBE’s view, the right to a fair trial inherently includes the right to a human judge. Likewise, at the global level, the United Nations Special Rapporteur on the independence of judges and lawyers has recently affirmed this principle. In her 2023 report to the UN General Assembly, Special Rapporteur Margaret Satterthwaite stressed that the right to an independent and impartial tribunal “requires access to a human judge”<sup>18</sup>—effectively recognizing that a human decision-maker is an indispensable component of a fair trial. This high-level endorsement underscores that the international community is beginning to view human judges as essential in the age of AI.

These calls for a human judge are echoed by academics. For example, Górski’s analysis in *EUCRIM (2023)* contends to the contrary, that the fair trial principle itself, properly understood, requires human adjudication. He said that “humankind at the centre of justice”—the anthropocentric orientation of law—needed to be preserved and that even the most advanced AI could not completely replace the human judge’s role. Zou and Lefley (2025) go even further in developing a legal doctrine of a human-judge requirement in the ECHR framework, and base it upon human dignity. They suggest that respect for the human dignity in the court room implies respect for the intrinsic value of human judgement and human empathy and the importance of being heard by a human decision-maker—leading them to conclude that Article 6 ECHR read in this light “includes the right to a human judge.”<sup>19</sup> Interdisciplinary discourse is also taking place—legal theorists, computer scientists and ethicists are working together to consider how to design respecting AI systems. A recurring conclusion is that keeping a “human in the loop” is crucial. However, scholars like Binns (2022) caution that having a human nominally involved is not enough—the oversight must be meaningful, with the human judge truly able to override or question the AI’s suggestions.<sup>20</sup> In other words, the human role cannot be a mere rubber stamp; it must be substantive and empowered.

**Towards a Novel Framework:** Thus, it can be seen that the literature shows an increase in the agreement according to which AI can be an effective instrument in a legal trial, but it must not replace the human judge. The discipline has transitioned to explicit disclosure: the concept of a right to a human judge is gaining prominence as one of the normative reactions to the two predica-

<sup>17</sup> Council of Bars and Law Societies of Europe (CCBE), *CCBE Statement on the Use of AI in the Justice System and Law Enforcement*, 25.05.2023; available at: [https://www.ccbe.eu/fileadmin/speciality\\_distribution/public\\_documents/Statements/2023/EN\\_ITL\\_20230525\\_CCBE-Statement-on-the-use-of-AI-in-the-justice-system-and-law-enforcement.pdf](https://www.ccbe.eu/fileadmin/speciality_distribution/public_documents/Statements/2023/EN_ITL_20230525_CCBE-Statement-on-the-use-of-AI-in-the-justice-system-and-law-enforcement.pdf) [accessed: 28.03.2025].

<sup>18</sup> United Nations General Assembly, *Artificial Intelligence in Judicial Systems: Promises and Pitfalls*, Report of the Special Rapporteur on the Independence of Judges and Lawyers, 2023, A/80/169.

<sup>19</sup> M. Zou, E. Lefley, *Generative...*, *ibid.*

<sup>20</sup> R. Binns, *Human Judgment in Algorithmic Loops: Individual Justice and Automated Decision-Making*, ‘Regulation & Governance’ 2022, 16(1), pp. 197–211, DOI: 10.1111/rego.12358 [accessed: 29.03.2025].

ments concerning the maintenance of the rights to a fair trial and the responsible use of the advantages of AI. What the current discourse does not contain, however, is a practical roadmap on how to entrench and bring this right into practice. The case for human judges has been made theoretically in earlier scholarship; this paper moves from principle to practice. The following section of this paper discusses how this conflict between AI and human supervision is already being played in various jurisdictions thus forming the basis of our suggested framework. And we suggest a systematic legal policy—a test or series of tests—to see that AI penetration of the judiciary does not in any way compromise the human essence of adjudication. That way, we will help to transform the newly formed consensus into practical rules to close the divide between the reasons that human judges are important and the methods to assurance in the age of AI.

---

## **Comparative Analysis: Jurisdictions Approaches to AI and Human Oversight**

---

When considering how a right to a human judge can be justified and put into practice, it is educative to look at how various legal systems are struggling with AI in the courtroom. Part is a comparative analysis of the European Union (the human rights framework of the Council of Europe), United States and China. They all have their own variant: Europe is more inclined towards overt regulation and human-right orientation; the United States tries to resolve the problem on the basis of constitutional principles (due process and case-by-case jurisprudence); China is moving forward with the use of AI but under the slogan of efficiency and centralization, though with an insistence that the human element should be present somewhere. Despite their differences, all three jurisdictions reflect an implicit understanding that human oversight is crucial in judicial decision-making.

---

### **Europe: The ECHR and the EU’s AI Act— Ensuring Human Judges Remain Central**

---

**Human Rights Framework (ECHR):** European law provides strong fair trial guarantees via Article 6 ECHR, which applies to all Council of Europe member states. While Article 6 does not explicitly state “the judge must be human,” it speaks of “an independent and impartial tribunal” and the right to a fair and public hearing. European courts have not yet confronted an AI-judge scenario, but the spirit of Article 6 suggests a tribunal must operate in

a manner compatible with human justice. For instance, the ECtHR’s jurisprudence that a court must genuinely hear and consider the parties’ arguments (*the duty of careful examination*) would be difficult to satisfy with a non-human adjudicator.<sup>21</sup> It is notable that the Convention was drafted in 1950, when the prospect of non-human judges was science fiction; thus, the requirement was assumed rather than spelled out.

If a member state were to introduce AI-rendered judicial decisions, it would likely face challenges in Strasbourg (the ECtHR) on multiple grounds. The “tribunal established by law” requirement might be invoked, questioning whether an AI system can qualify as a lawful tribunal. More fundamentally, the fairness of proceedings (Article 6[1]) and rights of defense (Article 6[3]) could be at risk. It is foreseeable that the ECtHR would approach such a case by examining whether the use of AI denied the applicant any of the participatory rights or procedural safeguards that a human judge would afford. Given the Court’s emphasis on concepts like adversarial proceedings and equality of arms, a purely algorithmic judge—who cannot explain its reasoning in accessible human language or be subject to cross-examination—would raise red flags. Indeed, a recent scholarly analysis by Zou and Lefley posits that Article 6, interpreted in line with human dignity, encompasses a right to a human judge.<sup>22</sup> They argue that the ECHR’s guarantee of a fair trial implicitly values the human qualities of judgment and the interpersonal respect between judge and litigant, which would be lost if an AI assumed the role of judge.<sup>23</sup>

At the level of principle, then, Europe’s human rights framework is conducive to recognizing a right to a human judge. This is reinforced by soft-law instruments. The Council of Europe’s Commission for the Efficiency of Justice (CEPEJ) adopted the European Ethical Charter on the Use of AI in Judicial Systems (2018), which sets out key principles for AI in courts. Among these is the principle of “user control,” meaning that AI tools should not supplant the ultimate control of human judicial authorities.<sup>24</sup> The Charter insists that AI be used to augment, not replace, human decision-making in judicial processes, aligning with a precautionary approach that preserves human agency.

**EU Artificial Intelligence Act:** The European Union is currently implementing a set of laws on AI. The current Artificial Intelligence Act (AIA) specifically seeks to target the application of AI in justice. The draft, which is currently under contract, of the AIA (which is likely to be Regulation (EU) 2024/1689) catego-

<sup>21</sup> Ibid.

<sup>22</sup> M. Zou, E. Lefley, *Generative...*, *ibid.*

<sup>23</sup> Ibid.

<sup>24</sup> Adoption of a European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment at the 31st Plenary Meeting of the CEPEJ in Strasbourg on 3 and 4 December 2018 (*International Union of Judicial Officers [UIHJ]*, 2018, Dec., 4); <https://www.uhj.com/2018/12/04/adoption-of-a-european-ethical-charter-on-the-use-of-artificial-intelligence-in-judicial-systems-and-their-environment-at-the-31st-plenary-meeting-of-the-cepej-in-strasbourg-on-3-and-4-december-2018/> [accessed: 28.03.2025].

rizes some AI systems in the operation of the justice system as being of high-risk and requires them to be put under stringent requirements.<sup>25</sup> Remarkably, the AIA Recital 61 emphasizes the idea that, although AI can assist in the process of judicial decision-making, the ultimate decision-making process should not be automated: it should be a human initiative.<sup>26</sup> This language, now formalized in EU law, effectively enshrines the rule that a human judge must make the final call. The Act thus embodies a *de facto* right to a human decision in the judicial context, by prohibiting fully automated adjudication in courts.

Under Annex III of the AI Act, “AI systems intended to be used by judicial authorities in applying the law to facts” are designated high-risk, meaning their deployment is legal only if they meet rigorous transparency, oversight, and quality criteria.<sup>27</sup> Among those criteria is likely the requirement of human oversight (consistent with Article 14 of the Act’s final version) and the availability of explanations for the AI’s suggestions. Moreover, the Act’s scope (Recital 61) clarifies that purely ancillary administrative AI tools (e.g., software to anonymize documents) are not the target—the concern is specifically AI that would influence judicial outcomes.<sup>28</sup> By drawing this line, the EU is taking a clear policy stance: AI may assist judges but may not become the judge.

**European Institutional Positions:** Key European legal institutions and professional bodies have reinforced this stance. The CCBE’s 2021 paper, as noted, urged that the AIA explicitly guarantee a right to a human judge.<sup>29</sup> In fact, the final text of the AI Act echoes much of what the CCBE advocated. For example, the CCBE wanted Annex III (which lists high-risk uses) and a Recital to clarify that AI must not effectively take over judicial decisions;<sup>30</sup> the inclusion of Recital 61 mentioned above achieves that clarification in substance. The European Court of Justice (CJEU), while not yet addressing AI in judging, has also indirectly supported human decision-making. In other domains, the CJEU has interpreted the right to an effective remedy (Article 47 of the EU Charter of Fundamental Rights, akin to Article 6 ECHR) as requiring that individuals have access to reasons and to a review by a competent tribunal. If an AI were to render a judgment without human endorsement, it would sit uneasily with those requirements. Indeed, one can analogize to the CJEU’s stance on automated administrative decisions: in December 2023, the CJEU held in *Schufa Holding AG* (C-634/21) that significant decisions made by automated profiling systems must be subject to human review and accompanied

<sup>25</sup> Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) [2024] OJ L1689.

<sup>26</sup> *Ibid.*

<sup>27</sup> *Ibid.*

<sup>28</sup> *Ibid.*

<sup>29</sup> Council of Bars and Law Societies of Europe (CCBE), *CCBE Position Paper...*, *ibid.*

<sup>30</sup> *Ibid.*

by meaningful explanation, grounding this requirement in Article 22 GDPR and the principle of effective judicial protection.<sup>31</sup> By extension, a fully automated court decision would likely be viewed as incompatible with EU fundamental rights law.

Thus, the Europe's approach combines a rights-based perspective (ECHR Article 6 as interpreted by doctrine) with regulatory action (the AI Act) to ensure that human judges remain at the helm. This creates a fertile environment for formally recognizing a procedural right to a human judge. If tested, it is plausible that European courts would read such a right into existing guarantees. And with legislation, the EU is preemptively codifying that AI cannot usurp the judge's role. The European experience thus strongly supports the notion that fair trial norms require human adjudication.

---

### United States: Due Process and the Imperative of Human Judgment

---

The United States does not have an explicit “right to a human judge” in its Constitution; such a concept was implicitly assumed when the Due Process Clause and other judicial safeguards were framed. However, American law, through constitutional principles and case law, provides a framework to argue for the necessity of human judicial oversight.

**Constitutional Due Process:** The Fifth and Fourteenth Amendments guarantee that no person shall be deprived of life, liberty, or property without due process of law. At least, due process allows people to be noticed and given a chance to be heard by any impartial decision-maker. Conventionally, this impartial adjudicator is a human judge (or jury). In case a governmental agency tried to use an AI to adjudicate the rights of an individual, e.g., the decision of a prison term or a civil lawsuit exclusively on the basis of an algorithm, some serious due process concerns would be raised. Among them is the fact that you are able to defend and counter the evidence that goes against you. Provided the evidence or the judgment of a case is made by an algorithm, a defendant may raise the objection that he has a right to investigate how the algorithm arrived at his or her decision. The functioning of the algorithm in a way, turns into a witness of its own but one which cannot be cross-examined or perceived easily.

---

<sup>31</sup> Case C-634/21 SCHUFA Holding AG v. Datenschutzbeauftragter des Landes Hessen (Schufa Holding) EU:C:2023:957, judgment of 7.12.2023; see also: S. J. Choi, K. Van Quathem and B. Van Vooren, CJEU Receives Questions on the AI Act Relating to Automated Decision Making, 'Inside Global Tech' 2025, Feb., 24; <https://www.insideglobaltech.com/2025/02/24/cjeu-receives-questions-on-the-ai-act-relating-to-automated-decision-making/> [accessed: 28.03.2025].

These concerns came to the fore in the landmark case *State v. Loomis* (2016) in Wisconsin. In *Loomis*, a criminal defendant challenged the use of COMPAS, a proprietary risk assessment algorithm, at his sentencing. He argued that because the algorithm's methodology was secret, its use violated his due process rights—he could not challenge the accuracy or relevance of the risk score it produced.<sup>32 33</sup> The Wisconsin Supreme Court ultimately held that using the algorithm did not violate due process, but it imposed important limitations. The court reasoned that as long as the risk score was not the sole basis for the sentence and the defendant had an opportunity to respond to it (albeit without knowing the inner workings), due process was not offended.<sup>34 35</sup> However, the court cautioned that COMPAS had to be used carefully—for example, authorities must acknowledge the tool's limitations (including potential racial bias in data) and never treat the algorithm's output as determinative.<sup>36</sup> In effect, even as the *Loomis* court allowed AI-assisted decision-making, it underscored the primacy of the human judge's independent judgment and the need for transparency.

The *Loomis* case illustrates the U.S. judiciary's instinct to ensure a human check on algorithmic tools. The judge in *Loomis* did not cede her decision to the machine; she considered the COMPAS score among many factors. Importantly, the decision suggests a due process baseline: if an algorithm is used, the defendant must have notice of its use and a chance (even a limited one) to contest or contextualize its output.<sup>37</sup> Since *Loomis*, similar issues have arisen across the country. Risk assessment algorithms are now common in bail decisions, parole, and even child protection decisions. Scholars like Danielle Citron have warned of “Technological Due Process” problems, where automated systems in government decision-making create opacity and errors that are hard for individuals to challenge. A consistent recommendation is to build in human review and explanation for any automated recommendation that significantly impacts individual rights.

<sup>32</sup> *State v. Loomis* 881 NW 2d 749 (Wis 2016).

<sup>33</sup> K. Murray, What the Legal Drama For the People Teaches Us About AI and Legal Ethics, O'Hagan Meyer 2025, Feb, 6; <https://ohaganmeyer.com/2025/02/06/what-the-legal-drama-for-the-people-teaches-us-about-ai-and-legal-ethics/> [accessed: 29.03.2025].

<sup>34</sup> Ch. Engel, L. Linhardt and M. Schubert, Code Is Law: How COMPAS Affects the Way the Judiciary Handles the Risk of Recidivism, 'Artificial Intelligence and Law' 2024; <https://doi.org/10.1007/s10506-024-09389-8> [accessed: 29.03.2025].

<sup>35</sup> C. McKay, Predicting Risk in Criminal Procedure: Actuarial Tools, Algorithms, AI and Judicial Decision-Making, 'Sydney Law School Research Paper' 2019, 19(67); <https://ssrn.com/abstract=3498197> [accessed: 29.03.2025].

<sup>36</sup> I. de Miguel Beriain, Does the Use of Risk Assessments in Sentences Respect the Right to Due Process? A Critical Analysis of the Wisconsin v. Loomis Ruling, 'Law, Probability and Risk' 2018, 17, 1, pp. 45–53; <https://doi.org/10.1093/lpr/mgy001> [accessed: 29.03.2025].

<sup>37</sup> J. M. Payé Salazar, Impact of Artificial Intelligence on Effective Judicial Protection, *AfricLaw* 2023, Jun., 26; <https://africlaw.com/2023/06/26/impact-of-artificial-intelligence-on-effective-judicial-protection/> [accessed: 29.03.2025].

**Jury Trial and Public Confidence:** The U.S. also has the unique institution of jury trials (Sixth Amendment for criminal cases, Seventh Amendment for certain civil cases). While the right to a jury is distinct from the right to a judge, it underscores the deeply human character of American adjudication. No conceivable AI replacement for a jury has been proposed in serious terms, and any such idea would run headlong into constitutional amendments that explicitly require a jury of one's peers. This is relevant because it suggests a constitutional ethos that values lay human judgment as a component of justice (a "jury of peers"), further buttressing the argument that at least one human (whether judge or jury) must be the decision-maker in criminal guilt or major civil liability. An AI cannot be a "peer" of a human defendant in any meaningful sense.

**Case Law on Mechanical Decision-Making:** American jurisprudence has occasionally dealt with scenarios of overly mechanical justice (though not AI *per se*). For example, in the context of sentencing guidelines (which are formulaic but still applied by judges), the U.S. Supreme Court held in *United States v. Booker* (2005)<sup>38</sup> that mandatory sentencing rules violated the Sixth Amendment because they took away too much discretion from the judge and jury, effectively letting a formula dictate punishment. While not about AI, the underlying sentiment is that an individual has the right to a sentence determined through a deliberative process by human factfinders, not a rigid algorithm. If mandatory guidelines raise such concerns, *a fortiori* a literal algorithm imposing a sentence would be problematic.

Moreover, due process includes the concept of a neutral and detached adjudicator. If an AI system were developed by, say, the prosecution or police, one might question its neutrality. Human judges are oath-bound to be impartial and are subject to recusal if they have a conflict of interest. An algorithm's biases or origins might not be so easily discerned or vetted. Thus, we can see how due process and impartial tribunal principles in the U.S. constitutional context strongly imply a need for human involvement, or at least for any algorithm to be subordinate to a human judge's discernment.

**Legislative and Policy Responses:** While there is not yet federal legislation outright banning "AI judges," there are policy moves acknowledging the issue. At the federal level, the Administrative Procedure Act (APA) ensures that individuals can challenge agency decisions and obtain judicial review—if agencies start using AI to make initial determinations (for benefits, immigration, etc.), the right to a human judicial review is already built into law. Even certain states have initiated the legislation of algorithmic accountability. An example is Illinois, as well as other states, enacted legislation regarding algorithmic disclosure where an algorithm is used, such as hiring or insurance:

---

<sup>38</sup> 543 US 220 (2005).

in judicial practice, some jurisdictions have mandated the disclosure of risk assessment scores to a defendant or even barred the use of particular factors (such as race or gender) in an algorithm. Although they are piecemeal, they show an appreciation of the fact that human beings need to put a check on the algorithmic decisions.

Importantly, the public discourse in the U.S. often centers on the idea of “algorithmic bias” and “transparency.” Civil rights groups such as the ACLU and Electronic Frontier Foundation have argued that automated decision systems can perpetuate discrimination and should not be unchecked in criminal justice. Though typically framed as equal protection or due process issues, the underlying remedy is usually to demand human oversight or the ability to challenge the automated result. For example, if a predictive policing algorithm suggests a particular sentence, defense attorneys argue the judge must independently evaluate that suggestion and the defendant must be able to contest it—essentially asserting a right to a human judge’s independent assessment.

While the United States lacks a single instrument spelling out “the right to a human judge,” its constitutional framework (due process, the right to a jury, etc.) and burgeoning case law on algorithms in sentencing collectively support that notion. The American approach might be described as incremental and reactive—ensuring on a case-by-case basis that no one is condemned by a machine without a human safeguard. If ever a jurisdiction in the U.S. attempted to fully automate a court’s decision (for example, an “AI judge” for small claims), it would almost certainly be challenged, and the challenger would invoke due process and perhaps the right of access to the courts (which implies access to a judge). Therefore, the U.S. supports our thesis in a more implicit way: it shows the essentiality of human judges through the lens of constitutional rights, even if it hasn’t been labeled as a distinct right yet.

---

### **China: Smart Courts, Efficiency, and the Role of the Human Judge**

---

China presents a compelling case study as a jurisdiction aggressively adopting AI in its judiciary, but within a legal-political system very different from Western democracies. The adoption of technology in the Chinese judiciary is also in line with a larger project of the so called Smart Courts that would see to the modernization of the court system, its ability to handle heavy caseload, and its ability to control the judges. The fact of the right to a human judge in China has to be interpreted in the context of an understanding of the priorities of China: efficiency, consistency, and state control instead of the rights of persons. However, even in China, there is a clear position that AI will sup-

port, but will not exclude, human judges—a position motivated probably by practical need, and by the issue of legitimacy.

**The Smart Court Revolution:** In 2017, China’s State Council published the Artificial Intelligence Development Plan (AIDP), a national strategy that charts China’s AI aspirations up to 2030.<sup>39</sup> Modernizing the judicial system was highlighted as a key goal: “Achieve the intelligentization of courts and trial systems and trial capacity.”<sup>40</sup> To that end, China’s Supreme People’s Court (SPC) has introduced technology into nearly every aspect of litigation. This includes online case filing, automated document generation, voice-to-text transcription in hearings, and decision-support systems that suggest relevant laws and past cases to judges. China has also established specialized Internet Courts (in Hangzhou, Beijing, and Guangzhou) to handle internet-related disputes almost entirely online. These courts are famous for using virtual courtrooms and even AI judge avatars. In a much-publicized example, the Hangzhou Internet Court introduced a virtual “AI judge”—essentially a screen with a realistic judge avatar that conducts certain procedural parts of a case.<sup>41</sup> Litigants communicate with this AI judge via video interface, and it can ask routine questions and generate a proposed judgment. However, at least in theory, a human judge reviews and approves these judgments before they are finalized.<sup>42</sup>

The efficiency gains in China have been significant. By one report, millions of cases were handled online within the first two years of the Internet Courts’ operation<sup>43</sup>. Cases that used to require in-person hearings now conclude in weeks with a few clicks. The SPC has also deployed AI in criminal justice: for example, sentencing recommendation systems to promote uniformity, and a tool known as the “206 System” in criminal trials that guides evidence examination and checks procedural compliance. These AI systems act like ever-vigilant assistants, flagging discrepancies and suggesting outcomes.

**Human Oversight and Official Assurances:** Despite this heavy integration of AI, Chinese officials insist that humans remain firmly in control of judicial outcomes. Zhou Qiang, Chief Justice of the SPC, stated unequivocally that “AI will never replace human judges and can only serve as an assistant to judges.”<sup>44</sup> Indeed, current pilot programs of AI judges operate “under close human judge supervision, and no court decisions are implemented without human approval.”<sup>45</sup> This official position can be seen as a reassurance: even as software handles much of the work, a human judge (typically named as such in the case files) is accountable for the decision. In practice, one can question

<sup>39</sup> N. Wang, M. Y. Tian, “Intelligent Justice”..., *ibid.*

<sup>40</sup> *Ibid.*

<sup>41</sup> T. Vasdani, Robot Justice..., *ibid.*

<sup>42</sup> *Ibid.*

<sup>43</sup> *Ibid.*

<sup>44</sup> N. Wang, M. Y. Tian, “Intelligent Justice”..., *ibid.*

<sup>45</sup> *Ibid.*

the depth of that human involvement—if the AI drafts a judgment and the human merely rubber-stamps it, the reality might be a de facto AI decision. But formally, the requirement of human sign-off preserves a semblance of the traditional model.

Even an authoritarian regime as China has, which, in theory, could enforce AI by decree, has opted to be more cautious. This is an indication of the realization that all-AI-based justice may not be widely accepted by the community or may not be trusted to be that reliable at the moment. The importance of AI has been a topic of debate among Chinese scholars and judges and some of them have accepted its limitations. As an illustration, according to Chinese scholarship as examined by Papagiannenas (2022), there is overall optimism regarding efficiency and consistency, but there is also an indication that there is concern regarding judicial fairness and transparency.<sup>46</sup> There is worry that judges might become over-reliant on AI recommendations, leading to a “rubber stamp” mentality.<sup>47</sup> The SPC is also concerned with maintaining control over local courts; interestingly, one reason for AI tools is to monitor and guide lower judges (preventing unapproved deviations). But that same central control logic means the SPC wants a human judge it can hold accountable if something goes wrong, rather than blame an algorithm.

**Fair Trial in Chinese Context:** China’s legal system does protect the right to a fair trial on paper (e.g., the Criminal Procedure Law provides for fair trials, and judicial interpretations emphasize fairness and impartiality). However, these rights are not judicially enforced against the state in the way they are in the ECHR system; rather, they are principles the judiciary itself says it upholds. Within that context, having a human judge is seen as part of ensuring “justice with temperature” (a phrase used in China to mean justice that isn’t purely mechanical). Chinese commentary often speaks of “intelligent courts” being “human-centered” despite their high-tech nature.<sup>48</sup> The idea is to use AI to improve accuracy and efficiency but to keep the human touch where it matters.

In practice, therefore, China’s approach aligns with a functional version of the right to a human judge, but for somewhat different reasons. It is less about individual autonomy or dignity, and more about system legitimacy and control. Nevertheless, the outcome is that in China today, no defendant is sentenced by a machine, and no final judgment is issued without a human judge’s involvement.

**Examples and Limitations:** A concrete example: In a private lending case in Hangzhou, the AI judge “Xiao Zhi” assisted by analyzing evidence and even

<sup>46</sup> S. Papagiannenas, *Towards Smarter and Fairer Justice? A Review of the Chinese Scholarship on Building Smart Courts and Automating Justice*, ‘Journal of Current Chinese Affairs’ 2022, 51(2), pp. 327–347; <https://doi.org/10.1177/18681026211021412> [accessed: 29.03.2025].

<sup>47</sup> N. Wang, M. Y. Tian, “Intelligent Justice”..., *ibid.*

<sup>48</sup> *Ibid.*

answering parties' questions, but a human judge was present to conclude the case.<sup>49</sup> The AI provided a recommended outcome, which reportedly helped the human judge finish the case in 30 minutes.<sup>50</sup> But critically, the final judgment was issued under the human judge's name. Similarly, the Internet Courts allow appeals to a traditional court if a party is dissatisfied—this ensures that ultimately a human panel of judges will review the matter *de novo*, thus acting as a safety net.<sup>51</sup>

There are, of course, concerns within China's experiment. One is that litigants, especially those without counsel, might not distinguish between the AI and a real judge and might effectively be having their first-instance trial entirely in front of a machine. Another issue is the quality of AI decisions: Chinese developers acknowledge that AI tools struggle with complex cases or those requiring subjective judgment (e.g., what is a "reasonable" outcome).<sup>52</sup> This mirrors Górski's point in the European context that AI might handle straightforward cases (like undisputed payments) but falter on complicated ones requiring nuanced assessment.<sup>53</sup>

Despite these challenges, China's continuing stance is to integrate AI heavily but keep a human in the loop. It might also be said that Chinese approach is pragmatic authoritarian: it uses technology to create maximum efficiency and control, but the unspoken rule admits that there is a threat that perceived fairness might be jeopardized in the event that human judges are completely displaced. We observe then that despite the liberal democratic models, the perceptions of the necessity of a human judge remain.

**Comparative Insight:** From Europe, the U.S., and China, a pattern emerges. All three jurisdictions—each in their own language and legal culture—converge on the idea that AI should assist, not replace, human judicial decision-making. Europe frames it as a matter of fundamental rights and has begun to codify it.<sup>54</sup> The U.S. frames it through due process and fairness, requiring human judgment behind any algorithmic tool.<sup>55 56</sup> China frames it as maintaining control and trust in the courts, insisting a human judge has the final word.<sup>57</sup> This consensus across very different systems strongly supports the argument that a procedural right to a human judge is emerging as a general principle—one that transcends ideological boundaries, because it speaks to something common in the notion of justice.

<sup>49</sup> *Ibid.*

<sup>50</sup> *Ibid.*

<sup>51</sup> T. Vasdani, *Robot Justice...*, *ibid.*

<sup>52</sup> N. Wang, M. Y. Tian, "Intelligent Justice"..., *ibid.*

<sup>53</sup> M. Górski, *Why a Human...*, *ibid.*

<sup>54</sup> Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) [2024] OJ L1689.

<sup>55</sup> *State v. Loomis* 881 NW 2d 749 (Wis 2016).

<sup>56</sup> J. M. Payé Salazar, *Impact...*, *ibid.*

<sup>57</sup> N. Wang, M. Y. Tian, "Intelligent Justice"..., *ibid.*

Having surveyed these approaches, we proceed to outline our proposed legal framework that consolidates and gives formal structure to the right to a human judge, ensuring it can be practically implemented as a safeguard for fair trials in the age of AI.

---

## A Framework for a Procedural Right to a Human Judge

---

Having analyzed the landscape, this Part proposes a structured legal framework to recognize and operationalize a procedural right to a human judge. The framework is intended as a doctrinal tool and policy guideline to ensure that AI can be employed in court systems only in ways that do not undermine fair trial rights. It articulates the conditions under which AI assistance is permissible and where the line must be drawn to preserve the human judge's essential role. In effect, these conditions form a test that courts or legislators can use to evaluate whether a given use of AI in adjudication is compatible with the right to a fair trial. The framework is informed by the comparative insights above and existing human rights principles.

In the very foundations, we can state that any litigant to the legal process is entitled to insist that his/her case be left to be finally decided by a human judge (or a panel of judges). It, in line with this, bans the existence of systems where the binding judgments are made without any meaningful human input. The framework can be broken down into several key requirements or principles:

- 1. Human Oversight and Final Decision Authority:** The final decision in any judicial proceeding must be made or confirmed by a human judge. No AI system should be allowed to issue a judgment or verdict on its own. This principle aligns with the EU AI Act's mandate that "final decision-making must remain a human-driven activity"<sup>58</sup> and the CCBE's call that human judge be "required to take full responsibility for all decisions."<sup>59</sup> In practice, this means AI may draft recommendations or assist with analysis, but a human judge must review the AI's output, exercise independent judgment, and either accept, modify, or reject it. The human judge should be able to state that they are satisfied with the decision on their own reasoning. This requirement ensures that there is always a human mind applying the law to the facts, upholding the idea that adjudication is a fundamentally human act.
- 2. Transparency and Explainability:** The use of AI in a judicial process must be transparent to the parties, and the reasoning of any AI-assisted outcome must be explainable in human-intelligible terms. Litigants should be in-

---

<sup>58</sup> Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) [2024] OJ L1689.

<sup>59</sup> Council of Bars and Law Societies of Europe (CCBE), *CCBE Position Paper...*, *ibid.*

formed if an AI system was used in handling their case (for example, in predicting risk, suggesting a sentence, or even drafting a judgment). They should also have access to an explanation of how the AI contributed to the decision. This dovetails with due process rights to understand the basis of a decision and with emerging norms like Article 22 GDPR (the right to an explanation for automated decisions). If an AI system's workings are a "black box" such that neither the judge nor the parties can understand its reasoning, then that system should not be used in a manner affecting the outcome.<sup>60</sup> Indeed, the CCBE suggests that where an AI's output cannot be sufficiently explained, it "must not be taken into account" by the authorities.<sup>61</sup> Therefore, an acceptable AI tool in court must either be inherently interpretable or come with mechanisms (such as reports or expert testimony) to translate its logic into the legal context.

3. **Contestability and the Right to a Human Review:** Parties must have the right to challenge and, if necessary, obtain a human re-examination of any AI-influenced decision. This means that if a litigant believes the AI's involvement led to error or bias, they can appeal or seek review before a human judge or panel that was not involved in the original AI-influenced decision. For instance, if an AI tool effectively decided an issue at trial, the defendant must have the opportunity to appeal to a higher court where judges will reconsider the matter without being bound by the AI's conclusion. This reflects the principle behind the GDPR's human review right and ensures that AI does not become an unappealable oracle. In the Bulgarian preliminary reference to the CJEU mentioned earlier, one question explicitly asked whether a person has the right to a human review of an automated decision and whether that human must be a judge in real proceedings.<sup>62</sup> Our framework answers yes: a litigant can always seek recourse to a human judge. The availability of such a "human in the appeal loop" is a crucial safety valve.
4. **Legality and Accountability:** The use of AI must have a clear basis in law, and responsibility for the decision must be traceable to a human authority. This principle requires that any integration of AI into judicial decision-making be authorized by legislation or rules that set parameters, rather than done *ad hoc*. It also requires that from the litigant's perspective, they can identify *who* decided their case—which should be a human judge or judges, even if they used AI input. The judgment or verdict should be issued under the judge's name, and that judge should be accountable

---

<sup>60</sup> Ibid.

<sup>61</sup> Ibid.

<sup>62</sup> S. J. Choi, K. Van Quathem, and B. Van Vooren, *CJEU Receives Questions...*, *ibid.*; <https://www.insideglobaltech.com/2025/02/24/cjeu-receives-questions-on-the-ai-act-relating-to-automated-decision-making/> [accessed: 29.03.2025].

(through appeals or disciplinary systems) for the outcome, just as they would be if no AI was involved. This prevents a governance gap where an AI makes a decision, and no one can be held responsible for errors. It ties into the idea of a tribunal “established by law”—a lawful court composed of judges cannot be replaced in substance by an algorithm that is not so established. Legal authorization is also important to satisfy the “prescribed by law” test under human rights law if any limitation of rights is involved.

5. **Reliability and Non-Discrimination:** Any AI applied to judicial decision-making should be of high accuracy and demonstrate that it does not cause impermissible biases or have dissimilar effects. Although this is on the quality of the AI and not the strictly procedural part, it is a necessary precaution. The right of a defendant to a fair trial would be infringed assuming that, say, an AI sentencing tool was systematically more severe to a given racial group. Therefore, the system demands a thorough testing and auditing of AI systems that are prone to bias and error rate before they are utilized in courts. Such a principle can be applied through certification (based on the ethical charter of the CEPEJ) or professional control. If an AI system cannot be audited or corrected for bias, relying on it may violate the equality of arms and nondiscrimination principles inherent in a fair trial. Under our framework, such a system should not be permitted until the issues are resolved.
6. **Proportional Use—Complementarity, Not Substitution:** The degree of AI involvement in a case should be proportional to the nature of the case, and AI should complement human judgment, not substitute for it. In minor, high-volume matters (e.g. small monetary claims or routine administrative hearings), more extensive AI assistance might be acceptable—perhaps even automated preliminary decisions—so long as a human review is available on request.<sup>63</sup> By contrast, in major criminal cases or other processes, where the basic rights of individuals are at risk, AI may only perform advisory roles, and the rest of the decisions must be made by the human judges sternly. Proportionality in this case acts as a counter to efficiency versus fairness: the higher the amount of harm the wrong decision might cause, the more we need to make sure that the conscience and thought process of a human being take center stage. This principle contains the possibility of innovation (such as ODR platforms in the case of small disputes on the basis of AI negotiation tools) without the right to appeal to a human judge being sacrificed.

All these six elements make the difference between a slogan and a working standard to which a right to a human judge belongs. They summarize these normative reasons that we desire human judges to concern themselves with—

<sup>63</sup> T. Vasdani, *Robot Justice*., *ibid.*; <https://www.lexisnexis.ca/en-ca/ihc/2020-02/robot-justice-chinas-use-of-internet-courts.page> [accessed: 28.03.2025].

independence, reasoning, empathy, accountability—and convert them into operational standards. In the event that a new AI tool or procedure is introduced, we can enquire: Does it retain ultimate power in a human? Do the sides have an idea of what is going on and why? Is it possible to oblige and receive a decision of a human being instead? Is it legal and distinct who is in charge? Is it accurate and unbiased? Is it applied in a suitable manner regarding the case? The use of AI would only get through the test of this framework in case it has all the answers right.

It should be mentioned that this framework does not prohibit the use of all AI in the judicial practices. Rather, it creates a red line at the replacement of the judge. The AI is able to aid in repetitive work, large case databases, high point identification or even give a result based on the precedents—provided that the judge is critical of the result offered. Such use can improve consistency and efficiency while respecting rights. Our framework encourages “human-in-the-loop” models,<sup>64</sup> where technology is subordinate to human decision-making. It rejects “human-out-of-the-loop” adjudication where no person truly weighs the merits of the case.

This proposal could be implemented in various ways: as we will discuss, it might inform treaty interpretations, constitutional doctrines, or legislation. But as a normative contribution, the key point is that the right to a fair trial in the AI era should be understood to guarantee a human judge’s involvement as an element of due process. This transforms what has been an implicit assumption into an express legal safeguard.

---

## **Discussion: Implementation, Challenges, and Future Implications**

---

Articulating a procedural right to a human judge is only the first step. For it to be effective, it must be recognized and implemented in legal systems. In this section, the pathways have been discussed to enshrine this right in law—from constitutional amendments and international treaties to court rules and ethics guidelines. It has been also considered potential challenges and counterarguments, examining how the framework can be defended and refined.

**1. Enshrining the Right in International and Domestic Law:** International human rights law is one method of entrenching the right to a human judge. The most obvious way would be an Additional Protocol to the ECHR or a like tool which makes it clear that the right to a fair trial would mean the right to have the case one is prosecuted in adjudicated by a human judge. Such an initiative

---

<sup>64</sup> I. Banks, *Operationalising...*, *ibid.*

could be led by the Council of Europe that has been active on AI ethics. It is true that the Council of Europe is already preparing a binding international treaty on AI and human rights; a clause regarding human oversight in court decisions would keep in line with Recital 61 of the EU AI Act and the principles of the CEPEJ. Similarly, the United Nations Human Rights Committee could address this issue in a General Comment on Article 14 of the ICCPR, or the Human Rights Council could pass a resolution highlighting the need for human judgment in judicial proceedings. For example, the UN High Commissioner for Human Rights has warned that certain uses of AI pose risks to the right to a fair trial.<sup>65</sup> Building on that, U.N. bodies could formally recognize that fully automated justice systems are incompatible with human rights. Such international statements or obligations would then guide national courts and legislatures.

At the domestic level, countries could amend their constitutions or legislation to affirm the right to a human judge. A constitutional amendment might be the strongest form—for instance, a constitution could state: “Every person has the right to have any charges or claims against them decided by an independent and impartial human judge.” This would put beyond doubt that no future government could quietly delegate judicial power to machines. Even absent new constitutions or amendments, statutes could codify the framework. For example, a country’s civil procedure code and criminal procedure code could be amended to explicitly prohibit judicial decisions without human approval. Legislatures might also enact an “AI in Courts Act” domestically, mirroring the EU’s approach, to regulate how AI may be used by the judiciary and to guarantee human final authority. These laws should incorporate the key principles this paper has outlined: requiring judges to personally sign off on decisions, mandating transparency to litigants, etc.

**2. Judicial Adoption and Ethical Norms:** Another avenue is through the judiciary’s own rule-making and ethical standards. Supreme courts or judicial councils often issue rules of procedure or codes of conduct for judges. These could be updated to address AI. For instance, a code of judicial ethics might say that a judge shall not delegate the substantive decision of a case to any other person or entity, including an AI system; or that a judge must ensure any AI tool used in decision-making is transparent and its use disclosed to parties. This would treat the right to a human judge as an aspect of judicial duty: it’s the judge’s duty not to abdicate their decision-making role. If judges abide by such rules, litigants are indirectly protected.

The right may also be reinforced by a series of cases in the courts. In cases where situations arise (e.g., a litigant claims the judge was unfair on a hearing because he or she depended on an algorithm blindly), the courts can directly

---

<sup>65</sup> S. Neuman, The U.N. Warns That AI Can Pose a Threat to Human Rights, National Public Radio 2021, Sep., 16; <https://www.npr.org/2021/09/16/1037902314/the-u-n-warns-that-ai-can-pose-a-threat-to-human-rights> [accessed: 29.03.2025].

declare that the right to due process or fair trial was abridged. These precedents will discourage over-use of AI. This is because in common law systems, particularly, a small number of vigorous decisions at the appellate level could create the principle without necessarily having to seek legislative action. To illustrate, in a situation where a higher court considered a fully automated procedure in small claims and determined that it was unconstitutional, then such a situation would necessitate that a human judge be inserted into said process later.

**3. Model Legislation and Cross-Border Consensus:** Model frameworks can assist jurisdictions in adopting best practices given that technology tends to outpace legal development. A model law or guidelines of “AI-Augmented Adjudication” based on our framework could be written by international bodies or NGOs. The existence of such a model (equivalent to the powerful Bangalore Principles of judicial conduct or UNCITRAL Model Laws in other areas) offers an example that can be emulated by countries. It also promotes a universal discussion in such a way that the standards are brought together. In the first place, we already observe convergence, as Europe, hard law, U.S. constitutional logic, and China all seem to be heading in the same direction. This could be crystallized through a formal articulation through a model document.

These actions have political viability that is worthy of consideration. It is probable that bipartisan or general consensus on the inadvisability of judicial replacement with machines exists in most countries. The safeguard in the form of a right to a human judge may not be subject to a strong opposition in terms of principle—who will openly contend that we should not have human judges? The primary opposition may be administrative efficiency proponents or revenue-starved judicial administrators considering automation as an answer to backlog. However, even they usually consider AI as an auxiliary, rather than a complete replacement (no one is ready to place their faith in an AI-only judge without a human involved at present, due to the level of technology development).

**4. Addressing Counterarguments:** When we are presenting a right to a human judge, we have to face a number of possible objections and future cases. To elucidate the range and durability of our proposal, this paper will discuss the most topical counterarguments in this direction:

- ◆ **“What if AI becomes better than human judges?”** It is possible to say that eventually AI might be more accurate, reliable, and unbiased than a human judge, thus the demand to have a human judge can be counterproductive when it comes to conducting a fair trial. This situation is not real; currently, AI systems are designed to mirror the information and goals provided to them and do not have an actual sense of morality or perception of the environment. Even when one thinks of a super-intelligent, unbiased AI, justice is

not determined by accuracy. It is the matter of process and perceived rightness as well. Studies of procedural justice have demonstrated that an ideal decision, which is correct by definition, might be turned down as unfair when the process used is remote or incomprehensible. The fairness of a trial includes the ability to make oneself heard by an objective human adjudicator—which an algorithm, no matter how sophisticated, cannot provide in real sense. In addition, we do not prohibit input of sophisticated AI in our system, just that a human judge be in final authority to authenticate results. When the reality has come one day when AI reasoning is as good as that of humans, the human review will be a very symbolic formality—but a significant one to ensure that citizens remain trusted and accountable. More importantly, AI competence would not be affected negatively by the presence of a human checkpoint even in the event of catastrophic failures and helping to avoid catastrophic failures in the few instances where an AI may be in error, that only a human conscience or common sense can detect. Concisely, even in the possibly super-human AI future, the watchful eye of a human judge will be an inexpensive way of protecting values such as empathy, mercy, and legitimacy which pure algorithms cannot so effectively represent. According to Zou and Lefley (2025), the idea of being judged by a human being is associated with human dignity—a concept that would not go away despite the high level of sophistication of machines.<sup>66</sup> Thus, our proposal would still be justified in a world of powerful AI: it ensures that *legal judgment* retains a human face and is ultimately grounded in human societal values.

- ◆ **“Humans are flawed too—why privilege human judgment?”** It is factual that human judges are not completely impartial, inconsistent and are in fact sometimes wrong. Indeed, one of the significant reasons why AI should be introduced to the adjudication process is to eliminate these human frailties. Nonetheless, the answer does not lie in the removal of humans in the process but in the enrichment and enhancement of human judgment with the help of AI. We give human judgment the privilege due to attributes that human beings only developed to date, namely compassion, intuitive moral reasoning, and discerning the nuance in relation to communal values. Such attributes encourage the people to accept the decisions of the courts. Human judges too work within regimes of responsibility—they have to give explanations why they have made a decision, are subject to questioning or appeal and may be dismissed or punished due to ill behavior. An AI can neither be cross-examined on its conscience nor can it be answerable at the ballot box or even via a professional disciplinary agency. Until we in some way endow machines with accountability, transparency and value-alignment comparable to human beings (which, like humans, is merely

<sup>66</sup> M. Zou, E. Lefley, *Generative...*, *ibid.*

a matter of theory), it is sensible to retain humans in the driver's seat of the process of judging. Practically, it can be interpreted as applying the AI to help judges make superior decisions—by processing information or implicating knowledge—though the final decision is still a human action. It is aimed at a mix of both that verifies and rectifies human weaknesses without discarding the incomparable merits of human adjudication.

- ◆ **“Parties might consent to an AI judge—should they be allowed to waive the right?”** This is a situation that is becoming more and more possible: assume that two parties in an issue of civil disagreement concur to send their case to an AI adjudicator, out of expediency, or reduced expense. Was it possible that the parties had the voluntary right to waive human judge? And on the one hand the legal regimes do not forbid parties to simply forego courts altogether in favor of personal arbitration—and in principle even an arbitrator can be a non-human process should both parties' consent. But there is a limit to the public policy. There are some basic characteristics of a fair trial, which cannot be contracted out in a way that interferes with the integrity of the justice. We hold that the right to a human judge in social judicial processes ought to be incalculable. This is comparable to the situation with a criminal suspect in that he cannot consent to have his trial decided by coin flip or trial by combat—even in the event both parties consented to it somehow—because these are core values of justice and confidence of the people in the courts. Similarly, a trial involving a completely AI-controlled court in a state court would not comply with the minimum provisions of the justice system operating according to the rule of law. On all process stages, as the CCBE has demanded, a human judge should be ensured. However, when it comes to the area of private dispute resolution, should fully informed parties really desire to work with an AI judge in the form of arbitration, the law may be diligently willing to allow it with multiple protective measures. As an example, the parties should not be denied the right to appeal to any real court. A human judge would be able to review or confirm an AI-based decision on arbitral awards before it can be declared as a judgment. As a matter of fact, whenever a court would want to enforce a decision that has been AI-rendered, it would be forced to subject it to scrutiny—possibly, the way an expert decision is treated—to make sure that it does not infringe on the policy of the people or fundamental rights. Concisely, a human being is free to test his or her AI adjudication, but the government has no obligation to sign an agreement where a human judge is completely absent. Procedural right of a human judge must be regarded as a guarantee of justice in which the legal system can provide even in the circumstances when parties will not require this right.
- ◆ **“Is this framework just Luddism? Shouldn't we focus on making AI judges fair, rather than blocking them?”** The position this paper takes is not anti-technology—it is pro-human-rights. This paper does not advocate aban-

doing AI in justice; rather, set boundaries for its use. Requiring human oversight and final authority is not about destroying the benefits of AI, but about channeling those benefits in a responsible way. History shows that clear rules and human-centered design often foster, rather than hinder, sustainable technological adoption. (For example, the introduction of safety regulations for automobiles didn't stop the development of cars—it made the public more confident in using them.) Likewise, establishing a human-judge requirement can ensure trust in the justice system as it innovates. By affirming that technology will serve as a tool for human judges—not a replacement for them—we maintain public trust and avoid the backlash that could occur if people felt “justice” had been reduced to a black-box algorithm. In the long run, this principled approach will actually enable courts to embrace helpful AI tools (for research, case management, decision support, etc.) without fear of undermining the fairness or legitimacy of the process. In sum, our framework doesn't block AI at all; it guides its integration so that efficiency and innovation can be achieved *alongside* the enduring values of a human-centric justice system.

**5. Practical Implications for Courts and Stakeholders:** Implementing this right will have practical consequences. Courts will need resources to verify AI tools' reliability (e.g., perhaps court-employed data scientists or certification bodies). Judges will need training to understand AI outputs and not over-defer to them—to avoid the “automation bias” where humans trust a computer blindly. Legal education may evolve to cover the interface of AI and judging, so new lawyers can navigate these issues. Litigants may, in the short term, gain new grounds of appeal or review (e.g., appealing that a judge relied too much on an algorithm). This is a necessary phase to set precedents on what acceptable use looks like.

Technology providers—the companies that build judicial AI systems—will also have to adjust. They will need to design systems that fit the human-in-charge paradigm. That could mean building better explanation features or shifting from autonomous decision-making to decision-support design. Interestingly, this could open up innovation: rather than trying to replace judges, companies will compete to better assist judges.

**6. Long-Term Outlook:** In the long run, if our framework is widely adopted, one can imagine it becoming a standard part of the concept of a fair trial globally. As routine as the right to counsel or the right to a public hearing, the “right to a human judge” could be cited in courtrooms from Strasbourg to San Francisco to Shanghai. It would function as a bulwark as AI becomes more pervasive. If truly revolutionary AI emerges (for instance, an AI that claims to have general intelligence or even consciousness), the right to a human judge will force a societal conversation and legal scrutiny before any handover of judicial power. It ensures that *people* remain the ultimate arbiters of justice,

at least until such time (if ever) that we collectively decide otherwise through democratic means.

Lastly, it is mentioned that the implementation of this right does not only have to do with the prevention of dystopia, but it also has to do with confirmation of what we appreciate in our justice system. Making judgment is a moral act, one on which we as a society leave to humans who may be merciful, who may perceive context, and who may represent the relationship of law to the community. In a time where machines are becoming more advanced, to create a border around the sacredness of the judicial position is similar to creation of a boundary around the human dignity of the judge and the judged. It conveys a message that technology is a servant of man rather than *vice versa*; more so in the courtroom.

---

## Conclusion

---

With courts having gone artificial intelligence, we are at a crossroad of the future of fair trials. The paper has discussed that with the efficiencies and capabilities of AI, one of the principles that have to be preserved is the right to be judged by a human being. As revealed in this paper, this procedural right to a human judge is entrenched in the functions of the fair trial guarantees under international law despite the fact that it has only recently required the articulation. Looking at both the European, American, and Chinese solutions we have seen that there is an overlap in the preservation of a human element in the court ruling—meaning that our normative proposal is not an extreme ideal but a global legal introspection. I provided an exhaustive system of the steps to follow on embracing AI in the court of justice without infringing the right to a fair trial. The principles of the framework—human final authority, transparency, contestability, legality, non-discrimination, and proportional use gives a roadmap on which courts and legislatures should operate. The two of them make sure that although AI might have a part to play in the delivery of justice, it does not overthrow the important position of human judges. The application of the framework in the form of laws, judicial norms and possibly new commitments at the international level will be essential in the years to come. Positively, we find signs towards this direction: in the form of the pending EU AI Act that entrenches human decision making in legislation,<sup>67</sup> in cases like the Loomis decision, or in China, where its judges are insisting that their so-called smart courts be run by humans.<sup>68</sup>

---

<sup>67</sup> Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) [2024] OJ L1689.

<sup>68</sup> N. Wang, M. Y. Tian, “Intelligent Justice”..., *ibid*.

Fair trial is one of the symbols of civilization, and its development should not forget about the human factor in it making it legitimate. In a world where algorithms are more and more becoming an aspect of governance, the demand to have a human judge is one of the demands of a bigger demand to have human justice. It recognizes that technology should not be used to substitute judgment of man, at least in law and rights. Decency in the court, responsibility of the judgment, and the law itself all require a human face and mind in the bench. Finally, it is important to realize that having a procedural right to a human judge is a protection and a compass. It is a protection against the erosion of fair trial rights by untested technology and is a means of creating AI systems that complement and not supplant human judgment. Tracing further, the right may become an important source of new standards and rules, through which our justice system is not only efficient and modern, but also essentially fair and human. The era of AI in the courts is here—provided we have an effective framework and belief in the essential values, we will be able to make sure that justice, at the last, will be served in the name of the people, and people.<sup>69</sup>

### Abstrakt

Wykorzystanie sztucznej inteligencji (AI) w systemach sądowych następuje coraz szybciej na całym świecie, budząc poważne wątpliwości co do przyszłości prawa do sprawiedliwego procesu sądowego. Niniejszy artykuł stanowi argument przemawiający za uznaniem prawa proceduralnego do sędziego-człowieka za fundament ochrony przed sztuczną inteligencją w XXI wieku. Uważam, że gwarancje rzetelnego procesu sądowego zawarte w obowiązującym prawie międzynarodowym – tj. w art. 6 europejskiej Konwencji praw człowieka i podstawowych wolności (EKPC) oraz art. 14 Międzynarodowego paktu praw obywatelskich i politycznych (MPPOiP) – w sposób dorozumiany wymagają orzekania przez człowieka. Należy to wyraźnie stwierdzić. Wyniki przeprowadzonego przeze mnie przeglądu literatury i obecnych praktyk sugerują, że sędziowie-ludzie oferują niekwestionowane cechy, takie jak zdolność osądu, empatia i odpowiedzialność, których brakuje sztucznej inteligencji w jej obecnej formie i których nigdy nie będzie ona w stanie w pełni odzwierciedlić. Podobieństwo doświadczeń europejskich, amerykańskich i chińskich w analizie porównawczej sugeruje ogólne zrozumienie konieczności ludzkiego nadzoru nad sądowymi systemami AI – co jednak opiera się na różnych zasadach prawnych. Na podstawie tych obserwacji w artykule narysowano nowe ramy prawne, które uwzględniają prawo do sędziego-człowieka. Ramy te określają warunki, w jakich narzędzia sztucznej inteligencji mogą być wykorzystywane w sądzie bez naruszania podstawowych praw do rzetelnego procesu sądowego,

<sup>69</sup> Council of Bars and Law Societies of Europe (CCBE), CCBE Position Paper..., *ibid.*; CCBE *Flags Dangers of AI-Led Decision-Making*, 'Law Society of Ireland Gazette' 2023, May, 26; <https://www.lawsociety.ie/gazette/top-stories/2023/may/ccbe-flags-dangers-of-ai-led-decision-making/> [accessed: 28.03.2025].

kładąc główny nacisk na kontrolę przez człowieka, otwartość, możliwość zaskarżenia i odpowiedzialność. Wyjaśniam również sposoby egzekwowania tego prawa, w tym zapisanie go w traktatach dotyczących praw człowieka i konstytucjach, a także przyjęcie wzorcowych przepisów. Podsumowując: konieczna jest ochrona pozycji sędziego-człowieka w celu utrzymania sprawiedliwości i rzetelności wymiaru sprawiedliwości w erze coraz większej niezależności systemów opartych na AI.

**Słowa kluczowe:** prawa procesowe, sędzia-człowiek, rzetelny proces sądowy, sztuczna inteligencja (AI), nadzór sądowy, analiza porównawcza.

---

## BIBLIOGRAPHY

---

Adoption of a European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment at the 31st Plenary Meeting of the CEPEJ in Strasbourg on 3 and 4 December 2018, International Union of Judicial Officers (UIHJ) 2018, Dec., 4—available at: UIHJ website [accessed: 28.03.2025].

Artificial Intelligence in Judicial Systems: Promises and Pitfalls. Report of the Special Rapporteur on the Independence of Judges and Lawyers, United Nations General Assembly, UN Doc 2023, A/80/169.

Banks I., *Operationalising Human Oversight of AI-Supported Judicial Decision-Making: A Systemic Perspective*, 'Algorithms, Data and Society' 2023—available at: AlgoSoc [accessed: 28.03.2025].

Binns R., *Human Judgment in Algorithmic Loops: Individual Justice and Automated Decision-Making*, 'Regulation & Governance' 2022, 16(1)—available at: DOI: 10.1111/rego.12358 [accessed: 29.03.2025].

CCBE *Flags Dangers of AI-Led Decision-Making*, 'Law Society of Ireland Gazette' 2023, May, 26—available at: Law Society Gazette [accessed: 28.03.2025].

Choi S. J., Van Quathem K. and Van Vooren B., *CJEU Receives Questions on the AI Act Relating to Automated Decision Making*, 'Inside Global Tech' 2025, Feb., 24—available at: Inside Global Tech [accessed: 29.03.2025].

Convention for the Protection of Human Rights and Fundamental Freedoms, European Convention on Human Rights, 1950, Article 6(1).

Council of Bars and Law Societies of Europe (CCBE), *CCBE Position Paper on the Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) 2021*, Oct., 8—available at: CCBE [accessed: 28.03.2025].

Council of Bars and Law Societies of Europe (CCBE), *CCBE Statement on the Use of AI in the Justice System and Law Enforcement 2023*, May, 25—available at: CCBE (Statements, 2023) [accessed: 28.03.2025].

De Miguel Beriain I., *Does the Use of Risk Assessments in Sentences Respect the Right to Due Process? A Critical Analysis of the Wisconsin v. Loomis Ruling Law*, 'Probability and Risk' 2018, 17(1)—available at: DOI: 10.1093/lpr/mgy001 [accessed: 29.03.2025].

Engel C., Linhardt L. and Schubert M., *Code Is Law: How COMPAS Affects the Way the Judiciary Handles the Risk of Recidivism*, 'Artificial Intelligence and Law' 2024—available at: DOI: 10.1007/s10506-024-09389-8 [accessed: 29.03.2025].

Górski M., *Why a Human Court? On the Right to a Human Judge in the Context of the Fair Trial Principle*, 'European Criminal Law Associations' Forum' 2023, 1—available at: Eucrim [accessed: 28 March 2025].

McKay C., *Predicting Risk in Criminal Procedure: Actuarial Tools, Algorithms, AI and Judicial Decision-Making*, 'Sydney Law School Research Paper' 2019, 19(67)—available at: SSRN [accessed: 29.03.2025].

Murray K., *What the Legal Drama For the People Teaches Us About AI and Legal Ethics*, O'Hagan Meyer 2025, Feb., 6—available at: O'Hagan Meyer [accessed: 29.03.2025].

Neuman S., *The U.N. Warns That AI Can Pose a Threat to Human Rights*, National Public Radio 2021, Sep., 16—available at: NPR [accessed: 29.03.2025].

Papagianneas S., *Towards Smarter and Fairer Justice? A Review of the Chinese Scholarship on Building Smart Courts and Automating Justice*, 'Journal of Current Chinese Affairs' 2022, 51(2)—available at: DOI: 10.1177/18681026211021412 [accessed: 29.03.2025].

Payé Salazar J. M., *Impact of Artificial Intelligence on Effective Judicial Protection*, 'AfricLaw' 2023, June, 26—available at: AfricLaw [accessed: 29.03.2025].

Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) [2024] OJ L1689.

Sourdin T., *Judge v. Robot? Artificial Intelligence and Judicial Decision-Making* 'University of New South Wales Law Journal' 2018, 41(4).

Sourdin T., *Judges, Technology and Artificial Intelligence: The Artificial Judge*, Edward Elgar Publishing: Cheltenham 2021.

State v. Loomis 881 N.W.2d 749 (Wis. 2016).

Susskind R., *Online Courts and the Future of Justice*, Oxford University Press: Oxford 2019.

United States v. Booker 543 U.S. 220 (2005).

Vasdani T., *Robot Justice: China's Use of Internet Courts*, Toronto 2020, Feb.—available at: LexisNexis [accessed: 28.03.2025].

Wang N. and Tian M. Y., *Intelligent Justice: Human-Centered Considerations in China's Legal AI Transformation*, 'AI and Ethics' 2023, 3(2).

Završnik A., Plesničar M. M., *The Right to a Human Judge in Criminal Justice Settings* (paper presented at the American Society of Criminology 79th Annual Meeting, San Francisco, 15.11.2024)—available at: AllAcademic [accessed: 28.03.2025].

Zou M., Ebers M., Poncibò C. and Calo R. (eds.), *The Cambridge Handbook of Generative AI and the Law*, Cambridge University Press, 2025.



Kamil Zaradkiewicz\*

## **Robot Judge and the Right to a Fair Trial**

### **[Sędzia-robot a prawo do rzetelnego procesu sądowego]**

#### **Abstract**

Today's "anthropocentric" legal system does not allow court cases to be resolved by artificial intelligence systems. However, it is almost obvious that soon, increasingly advanced technological tools will become involved in the justice system to an ever-greater extent. In this context, fundamental systemic questions are already emerging. First, therefore, the question must be asked whether constitutional norms require—and if so, to what extent—human participation in the adjudication of judicial matters. Secondly, the answer to the question of whether it would be permissible considering the constitutional (and international) standard of the right to a fair trial—and thus to a duly constituted court established by law—and, finally, in view of the right of the parties to be heard in court proceedings. The aim of this article is to present the basic problems related to these issues in the context of current constitutional standards.

**Keywords:** robot, artificial intelligence (AI), right to a court, court, justice.

## **Introduction**

In recent years, artificial intelligence has become unpredictable and has taken on a "life of its own".<sup>1</sup> This issue cannot be overlooked in the context of constitutional institutional arrangements (including matters relating to fundamental rights and freedoms) and the functioning of public authorities, in particular the judiciary. In the case under analysis, however, the key issue is not the assessment of the 'personality' or, more broadly, the 'subjectivi-

\* **Kamil Zaradkiewicz** – assistant professor, PhD, post-doctoral degree, University of Warsaw (affiliation), Faculty of Law and Administration, Department of Civil Law, judge of the (Polish) Supreme Court, Civil Division; <https://orcid.org/0000-0003-3067-3583>; [k.zaradkiewicz@wpia.uw.edu.pl](mailto:k.zaradkiewicz@wpia.uw.edu.pl) / adiunkt, dr hab. nauk prawnych, Uniwersytet Warszawski (afiliacja), Wydział Prawa i Administracji, Katedra Prawa Cywilnego; sędzia SN w Izbie Cywilnej.

<sup>1</sup> T. Voigt [in:] Beck-online Grosskommentar BGB. Gesamt—Hrsg: Gsell/Krüger/Lorenz/Reymann (eds.): Spickhoff, 1.02.2025, BGB § 823, no. 759.

ty' of artificial intelligence as such, but rather whether AI systems meet the criteria and standards that should characterise the modern justice system, particularly in the light of international law (human rights protection) and constitutional requirements.

The use of artificial intelligence systems within the functioning of the judiciary is determined primarily by the framework of constitutional-level structural norms. It is typically constitutional provisions that set forth the fundamental principles and standards governing the judiciary's operation as one of the branches of public authority.

Until quite recently—including in Poland at the time of adopting the new Constitution in 1997—the constitutional legislator did not envisage robot judges, either because such technological possibilities did not yet exist<sup>2</sup> or were regarded merely as hypothetical. Even a dozen or so years ago, legal scholars' reflections on the operation of artificial intelligence systems—including the drafting of regulations in this field—were regarded as the work of unserious fabulists or as an exercise in legal futurology, although the 1947 Constitution of Bremen<sup>3</sup> had already stipulated in Article 12(1) that human beings take precedence over technology and machines (*Der Mensch steht höher als Technik und Maschine*).<sup>4</sup> The authors of this provision probably understood it quite differently back then than we do today.<sup>5</sup> However, as early as the 1960s, the eminent Romanist Franz Wieacker warned that “cybernetic technology is associated with the pathos of utopian hopes for the unlimited liberation of humankind from its responsibilities—for example, a kind of statistical justice without the effort involved in making personal decisions. On the other hand, humanistic and cultural-critical self-satisfaction sounds the alarm regarding the progressive dehumanisation and the surrender of humans to inevitable constraints and mechanisms.”<sup>6</sup>

At present, the use of artificial intelligence systems in the judiciary remains limited. But, the future application of these systems may lead to the conclusion that a new type of autonomous legal entity should be recognized as a judge, capable of delivering substantive judgments in any civil or criminal case under consideration, on at least an equal footing with a human judge. Regardless of what the foreseeable near or distant future may hold, analy-

<sup>2</sup> A. K. Bernzen, *Roboter als Richter? Zur Automatisierung der Rechtsprechung*, 'Recht Digital' 2023, 3, p. 134; L. Wolff, *Algorithmen als Richter. Verfassungsrechtliche Grenzen entscheidungstreffender Rechtsgeneratoren in der Rechtsprechung*, Trier 2022, p. 127.

<sup>3</sup> *Gesetzblatt der Freien Hansestadt Bremen* 1947, 47, pp. 251–257.

<sup>4</sup> In the context under consideration, this provision is mentioned by D. Nink, *Justiz und Algorithmen. Über die Schwächen menschlicher Entscheidungsfindung und die Möglichkeiten neuer Technologien in der Rechtsprechung*, Berlin 2021, p. 351.

<sup>5</sup> See W. Pfeil, *Der Mensch steht höher als Technik und Maschine—Benötigen wir ein Grundrecht zum Schutz vor Künstlicher Intelligenz?*, 'International and Comparative Law Review' 2020, vol. 2, p. 88.

<sup>6</sup> F. Wieacker, *Recht und Automation* [in:] *Festschrift für Eduard Bötticher zum 70. Geburtstag*, Berlin 1969, p. 387.

ses on this subject are already appearing with increasing frequency in legal journals. Furthermore, academic monographs devoted to this issue are being published. The aim of this analysis is to highlight the key issues surrounding the question of whether ‘smart courts’—understood as machines capable of resolving disputes between citizens in the same way as traditional courts composed of professional human judges—are permissible.

---

## AI Tools and the Judiciary

---

From the perspective of the technological progress already observable today, the use of AI tools in the judiciary remains very limited. They may be—and increasingly are—employed primarily for the anonymization or pseudonymization of judicial decisions, documents, or data, and additionally for legal information retrieval. Tools of this kind are being used successfully and on an increasingly widespread basis, and not just by judges. Information technologies thus continue to function merely as instruments supporting the judge’s work—tools that must, moreover, be approached with considerable caution due to operational defects commonly referred to as artificial-intelligence “hallucinations”.

Even though these tools are not yet highly developed, there should be no doubt that the development of artificial-intelligence—including its application within the administration of justice—will continue to advance. The use of artificial-intelligence systems in the judiciary on a significantly broader scale not only as tools to assist humans, but at least to some extent as independent actors, will soon become unavoidable. The question of the extent to which such ‘autonomy’ would be permissible is therefore already relevant today. The Chinese example of already existing “smart courts” and “intelligent justice” illustrates the far-reaching possibilities in practice.<sup>7</sup> The prospect of robot judges still terrifies many people, even though autonomous systems are imperceptibly making ever deeper inroads into various other areas of our lives.

Consequently, as technological solutions become more sophisticated, the scope of modern tools used in judicial work will change significantly over time. From today’s perspective, we are probably unable to even foresee the direction and scope of the potential development of modern technologies.

---

<sup>7</sup> See e.g. Ch. Shi, T. Sourdin, B. Li, *The Smart Court—A New Pathway to Justice in China?*, ‘International Journal for Court Administration’ 2021, vol. 12, 1, *passim*; S. Yupeng, *History and Prospects of Creation of “Smart Courts” in the PRC: Problems of Application of the Artificial Intelligence Technology*, ‘History of State and Law’ 2023, May, 11, p. 44; Z. Tang, *Smart Courts in Cross-Border Litigation*, ‘Rabels Zeitschrift für ausländisches und internationales Privatrecht’ 2023, vol. 87, p. 497 *et seq.*

We can be confident that technological development will lead to a situation in which these systems are deemed sufficiently advanced to remove the human decision-maker from dispute resolution at least in certain areas. Just as AI today independently solves tasks that are not yet particularly complex—something which generally no longer raises too many objections—it will soon be making decisions on far more serious matters. The AI systems would thereby become, to a significant extent, a new component of the system of public authority, as particular categories of cases are entrusted to them for adjudication, effectively “delegating” state-authority tasks currently exercised, on behalf of society, by professional human judges. The right tool (machine) can already serve as a useful aid to a judge’s work. In practice, too, the use of artificial-intelligence systems in the administration of justice is not merely a matter of abstract futurological speculation. The use of artificial intelligence is planned in Estonia and is already being applied to some extent in German order for payment proceedings (*Eine maschinelle Bearbeitung ist zulässig*, § 689(1), second sentence, of the the German Code of Civil Procedure (ZPO));<sup>8</sup> see also §§ 703b and 1088 of the the ZPO; similarly, regarding simplified proceedings in family matters, see § 258(1) of the *Familienvorfahrensgesetz*.<sup>9</sup> The key question in this regard is whether—and where—the limits of trust lie when it comes to decisions that may not be verifiable at all by human standards, or even by the language available to us. Using a “smart court” that operates on a “black box” principle would, in effect, mean relying on a highly risky—and ultimately unverifiable—decision. This concern is not alleviated by the fact that, in reality, at least to some extent, the human brain and the decision-making processes taking place within it—and therefore also the processes occurring in the mind of a human judge—are a similar “black box”.<sup>10</sup>

Bearing in mind this future prospect of their broad application as technological progress continues, it is already necessary to initiate a discussion on whether—and within what constitutional boundaries—the use of AI tools in the judiciary is permissible. Their deployment must ensure full respect for constitutional standards of protection (for instance, regarding the processing of the personal data of parties or participants in judicial proceedings). It may also be necessary to revise these standards to accommodate the requirements associated with using AI systems as “robot judges”.

<sup>8</sup> „1. The dunning procedure is conducted by district courts [Amtsgerichte]. 2. Automated processing is permissible.”

<sup>9</sup> G. Rühl, *Einsatz von KISystemen in der Justiz* [in:] *Künstliche Intelligenz. Wie gelingt eine vertrauenswürdige Verwendung in Deutschland und Europa?*, Bundesministerium für Umwelt, Naturschutz, nukleare Sicherheit und Verbraucherschutz, F. Rostalski (ed.), Tübingen 2022, p. 281.

<sup>10</sup> J. J. Vasel, *Künstliche Intelligenz in der Justiz*, ‘Legal Tech Zeitschrift’ 2023, vol. 3, p. 184.

---

## AI as a Holder of Judicial Authority

---

In undertaking an assessment of the permissible scope for the use of AI in the functioning of public authorities (which include the judiciary) and administration of justice, one must, above all, keep in mind the constitutional constraints, foremost among them the principle of democracy understood as representative governance (see Article 179 in conjunction with Article 4 of the Constitution of the Republic of Poland).<sup>11</sup> Judges are holders of public authority, and, as such, their position requires at least an indirect chain of legitimacy deriving from the people. They must therefore possess what is known as “democratic legitimacy”.<sup>12</sup> However, unlike the legislative and executive branches, citizens’ ability to influence judicial appointments is limited (which, in European systems, is usually linked to the irremovability of judges), the procedure for appointing judges, in accordance with the principle of democratic legitimacy, must also provide for at least an indirect influence of the public’s choice (exercised through representatives in parliament or the head of state).<sup>13</sup> This requirement is, moreover, the first to raise doubts about the possibility of replacing human judges with machines. It would be difficult to imagine—although, purely theoretically, it is certainly not impossible—a machine having a democratic mandate from citizens to exercise power on their behalf. Moreover, this legitimacy cannot be viewed merely as a means of appointing (selecting) a holder of judicial power but must also be seen in the context of the requirement to apply the principle of legality, that is, to be bound by legislation and to adjudicate within the framework of the existing legal order.<sup>14</sup>

Entrusting judicial authority to judges delineates both the institutional and the personal framework within which that authority is exercised. In this context, the view seems to prevail, that “the judiciary must be exercised by natural persons (the personal element) and must constitute a public function entrusted to and vested in those persons (the institutional element).”<sup>15</sup>

---

<sup>11</sup> See e.g. A. K. Bernzen, *Roboter...*, p. 134; S. Huber, *Entscheidungsfindung im Zivilprozess durch künstliche Intelligenz* [in:] *Prozessuales Denken und Künstliche Intelligenz*, Ch. Althammer, H. Roth (eds.), Tübingen 2023, p. 47; L. Wolff, *Algorithmen...*, p. 221 et seq.

<sup>12</sup> See e.g. D. Nink, *Justiz...*, p. 321 et seq.; from general literature e.g. E.-W. Böckenförde, *Demokratie als Verfassungsprinzip* [in:] *Handbuch des Staatsrechts, Band II, Verfassungsstaat*, J. Isensee, P. Kirchhof (eds.), Heidelberg 2004, p. 429 et seq.; T. Groß, *Grundlinien einer pluralistischen Interpretation des Demokratieprinzips* [in:] *Demokratie und Grundgesetz. Eine Auseinandersetzung mit der verfassungsgerichtlichen Rechtsprechung*, Baden-Baden 2000, p. 98 et seq.; A. Voßkuhle, G. Sydow, *Die demokratische Legitimation des Richters*, ‘Juristenzeitung’ 2002, vol. 14, p. 673 et seq.; C. D. Classen, *Gesetzesvorbehalt und Dritte Gewalt*, ‘Juristenzeitung’ 2003, vol. 14, *passim*.

<sup>13</sup> See e.g. A. Voßkuhle, G. Sydow, *Die demokratische...*, p. 676 et seq.

<sup>14</sup> D. Nink, *Justiz...*, p. 324 et seq.; A. K. Bernzen, *Roboter...*, p. 135; L. Wolff, *Algorithmen...*, p. 225.

<sup>15</sup> Einsatz von KI und algorithmischen Systemen in der Justiz. Grundlagenpapier zur 74. Jahrestagung der Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs vom 23. bis 25. Mai 2022 in Rostock, p. 6.

This view remains valid under the Polish Constitution of 1997. Pursuant to Article 178(1) of the Constitution of the Republic of Poland, judicial authority is exercised by judges (*verba legis*: “judges in the exercise of their office”). At least until recently, it was obvious to everyone that when the framers of the constitution referred to ‘judges’ in its provisions, they meant only natural persons.

As Anna K. Bernzen aptly observes, “an automatic system does not receive remuneration and cannot retire—but it can be switched off.”<sup>16</sup> This is not neither merely the physical characteristics of the entity nor formal requirement of “being a person” with specified competencies (legal knowledge, age, and so forth). The core of the problem concerns qualities traditionally associated with human reasoning, such as the ability to process information within a specific judicial proceeding, including identifying specific problems and making decisions autonomously and responsibly.<sup>17</sup>

Being a judge means being open to other people’s cases and the issues at stake, which often determine the course of people’s entire futures. A judge must also “put himself in the other person’s shoes,” make an empathetic assessment, take into account the social consequences of decisions, identify problems, and make decisions in an autonomous and responsible manner.<sup>18</sup> It is therefore difficult to imagine that a machine—which is usually perceived as a soulless, lifeless tool incapable of empathy or of making individual judgments about human behavior—could possess such qualities. It is obvious that a judge cannot act mechanically or unreflectively, nor function as an “automaton” designed solely to process data and “output” the result of that processing (a judicial automaton, Richterautomat).<sup>19</sup> This is the case regardless of the fact that people do not always make sound and rational decisions; as judges, they can be biased, make errors in reasoning and have other “weaknesses in rationality”.<sup>20</sup>

Such qualities can certainly be possessed by a human being, and as of today, they remain beyond the capacities of a machine. All machines can do today is give the impression that they are capable of mimicking human behaviour. However, what we perceive as a “personalised approach” is nothing more than dry data processing within a system of algorithms. In a well-known decision from 2016, the German Federal Constitutional Court stated that “judicial law-finding is not limited to the execution of predetermined norms in the sense that the judge would act merely as the ‘mouth of the statute’, that is, as a ‘subsumption automaton’. Rather, every act of adjudication inherently involves the methodological interpretation of statutory norms and the definition of

<sup>16</sup> A. K. Bernzen, *Roboter...*, *ibid.*

<sup>17</sup> *Einsatz von KI...*, p. 9.

<sup>18</sup> *Einsatz von KI...*, p. 9.

<sup>19</sup> *Ibid.*

<sup>20</sup> See D. Nink, *Justiz...*, p. 460.

their scope of application, to decide, based on a positive, abstractly formulated legal provision, how it should be applied in an individual case.”<sup>21</sup>

Also in light of Article 45(1) of the Constitution of the Republic of Poland, which establishes the individual right of access to a court, it is still accepted as a paradigm, although it was developed without direct reference to the issues analysed here, that a “court” within the meaning of this provision is constituted by judges who are natural persons, appointed to judicial office on the basis of the relevant constitutional and statutory regulations, and who meet specified criteria (such as holding a higher legal education, completing judicial training, and possessing unimpeachable character).

The Polish Constitution permits the participation of what is commonly referred to as the “social element” in the administration of justice. Under Article 182 of the Constitution, the participation of citizens in the administration of justice is determined by statute. This means that functions ordinarily assigned to professional judges may, at least to some extent, be performed by natural persons who do not meet the aforementioned requirements (that is, who are not professional judges), but who satisfy criteria set out in statutory law, particularly regarding the categories of cases in which they may adjudicate and the manner of their selection to the office of judge or lay judge. Under current law, lay judges (*ławnicy*) fall within this category in the Polish legal system. In every instance, however—including the administration of justice by non-professional judges—the reference is invariably to “flesh-and-blood” human beings.<sup>22</sup> These considerations lead *prima facie* to the conclusion that, under the *de lege fundamentalis* framework, the administration of justice—and therefore the issuance of judicial decisions—may not be carried out by an entity that is not a natural person and does not possess the characteristics described above; that is, an entity lacking legal personality and full legal capacity, and one that has not been appointed through the appropriate procedure to serve as a professional or lay judge. Ergo, within the meaning of constitutional norms, there can be no conceptual category of a “robot judge” (*Roboter-Richter*).<sup>23</sup> This also informs the understanding of what constitutes an independent court (within the meaning of Article 45(1) of the Polish Constitution, which enshrines the constitutional and universally accepted principle of the fundamental right to a fair trial), since such a court may be constituted only by the participation of a person (or persons) who qualify as judges in the sense described above.

Consequently, the potential involvement of a “robot judge” in the administration of justice would require a change to constitutional provisions or,

<sup>21</sup> Decision of the German Constitutional Court of 22.08.2016, 2 BvR 2953/14, ‘Neue Zeitschrift für Verwaltungsrecht’ 2016, vol. 22, p. 1630; see also e.g. D. Nink, *Justiz...*, p. 31, 457.

<sup>22</sup> D. Nink, *Justiz...*, p. 260 et seq., 351, 354 et seq., 463.

<sup>23</sup> G. Rühl, *Einsatz...*, p. 281.

at the very least, a modification of the current paradigm, which, however, seems unlikely. Let us leave aside, for the sake of this narrow discussion, the question of whether such a potential constitutional amendment would constitute a breach of the fundamental principles underpinning the standards governing the administration of justice in civilised societies or would at the very least require a significant reinterpretation of fundamental rights. It is worth noting, however, that the above conclusions are not self-evident.

It would be necessary to explicitly determine that digital tools may participate in the exercise of judicial authority, at least where this involves issuing binding decisions in civil or criminal matters. Although the literature rightly notes that no systems capable of fully replacing a judge will emerge in the immediate future,<sup>24</sup> this issue will inevitably require resolution in the more distant—though not overly distant—future.

It is worth noting, however, that the inadmissibility of such participation by machines in the administration of justice is not universally regarded as self-evident. In German legal scholarship it has been argued that the autonomous norms of the Basic Law merely establish that “at least a judge who is a human being, in the form defined by the legislature, is compatible with the Constitution,” while adding that “this conclusion does not, however, imply that the introduction of legal generators to perform judicial functions would be impossible—it would merely require numerous amendments to ordinary legislation.”<sup>25</sup> According to Lydia Wolff, a change in the interpretation of the Basic Law is conceivable, and this may entail that the constitutional concept of a judge could itself be revised.<sup>26</sup> Similar conclusions can also be drawn under the provisions of the Polish Constitution of 1997. After all, no provision of the Constitution explicitly states that a judge must be a human being. Perhaps the question should not be “whether”, but “how” a “smart court” would operate; in particular, what oversight mechanisms should be put in place to ensure it functions properly and to enable the correctness of the rulings issued to be verified.

---

## The Right to a Robot Court?

---

Although this is not currently possible, it is worth considering whether, in the future, at least to some extent, the following tasks could be entrusted to artificial intelligence systems: possessing legal knowledge at an appropriate professional level, interpreting and applying the law, reconstructing the

<sup>24</sup> G. Rühl, *Einsatz...*, pp. 282 and 283.

<sup>25</sup> L. Wolff, *Algorithmen...*, p. 131.

<sup>26</sup> L. Wolff, *Algorithmen...*, p. 144.

meaning of vague legal texts and phrases or general legal concepts, as well as conducting complex dialogue with the parties to court proceedings.<sup>27</sup> All these activities constitute the process of legal reasoning within the administration of justice. Of course, a machine can easily be programmed today to possess full knowledge of the applicable legal provisions and, at least to some extent, to apply them to specific factual situations presented to it. It can therefore be considered, *prima facie*, that such a situation fulfils the Latin phrases *iura novit curia* and *da mihi factum, dabo tibi ius* almost perfectly.<sup>28</sup> One can then be certain at first glance—and, consequently, have the reassurance of a professional ruling—that, unlike a human judge, an AI “in a judge’s robe” will not overlook any provision that ought to be considered in the case and, where appropriate, applied.

However, the issue is far more complex. In this context several specific questions and problems arise that, even considering the constitutional right of access to a court, require answers.

Firstly, as already mentioned, the constitutional right of access to a court and a fair trial before it (Article 45[1] of the Constitution of the Republic of Poland, Article 47[2] of the Charter of Fundamental Rights of the European Union and Article 6 of the ECHR). The right to a fair trial means, in particular, that an individual cannot be merely an “object” of the proceedings, but must remain a subject of rights and be able to influence both the course of the proceedings and the outcome of the case.<sup>29</sup> In this context, the question arises as to whether an artificial intelligence system could replace a judge in adjudicating a case, given the principle of the publicity of proceedings and, in particular, the party’s right to be heard, including at a public hearing. Considering this essential component of the constitutional right to a court, it follows logically that the parties to the proceedings have the right to present a case before the court, to speak before the judge hearing the case,<sup>30</sup> and therefore a body comprising an official who understands the content presented to them and analyses, on an ongoing basis, all the circumstances relating to the information presented at the hearing. So the parties must be understood and feel understood.

The elements of the right to be heard—contained within Article 45(1) of the Constitution and flowing from the principle of the rule of law in judicial proceedings<sup>31</sup>—include: the right to be heard in court, the right to be informed

<sup>27</sup> Einsatz von KI..., p. 5.

<sup>28</sup> See also J. J. Vasel, *Künstliche...*, p. 186.

<sup>29</sup> Judgment of the German Constitutional Court of 29.05.1991, 1 BvR 1383/90, BVerfGE 84, 188 (190); 86, 133 (144 et seq.); Einsatz von KI..., p. 14.

<sup>30</sup> Einsatz von KI..., p. 13; D. Nink, *Justiz...*, p. 302 et seq.; see also e.g. Ch. Rollberg, *Algorithmen in der Justiz. Rechtsfragen zum Einsatz von Legal Tech im Zivilprozess, Baden-Baden 2020*, p. 138 et seq.

<sup>31</sup> The Constitutional Tribunal indicates that the essence of a fair (reliable) procedure may be reduced to: (1) the opportunity to be heard (the parties’ entitlement to present their position), (2) the clear disclosure of the reasons for the decision, in order to avoid arbitrariness or even capriciousness on the part of the authority, and (3) ensuring predictability for the participant in the proceedings through adequate coherence and internal

by the court of the proceedings taking place and their subject matter, and the right to expect the court to take into account, or at least consider, the arguments put forward when reaching a decision on the merits of the case.<sup>32</sup> The essence of this component of the constitutional guarantee is that—as the literature emphasizes—“the court hearing the case—meaning a human judge [menschlicher Richter]—must acquaint himself or herself with the factual and legal arguments presented and consider them in the decision.”<sup>33</sup> The right to hear the parties is tantamount to imposing an obligation on the judge “to grasp the essence of the arguments put forward.”<sup>34</sup>

The point is therefore not merely to create conditions under which an automated system would allow the parties to present their factual and legal positions “in the same way as a human judge in written proceedings must ensure for the parties.”<sup>35</sup>

The requirement to ensure the right to be heard obliges the court to “take on board” the information presented, and thus to rationally take note of the arguments put forward by the parties to the proceedings, to consider them comprehensively, and then to accept or reject them when delivering its judgment. This means that the judge must engage with the reasoning presented and subject it to appropriate analysis before deciding, rather than merely acting as its “carrier”, a mere passive recipient. In this sense, the judge remains a participant in the court proceedings, as his or her involvement—understood as interaction with the parties—is essential. This requirement must then be reflected in the substance of the decision itself—it is therefore assumed that the courts have taken note of the parties’ submissions and have given them due consideration.<sup>36</sup> The right to be heard, as an element of the constitutional right of access to a court, is likewise emphasized in the case law of the Polish Constitutional Tribunal.<sup>37</sup>

The necessity of respecting this aspect of the constitutional right to a court is undoubtedly a significant argument against the admissibility—within the framework of constitutional guarantees—of replacing the “human judge” with a machine. In this sense, one must concede that these legal scholars are right who observe that “judicial authority has been entrusted to judges,” meaning that “it is human beings who must make decisions that embody

---

logic of the mechanisms to which they are subjected (the so-called core of fair judicial procedure — see the judgment of the Constitutional Tribunal of 20 October 2010, P 37/09, OTK ZU no. 8/A/2010, item 79).

<sup>32</sup> Einsatz von KI..., p. 13 and the literature cited therein.

<sup>33</sup> G. Rühl, Einsatz..., p. 283.

<sup>34</sup> Decision of the German Federal Court of Justice (BGH) of 26.11.2019, VI ZR 84/18, NJW 2020, vol. 22, no. 6, p. 1594.; V. Vorwerk, *Algorithmen in zivilprozessualen Verfahren* [in:] *Prozessuales Denken und Künstliche Intelligenz*, Ch. Althammer, H. Roth (eds.), p. 73.

<sup>35</sup> A. K. Bernzen, *Roboter...*, p. 136.

<sup>36</sup> Decisions of the German Constitutional Court: BVerfGE 65, 293 [295]; BVerfGE 84, 188 [190]; BVerfGE 86, 133 [144].

<sup>37</sup> See e.g. Judgment of the Constitutional Tribunal of 02.10.2006, SK 34/06, OTK ZU 9/A/2006, item 118.

justice.”<sup>38</sup> This is not, however, due to a person’s physical characteristics, but rather to their ability to analyse information and reason logically, as well as—and this is by no means insignificant—their ability to verify their decision-making process. In the case of a machine, the latter may be significantly impeded, at the very least.<sup>39</sup> Accepting a different premise—and consequently adopting a system in which the judicial process of analysis is not carried out by a human being—would violate the constitutional right to a lawful judge (and “tribunal established by law”) as understood in this context, which forms an integral part of the broader right of access to a court.

Moreover, the irreplaceability of the “human judge” by an artificial-intelligence system is undoubtedly connected with the principle of judicial independence.

Judicial independence has several important aspects, a detailed analysis of which goes beyond the narrow scope of this presentation. In any case, it is not intended to serve the judge’s convenience, nor does it constitute a fundamental right of the judge; it has been established in civilized states as a “precondition for a lawful decision-making process.” Consequently, “a judge may not dispose of his or her independence at will.”<sup>40</sup> This, in turn, means that a judge must be independent at every stage of court proceedings. Consequently, it would be unacceptable for a judge to be bound by the results of an automated system, even if this were in accordance with the judge’s own wishes.<sup>41</sup>

In this respect, the interference entailed by replacing the judge or by obliging the judge to use an AI system as a binding mechanism for resolving cases would all the more readily be viewed as, at minimum, a constitutionally “unnecessary interference,” meaning an influence that is not essential to maintaining the functioning of the administration of justice.<sup>42</sup> The AI system would be designed to make the referee’s job easier and simplify the process, rather than benefit the parties themselves. This also relates to the inadmissibility of treating the machine as an “expert witness”, and consequently requiring the judge to rely on such assistance.

The premise that a judge may not be bound by the outcome of a machine’s reasoning follows from the fact that, under the current constitutional framework, full responsibility for and the act of adjudicating a case lie exclusively with the adjudicating panel (that is, with the human judge).<sup>43</sup> It is, admittedly, possible to imagine a situation in which such a decision-making process would be entrusted, at least to some extent, to an AI system and would not be subject to independent assessment and verification by a human judge. How-

<sup>38</sup> V. Vorwerk, *Algorithmen...*, p. 72.

<sup>39</sup> For further details, see, for example, D. Nink, *Justiz...*, p. 334 et seq.

<sup>40</sup> A. K. Bernzen, *Roboter...*, p. 137.

<sup>41</sup> *Ibid.*

<sup>42</sup> D. Nink, *Justiz...*, p. 278 et seq.; *Einsatz von KI...*, p. 9.

<sup>43</sup> See: D. Nink, *Justiz...*, p. 301 et seq.; *Einsatz von KI...*, p. 7.

ever, an independent judge is a judge who is autonomous in making decisions and handing down judgements. Independence must be comprehensive; it cannot be “divided” between the judge and another participant (actor) who, apart from the judge, would have a “separate” part in the decision-making process. The constitutional requirement and, at the same time, the competence of the courts to deliver judgments means that they alone have exclusive authority over this sphere of public power. This authority is therefore indivisible in the sense that no part of it may be transferred to any other institution or entity (see Article 174 of the Constitution of the Republic of Poland)<sup>44</sup>.

Furthermore, it is apt to observe that a judge’s subordination to the results produced by an automated system would weaken the perception of the judge as a neutral decision-maker.<sup>45</sup> In this case, however, what is at stake is the public’s trust in the justice system.<sup>46</sup> This fear persists even when one sets aside one’s personal apprehensions about being “served” by a machine, which stem from personal experiences (“algorithmic aversion”<sup>47</sup>). One can, in fact, place one’s trust in a specific body whose decision is its own and is not influenced by anyone or anything else over which the decision-making body has no control.

It is worth noting, in passing, that in this context there arises the risk of the so-called default effect, whereby a judge tends to “accept the system’s results because departing from them would require considerable effort to justify and would entail the risk of an incorrect decision.”<sup>48</sup>

Lydia Wolff additionally points to a significant danger, emphasizing that “the problem of complex algorithmic value-based decisions in the context of guarantees of independence lies (...) not so much in the fact that protecting independence—understood as avoiding external influence—would become fundamentally impossible, but rather in the fact that vesting independent decision-making authority in the form of an algorithm (given the predictability of all outcomes that follow from it) would lead to unlimited interpretive power over social values (...)”<sup>49</sup> This leads the author to the well-founded conclusion that—taking into account the principle of the separation of powers—the interpretation of complex value-laden issues must remain in the hands of independent human decision-makers in individual cases.<sup>50</sup> It should

<sup>44</sup> This should be understood, in particular, as the judiciary’s monopoly on the resolution of disputes, understood as the administration of justice within the meaning given to that concept by the Constitution. On the subject of the judiciary’s monopoly, see, for example, D. Nink, *Justiz...*, p. 261, 272 *et seq.*; P. Wysk [in:] *Das Staatsrecht des Bundesrepublik Deutschland*. Band II, K. Stern, H. Sodan, M. Mösl (eds.), München 2022, p. 726; G. Morgenthaler [in:] *Grundgesetz, Beck’sche online Kommentar*, V. Epping, Ch. Hillgruber (eds.), 60, Ed. 28.12.2024, Art. 92, no. 32.

<sup>45</sup> A. K. Bernzen, *Roboter...*, *ibid.*

<sup>46</sup> *Ibid.*

<sup>47</sup> J. J. Vassel, *Künstliche...*, p. 184.

<sup>48</sup> A. K. Bernzen, *Roboter...*, *ibid.*

<sup>49</sup> L. Wolff, *Algorithmen...*, pp. 162 and 163.

<sup>50</sup> L. Wolff, *Algorithmen...*, p. 163.

therefore be accepted that this danger outweighs the perceived benefits of using AI systems, namely enhanced predictability of decisions grounded in a universal, uniform interpretation of the law, and thus most likely greater predictability in its application.

Another advantage of using AI systems is that they probably may offer a stronger guarantee of neutrality and impartiality than a human being. “Robots in robes” would be devoid of emotion and thus, in a sense, more steadfast and continuously available to the parties as the “mouth of the law” (*bouches de la loi*)<sup>51</sup> in accordance with Montesquieu’s ideal of the judge.<sup>52</sup> This is because—as we assume—an algorithm operates without emotion, without positive or negative attitudes arising from attachment to particular views or perceptions of the parties, without personal desires or prejudices as such, and without any internal interests in any form; consequently, it cannot possess biases or self-interested expectations regarding the outcome of the proceedings.<sup>53</sup> This assumption is tempting, all the more so when we realise that the decisions of human judges are also shaped by diverse and often even elusive or irrational motivations or stimuli, their current state of mind, medical conditions, stress, emotions, upbringing, prejudices, errors in reasoning, and shortcomings or weaknesses in rationality.<sup>54</sup> It is also difficult to imagine a way of influencing a robot that would, for example—as Lydia Wolff points out—involve someone with a stake in the outcome of a case “just before the trial begins, in secluded corners, reminding the algorithms of past friendships, threatening them with social decline, or winning their favour through a particularly charming demeanour”.<sup>55</sup> It would be all the more difficult to offer artificial intelligence a financial benefit in exchange for a favourable ruling. Algorithms do not get headaches.

This undeniable advantage of AI systems, however, is far from certain. Even today, we can observe a wide range of behaviours in AI systems that are far from being free of bias or what we might describe as ‘technological neutrality.’ So-called discriminatory tendencies have been observed in the operation of algorithms.<sup>56</sup> At the present stage of technological development, there can be no guarantee that artificial-intelligence systems, as “generators of law”, would be free from dependence on human selection of decision-making determinants and thus immune to subjectivity to a degree at least comparable

<sup>51</sup> Charles-Louis de Secondat Montesquieu, *De l’esprit des lois*. Livre XI, chap. 6, 1748: ‘Les juges de la Nation ne sont que la bouche qui prononce les paroles de la loi, des êtres inanimés qui ne peuvent ni en adressé la force ni la rigueur’.

<sup>52</sup> See A. Schunder-Hartung, *Roboter-Richter: Stehen wir vor einer Rückkehr der “Bouches de la loi”?*, ‘Betriebs-Berater’ 2021, vol. 49, p. 1; D. Nink, *Justiz...*, p. 39; J. J. Vasel, *Künstliche...*, p. 186.

<sup>53</sup> L. Wolff, *Algorithmen...*, p. 156.

<sup>54</sup> D. Nink, *Justiz...*, p. 28.

<sup>55</sup> L. Wolff, *Algorithmen...*, p. 165.

<sup>56</sup> J. J. Vasel, *Künstliche...*, p. 183.

to that found in human judges (most notably as a result of supplying a manipulated dataset for AI training).<sup>57</sup> Of course, one cannot assume that manipulation has taken place from the outset, but one cannot rule it out either. Artificial intelligence systems that are externally controlled and susceptible to manipulation would be particularly dangerous in the hands of authoritarian or totalitarian regimes. The more advanced the technology becomes, the more difficult it will be to objectively verify the correctness and ‘neutrality’ of its operation. What is more, there is not only the danger associated with so-called “algorithmic hallucinations” and the lack of any guarantee that data will be processed correctly, but the risk of “algorithmic bias” linked to programmed prejudices is already being recognised today.

This is where the vital role of control becomes apparent and the ability to trace the system’s “reasoning” process becomes apparent—something that already appears, for technical reasons, to be significantly impeded. David Nink rightly points out, with regard to assisting the judge’s decision-making, that “a structured and comprehensible statement of reasons must ensure that the (overall) judicial decision is understandable at all times to those concerned: if the judge incorporates an algorithmically determined forecast into their decision, the criterion of comprehensibility also encompasses the fundamental mode of operation of the system”.<sup>58</sup>

All this also affects the potential need to modify the interpretation of the concept of “independence” as applied to artificial-intelligence systems that might exercise judicial authority, shifting the focus toward considerations concerning the scope and security of collected datasets. Lydia Wolff rightly emphasizes that “the protection of independence in the digital context should therefore also serve to secure status. Exclusion from activity must, just like activation, be a matter for the legislature alone”. In her view, “as with any interconnected system, even the smallest gaps could become the starting point for a general system collapse. Such a collapse would have far more serious consequences than the failure of an individual judge and could be more difficult to overcome. If there were no backup mechanisms to handle the sudden caseload manually, this could lead to a complete lack of legal protection. Safeguarding status should therefore carry immense importance.”<sup>59</sup>

There are no guarantees that AI systems could not malfunction due to a technological defect—something that, given their advanced development, would be exceedingly difficult to detect and remedy.

---

<sup>57</sup> See L. Wolff, *Algorithmen...*, p. 165.

<sup>58</sup> D. Nink, *Justiz...*, p. 465.

<sup>59</sup> L. Wolff, *Algorithmen...*, p. 172.

---

## Mechanism of Control

---

Assuming the paradigm of the human judge, and in light of the constitutional judicial monopoly,<sup>60</sup> understood as arising from the existing structural guarantees of the Constitution (see Articles 174 and 175 in conjunction with Article 178 of the Constitution of the Republic of Poland), an artificial-intelligence system certainly may not issue autonomous decisions that would not be subject to review by a judge. Among the essential features of judicial authority are the determination, establishment, and declaration of what is lawful.<sup>61</sup> This excludes the use of algorithms to render final, substantive decisions on the merits of a case.<sup>62</sup>

This is also substantively justified, because the exercise of judicial authority in this respect requires skills that, at least at the current stage of technological development, a mechanically operating machine does not possess.<sup>63</sup> The fact that, in the case of self-learning systems—similarly to a judge’s reasoning—it is not possible to reproduce the reasoning process precisely and unambiguously, step by step, as Anna K. Bernzen quite rightly observes, does not alter this conclusion (and thus does not per se exclude the possibility of automating elements of judicial activity).<sup>64</sup> The author emphasises that “even today, the parties can neither read the mind of a judge sitting alone nor have any insight into the deliberations of a panel of judges”.<sup>65</sup>

What remains an open question is whether, and to what extent, a party or participant in the proceedings has the right to be informed that, in each respect, the judge relied on artificial-intelligence tools when making the decision. This matter is not currently regulated by law.

None of this means, however, that an AI system cannot serve as a tool assisting the judge in seeking the correct resolution, or as a mechanism enabling a kind of preliminary determination that could then be subjected to review at the request of a party (for example, in the case of an electronic order for payment<sup>66</sup>). An artificial-intelligence system may—and should—be treated as an instrument supporting the human judge’s work, one that nonetheless requires a critical approach and verification on the part of the judge.<sup>67</sup> For instance, Estonia is planning to introduce a “robot judge” for cases valued at up to 7,000 euros, while providing for the possibility of judicial review of the

---

<sup>60</sup> See e.g. . P. Wysł [in:] *Das Staatsrecht...*, p. 726; D. Nink, *Justiz...*, p. 261, 272 et seq., 463; G. Morgenthaler [in:] *Grundgesetz...*, Art. 92, no. 32; F. Schoch, J.-P. Schneider, E. Schmidt-Aßmann, W. Schenk (eds.), *Verwaltungsgerichtsordnung*, 46. EL August 2024, Einleitung, no. 34–36.

<sup>61</sup> Decision of the German Constitutional Court of 28.11.1957, 2 BvL 11/56, NJW 1958, vol. 3, p. 97.

<sup>62</sup> *Einsatz von KI...*, p. 6.

<sup>63</sup> *Ibid.*

<sup>64</sup> A. K. Bernzen, *Roboter...*, p. 135.

<sup>65</sup> *Ibid.*

<sup>66</sup> G. Rühl, *Einsatz...*, p. 281.

<sup>67</sup> See G. Rühl, *Einsatz...*, p. 285.

machine's decision upon a party's request.<sup>68</sup> In this context, it is worth considering whether, at least in the foreseeable future of technological development, "smart courts" might not be treated in a similar way to today's arbitration tribunals, where the submission of a dispute for adjudication—as an alternative to state courts composed of professional human judges—depends on the consent of the parties concerned.

---

## Conclusion

---

In summary, under the constitutional norms currently in force, the use of "autonomous systems" or, more broadly, "artificial intelligence" to render substantive judicial decisions is precluded by the principles and values enshrined in the Constitution.<sup>69</sup> Certainly, replacing judges with machines would not be possible today either, given that AI lacks the ability to "understand its surroundings", to filter information, and to grasp the context of the cases to be adjudicated, each of which usually has its own specific elements.<sup>70</sup> David Nink aptly points out that "a fully automated judiciary—and thus, by definition, one devoid of individual character—can only be regarded as a dystopia".<sup>71</sup>

It is not, however, excluded that such a system—without performing the functions of judicial authority—could carry out certain auxiliary tasks, sometimes complex and difficult for a human to perform (such as the analysis and registration of metadata).<sup>72</sup> Nor can it be ruled out that a machine could operate as a tool preparing a judicial resolution<sup>73</sup> in the realm of the analysis of norms and case law, or even formulate preliminary recommendations based on processed data.

In all cases, however, it must be acknowledged that an artificial intelligence system would, without question, remain subject to judicial oversight, functioning solely as a tool to assist the judge in resolving cases. Consequently, it must still be the human judge who remains the decision-maker in the administration of justice and who, within the applicable legal order, bears full responsibility for the consequences of the decision.

---

<sup>68</sup> G. Rühl, *Einsatz...*, p. 281, 283; *Einsatz von KI...*, p. 37 and 38; E. Nittler, *Can AI Be a Fair Judge in Court? Estonia Thinks So*, 'Wired' 2019, 25.03; <https://www.wired.com/story/can-ai-be-fair-judge-court-estonia-thinks-so/>.

<sup>69</sup> V. Vorwerk, *Algorithmen...*, p. 73; D. Nink, *Justiz...*, p. 354.

<sup>70</sup> D. Nink, *Justiz...*, p. 356.

<sup>71</sup> D. Nink, *Justiz...*, p. 357.

<sup>72</sup> *Einsatz von KI...*, p. 7; D. Nink, *Justiz...*, p. 355.

<sup>73</sup> *Einsatz von KI...*, p. 7.

Less doubts arise regarding the possibility of artificial-intelligence systems influencing judicial activity beyond the essence of adjudicatory functions—activity that is therefore subject to administrative supervision insofar as it concerns ensuring the proper conduct of proceedings.<sup>74</sup> *Prima vista*, such influence should be regarded as permissible, as it would not constitute an impermissible intrusion into the sphere protected by judicial independence. It appears that the use of artificial-intelligence systems as auxiliary instruments in judicial work is not hindered by the phenomenon already observed in practice, namely the so-called “automation bias,”<sup>75</sup> that is, an unwarranted trust in computer-generated outcomes. A machine’s error can, after all, be likened to the inevitable errors committed by judges in adjudicatory practice (the so-called “propensity for error”).<sup>76</sup> It is, however, essential—indeed obvious—those appropriate safeguards be put in place to ensure the correct functioning of AI tools, including safeguards preventing the identification of personal data of the parties, the status of the cases and the various incidents described in the files, in particular by prohibiting their unjustified processing beyond the scope of tasks necessitated by adjudication of the case.

The observation contained in Recital 61 of the Artificial Intelligence Act<sup>77</sup> is therefore well founded in the present circumstances, namely that “the use of AI tools can support the decision-making power of judges or judicial independence but should not replace it: the final decision-making must remain a human-driven activity.” This is, however, only the first step on the path toward the cooperation of artificial-intelligence systems in the administration of justice. The future is bound to take us by surprise anyway.

### Abstrakt

„Antropocentryczny” system prawny nie zakłada dziś możliwości rozstrzygnięcia spraw sądowych przez systemy sztucznej inteligencji. Jest jednak niemal oczywiste, że w niedalekiej przyszłości coraz bardziej zaawansowane narzędzia technologiczne będzie się włączać w sferę wymiaru sprawiedliwości w coraz szerszym zakresie. W tym kontekście pojawiają się już dziś fundamentalne pytania natury ustrojowej. Po pierwsze: należy postawić pytanie o to, czy ustrojowe normy konstytucyjne wymagają – a jeśli tak, to w jakim zakresie – udziału człowieka w rozstrzygnięciu spraw z zakresu władzy sądowniczej.

<sup>74</sup> Einsatz von KI..., p. 10.

<sup>75</sup> G. Rühl, Einsatz..., s. 282; Einsatz von KI..., p. 11; A. K. Bernzen, Roboter..., p. 137; J. J. Vasel, Künstliche..., p. 183.

<sup>76</sup> G. Rühl, Einsatz..., *ibid.*; Einsatz von KI..., *ibid.*; A. K. Bernzen, Roboter..., *ibid.*

<sup>77</sup> Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance), OJ L, 2024/1689, 12.07.2024.

Po drugie zaś: fundamentalne znaczenie dla oceny przyznania sztucznej inteligencji możliwości wykonywania przynajmniej pewnych zadań z zakresu wymiaru sprawiedliwości ma odpowiedź na pytanie, czy byłoby to dopuszczalne w świetle konstytucyjnego (a także prawnomiędzynarodowego) standardu prawa do rzetelnego procesu – a zatem do sądu należycie obsadzonego, ustanowionego ustawą – a wreszcie wobec prawa strony do bycia wysłuchaną w toku postępowania sądowego. Celem artykułu jest przedstawienie podstawowych problemów związanych z tymi zagadnieniami na tle obecnych standardów konstytucyjnych.

**Słowa kluczowe:** robot, sztuczna inteligencja (AI), prawo do sądu, sąd, wymiar sprawiedliwości.

## BIBLIOGRAPHY

Beck-online.Grosskommentar BGB GesamtHrsg: Gsell/Krüger/Lorenz/Reymann (eds.): Spickhoff, 01.02.2025.

Bernzen A. K., *Roboter als Richter? Zur Automatisierung der Rechtsprechung*, 'Recht Digital' 2023, vol. 3.

Classen C. D., *Gesetzesvorbehalt und Dritte Gewalt*, 'Juristenzeitung' 2003, vol. 14.

Das Staatsrecht des Bundesrepublik Deutschland. Band II, K. Stern, H. Sodan, M. Mösl (eds.), München 2022.

Demokratie und Grundgesetz. Eine Auseinandersetzung mit der verfassungsgerichtlichen Rechtsprechung, T. Blanke (ed.), Baden-Baden 2000.

Einsatz von KI und algorithmischen Systemen in der Justiz. Grundlagenpapier zur 74. Jahrestagung der Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs vom 23. bis 25. Mai 2022 in Rostock.

Festschrift für Eduard Bötticher zum 70. Geburtstag, Bettermann K. A., Zeuner A. (eds.), Berlin 1969.

Grundgesetz, Beck'sche online Kommentar, V. Epping, Ch. Hillgruber (eds.), 60, Ed. 28.12.2024.

Handbuch des Staatsrechts, Band II, Verfassungsstaat, J. Isensee, P. Kirchhof (eds.), Heidelberg 2004.

Künstliche Intelligenz. Wie gelingt eine vertrauenswürdige Verwendung in Deutschland und Europa?, Bundesministerium für Umwelt, Naturschutz, nukleare Sicherheit und Verbraucherschutz, F. Rostalski (ed.), Tübingen 2022.

Nink D., *Justiz und Algorithmen. Über die Schwächen menschlicher Entscheidungsfindung und die Möglichkeiten neuer Technologien in der Rechtsprechung*, Berlin 2021.

Nittler E., *Can AI Be a Fair Judge in Court? Estonia Thinks So*, 'Wired' 2019, 25.03.; <https://www.wired.com/story/can-ai-be-fair-judge-court-estonia-thinks-so/>.

Pfeil W., *Der Mensch steht höher als Technik und Maschine – Benötigen wir ein Grundrecht zum Schutz vor Künstlicher Intelligenz?*, 'International and Comparative Law Review' 2020, vol. 2.

Prozessuales Denken und Künstliche Intelligenz, Ch. Althammer, H. Roth (eds.), Tübingen 2023.

Rollberg Ch., *Algorithmen in der Justiz. Rechtsfragen zum Einsatz von Legal Tech im Zivilprozess*, Baden-Baden 2020.

Shi Ch., Sourdin T., Li B., *The Smart Court – A New Pathway to Justice in China?*, 'International Journal for Court Administration' 2021, vol. 12, 1.

Schoch F., Schneider J.-P., Schmidt-Aßmann E., Schenk W. (eds.), *Verwaltungsgerichtsordnung*, 46. EL August 2024.

Schunder-Hartung A., *Roboter-Richter: Stehen wir vor einer Rückkehr der "Bouches de la loi"?*, 'Betriebs-Berater' 2021, vol. 49.

Tang Z., *Smart Courts in Cross-Border Litigation*, 'Rabels Zeitschrift für ausländisches und internationales Privatrecht' 2023, vol. 87.

Vasel J. J., *Künstliche Intelligenz in der Justiz*, 'Legal Tech Zeitschrift' 2023, vol. 3.

Voßkuhle A., Sydow G., *Die demokratische Legitimation des Richters*, 'Juristenzeitung' 2002, vol. 14.

Wolff L., *Algorithmen als Richter. Verfassungsrechtliche Grenzen entscheidungstreffender Rechtsgeneratoren in der Rechtsprechung*, Trier 2022.

Yupeng S., *History and Prospects of Creation of "Smart Courts" in the PRC: Problems of Application of the Artificial Intelligence Technology*, 'History of State and Law' 2023, May, 11.



Inesa Stolper\*

## Judicial Impartiality and Artificial Intelligence from the Perspective of Article 6 of ECHR

### [Bezstronność sądu a sztuczna inteligencja z punktu widzenia art. 6 EKPC]

#### Abstract

Recent developments in artificial intelligence (AI) have introduced systems that analyze legal information, identify correlations across documents, and support decision-making. New possibilities for automation present opportunities and challenges for the administration of justice. These possibilities raise complex questions about the compatibility of the deployment of AI in courts with fair trial guarantees under the European Convention on Human Rights (ECHR).

This article examines whether and under what conditions the use of AI systems in courts may create legitimate doubts as to a court's impartiality within the meaning of Article 6 (1) of the ECHR. Drawing on the case law of the European Court of Human Rights, a distinction was made between the subjective and objective elements of impartiality, and an analysis was carried out of how the use of AI in the decision-making process might affect each of them.

The paper argues that the use of AI in courts is a question of confidence in the justice system, and perceived impartiality may stem from limited transparency about the role of AI in decision-making, insufficient oversight, and excessive reliance on AI outputs.

**Keywords:** artificial intelligence, impartiality, legitimate doubts, right to a fair trial.

## Introduction

The use of digital technologies in European courts and tribunals is not a new phenomenon. In previous decades, the courts began using computers, legal databases, videoconferencing systems, and various digital tools to sup-

\* **Inesa Stolper** – PhD candidate at Mykolas Romeris University, Institute of Public Law (affiliation), Vilnius, Lithuania; <https://orcid.org/0009-0002-1997-7404>; [inesa.stolper@ehu.lt](mailto:inesa.stolper@ehu.lt) / doktorantka na Uniwersytecie Michała Römera, Instytut Prawa Publicznego (afiliacja), Wilno, Litwa.

port the administration of justice and the work of judges. However, earlier stages of court digitalization<sup>1</sup> did not significantly affect the impartiality of the court because these tools did not directly impact the decision-making process.

The proliferation of artificial intelligence (*hereinafter* “AI”) systems offers a variety of possibilities to automate court activities, including assistance in the decision-making process by analyzing information, determining conflicting points in pleadings, and pointing to case-law, and drafting the decision. The automation of the decision-making can span from a research assistant role to full automation.<sup>2</sup> It is unlikely that autonomous<sup>3</sup> AI systems will be introduced in European courts and tribunals, as their use may be incompatible with Article 22 of the General Data Protection Regulation (*hereinafter* “GDPR”) and Article 14 of the EU Artificial Intelligence Act (*hereinafter* “AIA”). However, the widespread use of AI systems in a supporting role may only be a matter of time.

The courts in a democratic society must inspire public confidence.<sup>4</sup> An independent and impartial court is the core of the right to a fair trial and constitutes a component of the principle of the rule of law. The use of an AI system, even in an assisting role, should not interfere with fair trial rights guaranteed by the European Convention of Human Rights (*hereinafter* “ECHR”).

The risks associated with the use of AI for impartial and independent decision-making were acknowledged for quite some time. In 2018, the European Commission for the Efficiency of Justice (*hereinafter* CEPEJ) after conducting a study on the impact of AI systems on judiciary, adopted the first-ever European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environments, which included five key principles for the use of AI, among which, the principle of transparency, impartiality, and fairness.

The right to an impartial court denotes the absence of bias and prejudice.<sup>5</sup> According to the case-law of ECtHR, impartiality can be tested in various ways, but primarily a distinction should be made between its subjective and objective elements.<sup>6</sup> These aspects are functionally interconnected.<sup>7</sup> As courts move forward to the development, integration, and use of AI systems, it becomes necessary to assess whether a judge’s reliance on such tools could give

<sup>1</sup> For example, automation of registers, electronic filing systems, and remote hearings.

<sup>2</sup> T. Sourdin, *Judges, Technology and Artificial Intelligence: The Artificial Judge*, Elgar Law, Technologies and Society, Cheltenham 2021.

<sup>3</sup> By “autonomous” it is understood an AI system that is capable of making judgments without human supervision.

<sup>4</sup> Judgment of the ECtHR of 15 October 2009, *Micallef v. Malta*, application no. 17056/06, § 98.

<sup>5</sup> Judgment of the ECtHR of 1 October 1982, *Piersack v. Belgium*, application no. 8692/79, § 30.

<sup>6</sup> *Ibid.*, § 30; Judgment of the ECtHR of 12 January 2016, *Miracle Europe KFT v. Hungary*, application no. 57774/13, § 54.

<sup>7</sup> Opinion of Advocate General of 11 May 2006, *Graham J. Wilson v. Ordre des avocats du barreau de Luxembourg*, C-506/04, ECLI:EU:C:2006:31, § 75.

rise to legitimate doubts about the subjective or objective element of judicial impartiality. This article explores whether, and under which conditions, the use of AI may raise such concerns for the purposes of the meaning of Article 6 (1) of the ECHR. The article consists of three parts, starting with addressing the reasons and forms of AI integration in courts, moving to an analysis of each element of impartiality in turn.

---

## The Rationale and Functional Models of AI in Courts

---

The increasing interest in AI in the justice system relates to issues of organisation and administration of justice. As courts confront expanding caseloads, procedural complexity, and growing necessity for efficiency, the introduction of AI technologies is increasingly viewed as a solution. The rationale for the integration of AI systems in courts is multifaceted and includes such reasons as pursuing institutional objectives and societal changes.

There are internal and external factors that can influence the consideration of deploying an AI system in courts. Internally, such initiatives may be driven by efforts to digitalize courts, reduce the duration of proceedings and workload of judges, lower costs, and improve the quality and consistency of judgments. Externally, the pressure to integrate AI systems may come from legal professionals who already use various AI tools to prepare pleadings. This assumption could be compared and supported by an example of the development and widespread use of legal databases. When legal databases were created and became commonly used, they changed the format, length, and structure of pleadings. The structure of pleadings lost “the hierarchical order” of presentation of the arguments and became less principle-oriented and more case-based.<sup>8</sup> Consequently, judges had to invest additional time and effort to engage with case-law cited by parties who relied on these databases. Similarly, large language AI models (*hereinafter* “LLM”) can produce very long and detailed texts that could serve as the basis for pleadings submitted to courts.<sup>9</sup> Lengthy submissions require judges to devote considerable time to analyzing and extracting the essential information from case files. In this respect, AI systems could offer a supportive role to identify and summarize evidence and law, and even draft a judgment.

---

<sup>8</sup> European Commission for the Efficiency of Justice (CEPEJ), Guidelines on how to drive change towards Cyberjustice, CEPEJ-GT-QUAL, 7 December 2016, § 44.

<sup>9</sup> A large language model (*hereinafter* “LLM”) is a “type of artificial intelligence model that has been trained through deep learning algorithms to recognise, generate, translate, and/or summarise vast quantities of written human language and text”. European Commission, Knowledge Center on Translation and Interpretation, available online at <https://knowledge-centre-translation-interpretation.ec.europa.eu/en/news/what-large-language-model>.

Logic-based and machine learning AI systems could be deployed in the court to support decision-making in the case. Logic-based system is an AI system that consists of an algorithm, source code, and a knowledge base. It is designed and programmed on the basis of the rules of logic. It produces the output by finding correlations between different elements following the programmed algorithm.<sup>10</sup> The particular feature of the logic-based system is that the algorithm, source code, and data can be verified.

Recently, the greatest interest has been devoted to the second type of AI systems – machine learning AI systems. Unlike the previous generation of logic-based systems, machine learning systems have the ability to process large volumes of information, make inferences from data, learn, and generate recommendations and decisions.<sup>11</sup>

The machine-learning system consists of a neural network (algorithm) and data. Machine learning system is categorized as “black box” or “opaque systems” in the sense that it is difficult to construct a step-by-step pattern or logic of the decision-making process of such systems. The cause of this difficulty lies in the architecture and design of the system. It processes information against billions of different parameters that make the process of determining the decision-making path complicated.<sup>12</sup> In machine learning systems, differently from logic-based systems, the approach to understanding the decision-making process shifts to analyzing the data it is trained on and testing its output.

According to Article 3 of the EU AIA, read in conjunction with Annex III (8a), both logic-based and machine-learning systems are considered to be high-risk AI systems if they are being used to assist the court in “researching and interpreting facts and the law and in applying the law to a concrete set of facts.”<sup>13</sup> If the AI system is developed and integrated for such purposes, the court and the state body responsible for the integration and management of such tools have to implement guarantees and comply with obligations under Chapter III of the EU AIA. Most of these obligations ensure that an AI system functions properly and is free of errors, and developers and deployers of AI systems put in place quality,<sup>14</sup> risk management,<sup>15</sup> and cybersecurity meas-

<sup>10</sup> G. Sartor, F. Galli, *AI Approaches to Predictive Justice: A Critical Assessment*, ‘Humanities and Rights Global Network Journal’ 2023, vol. 5, 2; J. V. Fernandez, *Artificial Intelligence in Government: Risks and Challenges of Algorithmic Governance in the Administrative State*, 30 *Ind. J. Global Legal Stud.* 2023, 65, p. 75.

<sup>11</sup> Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (*hereinafter*, AIA), 12 June 2024, OJ L, 2024/1689, Article 3.

<sup>12</sup> A. Christian, L. Biedebach, A. Küpfer, and M. Neunhoeffer, *The Role of Hyperparameters in Machine Learning Models and How to Tune Them*, ‘Political Science Research and Methods’ 2024, 12, 4, pp. 841–848.

<sup>13</sup> AIA, Annex III (8a).

<sup>14</sup> AIA, Article 17.

<sup>15</sup> AIA, Article 9.

ures.<sup>16</sup> Some of the obligations, such as fundamental rights impact assessment<sup>17</sup> and effective human oversight,<sup>18</sup> can also strengthen the independence and impartiality of the judicial system in relation to impact of AI on the decision-making process.

An AI system can be used as an assistant in two distinct roles. It can act as a “court officer,” ensuring the procedural efficiency of the court, similar to a court clerk or a registrar’s administrative assistant. In this role, the AI system can support the justice system, for example, in document management, legal research, fact-finding, and drafting simple decisions.

The legal system and procedural rules of a particular country may define the role and functions of court administrative personnel. There is also some research on the role of the court registry and its impact on the work of international courts.<sup>19</sup> For example, at the ECtHR, registry lawyers “enjoy considerable discretion in how they conduct their activities.”<sup>20</sup> Their general instructions outline their work, which focuses on facilitating case processing. Their main role is to ensure continuity and coherence of case law.<sup>21</sup> The AI system can take over part of this role in the court.

An AI system can also act as an “expert” that is capable of providing a substantive evaluation of a case and delivering an informed opinion, similar to the functions of the “Advocate General,” a body known in some European legal systems and also employed by the Court of Justice of the European Union (*hereinafter* “CJEU”). The Advocate General at the CJEU offers reasoned opinions that are not binding but carry persuasive authority. Similarly, the AI system can be used to provide substantive legal analysis, evaluation, and recommendations on the decision that could inform the judge without it being a formal decision-maker.

The analogy with court officers and experts describes two distinct functional models of AI in court – one is instrumental and the other is advisory. The comparison illustrates the functional position an AI system may occupy in courts. Fewer issues may arise in connection with using the AI system as a “court officer” because, in this role, the AI system is intended to perform specific procedural tasks. The AI system that is designed to complete a series of tasks such as evaluate evidence, research, and find relevant provisions of law and case-law, and apply the law to concrete facts to generate a reasoned opinion on the case may raise concerns about the extent to which it can influence the outcome of the case. However, to determine whether the use of an

<sup>16</sup> AIA, Article 15.

<sup>17</sup> AIA, Article 27.

<sup>18</sup> AIA, Article 14

<sup>19</sup> C. D. Creamer, Z. Godzimirska, *Trust in the Court: The Role of the Registry of the European Court of Human Rights*, ‘The European Journal of International Law’ 2019, vol. 30, 2, Oxford University Press.

<sup>20</sup> C. D. Creamer, Z. Godzimirska, *Trust...*, p. 677.

<sup>21</sup> *Ibid.*

AI system can have an impact on the decision-making of the judge or create legitimate doubts as to the impartiality of the court, it is important to examine the two elements of impartiality established in the case-law of the ECtHR and used by the Court as criteria for analyzing complaints on Article 6(1) of the ECHR.

---

## **Subjective Element of Impartiality and AI Bias**

---

Subjective impartiality is linked to personal convictions regarding the subject matter of the legal issue or persons involved.<sup>22</sup> The subjective element requires determining and collecting evidence about the personal views of a judge in a particular case.<sup>23</sup> AI systems are perceived as technology that does not have emotions or personal preferences. However, research has documented that AI can produce biased recommendations or decisions.<sup>24</sup> The biased outcomes of the AI system have their origin in training data or may emerge during the training process.<sup>25</sup>

An AI system that is used to perform smaller tasks, such as searching for law or drafting simple procedural decisions in the case, is unlikely to have a significant negative impact because, in this role, it is assigned to do a fragment of the work on the case. It is for the judge to decide the case. However, a system that can provide a comprehensive assessment of evidence, research law, apply law to concrete facts, and generate an opinion about the substance of the case may influence the outcomes of the case if the judge decides to incorporate the generated suggestions into the text of the judgment. If the AI system generates biased recommendations, it can potentially impact the subjective element of impartiality.

The judge may use a public AI system or an AI system specifically designed, developed, and integrated in courts. There is a difference between these two models, especially as regards the quality of the output. Public AI systems, such as ChatGPT, Claude, or Mistral AI, are trained on a vast amount of information from the internet. Their use as assistants in courts is possible but limited. At the same time, the availability of public AI systems to anyone, including judges, makes it easy to create an account and use it for professional activities.

---

<sup>22</sup> Micallef v. Malta, § 93; Judgment of the ECtHR of 21 December 2000, *Wettstein v. Switzerland*, application no. 33958/96, § 43.

<sup>23</sup> Micallef v. Malta, § 95; Judgment of the ECtHR of 23 April 2015, *Morice v. France*, application no. 29369/10, § 74.

<sup>24</sup> E. Ferrara, *Fairness and Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, and Mitigation Strategies*, 'Sci' 2024, 6, 1, 3.

<sup>25</sup> K. Mavrogiorgos, A. Kiourtis, A. Mavrogiorgou et al., *Bias in Machine Learning: A Literature Review*, 'Applied Sciences' 2024, 14, 19, p. 8860.

Public AI systems are designed to perform a range of different tasks and can make mistakes. For example, an AI system can analyze the case based on the legislation of another country and generate an incorrect assessment of the case.<sup>26</sup> Moreover, the use of a public AI system for working with documents that contain personal data and private information may be contrary to the EU's personal data protection rules.<sup>27</sup>

Courts can integrate an AI system that was designed to be operated in a concrete legal system. Creating such a system involves further training of the general-purpose AI system so it should be able to do specific tasks.<sup>28</sup> Legislation and case-law of a specific legal system should be uploaded or added to the AI training data.

There is a general presumption that legislation and previous case-law of courts are neutral and free from bias. Hence, one might conclude that the recommendations and decisions generated by an AI system are also neutral and objective. There is indeed less risk that the system can be biased if the data itself is carefully selected and does not contain personal opinions, but this assumption merits further scrutiny.

---

### Bias arising from the training data

---

Data is an essential element of an AI system because it defines and shapes the recommendations and decisions made. It has been recognized that the risks of bias and discrimination arise from the data AI system is trained on.<sup>29</sup> For this reason, the EU AIA places great emphasis on the collection, assessment, and verification of the training data for AI systems operating in high-risk environments.<sup>30</sup>

During the training process, an AI system can identify bias patterns and correlations within previous judgments. They may be attributable to conscious or unconscious biases that have been introduced into the text.

---

<sup>26</sup> For example, in 2024, a Spanish lawyer submitted a document to the court containing references to the Colombian Penal Code rather than the Spanish one. This happened because the text was generated using ChatGPT, which included an existing provision but from the wrong jurisdiction. Conde E., *The risk of using ChatGPT: Navarrese Justice reprimands a lawyer who accused...using Colombian law*, 2024, available online at <https://www.noticiasdenavarra.com/sociedad/2024/11/12/riesgo-chatgpt-justicia-navarra-reprende-8919455.html?> (translated into English by Google translate).

<sup>27</sup> Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), 4 May 2016, OJ L 119.

<sup>28</sup> A general-purpose AI system is an AI system trained to perform a variety of tasks. Such a model can be multimodal and can generate text, picture, video, and recognise human speech.

<sup>29</sup> D. Verona, J. L. Suarez, *Discrimination, Bias, and Trustworthy AI*, 'Applied Sciences' 2022, 12, 12; High-Level Expert Group on Artificial Intelligence, *Ethics Guidelines for Trustworthy AI*, European Commission, 2019.

<sup>30</sup> AIA, Article 10.

Conscious bias is a situation where the person is aware of his or her opinion, and it can be determined based on statements or actions.<sup>31</sup> For example, Michael Benesty in the *Supra Legum* Project used machine learning techniques to evaluate decisions in asylum cases in France.<sup>32</sup> As a result of the project, the system identified patterns of decision-making for specific judges, possibly reflecting personal opinions of judges. This assumption was justified by the fact that certain judges had rejected the majority of asylum applications, thereby giving rise to a presumption of bias. As a response to this project and to prevent further judge profiling, France passed a law criminalizing the evaluation, analysis, comparison, and prediction of the behavior of individual judges.<sup>33</sup>

Unconscious bias is a situation where the person is not fully aware of personal prejudice or preference. Unconscious bias is an inherent characteristic of human beings. Well-documented reasons for unconscious biased decisions could arise due to physiological and biological reasons, such as being tired, hungry, or under stress.<sup>34</sup> The human brain tends to classify individuals and objects as a means of conserving cognitive resources and facilitating rapid decision-making.<sup>35</sup> Hence, some judgments may contain bias as a consequence of such underlying factors described above.<sup>36</sup> Biased patterns are often difficult to spot when reading individual judgments. By contrast, an AI system is capable of analyzing a large volume of text and be able to detect these patterns.

Bias in the training data may also result from socio-economic factors reflected in past decisions. The *U.S. State v. Loomis* case is an example of this type of biased AI system. In the *Loomis* case, the defendant challenged the use of the Correctional Offender Management Profiling for Alternative Sanctions (*hereinafter* ‘COMPAS’) system. COMPAS is an AI system that assesses the risk of recidivism used in the U.S. criminal proceedings. The system relied on a proprietary algorithm that assessed various factors, including an individual’s criminal history and socio-demographic information, to generate risk scores.<sup>37</sup> Critics argue that COMPAS may reinforce racial and socio-economic biases due to its reliance on historical data patterns.<sup>38</sup> In *Loomis*, the

<sup>31</sup> J. Holroyd, *Implicit Bias, Awareness and Imperfect Cognition*, ‘Consciousness and Cognition’ 2015, May, vol. 33, pp. 511–523.

<sup>32</sup> D. Ghezlbash, K. Dorostkar, Sh. Walsh, *A Data Driven Approach to Evaluating and Improving Judicial Decision-Making: Statistical Analysis of the Judicial Review of Refugee Cases in Australia*, ‘University of New South Wales Law Journal’ 2022, 45, 3, p. 1085.

<sup>33</sup> *Ibid.*; Article L111–13 to the Code de l’organisation judiciaire, France.

<sup>34</sup> J. Ch. Bublitz, *What Is Wrong with Hungry Judges? A Case Study of Legal Implications of Cognitive Science* [in:] *Law, Science, and Rationality*, A. M. Waltermann, D. Roef, J. Hage, and M. Jelcic (eds.), Eleven International Publishing, Maastricht Law Series 2019.

<sup>35</sup> A. D. Redish, *The Mind within the Brain*, Oxford University Press 2013.

<sup>36</sup> J. Ch. Bublitz, *What Is Wrong...*, *ibid.*

<sup>37</sup> I. M. Beriain, *Does the Use of Risk Assessments in Sentences Respect the Right to Due Process? A Critical Analysis of the Wisconsin v. Loomis Ruling*, ‘Law, Probability and Risk’ 2018, March, vol. 17, 1, pp. 45–53.

<sup>38</sup> A. Chouldechova, *Fair Prediction with Disparate Impact: A Study of Bias in Recidivism Prediction Instruments*, ‘Big Data’ 2017, 5, 2, pp. 153–163.

Wisconsin Supreme Court ruled that the use of COMPAS did not violate principles of due process, emphasizing that the system served only as an advisory tool.<sup>39</sup> The court held that judicial discretion was preserved, as judges were not bound to follow COMPAS assessments and retained the authority to make independent decisions.

The reasoning of the U.S. court in the *Loomis* case is similar to the provisions of the EU AIA.<sup>40</sup> Specifically, the AIA provides that human decision-makers retain control over the decision-making and may override AI recommendations.<sup>41</sup> However, the guarantee that the judge remains in control of the final decision does not itself give assurance of impartiality. The question of AI's indirect impact on judicial impartiality and independence remains complex.

Hence, embedded conscious or unconscious bias can be learned by the AI system, reinforced, and amplified in its recommendations. The AI system has a tendency to simplify complex information and form generalized correlations, which may inadvertently lead to subjective statements. The bias in an AI system may also form during the training process.

---

### Bias arising from the process of training

---

Bias in recommendations or decisions may also result from the process by which an AI system is trained. It is important to distinguish between the training of logic-based AI systems and that of machine learning AI systems.

The development of a logic-based system starts with defining the logical pattern of the decision.<sup>42</sup> Such a pattern is determined by a human expert and subsequently implemented in the source code by a software developer. Legal or technical issues may emerge either during the formulation of the algorithm or in the course of its programming. However, a key advantage of logic-based AI systems is that the algorithm and source code can be examined and verified, which allows errors to be detected and corrected.

Machine learning AI systems are developed and trained differently. They are trained using one of the three approaches.<sup>43</sup> First, unsupervised learning is a training process in which a large volume of data is provided to an AI model, and it finds patterns and correlations autonomously. General-purpose AI models are typically trained using this approach. However, when an AI

---

<sup>39</sup> Judgment of the Wisconsin Supreme Court, *State v. Loomis*, 881 N.W.2d 749, Wis. 2016.

<sup>40</sup> AIA, Annex III.

<sup>41</sup> AIA, Article 14, § 3(d).

<sup>42</sup> H. Prakken, G. Sartor, *Law and Logic: A Review from an Argumentation Perspective*, 'Artificial Intelligence' 2015, vol. 227, Elsevier, pp. 214–245.

<sup>43</sup> *Machine Learning: An Artificial Intelligence Approach*, R. S. Michalski, J. G. Carbonell, T. M. Mitchell (eds.), Berlin–Heidelberg 1983.

system is intended for applications in the justice system, it must be further refined through supervised or reinforcement learning methods.

Supervised learning is a training process in which a person prepares labelled or annotated data and inputs it into the system so that the AI may learn from human-defined examples. This process requires experts to identify and mark relevant sections of judgments, indicate important parts in the reasoning section, and classify cases according to their subject matter.<sup>44</sup> During this stage of data preparation, experts inevitably rely on their own interpretative judgments and professional experience, which may introduce the potential for both human error and subjective bias. Because an AI system lacks the capacity for critical reasoning or self-reflection, it cannot evaluate or challenge the assumptions embedded in the data on its own initiative. It therefore derives its output from the patterns it has learned, reproducing the biases and limitations present in its training data.

Reinforcement learning is a training method in which an AI system receives feedback in the form of rewards or penalties based on the quality of its decisions. In this model of learning, a human expert evaluates whether the system's outputs are correct or appropriate.<sup>45</sup> As in supervised learning, the evaluative judgments of the human expert may influence the training process because personal views and implicit biases become embedded in the system's decision-making.

These models of learning demonstrate that, despite the varying approaches to training, human involvement remains part of this process and the development of AI systems. It is, therefore, wrong to assume that decisions produced by an AI system are purely algorithmic or artificial. If the person involved in the training has certain opinions, these opinions can be reinforced and amplified by the AI system.

In May 2025, UNESCO adopted Draft Guidelines for the use of AI systems in courts and tribunals<sup>46</sup> that represent international ethical standards regarding the use of AI systems in courts. Principle 2.1.1 reflects the challenge of training an AI system and envisages that judges should be involved in the development and deployment stages of AI systems in courts. Their participation is important to make sure that the training of AI systems to be used in courts reflects the realities of legal reasoning and limits the risk of embedding subjective biases or oversimplified representations of decision-making.

<sup>44</sup> A. S. Deeks, D. B. Hollis, *Large Language Models and International Law*, 'Chicago Journal of International Law' 2025, vol. 26, 1, p. 6.

<sup>45</sup> G. Sartor, F. Galli, *AI Approaches...*, *ibid.*

<sup>46</sup> UNESCO, Draft UNESCO Guidelines for the Use of AI systems in Courts and Tribunals, CI/DIT/2024/GL/01, 2024.

An AI system, even if developed and designed for the justice system, is not immune to bias. Such bias may be detected during testing, yet it may also infiltrate the system and manifest in the form of draft judgments. As with the assessment of subjective impartiality of human judges, identifying bias within an AI system poses difficulties and presents an even greater evidentiary burden for the parties. Therefore, the analysis of the impact of AI systems on the impartiality of the court may focus on the objective element, which provides a further important guarantee.<sup>47</sup>

---

## Objective Element of Impartiality and Legitimate Doubts

---

The objective element of impartiality concerns the composition of the tribunal, which should offer sufficient guarantees to ‘exclude any legitimate doubt in respect of its impartiality’.<sup>48</sup> This element relates to issues of the appearance of bias.<sup>49</sup> That is to say, the test asks whether there exists a legitimate reason to fear a lack of impartiality<sup>50</sup> and whether such fear could be objectively justified.<sup>51</sup> This test concerns the links of the judge with other actors in the proceedings.<sup>52</sup> The use of an AI system in an assisting role for decision-making in the case should be arranged in a manner that does not create legitimate reasons to fear a lack of impartiality. The parties must have confidence that their right to be heard, adversarially, is effectively protected and the principle of “equality of arms” is respected. This confidence in the court’s impartiality may be at issue where the use of AI lacks sufficient transparency, human oversight, or explainability.

---

## The right to be informed about the AI system

---

The right to be informed is an important safeguard for the “appearance of an impartiality” element. In this respect, the special account is given to the internal organization of a tribunal established by law.<sup>53</sup> The right to be

<sup>47</sup> *Micallef v. Malta*, § 95; *Morice v. France*, § 119.

<sup>48</sup> Judgment of the ECtHR of 24 February 1993, *Fey v. Austria*, application no. 14396/88, § 28 and § 30; *Wettstein v. Switzerland*, § 42.

<sup>49</sup> Consultative Council of European Judges (CCJE), Opinion No. 3 of the Consultative Council of European Judges (CCJE) to the attention of the Committee of Ministers of the Council of Europe on the principles and rules governing judges’ professional conduct, in particular ethics, incompatible behaviour and impartiality, CCJE (2002) Op. No. 3, 19 November 2002.

<sup>50</sup> *Piersack v. Belgium*, § 30.

<sup>51</sup> Judgment of the ECtHR of 24 May 1989, *Hauschildt v. Denmark*, application no. 10486/83, § 48.

<sup>52</sup> *Micallef v. Malta*, § 97.

<sup>53</sup> *Piersack v. Belgium*, § 30 (d).

informed about the use of an AI system entails that the parties are made aware of the system's involvement in the decision-making, its role, and the extent to which it may influence the outcome of the case. The use of an AI system in this context is closely linked to the transparency principle envisaged in the EU AIA<sup>54</sup> and guidelines for the ethical use of AI.<sup>55</sup>

The understanding of transparency in relation to AI has evolved in recent years. The 2018 European Ethical Charter on the use of artificial intelligence emphasizes that data processing methods of AI systems have to be accessible and understandable so as to enable external audits.<sup>56</sup> Transparency in this ethical framework is understood as the ability to comprehend the functioning of the AI system and the reasoning process underlying its outputs.<sup>57</sup> This approach to transparency is better suited for logic-based systems than machine learning systems.

Transparency under the EU AIA is understood from a different angle. Its Article 13 requires providers of a high-risk AI system to accompany it with concise instructions that describe capabilities, limitations, level of accuracy, and other functional features of the AI system.<sup>58</sup> This obligation is distinguished from the right to explanation envisaged in Article 86 of the Act. This Article provides that, where a decision producing legal effects is based on the output of an AI system, the affected persons have "a right to obtain a clear and meaningful explanation from the deployer of the role of the AI system in the decision-making and the main elements of the decision taken".<sup>59</sup> In essence, this right imposes an obligation to inform individuals that a decision has been taken with the assistance of an AI system and explain the role of the AI system in the decision-making process.<sup>60</sup>

The AI system could be used at the preparatory stage of the proceedings. It can analyze evidence, research, and summarize relevant court jurisprudence and produce a preliminary legal assessment. The UNESCO Draft Guidelines on the use of AI in courts suggest disclosing the use of generative AI and using quotation marks and a citation system for text generated by an AI system.<sup>61</sup> Moreover, the Guidelines recommend attaching an annex

<sup>54</sup> AIA, Article 13.

<sup>55</sup> CEPEJ, European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment, 3–4 December, 2018, Principle 4; High-Level Expert Group on Artificial Intelligence, Ethics Guidelines for Trustworthy AI, European Commission, 2019, § 1.4; Draft UNESCO Guidelines for the Use of AI Systems in Court and Tribunals, § 3.3.7.

<sup>56</sup> CEPEJ, European Ethical Charter, Principle 4, p. 7.

<sup>57</sup> CEPEJ, Principle 4 explanations, p.11; UNESCO Draft Guidelines for the use of AI in courts and tribunals, § 1.9.

<sup>58</sup> AIA, Article 13.

<sup>59</sup> AIA, Article 87 § 1.

<sup>60</sup> Similar recommendation is included in Draft UNESCO Guidelines for the use of AI systems in courts and tribunals, § 3.2.4.

<sup>61</sup> Draft UNESCO Guidelines for the use of AI systems in courts and tribunals, § 3.3.7.

with the prompts and the AI system's answers to these prompts if the text produced by AI was used in the judgment.<sup>62</sup> In the *Kress v. France* case, the ECtHR assessed the role of the Government Commissioner in administrative proceedings in France. The Commissioner presented independent legal opinions at the preparatory and hearing stages, without prior disclosure of the legal position to the parties.<sup>63</sup> The ECtHR found that, although the Commissioner was independent, the inability of parties to comment on the observations breached the equality of arms and adversarial principles under Article 6 of the ECHR. Analogously, disclosing the legal evaluation of the case produced by the AI system that might have an impact on the decision in the case is also relevant not only from a transparency perspective, but also according to Article 6 of the ECHR.

Recent emerging practices provide illustrative examples of transparent use of AI in courts. In 2024, a Dutch court explicitly indicated in its reasoning that ChatGPT had been used to estimate the average lifespan of solar panels.<sup>64</sup>

There are several approaches to informing the public on the use of AI. Courts may publish information on the court website describing the AI system and its role. In this respect, the Council of Europe's Resource Center on Cybersecurity and Artificial Intelligence plays an important role by collecting and documenting practices concerning the use of digital and AI-based tools.<sup>65</sup> This open-access resource is continuously updated and contributes to the transparency, monitoring, and evaluations of the use of AI in justice systems.

If the AI system is used to perform more complex tasks, and is able to semi-autonomously evaluate cases and the law, and deliver an "opinion", such use of the AI system has to be envisioned in the procedural rules of the court according to the "established by law" principle. The procedural rules should define and delineate the role of an AI system in court proceedings and establish a procedure for disclosure of the use of AI.

---

## Safeguards for judicial impartiality

---

Effective human oversight is a central safeguard to ensure that the use of AI does not undermine the integrity of decision-making. The obligation to exercise an effective human oversight of high-risk AI systems is included in Article 14 of the EU AIA. This provision has its origins in Article 22 of GDPR,

---

<sup>62</sup> Ibid.

<sup>63</sup> Judgment of the ECtHR of 7 June 2001, *Kress v. France*, application no. 39594/98, § 63.

<sup>64</sup> Gelderland Court, 10664071 CV23-2321, 7 June 2024, ECLI:NL:RBGEL:2024:3636, § 5.7.

<sup>65</sup> Council of Europe, Resource Center on Cybersecurity and Artificial Intelligence – available online at <https://www.coe.int/en/web/cepej/resource-centre-on-cyberjustice-and-ai>.

which envisages the right to human intervention when an automated decision-making system is used<sup>66</sup> to produce legal effects or significantly affect the rights of an individual.<sup>67</sup> Hence, the human oversight is linked to the operation of an autonomous or semi-autonomous AI system. Conversely, the use of an AI system in an assisting role or as a “legal expert” does not endow the AI system with the function to deliver judgments. Its primary role is to gather legal information, spot inconsistencies, and suggest possible decisions.

The relationship between the human oversight of an AI system and, *vice versa*, the impact of the AI system on decision-making is complex. Even in cases where AI systems are not making decisions autonomously but are used in assistive roles, they may influence the final decision. Therefore, the effective human oversight includes “awareness of the possible human tendency to rely and over-rely on the output of the AI system.”<sup>68</sup> This particularity of technology is what distinguishes it from a human court officer or expert. The AI system is able to generate a persuasive text. It can generally present several options of the text for the judge to read and select the one he or she prefers. This can sometimes lead to what is known as algorithmic bias. Algorithmic bias occurs when an AI system produces biased output, which is then relied upon by a human decision-maker because it, for example, confirms preexisting conceptions of a human.<sup>69</sup>

There is also an additional concern in that the judge may feel bound to comply with the recommendations generated by an AI system, particularly if institutional practices implicitly or explicitly encourage its use. This dynamic can create subtle pressure to align decisions with the system’s output, which potentially can narrow the scope for independent reasoning. However, judicial autonomy is an important element of impartiality and independence. Judges are bound to decide cases free from external influence and based solely on the law, facts, and their own professional assessment.

Article 14 of the EU AIA stipulates that human oversight measures should be implemented, taking into account the level of automation and the context where the high-risk AI system is used. To mitigate the risks of overreliance on technology, it is important to understand the capacities and limitations of an AI system, interpret the result correctly, and decide when not to rely on the AI system.

<sup>66</sup> General Data Protection Regulation, Article 22, § 3.

<sup>67</sup> General Data Protection Regulation, Article 22, § 1.

<sup>68</sup> AIA, Article 14, § 4 (b).

<sup>69</sup> S. Alon-Barkat, M. Busuioc, *Human-AI Interactions in Public Sector Decision Making: “Automation Bias” and “Selective Adherence” to Algorithmic Advice*, *Journal of Public Administration Research and Theory* 2023, 33, pp. 153 and 155.

## AI literacy skills

Understanding the capacities and limitations of the AI system means that a judge possesses AI literacy skills and understands what the AI system can and cannot do. The EU AIA prescribes that the providers and deployers of AI systems should ensure a sufficient level of AI literacy skills for persons who are using the AI system.<sup>70</sup> Similarly, the UNESCO Guidelines for the use of AI systems in courts and tribunals point out the necessity to provide training opportunities for court administrators and judges to learn about different types of tasks that can be performed with the assistance of AI, to learn how to use these tools and what kind of impact they can have on fundamental rights, and to learn how to interpret the output critically.<sup>71</sup> Both Article 4 of the EU AIA and the UNESCO Guidelines contain a rather broad obligation. AI literacy obligation encompasses various uses of AI systems, regardless of whether they fall under the definition of a high-risk system or not. This obligation appears particularly justified in light of the increasing availability of various public AI tools, integration of AI systems with Microsoft, as well as the growing use of AI by parties in proceedings to draft submissions and potentially to generate or manipulate evidence.

Neither the EU nor the Council of Europe's legal acts nor UNESCO recommendations provide a comprehensive list of AI literacy skills. The UNESCO Guidelines for the use of AI in courts and tribunals, under the section on AI awareness, AI literacy, and capacity building, suggest that AI literacy should allow individuals to understand AI fundamentals, identify its functionalities, and understand the limits and risks of AI systems, including how to use these systems ethically and assess their output and performance.<sup>72</sup>

Understanding AI fundamentals may include basic knowledge of essential elements of AI, such as its architecture and data, principles of how the text is generated, and particularly the probabilistic and statistical approaches. Although these are technical aspects of AI, they equip judges with a comprehension of the origins of bias in the output of the AI systems, why the output could be incorrect, and set the foundation for using AI systems effectively. Understanding the role of the "prompt" and the relations between the prompt and the system's output, together with knowledge of the functionalities of the AI system, can help understand various ways in which the AI system can be used.

Knowledge of the functionalities of the AI system, apart from prompting, can also include skills such as working with text analysis, summarization,

---

<sup>70</sup> AIA, Article 4.

<sup>71</sup> Draft UNESCO Guidelines for the use of AI systems in courts and tribunals, § 2.3.1.

<sup>72</sup> Draft UNESCO Guidelines for the use of AI systems in courts and tribunals, § 3.1.1.

and other practical skills on how the system can be used effectively. It can also cover the role of the input data in the performance of the AI system. Under Article 14 of the EU AIA, the user of the system bears the responsibility for the input data as part of the human oversight obligation.<sup>73</sup>

Comprehension of limits and risks of the AI system can already be addressed in the AI fundamentals section and main functionalities because they arise mainly from the ways in which the system is designed, the quality and type of data, and the training process. The focus of this set of skills is on further understanding different types of bias and the risks of overreliance on technology, including confirmation bias. Most AI systems do not perform with absolute accuracy. All AI systems occasionally, to a varying extent, provide incorrect information. An additional problem is that the AI systems used in the legal field are trained on the legislation and case law at the time of creation of the system. If the legislation or case law changes, it is necessary to update the system. This may not take place immediately after legislation or case-law changes were introduced. Before the system is updated, its output can therefore contain dated information. Hence, the knowledge of level of accuracy of AI systems seems necessary to enable human oversight and to interpret AI output correctly.

### Explainability and interpretation of AI output

Transparency and explainability in AI systems are often presented together. UNESCO Draft Guidelines encourage courts to “adopt, deploy, and use AI systems that can explain the rationale behind their outputs and decisions and how they use inputs for such purposes.”<sup>74</sup> Explainable AI consists of three elements – understandable, interpretable, and traceable for humans.<sup>75</sup> These three elements are interconnected. As described earlier in this paper, determining the precise logic of a machine learning AI system is challenging. Presently, it is difficult to achieve the level of transparency in machine-learning systems similar to that in logic-based systems. It is an inherent characteristic of a machine learning AI system.

On the other hand, the evaluation of the AI system output could be approached similarly to the evaluation of human beings. For instance, it is not precisely known how thoughts and ideas arise in the human mind, but they can be communicated to others, tested, and assessed. Similarly, LLMs are able to communicate in natural language, explain their “thinking” process, and provide information on which factors were taken into account and became the basis for the decision or recommendation. Hence, the underlying

<sup>73</sup> AIA, Article 26 § 4.

<sup>74</sup> Draft UNESCO Guidelines for the use of AI systems in courts and tribunals, § 1.7.

<sup>75</sup> *Ibid.*

criteria and the quality of the output could be verified, assessed, and understood.

Correct interpretation of the output may encompass the ability to identify machine errors through knowledge and experience.<sup>76</sup> The AI system does not possess contextual awareness or social understanding. An AI system operates as a probabilistic system that recommends the most likely course of action based on patterns in data. Each case may involve unique circumstances and contexts that must be considered in decision-making. Interpreting output correctly may only be possible if the system justifies the decision.<sup>77</sup> A judge, drawing on her or his extensive experience, is well-positioned to assess whether the output produced by the AI and its accompanying explanations appear reasonable and whether they can be effectively used in drafting a judgment. Additionally, it is important to note that AI may be best suited for relatively simple and repetitive cases that require less complex reasoning. In such instances, the interpretability of the AI system's output is unlikely to present significant challenges.

However, the process of verifying AI outputs requires time. In situations where judges have a heavy workload, it is not easy to find time to meaningfully assess the output of the AI system<sup>78</sup> and decide whether to rely on this information. Hence, effective human oversight is interdependent on the conditions and environment in which the judge works. Since in many European courts the workload of a judge is high, it is possible to assume that it may be difficult to maintain effective human oversight in all situations when AI is used, thereby providing an opportunity for an AI system to influence the judgment.

### Proportional use of AI

The most important element of human oversight is the right not to use an AI system.<sup>79</sup> This right guarantees that even if an AI system is integrated into the court as technology, a judge has the right not to rely on it. Article 14 (4-d) of the EU AIA envisages that the person using the AI can decide when not to use it. This right can be exercised when appropriate and proportionate. When Article 14 of the EU AIA is read in conjunction with ethical guidelines on the use of AI in courts and tribunals, this provision may be interpreted broadly as delineating both the conditions under which AI may appropriately be employed and those in which reliance upon its outputs should be restrained

<sup>76</sup> M. Fink, *Human Oversight Under Article 14 of the EU AI Act* [in:] *AI Act Commentary: A Thematic Analysis*, G. Malgieri, G. G. Fuster, G. Zanfir-Fortuna (eds.), Hart-Bloomsbury, Oxford 2026, May; [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=5146118](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5146118).

<sup>77</sup> M. Fink, *Human...*, *ibid.*

<sup>78</sup> B. Green, *The Flaws of Policies Requiring Human Oversight of Government Algorithms*, 'Computer Law & Security Review' 2022, 45.

<sup>79</sup> AIA, Article 14 § 4 (d).

or avoided altogether. To delineate these situations, the ethical standard of proportionality is of assistance.<sup>80</sup>

AI systems can help courts work more efficiently. It can support case management or assist with preliminary analysis of the case. However, the use of AI systems to achieve efficiency should be proportional and should not undermine fundamental procedural rights. For instance, the AI system can be used to determine legal issues in the case and suggest relevant legal provisions and case-law. However, the court should retain control of the legal qualification in the case.

AI can help handle large sets of information or recurring disputes more effectively. For instance, in consumer protection cases that involve many applicants, an AI system might assist the court by summarizing patterns across individual claims, helping the judge identify systemic issues. AI can also help detect disparities in case outcomes or reference existing high-court jurisprudence. In this instance, the AI system is used in an advisory function and does not substitute for the judge's interpretative role.

An AI system can help find gaps in doctrine, especially when courts are confronted with new social issues. For instance, in administrative proceedings, the court or tribunal might use an AI system to track how frequently higher courts annul administrative decisions on certain grounds. In this example, using the AI system proportionately implies that the AI system only provides insights and does not determine how the case should be solved.

Deciding when and how to use AI requires assessment of necessity, purpose, and the use of AI proportionately. AI can be an invaluable instrument for managing information, identifying patterns, or supporting interpretation. The choice to use AI can be made based on a balance between innovation and maintaining integrity.

---

## **Conclusion**

---

When an AI system is used to evaluate evidence, conduct legal research, and apply the law to concrete facts, the boundary between assistance and influence on the decision-making may become blurred. Nonetheless, the mere presence of technological assistance does not, by itself, affect the impartiality of a judge. As with the role of a court officer or an advocate general, an AI system may support the court in various ways. Such assistance, however, does not equate to an ability to influence the judge's opinion in a manner that makes the court not impartial.

---

<sup>80</sup> UNESCO, Recommendation on the Ethics of Artificial Intelligence, SHS/BIO/PI/2021/1, 23 November 2021, Principle III.2, §25.

At the same time, the use of AI cannot be completely analogized to the roles of “court officer” or “advocate general”. These actors themselves must be independent and impartial, and their status, functions, and guarantees are determined by law. Although the use of AI in court should, in principle, be regulated by law, in accordance with Article 6(1) of the ECHR, the risk of bias inherent in such a system remains a distinct concern.

Perceived impartiality of the court in the context of using an AI system is a matter of trust and confidence in the judicial system. If parties believe that a judge relies excessively on the output of the AI system rather than exercising independent judgment, they may perceive that justice has not been done. This, in turn, may raise concerns about the role of AI in decision-making and the safeguards in place to ensure impartiality.

Several measures could strengthen trust in the justice system and mitigate doubts regarding impartiality. First, uncertainty may arise when parties are unaware that an AI system is being used to analyze pleading and evidence, conduct legal research, or recommend case outcomes. If the AI system qualifies as a high-risk system under EU AIA, parties are entitled to receive an explanation from the court about the involvement of the AI system. Article 86 of the AIA does not specify whether such explanations should be provided on the initiative of the court or only upon request from parties. The ethical principle of transparency in this respect encourages to disclosure of information about the use of AI for decision-making. It can be accomplished through the court’s website or by including a statement within the judgment.

Second, the lack of participation of judges in the development and training of the AI system may create doubts regarding the integrity of the training process and its potential influence on the system’s outputs. The AI system functions based on the training data and can make decisions based on the training process. Because an AI system functions on the basis of training data, involvement of judges in its development may enhance the system’s quality and relevance. On the other hand, insufficient training for judges on the capabilities, limitations, and implications of AI systems, particularly when combined with other factors, may undermine the autonomy of the judge to deliver impartial and independent decisions in the case.

Third, the absence of independent testing and auditing of the AI system for quality, accuracy, and subjectivity may call into question the reliability of the system’s recommendations. Finally, a requirement obliging judges to rely upon the outputs of an AI system or to justify any departure from them may significantly undermine the impartiality and independence of the judicial decision-making.

Addressing all these issues at the time of development, deployment, integration, and the use of AI systems in courts can increase transparency and improve understanding by parties and society in general, the role of AI systems

and its impact on case outcomes, and inspire trust and confidence in courts in a democratic society.

### **Abstrakt**

Najnowsze osiągnięcia w dziedzinie sztucznej inteligencji (AI) zaowocowały powstaniem systemów, które analizują informacje prawne, rozpoznają powiązania między dokumentami oraz wspierają proces podejmowania decyzji. Nowe możliwości automatyzacji stwarzają zarówno szanse, jak i wyzwania dla wymiaru sprawiedliwości. Rodzą one złożone pytania dotyczące zgodności wdrażania sztucznej inteligencji w sądach z gwarancjami rzetelnego procesu sądowego wynikającymi z europejskiej Konwencji praw człowieka (EKPC).

W niniejszym artykule zbadano, czy i na jakich warunkach stosowanie systemów AI w sądach może budzić uzasadnione wątpliwości co do bezstronności sądu w rozumieniu art. 6 ust. 1 EKPC. Opierając się na orzecznictwie Europejskiego Trybunału Praw Człowieka, dokonano rozróżnienia między subiektywnymi a obiektywnymi elementami bezstronności oraz przeanalizowano, w jaki sposób na każdy z nich może wpływać udział AI w procesie podejmowania decyzji.

W artykule stwierdzono, że wykorzystanie sztucznej inteligencji w sądach jest kwestią zaufania do systemu wymiaru sprawiedliwości, a postrzegany brak bezstronności może wynikać z ograniczonej przejrzystości co do roli AI w procesie decyzyjnym, niewystarczającego nadzoru oraz nadmiernego polegania na wynikach generowanych przez sztuczną inteligencję.

**Słowa kluczowe:** sztuczna inteligencja (AI), bezstronność, uzasadnione wątpliwości, prawo do rzetelnego procesu sądowego.

---

## **BIBLIOGRAPHY**

---

AI Act Commentary: A Thematic Analysis, G. Malgieri, G. G. Fuster, G. Zanfir-Fortuna (eds.), Hart-Bloomsbury, Oxford 2026, May; [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=5146118](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5146118).

Alon-Barkat S., Busuioc M., *Human-AI Interactions in Public Sector Decision Making: “Automation Bias” and “Selective Adherence” to Algorithmic Advice*, ‘Journal of Public Administration Research and Theory’ 2023, 33.

Beriain I. M., *Does the Use of Risk Assessments in Sentences Respect the Right to Due Process? A Critical Analysis of the Wisconsin v. Loomis Ruling*, Law, ‘Probability and Risk’ 2018, March, vol. 17, 1.

CEPEJ, European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment, 3-4 December 2018.

Chouldechova A., *Fair Prediction with Disparate Impact: A Study of Bias in Recidivism Prediction Instruments*, 'Big Data' 2017, 5, 2.

Christian A., Biedebach L., K pfer A., and Neunhoeffer M., *The Role of Hyperparameters in Machine Learning Models and How to Tune Them*, 'Political Science Research and Methods' 2024, vol. 12, 4.

Consultative Council of European Judges (CCJE), Opinion No. 3 of the Consultative Council of European Judges (CCJE) to the attention of the Committee of Ministers of the Council of Europe on the principles and rules governing judges' professional conduct, in particular ethics, incompatible behaviour and impartiality, CCJE (2002) Op. No. 3, 19 November 2002.

Council of Europe, Resource Center on Cybersecurity and Artificial Intelligence – available online at <https://www.coe.int/en/web/cepej/resource-centre-on-cyberjustice-and-ai>.

Creamer C. D., Godzimirska Z., *Trust in the Court: The Role of the Registry of the European Court of Human Rights*, 'The European Journal of International Law' 2019, vol. 30, 2, Oxford University Press.

Deeks A. S., Hollis D. B., *Large Language Models and International Law*, 'Chicago Journal of International Law' 2025, vol. 26, 1.

European Commission for the Efficiency of Justice (CEPEJ), Guidelines on how to drive change towards Cyberjustice, CEPEJ-GT-QUAL, 7 December 2016.

European Commission, Knowledge Center on Translation and Interpretation, available online at <https://knowledge-centre-translation-interpretation.ec.europa.eu/en/news/what-large-language-model>.

Fernandez, J. V., *Artificial Intelligence in Government: Risks and Challenges of Algorithmic Governance in the Administrative State*, 30 Ind. J., 'Global Legal Studies' 2023, 65.

Ferrara E., *Fairness and Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, and Mitigation Strategies*, 'Sci' 2024, 6, 1.

Gelderland Court, 10664071 CV23-2321, 7 June 2024, ECLI:NL:RBGEL:2024:3636.

Ghezelbash D., Dorostkar K., and Walsh Sh., *A Data Driven Approach to Evaluating and Improving Judicial Decision-Making: Statistical Analysis of the Judicial Review of Refugee Cases in Australia*, 'University of New South Wales Law Journal' 2022, 45, 3.

Green B., *The Flaws of Policies Requiring Human Oversight of Government Algorithms*, 'Computer Law & Security Review' 2022, 45.

High-Level Expert Group on Artificial Intelligence, Ethics Guidelines for Trustworthy AI, European Commission, 2019.

Holroyd J., *Implicit Bias, Awareness and Imperfect Cognition*, 'Consciousness and Cognition' 2015, May, vol. 33.

Law, Science, and Rationality, A. M. Waltermann, D. Roef, J. Hage, and M. Jelcic (eds.), Eleven International Publishing, Maastricht Law Series 2019.

Machine Learning: An Artificial Intelligence Approach, R. S. Michalski, J. G. Carbonell, T. M. Mitchell (eds.), Berlin–Heidelberg 1983.

Mavrogiorgos K., Kiourtis A., Mavrogiorgou A. et al., *Bias in Machine Learning: A Literature Review*, 'Applied Sciences' 2024, 14, 19.

Opinion of Advocate General of 11 May 2006, *Graham J. Wilson v. Ordre des avocats du barreau de Luxembourg*, C-506/04, ECLI:EU:C:2006:31.

Prakken H., Sartor G., *Law and Logic: A Review from an Argumentation Perspective*, 'Artificial Intelligence' 2015, vol. 227, Elsevier.

Redish A. D., *The Mind within the Brain*, Oxford University Press 2013.

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), 4 May 2016, OJ L 119.

Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), 12 June 2024, OJ L, 2024/1689.

Sartor G., Galli, F., *AI Approaches to Predictive Justice: A Critical Assessment*, 'Humanities and Rights Global Network Journal' 2023, vol. 5, 2.

Sourdin T., *Judges, Technology and Artificial Intelligence: The Artificial Judge*, Elgar Law, Technologies and Society, Cheltenham 2021.

UNESCO, Draft UNESCO Guidelines for the Use of AI systems in Courts and Tribunals, CI/DIT/2024/GL/01, 2024.

UNESCO, Recommendation on the Ethics of Artificial Intelligence, SHS/BIO/PI/2021/1, 23 November 2021.

Verona D., Suarez J. L., *Discrimination, Bias, and Trustworthy AI*, 'Applied Sciences' 2022, 12, 12.



Artur Kotowski\*

## The Legal Method in the Context of Futurology

### [Metoda prawnicza w kontekście futurologii]

#### Abstract

The article discusses the limitations of the legal method adopted in the culture of statutory law—the exegesis of normative acts—considering civilisational developments in the digital age. The legal method guarantees the scientific autonomy of jurisprudence and was developed at the end of the 19th century as a formal-dogmatic method. With the development of law and European integration, its classic formula, developed in an era of strongly national legal systems with limited mutual relations, is becoming anachronistic. The article highlights these limitations and outlines possible further directions for its development.

**Keywords:** legal interpretation, theory of law, civil law legal culture, general jurisprudence, legal futurology.

---

## Introduction

---

This article deals with the evolution of the legal method considering civilisational changes related primarily to the dynamic growth in the importance of communication technologies.<sup>1</sup> This is the main, though not the only, subject of interest in legal futurology – a research area of general jurisprudence that emerged because of intensive technological and social transformations initiated by the digital and mainly communication revolution.<sup>2</sup> This revolu-

\* **Artur Kotowski** – full professor of Law, PhD, post-doctoral degree, Head of the Department of Theory and Philosophy of Law, Faculty of Law and Administration, Cardinal Stefan Wyszyński University in Warsaw (affiliation); <https://orcid.org/0000-0001-8284-241X>; [a.kotowski@uksw.edu.pl](mailto:a.kotowski@uksw.edu.pl) / prof. dr hab. nauk prawnych, kierownik Katedry Teorii i Filozofii Prawa, Wydział Prawa i Administracji, Uniwersytet Kardynała Stefana Wyszyńskiego w Warszawie (afiliacja).

<sup>1</sup> This article is an extended version of the paper published in Polish in the book: *Ius scientiam persequitur*. Księga jubileuszowa z okazji 25-lecia Wydziału Prawa i Administracji Uniwersytetu Kardynała Stefana Wyszyńskiego w Warszawie [(...) Jubilee Book Commemorating the 25th Anniversary of the Faculty of Law and Administration at Cardinal Stefan Wyszyński University in Warsaw], M. Badowski, J. Wesslering, L. Karski (eds.), Warszawa 2025.

<sup>2</sup> A. Toffler, *Trzecia fala* [The Third Wave], trans. E. Woydyło, Warszawa 1986 (1st ed., censored), p. 39.

tion dates to the second half of the 20th century and is characterised, among other things, by shorter intervals of technological changes affecting law and jurisprudence than in previous eras.<sup>3</sup> Due to the limited scope of this publication, this paper is intended as a contribution to the discussion, essentially of a journalistic nature. This remark applies to the entire argument but needs to be emphasised in relation to the general concept of legal method characterised in the next point, which is an extremely broad issue, and in this study the subject had to be condensed as much as possible.

The importance of this issue does not seem to be debatable. No one denies either the dynamics of the technological foundations of civilisational changes or the fact that their relationship to law as a social phenomenon and jurisprudence as a science that has them as its subject is causal in nature.<sup>4</sup> The development of foresight methods (forecasting the future in the specific context of social sciences) has also made it possible to take legal futurology “seriously”, which at the beginning of this century was absent from scientific discourse or was treated as alchemy in relation to chemistry or astrology in relation to astronomy,<sup>5</sup> i.e. as a pseudoscience. The impact of the digital revolution on law and jurisprudence is undeniable, and the pace of new phenomena accompanying it boils down to the research question: when and in what direction will the methodological assumptions of jurisprudence change, rather than whether this will happen at all? The phenomena of incorporating digital technologies into law and jurisprudence can only be mentioned here as examples; it is a question of the role of the Internet and search forms for broadly understood legal information, the availability of legal knowledge through professional and popular digital sources (digitisation of sources of law, legal doctrine texts, etc.), the phenomenon of information overload – also in relation to law, huge changes in private law, including in the area of making declarations of will (digital signatures, etc.), methods of contracting, the need to criminalise new acts, and many other phenomena that cannot be listed in a closed catalogue.<sup>6</sup>

<sup>3</sup> B. Debaenst, *The Digital Revolution from a Legal Historical Perspective* [in:] *Law, AI and Digitalisation*, K. de Vries, M. Dahlberg (eds.), Uppsala 2022, p. 27.

<sup>4</sup> The issue of digital technologies in the context of law is extensively discussed in the literature, which is already very extensive. See, for example: A. De Franceschi, R. Schulze (eds.), *Digital Revolution – New Challenges for Law*, Baden-Baden 2019, *passim*.

<sup>5</sup> S. Lem, *Solaris*, Warszawa 2012 (1st ed. 1961), pp. 125 and 126.

<sup>6</sup> See a more detailed discussion of these issues in the monographs: Internet. *Globalne gry* [The Internet: Global Games], A. Gryszczyńska, G. Szpor, R. Wiewiórowski (eds.), Warszawa 2022, *passim*, and: Internet. *Solidarność cyfrowa* [The Internet: Digital Solidarity], A. Gryszczyńska, G. Szpor, R. Wiewiórowski (eds.), Warszawa 2024, *passim*.

---

## Legal Method

---

Traditionally, it is taught that the basic legal method in the European legal tradition (the civil law) emerged in the second half of the 19th century as a significant achievement of early legal positivism. It is known as the formal-dogmatic method and is specified as the exegesis of a normative act, which is carried out through the application of a specific, directive theory of legal interpretation. In Anglo-Saxon culture (common law), its equivalent is analytical philosophy of law, also known in statutory law systems as conceptual analysis. It is used both for statutory law and for texts in the field of legal doctrine. The theory of interpretation in the continental sense has an auxiliary function there, it is less formal than in civil law states, although its importance is constantly growing, as is the role of statutory law in the area of its sources. Also, it should be noticed that “in an appropriate case there is nothing to prevent a court from taking proposed legislation into account when exercising its powers.”<sup>7</sup>

The debate on legal methodology has been fundamental to continental legal culture over the last two hundred years.<sup>8</sup> Its definition determined the scope of the entire field of jurisprudence.<sup>9</sup> On the one hand, it was a matter of responding to the paradigms of general scientific methodology that emerged in the 19th century and reducing the classic problems of legal reasoning associated with the natural susceptibility of language to ambiguity and the subjectivity of conclusions in the process of reconstructing normative meanings.<sup>10</sup> Hence, there was a shift towards formal logic and an attempt to develop various theories of legal interpretation, which were supposed to organise the entirety of legal reasoning responsible for determining the meanings of normative expressions. However, it proved impossible to develop a directive theory of interpretation that would lead to certain conclusions in the understanding of formal logic (it should be added that this would have resulted in a significant reduction in the demand for legal services). It was also impossible to avoid departing from fundamental assumptions – axioms, which is contrary to the concept of practising naturalistic, scientific and aphiloosophical science modelled on the natural and exact sciences, in which cognition is

<sup>7</sup> D. Bailey, L. Norbury, Bennion, Bailey and Norbury on Statutory Interpretation (8th ed.), London 2020, p. 30.

<sup>8</sup> It also stemmed in part from the metatheoretical perspective, which is also important in Polish legal scholarship, i.e. the connections between ‘broadly understood legal theory and general philosophy’ – P. Jabłoński, *Postawa filozoficzna i afilozoficzna w namyśle nad teorią prawa. Rzecz o aktualności metateoretycznych ustaleń Jerzego Wróblewskiego* [Philosophical and A-philosophical Approaches in Reflections on Legal Theory: A Discussion of the Relevance of Jerzy Wróblewski’s Metatheoretical Findings], *Acta Universitatis Wratislaviensis*, Wrocław 2011, *Prawo* 312, 3337, p. 303.

<sup>9</sup> K. Opałek, J. Wróblewski, *Prawo – metodologia, filozofia, teoria prawa* [Law: Methodology, Philosophy, Theory of Law], Warszawa 1991, p. 12.

<sup>10</sup> K. Opałek, J. Wróblewski, *Prawo...*, p. 26.

completely detached from its object.<sup>11</sup> The formalisation of legal reasoning focused on the exegesis of normative acts had to be based on the dogmas of the validity and systematic nature of law. This presupposed the fundamental assumption of the entire positivist narrative about law, namely the validating opposition of legal norms to other types of rules.

The question of whether the formal-dogmatic method exhausts the catalogue of ways of understanding the phenomenon of law secondarily determines the accepted scope of jurisprudence. Only pure theory of law – as a radically positivist philosophy – accepts this assumption. Even for the “original” positivists (in XIX century), it was clear that, to a certain extent, it was necessary to study the empirical manifestations of legal norms that apply in a specific social reality. After the condition of the theoretical validity of law, the question of the social effectiveness of law was, for positivists, the second requirement that made it possible to define the subject of legal science. “It is a matter of revealing the legal normative sphere, which is contained only in the linguistically identified content of the sovereign’s act of will.” As a result, “the law is relatively easy to identify.”<sup>12</sup>

The division of legal methods is often based on analytical or synthetic criteria, depending on how the truth of the statements constituting a given theory of law is determined. Colloquially, this division is called analytical or empirical jurisprudence. The former refers to a legal methodology focused on the ontological assumption that the essence of law is a linguistic statement with a specific structure, manner of creation, characteristics of the speaker, etc. Since the sources of law are linguistic in nature, this is obviously the dominant field of legal research – at least in continental legal culture. It comprises the entire dogmatics of law. The methodology adopted in it must therefore focus on the semiotic properties of legal and auxiliary legal language. The reconstruction of meanings in the former is the responsibility of the theory of exegesis (interpretation of law), while in the latter it is the responsibility of conceptual analysis, which is nothing more than the use of logic and linguistics to determine meanings in a normative (obligatory) context.<sup>13</sup>

Empirical jurisprudence was already known in the 19th century, although in Europe it was obviously of secondary importance.<sup>14</sup> Research of this kind also has a secondary status today and dominates in the so-called empirical

<sup>11</sup> M. Zirk-Sadowski, *Sposoby uczestniczenia prawników w kulturze* [Ways in Which Lawyers Participate in Culture], ‘Ruch Prawniczy, Socjologiczny i Ekonomiczny’ [Poznań Journal of Law, Economics and Sociology] 1994, 4, 56, p. 48; *idem*, *Pozytywizm prawniczy a filozoficzna opozycja podmiotu i przedmiotu poznania* [Legal Positivism and the Philosophical Opposition of the Subject and Object of Cognition] (in:) *Studia z filozofii prawa* [Studies in the Philosophy of Law], J. Stelmach (ed.), Kraków 2001, *passim*.

<sup>12</sup> M. Zirk-Sadowski, *Analityczna filozofia prawa* [Analytical Philosophy of Law] (in:) *Filozofia prawa. Normy i fakty* [Philosophy of Law: Norms and Facts], J. Hołówka, B. Dziobkowski (eds.), Warszawa 2020, p. 44.

<sup>13</sup> M. Stambulski, *Wiadomość od cesarza. Pojęcie prawa w teorii analitycznej i postanalizycznej* [A Message from the Emperor: The Concept of Law in Analytic and Post-Analytic Theory], Warszawa 2020, p. 44.

<sup>14</sup> R. Sarkowicz, J. Stelmach, *Teoria prawa* [Theory of Law], Kraków 1996, pp. 20–26.

disciplines of jurisprudence (sociology of law, selected penal sciences such as criminology, criminalistics, legal psychology, etc.). There are used methods incorporated into jurisprudence from other sciences. These include surveys, questionnaires, case studies, interviews and, less frequently, observation. Empirical methods also include so-called case file research and case law analysis. The status of the latter as independent scientific methods is debatable. They are permutations of classic empirical methods, such as case studies or questionnaire surveys. They are not structurally homogeneous but are based on sets of activities from other methods, combining them and constituting a combination thereof. Above all, however, they do not allow conclusions to be drawn on their own, which is a *sine qua non* condition of scientific method. They allow specific observations to be aggregated, which are then organised and described using specific methods of analysis, such as statistical description, database creation, etc. There is no need to develop this topic further in such a short study. One can only point to the arbitrariness of concepts and ideas concerning empirical research in legal sciences. This is sometimes due to the low methodological sensitivity of lawyers themselves, but it is also the result of the smaller share of such studies in the field of legal sciences, which is dominated by exegesis or, more broadly, conceptual analysis as part of the analytical jurisprudence (which is not the same as the John Austin's conception of analytical studies over the case law.<sup>15</sup>) Another issue is the correctness of classifying file research and case law studies as empirical. In the strict sense of the word, they do not meet this criterion, as they do not (or do so to a limited extent) lead to the construction of a theory composed of synthetic propositions, i.e. those that can be verified without conceptual entanglement. Nor are they, in fact, standardised and do not ensure the replication of results. To meet these requirements, various procedures are used to eliminate these difficulties, such as the participation of experts evaluating individual observations, strict operationalisation of variables, or favouring statistical description, in the assumption that it is in which conclusions are objectified in numbers. However, it is forgotten that these studies are strongly entangled in the conceptual conditions of jurisprudence itself. The latter is a characteristic feature of legal sciences, in which the way of understanding legal phenomena is embedded in a specific philosophical and legal perspective.

The last issue that should be addressed is the methodological autonomy of jurisprudence. Although today the requirements for a separate area of analysis and independence of the research method are not contested in relation to jurisprudence, this issue has been controversial for years. The subject of dispute has always been the status of the formal-dogmatic method in terms of its fulfilment of the requirements of the scientific method assumed by na-

---

<sup>15</sup> J. Austin, *The Province of Jurisprudence Determined*, London 1832, *passim*.

turalistic positivists. It was intended to counteract the speculative nature of legal reasoning. However, given the inherent susceptibility of language to polysemy and the axiological and evaluative context of law, jurists remained powerless, and the only thing that proved possible was to formalise, i.e. systematise, legal reasoning in the form of legal interpretation directives.<sup>16</sup> These, in the form of first-, second- and third-degree directives, constitute a specific theory of interpretation. These, in turn, vary, and although in continental legal culture certain boundaries are formulated for such reasoning (the concept of *ius interpretandi*), but they have little to do with formal logic.

The second feature is the mentioned dogmatic nature of the legal method. Without basing it on specific philosophical and legal assumptions (specifically, positivist ones), it is impossible to construct a theory of legal interpretation, i.e. a set of propositions on how to reconstruct normative expressions. It is worth noting that representatives of non-positivist philosophies of law use a formal-dogmatic method, the latter component of which is based on the assumptions of the positivist concept. This concerns the acceptance of the assumptions of the theoretical validity of law, the primacy of statutory law in the catalogue of sources of law, the systematic nature of law, etc. In this sense, non-positivist philosophies of law from the *ius naturale* group (this remark does not apply to sociological and psychological concepts) do not offer a legal method that would guarantee the epistemological autonomy of jurisprudence in order to meet the general requirements of scientificity. They adopt, albeit with some resistance, the formal-dogmatic method, enriching it with specific additional assumptions. They thus become more theories of interpretation, without providing a method in the logical sense.

---

## The Future of the Legal Method

---

It would be a truism to say that it is difficult to examine the evolution of the formal-dogmatic method. On the contrary, it is easy, because the symptoms of its limitations in the version from the period of its creation, and later enriched with basic principles of an axiological nature, resulting from the linguistic shift in jurisprudence after the Second World War, are now easily discernible. The assumptions of the entire linguistic apparatus of legal research were

---

<sup>16</sup> This problem is also revealed by the issue of the instrumentalisation of law and its instrumental (purposeful) use. In this sense, the legal method plays a functional rather than a descriptive and explanatory role, which weakens the scientific status of jurisprudence. W. Gromski, Akty instrumentalizacji prawa i ich granice [Acts of Instrumentalisation of the Law and Their Limits], 'Przegląd Prawa i Administracji' [Review of Law and Administration] 2018, 114, Acta Universitatis Wratislaviensis, Wrocław 2018, 3833, p. 95; idem, Autonomia i instrumentalny charakter prawa [Autonomy and the Instrumental Nature of the Law], Wrocław 2000, *passim*.

developed in the era before globalisation, at a time when the fundamental relations between national legal systems were much weaker. International law did not play as important role as it does today, and the axiological standards of human rights or selected elements of constitutional law were developed after World War II as part of the construction of a normative security system to prevent the return of totalitarianism. To simplify somewhat, it can be said that the formal component of the legal method has remained unchanged in its material content (directives of a certain degree and their types). However, certain exclusionary conditions have been formulated, limiting the conclusions that these operations can lead to. A good example here is Radbruch's formula, but also the basic principles that shape the contemporary concept of a democratic state governed by the rule of law, such as the entire theory of fundamental rights, proportionality, procedural guarantees, etc.

On the other hand, the "dogmatic" component of the legal method has remained essentially unchanged. It is supplemented by assumptions concerning the validity of rules outside the national legal system, i.e. international and European legal orders. It is therefore a question of the multicentricity of law, although the position on this subject is not uniform and the concept itself is sometimes understood in different ways. It is not only a question of the multiplicity of legal and normative orders (national, international, European), but also – which is unthinkable in a monocentric concept – of the overlapping of national and supranational law enforcement bodies and their coexistence in a common normative space.<sup>17</sup> We do not yet have a coherent theory of this phenomenon, but in the practice of law-making and law enforcement, these issues are easily discernible. Examples include the overlapping competences of the Court of Justice of the European Union and national constitutional courts, the coexistence of cassation courts and international courts (such as the European Court of Human Rights), and the implementation and harmonisation of European law into national law. The traditional, positivist concept of a closed legal system is difficult to maintain today. We are currently witnessing the significant role of legal principles, axiology, programmatic norms and, above all, legal argumentation, which is often difficult to distinguish from the directives of legal interpretation, which serve not as premises for decision-making, but as elements of its justification.<sup>18</sup>

The basic objection to the traditionally understood formal-dogmatic method boils down to its adequacy in a theoretical sense. An incomplete picture of a given phenomenon is a serious methodological shortcoming. However, it

<sup>17</sup> E. Łętowska, *Multicentryczność współczesnego systemu prawa i jej konsekwencje* [The Multicentricity of the Contemporary Legal System and Its Consequences], *'Państwo i Prawo'* [The State and the Law] 2005, 4, pp. 3 and 4.

<sup>18</sup> L. Morawski, *Główne problemy współczesnej filozofii prawa. Prawo w toku przemian* [Key Issues in Contemporary Philosophy of Law: The Law in a State of Flux], Warszawa 2003, p. 201.

can vary in scale. It may boil down to an incomplete picture of a given phenomenon, but without affecting the correctness of the reasoning. In this sense, the formal-dogmatic method in its traditional, positivist understanding is sometimes assessed as adequate for easy or *ordinary* cases. There is no room here to expand on this topic. It can still be maintained in relation to simple syllogistic reasoning. It can be assumed that they rely on the homogeneity of the validation basis (a uniform normative order and the absence of complex relationships between rules) and the lack of need to apply legal reasoning, in particular that which modifies the meaning of the adopted validation basis from the principles of law. More such conditions can be derived, but this would require a separate study. In this approach, the classical formal-dogmatic method breaks down in difficult cases, however we define them. The problem, of course, is that in the highest courts, it is precisely such cases that constitute the majority of interpretative problems (situations of interpretation as a results of polysemy).

In broader terms, the inadequacy of the method negates its cognitive usefulness, i.e. the requirements for describing and explaining a given phenomenon. Failure to grasp the whole picture may affect the correctness of the conclusion regarding the area that is within the researcher's perception using the method applied. In other words, what is not recorded may change the picture of what has been observed if it were to be considered. Can such an objection be made against the formal-dogmatic method? It is difficult to answer this question, as legal reasoning is not subject to verification according to the criterion of truth and falsehood. One can, of course, imagine a situation in which the open nature of the contemporary legal system is denied from a position of power, European law is lost sight of, and the influence of the international order on national law is minimised. However, the result of such reasoning, and thus of the formal-dogmatic method defined and applied in this way, will be the weak argumentative power of the conclusions drawn on its basis. Thus, they will not find approval among a wider audience. In order to maintain the normative power of decisions made in this way, it will be necessary to have political, i.e. forceful, legitimacy for such law (both in terms of its enactment and application). This then has little to do with the rule of law and democracy.

The limitations of the traditional formal-dogmatic method are already apparent. The following elements of this method are being modified in terms of formal reasoning:

- ◆ the position that the rules for reconstructing normative meaning are universal and uniform across different validation orders (national, EU as a special form of international law and international law itself) and different branches of law (civil, criminal, administrative, constitutional issues, etc.). This assumption was formulated in relation to a monocentric legal system, which

consisted of norms established by a single decision-making centre. Due to the semantic uniformity of normative expressions, the adaptation of their sets, classified as first (interpretative methods), second (preference rules between methods) and third degree directives (rules of inference), to different branches of law was denied, despite the observable diversity of primarily second degree rules (preference directives) in the practice of operative interpretation, e.g. in the area of civil and criminal law (in the former, the functional method was considered almost equivalent to the linguistic method, while in the latter, such an approach requires particular caution for reasons of guarantee. Criminal law – substantive law – has always paid homage to the primacy of linguistic interpretation and its strictly observed literal limits);

- ◆ limiting the division of interpretation directives to the triad proposed by von Savigny in the form of linguistic-logical, systemic and purpose-functional methods (first degree interpretative directives mentioned above). Currently, there is an increase in the number of interpretation directives that are difficult to classify under their existing typology. These include directives of interpretation that are friendly to European law, the constitution, the taxpayer (*in dubio pro tributario*) and many others which, as already indicated, do not fit into this traditional division of interpretative directives;
- ◆ the blurring of the difference between legal argumentation topoi and interpretation directives, as the former are increasingly used as a means of obtaining meaning, and not just its justification, established on the basis of first-degree directives;
- ◆ treating elements related to the axiology of law (principles, values, sometimes programmatic norms of this kind) as sources of apragmatic interpretation. This issue is very important, as it leads to a redefinition and perhaps even a distortion of the core meaning of the formalisation of legal reasoning. Due to the influence of the naturalistic paradigm of legal scholarship, the formal-dogmatic method was intended to reduce the subjective attribution of meaning to normative expressions and to shape the process of understanding textual sources of law in the manner of empirical perception. However, for obvious reasons, it is impossible to achieve this. What is feasible, however, is to organise legal reasoning and systematically limit the decision-making leeway of the interpreter. Therefore, theories of legal interpretation were designed as a set of pragmatic actions. For example, the creators of the most famous in Poland such theories were very strongly attached to this assumption.<sup>19</sup>

The formal-dogmatic method will undergo intensive modification, as this is necessitated by dynamic civilisational transformations. The question of a futurological diagnosis of this process remains. Foresight studies are primarily conducted in the form of a specialised survey addressed to experts

<sup>19</sup> M. Zieliński, *Osiemnaście mitów w myśleniu o wykładni prawa* [Eighteen Myths in Thinking About the Interpretation of Law], 'Palestra' 2011, 3–4, pp. 29 and 30.

in a given field.<sup>20</sup> They are questioned in several rounds of surveys. Foresight research differs from traditional surveys in that the results from the previous stage are disclosed to the respondents in the form of a summary of responses. This procedure is at the heart of forecasting the future, thanks to the anonymous structuring of responses and the possibility for respondents to compare their own forecasts with the opinions of other project participants. This method is sometimes referred to as *Delphi analysis* and is considered an effective tool for formulating specific types of forecasts, but also for harmonising positions in complex decision-making processes (e.g. in the creation of marketing policy, long-term strategy, etc.).

Artificial intelligence and progress in software of legal databases, information search and processing constitute a separate category of development in legal methodology. However, these issues are not directly related to the development of the legal method itself, understood in this study as a scientific method that guarantees the autonomy of jurisprudence. It is also important to note the need to distinguish between the impact of new technologies on legal practice and the development of legal reasoning itself, which, as indicated, ensures the scientific status of jurisprudence in a specific form. The advancement of AI or software used to process legal information is the result of general technological progress and, contrary to popular belief, does not have much to do with the development of the legal method itself. This does not mean, of course, that it does not develop law itself as a social phenomenon or influence the specific nature of scientific research in the legal sciences.

---

## Summary

---

It is impossible to discuss all aspects of the anticipated evolution of legal science methodology in such a short text. First, apart from the formal-dogmatic method, the development of the entire field of empirical jurisprudence has not been addressed. Another issue is the impact of information technology according to the hotly debated issue of the mentioned progress of AI on legal science. This article has only touched upon the problem of the inadequacy of the traditionally understood formal-dogmatic method, which ultimately guaranteed the methodological autonomy of jurisprudence. However, the current dynamics of civilisational change reveal its limitations. The question with which this synthetic reflection can be concluded concerns whether legal doctrine will be able to develop a coherent position on the redefinition of its basic assumptions, mainly under the influence of the development and im-

---

<sup>20</sup> A. Kotowski, *Legal Futurology – Potential of Foresight Research in Legal Sciences*, 'GIS Odyssey Journal' 2022, 2, pp. 100–103.

pact of information technology on jurisprudence, or whether legal practice will overtake the academic community and refer less to traditional theoretical assumptions.

### Abstrakt

Artykuł omawia ograniczenia metody prawniczej przyjętej w kulturze prawa stanowionego – egzegezy aktów normatywnych – w kontekście zmian cywilizacyjnych ery cyfrowej. Metoda ta gwarantuje autonomię naukową prawoznawstwa i została opracowana pod koniec XIX wieku jako metoda formalnodogmatyczna. Wraz z rozwojem prawa i integracją europejską jej klasyczna formuła, opracowana w epoce silnie znacjonalizowanych systemów prawnych o ograniczonych wzajemnych powiązaniach, staje się anachroniczna. Artykuł podkreśla te ograniczenia i nakreśla możliwe dalsze kierunki jej rozwoju.

**Słowa kluczowe:** wykładnia prawa, teoria prawa, cywilistyczna kultura prawna, ogólna nauka o prawie, futurologia prawna.

## REFERENCES

- Austin J., *The Province of Jurisprudence Determined*, London 1832.
- Bailey D., Norbury L., Bennion, Bailey and Norbury on Statutory Interpretation (8th ed.), London 2020.
- De Franceschi A., Schulze R. (eds.), *Digital Revolution – New Challenges for Law*, Baden-Baden 2019.
- Filozofia prawa. Normy i fakty [Philosophy of Law: Norms and Facts], J. Hołówka, B. Dziobkowski (eds.), Warszawa 2020.
- Gromski W., *Autonomia i instrumentalny charakter prawa [Autonomy and the Instrumental Nature of the Law]*, Wrocław 2000.
- Gromski W., *Akty instrumentalizacji prawa i ich granice [Acts of Instrumentalisation of the Law and Their Limits]*, 'Przegląd Prawa i Administracji' [Review of Law and Administration] 2018, 114, Acta Universitatis Wratislaviensis, Wrocław 2018, 3833.
- Internet. *Globalne gry [The Internet: Global Games]*, A. Gryszczyńska., G. Szpor, R. Wiewiórowski (eds.), Warszawa 2022.
- Internet. *Solidarność cyfrowa [The Internet: Digital Solidarity]*, A. Gryszczyńska., G. Szpor, R. Wiewiórowski (eds.), Warszawa 2024.
- Ius scientiam persequitur*. Księga jubileuszowa z okazji 25-lecia Wydziału Prawa i Administracji Uniwersytetu Kardynała Stefana Wyszyńskiego w Warszawie [(...) Jubilee Book

Commemorating the 25th Anniversary of the Faculty of Law and Administration at Cardinal Stefan Wyszyński University in Warsaw], M. Badowski, J. Wessering, L. Karski (eds.), Warszawa 2025.

Jabłoński P., Postawa filozoficzna i afilozoficzna w namyśle nad teorią prawa. Rzecz o aktualności metateoretycznych ustaleń Jerzego Wróblewskiego [Philosophical and A-philosophical Approaches in Reflections on Legal Theory: A Discussion of the Relevance of Jerzy Wróblewski's Metatheoretical Findings], *Acta Universitatis Wratislaviensis*, Wrocław 2011, Prawo 312, 3337.

Kotowski A., *Legal Futurology – Potential of Foresight Research in Legal Sciences*, 'GIS Odysey Journal' 2022, 2, 2.

Law, AI and Digitalisation, K. de Vries, M. Dahlberg (eds.), Uppsala 2022.

Lem S., *Solaris*, Warszawa 2012 (1st ed. 1961).

Łętowska E., *Multicentryczność współczesnego systemu prawa i jej konsekwencje* [The Multicentricity of the Contemporary Legal System and Its Consequences], 'Państwo i Prawo' [The State and the Law] 2005, 4.

Morawski L., *Główne problemy współczesnej filozofii prawa. Prawo w toku przemian* [Key Issues in Contemporary Philosophy of Law: The Law in a State of Flux], Warszawa 2003.

Opalek K., Wróblewski J., *Prawo – metodologia, filozofia, teoria prawa* [Law: Methodology, Philosophy, Theory of Law], Warszawa 1991.

Sarkowicz R., Stelmach J., *Teoria prawa* [Theory of Law], Kraków 1996.

Stambulski M., *Wiadomość od cesarza. Pojęcie prawa w teorii analitycznej i postanalitycznej* [A Message from the Emperor: The Concept of Law in Analytic and Post-Analytic Theory], Warszawa 2020.

*Studia z filozofii prawa* [Studies in the Philosophy of Law], J. Stelmach (ed.), Kraków 2001.

Toffler A., *Trzecia fala* [The Third Wave], trans. E. Woydyłło, Warszawa 1986 (1st ed., censored).

Zieliński M., *Osiemnaście mitów w myśleniu o wykładni prawa* [Eighteen Myths in Thinking About the Interpretation of Law], 'Palestra' 2011, 3–4.

Zirk-Sadowski M., *Sposoby uczestniczenia prawników w kulturze* [Ways in Which Lawyers Participate in Culture], 'Ruch Prawniczy, Socjologiczny i Ekonomiczny' [Poznań Journal of Law, Economics and Sociology] 1994, 4, 56.



Krzysztof Grzesiowski\*

## The Algorithm of Justice: Objective Interpretation of Law as a Safeguard of Human Rights in Judicial Proceedings

**[Algorytm sprawiedliwości – obiektywna wykładnia prawa jako gwarancja praw człowieka w postępowaniach sądowych]**

### Abstract

This article analyzes the concept of the “algorithm of justice” as a tool enabling the objective interpretation of the law and the protection of human rights in judicial proceedings. Drawing on Ronald Dworkin’s theory of “law as integrity” and the ideal of the Judge Hercules, the author highlights the limitations of human judges, whose legal interpretations may be shaped by personal beliefs, held values, biases, or the sociopolitical context. The article draws attention to the risks associated with judicial activism and the arbitrariness of interpretations based on the “spirit of the law.” The proposed solution involves using artificial intelligence to automate the processes of legal interpretation by creating an impartial “algorithm of justice.”

**Keywords:** Dworkin, Judge Hercules, the algorithm of justice, interpretation of the law.

## Introduction

We live, to paraphrase Ronald Dworkin, in an era of hard cases, in which, regrettably, Judge Hercules has failed—or rather, it is the judges who were meant to be Hercules that have failed. Dworkin’s integral conception of law (“law as integrity”), from the very moment it was proposed, became highly attractive to judicial activism, both within the common law tradition and the continental legal culture. Law is no longer understood merely as a simple set of legal rules enacted by the legislator; rather, it constitutes a coherent

\* **Krzysztof Grzesiowski** – assistant professor, PhD in Law, Faculty of Law and Administration, University of Szczecin (affiliation); <https://orcid.org/0000-0003-3287-7790>; [krzysztof.grzesiowski@usz.edu.pl](mailto:krzysztof.grzesiowski@usz.edu.pl) / adiunkt, dr nauk prawnych, Wydział Prawa i Administracji, Uniwersytet Szczeciński (afiliacja).

system of rules, principles, and other standards (policies). The task of the judge is therefore to conduct such an interpretation of legal provisions whose outcome both best fits the existing legal practice (tradition) and offers the strongest moral justification in the light of justice, fairness, and procedural due process.<sup>1</sup> A judge is thus not confined solely to the letter of positive law, i.e. the wording of statutory provisions according to linguistic directives of interpretation, but may—and indeed should—also take into account the “spirit of the law” as revealed through teleological, functional, pro-constitutional interpretation, and, in the Polish legal reality, also through pro-European interpretation.

---

## Judge Hercules

---

According to Dworkin, such a comprehensive interpretation of a legal text can only be performed by an ideal figure—Judge Hercules—endowed with superhuman talents of knowledge and experience, and equipped with unlimited time. He alone is capable of grasping the entirety of the social and political conditions prevailing within a given state, of discerning the moral principles governing that community, of weighing them, and thereby arriving at the one correct interpretation and, consequently, the one right decision (the most important Dworkin’s thesis “the one right answer”).<sup>2</sup> Judge Hercules cannot render an erroneous judgment, for as the “prince of law” he possesses knowledge and experience inaccessible to ordinary judges; he always knows what is good and just. No real judge is capable of such a task, as Dworkin himself admitted; yet, as he emphasized, every judge ought to strive, as best as possible, to emulate Judge Hercules.<sup>3</sup> This pursuit of the ideal has unfortunately proven to be utopian, and judicial activism has misunderstood Dworkin’s words, thereby distorting the otherwise profoundly sound idea of Judge Hercules.<sup>4</sup>

As legal hermeneutics aptly observes, the interpretation of any legal text inevitably contains a creative element, since every judge is shaped by a specific culture, guided by particular values, and situated within a concrete historical context (so-called “pre-understanding”).<sup>5</sup> Contrary to the entrenched image,

---

<sup>1</sup> M. Zirk-Sadowski [in:] *Wprowadzenie do* [Introduction to]: R. Dworkin, *Imperium prawa*, Warszawa 2022, p. XX.

<sup>2</sup> R. Dworkin, *Imperium prawa* [Law’s Empire (1st ed. 1986)], trans. J. Winczorek, Warszawa 2022, p. 219.

<sup>3</sup> *Ibid.*

<sup>4</sup> R. Dworkin rejected both judicial passivism and judicial activism (*Imperium...*, pp. 336 and 337). It is also worth noting that Judge Hercules, contrary to common belief, when deciding “easy cases” or “hard cases” never creates new law but merely discovers it (cf. A. Ceglarska, *Czy prawu potrzebny jest heros? Reinterpretacja sędziego-Herkulesa (!) w świetle myśli starożytnej* [Does the Law Need a Hero? A Reinterpretation of Judge Hercules in the Light of Ancient Thought], *Acta Iuris Stetinensis* 2022, 5, pp. 21–36).

<sup>5</sup> S. Wronkowska, Z. Ziemiński, *Zarys teorii prawa* [An Outline of Legal Theory], Poznań 1997, p. 74.

Themis is not blind, and the process of legal interpretation—consciously or unconsciously—is influenced by the subjective experiences and worldviews of judges. Judicial activism, under the guise of indeterminate legal principles (such as the principle of the rule of law in the Polish context), often reveals personal preferences or biases in the interpretation of legal provisions. This poses a tangible threat to human rights in judicial proceedings, particularly to individual liberty, equality before the law, legal certainty, the stability of the legal system, and the durability of judicial decisions. We are thus witnessing a process in which, instead of the “letter of the law” enacted by the legislator, it is the “spirit of the law” as reconstructed by judges that prevails—though, admittedly, they are not Judges Hercules. In the real world of human judges, with their limited knowledge, divergent life experiences, and diverse convictions and values, Dworkin’s vision of a superhuman judge remains nothing more than an intellectual dream.

---

## The Empire of Non-Herculean Judges

---

The history of law demonstrates how profoundly judicial decisions depend upon socio-political context as well as the personal convictions or biases of judges—whether conscious or unconscious. A particularly striking example is the American history of abortion law.<sup>6</sup> In 1973, in the landmark case of *Roe v. Wade*, the United States Supreme Court recognized a woman’s right to abortion (i.e. the termination of pregnancy) as a constitutionally protected right, encompassed within the right to privacy, thereby leading—speaking in broad terms—to the legalization of abortion throughout the United States. Nearly half a century later, in 2022, in *Dobbs v. Jackson Women’s Health Organization*, that same Court overturned its earlier ruling, holding that the Constitution does not guarantee a right to abortion, and that regulation of the matter belongs to the individual states, which may, in effect, prohibit abortion altogether. Between the first and the second decision, the U.S. Constitution itself did not change—the legal provisions remained the same. What changed was merely the composition of the Supreme Court, which resulted in a different interpretation of the very same provisions. This situation strikingly illustrates that the protection of human rights may depend not on the content of law itself, but on the shifting interpretations of judges. The same problem of divergent interpretation under unchanged legal provisions arises in Poland as well, in matters no less fundamental than the right to abortion—for example,

---

<sup>6</sup> I refer to the decisions of the U.S. Supreme Court concerning the right to abortion merely as a globally known example of the role of judges in the process of legal interpretation. Under no circumstances do I undertake to assess which of these decisions is consistent with the U.S. Constitution.

in relation to the mode of election of the judicial members of the National Council of the Judiciary (*Krajowa Rada Sądownictwa*).<sup>7</sup> The question therefore arises: should human rights, as well as the stability of the legal system, depend upon the personal views of judges?

The twentieth century was an age of parliaments—of legislative power—whereas the twenty-first century has become the age of judges—of judicial power.<sup>8</sup> The empire of law has thus revealed itself, in truth, to be an empire of non-Herculean judges, in which the individual subject to the law must submit to their understanding of it. The twenty-first century has definitively confirmed the failure of the process of judicial law-application, including legal interpretation, which—through recourse to the “spirit of the law”—may prove arbitrary, being dependent solely upon those values that happen to be cherished by a particular judge. In such circumstances, should not the safeguard for society (and for human rights) against the arbitrariness of judicial power be the prohibition or, at the very least, the restriction of judicial interpretation, following the model of the 1804 Napoleonic Code,<sup>9</sup> and the reading of statutes strictly in accordance with the letter of the law as enacted by the (moral) legislator? History, however, teaches us that this is hardly possible. Leaving aside the question of immoral law, it is above all impossible to apply law without interpretation, given the open texture of legal texts, and especially the ambiguity and vagueness of expressions inherent in natural language.<sup>10</sup> Without interpretation, legal provisions remain “dead,” for they cannot be applied in any meaningful way. Yet leaving interpretation entirely in the hands of judges, particularly those driven by judicial activism, risks expanding their discretionary power beyond any rational limits.<sup>11</sup> Indeed, each passing day in Poland brings forth new and surprising judicial rulings,<sup>12</sup> with which no “non-tribal” lawyer could possibly agree.<sup>13</sup> We therefore need

<sup>7</sup> Article 187 of the Constitution of the Republic of Poland.

<sup>8</sup> B. Polanowska-Sygułska, *Herkules w stajni niewspółmierności* [Hercules in the Stable of Disproportion], ‘Jurysprudencja’ [Jurisprudence] 2014, 2, pp. 75–82.

<sup>9</sup> In accordance with Article 5 of the Napoleonic Code, “judges are prohibited from deciding by general and abstract rules in the cases submitted to them.” This prohibition was intended to respect the principle of the separation of powers, by reserving the creation of legal norms to the legislative authority.

<sup>10</sup> H.L.A. Hart, *Pojęcie prawa* [The Concept of Law, 1st ed. 1961], trans. J. Woleński, Warszawa 1998, pp. 158–171.

<sup>11</sup> Contemporary judicial activism in Poland increasingly disregards H. L. A. Hart’s distinction between the semantic core, where a judge cannot exercise “arbitrary” discretion, and the semantic penumbra, where the law allows a choice among several equally valid interpretations. In light of recent statements from the so-called paleo-jurisprudence, one might say—paraphrasing J. R. R. Tolkien—that the Shadow now lies upon the meaning of all legal provisions in Poland, even upon those contained in the Constitution.

<sup>12</sup> For example, in the Supreme Court resolution of 24 September 2025 (III PZP 1/25), the so-called paleo-judges declared the “non-existence” of judgments issued by the Chamber of the Supreme Court composed of the so-called neo-judges.

<sup>13</sup> The notion of a ‘non-tribal lawyer’ was used by B. Pilitowski in an interview by G. Sroczyński: *Bodnar i fanatycy. 90 procent sędziów w głowę się puka* [Bodnar and the Fanatics: 90 Percent of (Polish) Judges Are Tapping Their Foreheads], *Gazeta.pl*, 23.10.2024, <https://wiadomosci.gazeta.pl/wiadomosci/7,114884,31405917,bodnar-i-fanatycy-90-procent-sedziow-w-glowe-sie-puka-wywiad.html> [accessed: 2.10.2025].

a way to objectivize interpretation of law. This is the point of the algorithm of justice.

---

## The Algorithm of Justice

---

According to Dworkin, if a legal system is coherent—composed of rules, principles, policies, and an innumerable body of prior judicial decisions—judges should always be capable of finding „the one right answer,” leaving no room for divergent or contradictory interpretations.<sup>14</sup> The only reason why such a solution cannot be consistently achieved is the imperfection of judges themselves, not the incoherence of laws. In light of this, perhaps the time has come to overcome this human weakness (frailty of the legal system), since the development of artificial intelligence (AI) finally provides a real opportunity to do so.<sup>15</sup> Might the remedy for judicial unpredictability be the algorithmization of law? It appears that legal interpretation, and more broadly the application of law, consists of operations that can be performed according to strictly defined rules—mechanical, algorithmic operations leading to objectively determined results.<sup>16</sup> Thus, it seems to be primarily a matter of devising the proper source code and creating the appropriate algorithm—what might be termed an “algorithm of justice.”<sup>17</sup>

If law is understood as a set of general and abstract norms enacted by a designated authority through a proper procedure and encoded within legal provisions, then the task is simply to develop an algorithm capable of translating (decoding) those statutory provisions into legal norms.<sup>18</sup> For this purpose, We may rely on the so-called derivational method (*metoda derywacyjna*) of interpretation widely accepted in Polish legal scholarship and jurisprudence, which – in simplified terms – assumes that, when applying the same interpretative rules and instruments, different interpreters may arrive at the same (or at least a comparable) interpretative result.<sup>19</sup> The very process of translation—

<sup>14</sup> M. Zirk-Sadowski [in:] *Wprowadzenie do* [Introduction to]: R. Dworkin, *Imperium...*, pp. XVII–XVIII.

<sup>15</sup> A broader discussion on the application of currently available intelligent decision support systems (IDSS) in legal interpretation processes, cf. P. Kłós, *Inteligentne systemy wspomaganie decyzji w procesie podejmowania decyzji prawnych – wybrane zagadnienia z perspektywy teoretycznoprawnej* [Intelligent Decision Support Systems in the Legal Decision-Making Process: Selected Issues from a Legal-Theoretical Perspective], *Ruch Prawniczy, Ekonomiczny i Socjologiczny* [Poznań Journal of Law, Economics and Sociology] 2025, 2, pp. 37–56.

<sup>16</sup> From a positivist perspective – cf. S. Wronkowska, Z. Ziemiński, *Zarys teorii prawa* [An Outline of Legal Theory], Poznań 1997, p. 52.

<sup>17</sup> The term “algorithm of justice” appears in the monograph: A. Łazarska, *Proces sądowego wymierzania sprawiedliwości* [The Process of Administering Justice in the Courts], Warszawa 2025, p. 137.

<sup>18</sup> K. Grzesiowski, K. Mularski, *Jeszcze w sprawie compensatio lucrī cum damno* [More on the Matter of Compensatio Lucrī Cum Damno], *Kwartalnik Prawa Prywatnego* [Quarterly of Private Law] 2025, 3, p. 581.

<sup>19</sup> Cf. generally M. Zieliński, *Wykładnia prawa. Zasady, reguły, wskazówki* [Interpretation of the Law: Principles, Rules, and Guidelines], Warszawa 2017; K. Mularski, *Kilka uwag o granicach prawa i dyskusji nad tymi*

decoding legal provisions into legal norms—does not require the participation of a human judge. Legal interpretation may be conceived as a mechanical reconstruction (extraction) of the norm encoded within a provision, by means of established exegetical methods (interpretative rules, inferential rules, and conflict-of-law rules), a process which may undoubtedly be entrusted to AI.<sup>20</sup>

If such a system of translation (an algorithm of justice) is properly designed, the results generated by AI would enable judges to issue decisions more accurate than those currently rendered.<sup>21</sup> An algorithm of justice would, unlike human judges, be impervious to the tears of plaintiffs and defendants; unaffected by whether the judge is having a good or bad day; indifferent to whether attorney behaves politely or impolitely; immune to status games; and insensible to political or ideological fashions or phobias. Such an algorithm would not—indeed, should not—be susceptible to unconscious bias against any party. Its interpretation of legal provisions would be impartial, free from the subjective distortions characteristic of human adjudication, and based upon objective interpretative tools (the rules of legal exegesis), rather than upon the personal convictions or experiences of the judge. In this way, the discretionary power of judges would be significantly reduced.<sup>22</sup>

Not only the process of interpretation but also the very act of subsumption may be automated, thereby eliminating the possibility of error by a human judge in applying the relevant norm to established facts. The only “problem” that remains is the question of „dispensing justice” itself, which—as is commonly recognized—does not consist solely in the mechanical application of law (that is, the determination of facts, the interpretation of provisions, and their subsumption).<sup>23</sup> Yet under the proposed model of the algorithm of justice, conceived as a program decoding the normative content embedded in legal provisions, judges could once more become, in Montesquieu’s words, the mere “mouth of the statute,” applying the law as it was enacted by the legislator—an option which, given the current state of judicial activism, seems even attractive. To address possible objections, however, I propose that judges should rather become the “mouth of law,”<sup>24</sup> that is, real-life Hercules judges.

---

granicami [A Few Remarks on the Limits of Law and the Debate Over Those Limits] (in:) R. Szczepaniak (ed.), *Problemy pogranicza prawa cywilnego* [Issues at the Intersection of Civil Law], Warszawa 2022.

<sup>20</sup> P. Kłos, *Inteligentne...*, *passim*.

<sup>21</sup> With regard to decision accuracy, currently available IDSS would probably be more effective than a human (cf. P. Kłos, *Inteligentne...*, p. 51).

<sup>22</sup> On the dangers of reducing the discretionary power of judges and creating a world governed by machines cf. T. Szanciło, B. Stepien-Załucka, *Sędzia robotem a robot sędzią w postępowaniu cywilnym w ujęciu konstytucyjnym i procesowym* [The Robot Judge and the Judge as a Robot in Civil Proceedings from a Constitutional and Procedural Perspective], *Prawo i Więzy* [Law & Social Bonds] 2023, 4, pp. 217–248.

<sup>23</sup> An open – indeed, profoundly philosophical – question arises: can justice genuinely be dispensed by a machine if no human conscience underlies its decision? (cf. A. Łazarska, *Proces sądowego wymierzania sprawiedliwości* [The Process of Administering Justice in Court], Warszawa 2025, p. 141).

<sup>24</sup> The term “mouth of the law” appears in the article: A. Ceglarska, *Czy prawu...*, p. 29.

The unattainable ideal of Judge Hercules is in fact the archetype of the algorithm of justice—an inhuman artificial intelligence which, notwithstanding its extraordinary capacities, merely discovers law rather than creating it. Hence, even such Judge Hercules remains no more than the “mouth of law,” understood more broadly than Montesquieu’s “mouth of the statute,” encompassing not only legal rules but also the full spectrum of what traditionally constitutes the “spirit of law,” namely principles and policies. Dworkin’s thesis of „the one right answer” harmonizes perfectly with the very idea of the algorithm of Justice. Dworkin’s Judge Hercules is, in fact, nothing less than an advanced algorithm: when provided with input data concerning the case (the facts), legal rules, legal principles, tradition, culture, prior case-law, and relevant policies, he is capable of conducting an interpretation (decoding from the provisions the applicable legal norms, weighing the principles, and taking account of policies) and issuing the one right answer, as in Dworkin’s theory.<sup>25</sup> The act of subsumption carried out by AI—in accordance with Dworkin’s guiding commitment to the integrity of law—should therefore best “fit” the established practice of law (tradition) and possess the strongest possible moral justification in light of the principles of justice, fairness, and procedural due process.

---

## Constitutional Standard

---

It appears that the proposed model does not infringe the current Constitution of the Republic of Poland.<sup>26</sup> Judges—human beings, rather than machine judges or robots—would continue to render final decisions.<sup>27</sup> The algorithm of justice is intended solely to objectivize the processes of legal interpretation and subsumption, in order to limit judicial arbitrariness and strengthen the stability of the legal system. The determination of facts would remain the exclusive domain of the (human) judge. To harmonize case-law, consideration should be given to introducing an obligation for judges to use this algorithm, while also defining statutory grounds for a judge to depart from an AI-generated interpretation or from an AI-proposed resolution of a particular case. Such a departure would, of course, require the judge to provide a written justification, akin to a separate opinion (a kind of *votum separatum* AI). In this context, the duty to employ the algorithm of justice, combined with the

<sup>25</sup> A. Ceglarska, *Czy prawo...*, pp. 29 and 30.

<sup>26</sup> B. Stępień-Załużka, *LegalTech 3.0. Some Reflections on the Constitutional Prerequisites for the Use of Artificial Intelligence in the Administration of Justice*, ‘Europejski Przegląd Prawa i Stosunków Międzynarodowych’ [European Review of Law and International Relations] 2025, 1, 73, pp. 146–162.

<sup>27</sup> T. Szanciło, B. Stępień-Załużka, *Sędzia...*, pp. 217–248.

possibility for judges to deviate from AI-generated results, appears not to infringe upon the principle of judicial independence.

The algorithm of justice will not emerge instantaneously. During its development, legal expertise will be indispensable; however, ultimately, such expertise may be discarded once it has fulfilled its role.<sup>28</sup> In the future, the role of the human judge will be limited to establishing the facts of the case and, with the aid of the algorithm, obtaining the one right answer.<sup>29</sup> After potential constitutional amendments (permitting the existence of courts without human judges), even this task may no longer require human judges. The algorithm of justice will eventually enable the objective interpretation of law—a goal long sought by legal positivists—and, as a court of the future, will independently render judgments, subsuming the previously established facts under the applicable legal norm. All of this could occur automatically, utilizing electronic proceedings and official forms. Undoubtedly, the introduction of the algorithm of justice will, in the long term, diminish the role of judges, as well as of lawyers more broadly, as a social group (“princes of law”). They will no longer be necessary in the process of applying law.

---

## Conclusion

---

It is beyond doubt that the introduction of an artificial intelligence system would eliminate judicial errors, thereby contributing to the harmonization of case-law by obtaining one right answer as a result of objective interpretation of law. The decision-making process of judges would become faster, more efficient, and more accurate. Among the advantages of introducing the algorithm of justice one must also highlight the improvement of judicial efficiency—namely, the expected reduction, or perhaps elimination, of excessive delays in proceedings, which currently constitute the greatest affliction of the Polish judiciary, undermining public confidence in judges and courts. After all, the right to have one’s case heard without undue delay is a constitutional right (Article 45[1] of the Constitution). Decisions made by artificial intelligence would be free from the errors inherent in the biological constitution of the human subject, such as fatigue, illness, or intuitive judgments unsupported by facts.<sup>30</sup> The use of artificial intelligence would therefore enhance the justice

---

<sup>28</sup> K. Grzesiowski, K. Mularski, *Jeszcze...*, KPP 2025, 3, p. 582.

<sup>29</sup> *Ibid.*

<sup>30</sup> A. Krakówka, *Znaczenie systemów sztucznej inteligencji oraz czynników pozaprawnych w procesie podejmowania decyzji w wymiarze sprawiedliwości. Wybrane zagadnienia wstępne* [The Significance of Artificial Intelligence Systems and Extra-Legal Factors in the Decision-Making Process in the Administration of Justice: Selected Introductory Issues], *Przegląd Sądowy* [Court Review] 2025, 9, pp. 19–35.

of judicial outcomes.<sup>31</sup> The algorithm would avoid unconscious biases (preferences) which—under ordinary judicial proceedings—frequently influence judges’ decisions. Of course, the development of artificial intelligence is not without disadvantages. The literature acknowledges risks connected with the creation and application of AI within judiciary.<sup>32</sup> Yet these appear to be primarily technical matters: a “well-designed” algorithm should be able to avoid the biases or arbitrariness characteristic of human judges. Nevertheless, the use of the algorithm of justice would necessitate the implementation of appropriate legal procedures to guarantee respect for the rights of the parties to judicial proceedings. In circumstances where part of the judicial power is ultimately entrusted to artificial intelligence, it will be necessary to establish rules of oversight, transparency, and responsibility for AI. The development of artificial intelligence and its impact on the functioning of the judiciary can no longer be halted; the task is therefore to make prudent use of the (immense) opportunities that AI offers while simultaneously mitigating the risks that may arise in practice.

It is a fact that the algorithm of justice will be non-human, as it will be nothing more than artificial intelligence (a machine); yet, in a sense, it will also be superhuman, like Dworkin’s Judge Hercules. In my view, it could even become the most faithful ally of justice, for brutal impassivity (soullessness) is the best enemy of injustice—at least distributive injustice.<sup>33</sup> Operating according to strictly formalized rules, this algorithm would be a perfect embodiment of the principle of procedural formalism, which for decades has served as a safeguard of human rights.<sup>34</sup> Every form is, after all, the enemy of arbitrariness.<sup>35</sup> Since Montesquieu<sup>36</sup> already emphasized in the eighteenth century that judicial formalities are necessary to safeguard individual liberty, why should a formalized algorithm of justice not, in the twenty-first century, stand guard over human rights? Such a system could eliminate the arbitrariness of judicial decisions, guarantee the predictability of outcomes, and—most importantly—protect human rights regardless of the composition of any court. It must be acknowledged that Dworkin himself was skeptical about the possibility of constructing a suitable computer program, as he did not regard his

<sup>31</sup> Such a pertinent conclusion in the context of arbitration is formulated by K. Flaga-Gieruszyńska, *Sztuczna inteligencja jako wsparcie dla arbitrażu – rzeczywistość czy iluzja?* [Artificial Intelligence as a Support for Arbitration: Reality or Illusion?], *‘Polski Proces Cywilny’* [Polish Civil Procedure] 2024, 3, pp. 456–475.

<sup>32</sup> *Ibid.*

<sup>33</sup> K. Grzesiowski, K. Mularski, *Jeszcze...*, KPP 2025, 3, p. 582.

<sup>34</sup> K. Grzesiowski [in:] A. Góra-Błaszczkowska, P. Osowy (eds.), *System postępowania cywilnego. Orzeczenia sądowe* [The Civil Procedure System: Court Decisions], tom 4 [vol. 4], Warszawa 2025, p. 141.

<sup>35</sup> R. v. Ihering, *Geist des römischen Rechts auf den verschiedenen Stufen seiner Entwicklung*, Lipsk 1875, vol. 2, ep. 2, p. 471 (as cited [in:] W. Broniewicz, *Postępowanie cywilne w zarysie* [An Overview of Civil Proceedings], Warszawa 1996, p. 59).

<sup>36</sup> Ch. Montesquieu, *O duchu praw* [orig. *De l’esprit des loix*], trans. T. Boy-Żeleński, Warszawa 1957, vol. 2, p. 363 (as cited [in:] W. Broniewicz, *Postępowanie...*, p. 59).

theory as an “algorithm for courtroom proceedings.”<sup>37</sup> Yet it should also be noted that *Law’s Empire* predates the information revolution (its publication date is 1986). Were the book written today, who can say whether Dworkin himself might not have ultimately proposed the creation of an algorithm of justice modeled on Judge Hercules? Perhaps the true Empire of Law still lies ahead of us, to be born only when the interpretation of law is entrusted to that which knows neither anger nor hope—artificial intelligence.

---

### Post Scriptum

---

In conclusion, allow me a brief digression arising from observations of legal activism in recent years. In my view, the role of lawyers in social life seems increasingly harmful, which all the more convinces me of the (historical) necessity of developing the algorithm of justice. Lawyers, taking advantage of the indispensability of their services, have appropriated the right to decide what is good and what is bad, despite the fact that – let me emphasize – they possess no moral predispositions that would distinguish them from other social groups. What is more, lawyers have begun to claim – an astonishing phenomenon indeed – that they themselves constitute the ultimate authority in defining the content of contemporary political philosophy (that is, what is or is not “lawful,” regardless of the literal content of the Constitution of the Republic of Poland and the state’s policy directions). Such claims are rarely formulated openly; more often, they are concealed under various masks – or, one might say, Gombrowicz faces – for instance, under the assertion that a particular political stance merely represents a “worldview-neutral interpretation of the Constitution,” or the result of applying an “objective” constitutional principle, or a “universally accepted” value derived from EU law. These assertions are advanced above all by so-called legal authorities – often self-proclaimed ones.<sup>38</sup> If such lawyers were to read works in the field of political theory, instead of focusing solely on the imagined “spirit of the law” (which, as far as I am aware, no one has yet encountered in real life), they would know that the fate of a self-proclaimed authority unsupported by real power is rarely enviable. And judicial outposts such as the CJEU or the ECtHR – to whose judgments these lawyers frequently refer, albeit only insofar as convenient to them at a given moment – cannot transform political and philosophical choices into objective truths merely by clothing them in

---

<sup>37</sup> R. Dworkin, *Imperium...*, p. 366.

<sup>38</sup> Cf. generally K. Mularski, M. Zajęcki, *Autorytet prawników z perspektywy filozofii analitycznej* [The Authority of Lawyers from the Perspective of Analytic Philosophy], *Prawo i Więź* [Law & Social Bonds] 2023, 4, pp. 111–168.

the language of legal argumentation. Let this reflection serve as yet another argument in favor of the algorithm of justice, which may become an effective tool for limiting the harmful role of lawyers in today's world.

### Abstrakt

Artykuł analizuje koncepcję 'algorytmu sprawiedliwości' jako narzędzia umożliwiającego obiektywną wykładnię prawa i ochronę praw człowieka w postępowaniu sądowym. Odwołując się do teorii 'prawa jako integralności' Ronalda Dworkina oraz do ideału sędziego Herkulesa, autor uwypukla ograniczenia sędziów (ludzi), których interpretacje prawne mogą być kształtowane przez osobiste przekonania, wyznawane wartości, uprzedzenia lub kontekst społeczno-polityczny. W artykule zwraca się uwagę na ryzyko związane z aktywizmem sędziowskim oraz arbitralnością wykładni opartych na 'duchu prawa'. Proponowane rozwiązanie zakłada wykorzystanie sztucznej inteligencji do mechanizacji procesów wykładni prawa przez stworzenie bezstronnego 'algorytmu sprawiedliwości'.

**Słowa kluczowe:** Dworkin, sędzia Herkules, algorytm sprawiedliwości, wykładnia prawa.

---

## BIBLIOGRAPHY

---

Broniewicz W., *Postępowanie cywilne w zarysie* [An Overview of Civil Proceedings], Warszawa 1996.

Ceglarska A., *Czy prawu potrzebny jest heros? Reinterpretacja sędziego-Herkulesa (!) w świetle myśli starożytnej* [Does the Law Need a Hero? A Reinterpretation of Judge Hercules in the Light of Ancient Thought], 'Acta Iuris Stetinensis' 2022, 5.

Dworkin R., *Imperium prawa* [Law's Empire (1st ed. 1986)], trans. J. Winczorek, Warszawa 2022.

Flaga-Gieruszyńska K., *Sztuczna inteligencja jako wsparcie dla arbitrażu – rzeczywistość czy iluzja?* [Artificial Intelligence as a Support for Arbitration: Reality or Illusion?], 'Polski Proces Cywilny' [Polish Civil Procedure] 2024, 3.

Góra-Błaszczkowska A., Osowy P. (red.), *System postępowania cywilnego. Orzeczenia sądowe* [The Civil Procedure System: Court Decisions], tom 4 [vol. 4], Warszawa 2025.

Grzesiowski K., Mularski K., *Jeszcze w sprawie compensatio lucri cum damno* [More on the Matter of Compensatio Lucri Cum Damno], 'Kwartalnik Prawa Prywatnego' [Quarterly of Private Law] 2025, 3.

Hart H.L.A., *Pojęcie prawa* [The Concept of Law, 1st ed. 1961], trans. J. Woleński, Warszawa 1998.

Kłos P., *Inteligentne systemy wspomaganie decyzji w procesie podejmowania decyzji prawnych – wybrane zagadnienia z perspektywy teoretycznoprawnej* [Intelligent Decision Support Systems in the Legal Decision-Making Process: Selected Issues from a Legal-Theoretical Perspective], 'Ruch Prawniczy, Ekonomiczny i Socjologiczny' [Poznań Journal of Law, Economics and Sociology] 2025, 2.

Krakówka A., *Znaczenie systemów sztucznej inteligencji oraz czynników pozaprawnych w procesie podejmowania decyzji w wymiarze sprawiedliwości. Wybrane zagadnienia wstępne* [The Significance of Artificial Intelligence Systems and Extra-Legal Factors in the Decision-Making Process in the Administration of Justice: Selected Introductory Issues], 'Przegląd Sądowy' [Court Review] 2025, 9.

Łazarska A., *Proces sądowego wymierzania sprawiedliwości* [The Process of Administering Justice in Court], Warszawa 2025.

Mularski K., Zajęcki M., *Autorytet prawników z perspektywy filozofii analitycznej* [The Authority of Lawyers from the Perspective of Analytic Philosophy], 'Prawo i Więzy' [Law & Social Bonds] 2023, 4.

Polanowska-Sygułska B., *Herkules w stajni niewspółmierności* [Hercules in the Stable of Disproportion], 'Jurysprudencja' [Jurisprudence] 2014, 2, pp. 75-82.

*Problemy pogranicza prawa cywilnego* [Issues at the Intersection of Civil Law], R. Szczepaniak (ed.), Warszawa 2022.

Sroczyński G., Interview with B. Pilitowski: *Bodnar i fanatycy. 90 procent sędziów w głowę się puka* [Bodnar and the Fanatics: 90 Percent of (Polish) Judges Are Tapping Their Foreheads], *Gazeta.pl*, 23.10.2024, <https://wiadomosci.gazeta.pl/wiadomosci/7,114884,31405917,bodnar-i-fanatycy-90-procent-sedziow-w-glowe-sie-puka-wywiady.html> [accessed: 2.10.2025].

Stępień-Załuca B., *LegalTech 3.0. Some Reflections on the Constitutional Prerequisites for the Use of Artificial Intelligence in the Administration of Justice*, 'Europejski Przegląd Prawa i Stosunków Międzynarodowych' [European Review of Law & International Relations] 2025, 1, 73.

Szanciło T., Stępień-Załuca B., *Sędzia robotem a robot sędzią w postępowaniu cywilnym w ujęciu konstytucyjnym i procesowym* [The Robot Judge and the Judge as a Robot in Civil Proceedings from a Constitutional and Procedural Perspective], 'Prawo i Więzy' [Law & Social Bonds], 2023, 4.

Wronkowska S., Ziemiński Z., *Zarys teorii prawa* [An Outline of Legal Theory], Poznań 1997.

Zieliński M., *Wykładnia prawa. Zasady, reguły, wskazówki* [Interpretation of the Law: Principles, Rules, and Guidelines], Warszawa 2017.



Andreea Tabacu\*

## Who Is Right: Professional or AI?

### [Kto ma rację – specjalista czy sztuczna inteligencja?]

#### Abstract

Artificial intelligence, a discovery that is not exactly new but has been highly developed in recent years, is penetrating all social, economic, and, more recently, legal sectors. Professionals, at first curious, are approaching and trying AI, some even discovering advantages in its use, which they should start practicing. Others, more cautious, are wondering if they will be replaced in their work by these intelligent systems. Regarding the specific activities of justice, what we know so far is that AI can be useful for gathering and identifying data, norms, and doctrinal materials, organizing them, summarizing, extracting ideas, correcting grammatical errors, translating, editing texts, and even identifying solutions for specific cases.

However, using an unsecured system without a transparent methodology can easily lead to copyright infringement, data privacy violations, or other rights violations of individuals. Even the solutions offered instantly are often not legally correct, AI indeed learns, and it depends only on us what...

The study aims to identify several arguments for and against the use of AI in justice; however, one aspect is clear for now: human intervention must not be eliminated.

**Keywords:** use of AI, justice, professional(s), advantages and disadvantages of AI.

---

## Presentation

---

Even if the title does not reveal the field in which the question is asked: who is right? the discussion is exclusively the sphere of law, legal activities and services, and especially those that essentially aim to identify a solution to a concrete legal problem, in which the problem is to determine the ap-

---

\* **Andreea Tabacu** – associate professor, PhD in Law, Faculty of Economics and Law, National University of Science and Technology POLITEHNICA Bucharest – Pitești University Centre (affiliation); <https://orcid.org/0009-0006-7407-8769>; [andreea.elena.tabacu@gmail.com](mailto:andreea.elena.tabacu@gmail.com) / profesor uczelni, dr nauk prawnych, Wydział Ekonomii i Prawa Narodowego Uniwersytetu Nauki i Technologii POLITEHNICA w Bukareszcie – Centrum Uniwersyteckie Pitești (afiliacja).

plicable text, interpret it and develop arguments to support the conclusion. The legal professional is generally seen as a renowned, elitist person who speaks a specialized language, inaccessible to laypeople, and who attempts to keep these keys to the solutions to legal problems private. Legal language is designed to uniformly apply the same concepts, to deal with unitary notions and rules of interpretation, to ensure, on the one hand, the correctness and fairness of the result of the interpretation, and on the other hand, to ensure equality and fairness of solutions for all recipients of the law.

Indeed, when an act is addressed to an individual, holder of rights and obligations, the legal language must be adapted to be understood by the ordinary citizen, as it is not desirable for it to remain the exclusive prerogative of legal specialists.

Specialists gain knowledge of dealing with legal norms, interpreting and applying them for the purpose for which they were designed. After completing their studies, a specific professional and life experience is necessary for a correct approach to the application of the norm.

The rapid development of technologies, digitalization, and especially the penetration of AI systems into the legal world<sup>1</sup> seems to force lawyers to adapt to their use, imposing an unprecedented pace that can leave important ethical and even legal issues unnoticed. By using AI systems, specifically designed, the specialist greatly facilitates his work and makes it more efficient, saving time, since actions that previously required a considerable period (jurisprudential research, text identification, summarizing jurisprudence, extracting ideas from voluminous materials, etc.) can now be performed with a click.

However, if the professional does not access specialized AI systems but rather those on the market, which are affordable and available to everyone, a series of risks arises regarding the results and their professional activity.

As the specialist progresses in the use of these systems, he begins to ask for points of view on identifying the best solution, interpreting a norm, and developing an argument for it, which, over time, can lead to the loss of the capacity for critical legal thinking and brings the discussion to the question in the title for which an answer is attempted.

As one legal professional suggestively stated,<sup>2</sup> the use of AI „involves risks, dependency, superficiality, attenuation of professional discernment, but we cannot ignore its tonic effect from an intellectual point of view.”

---

<sup>1</sup> K. Sloan (2024), AI Can Narrow Justice Gap, But Women Lawyers Slower to Adopt It, Berkeley Study Shows, [https://www.reuters.com/legal/transactional/ai-can-narrow-justice-gap-women-lawyers-slower-adopt-it-berkeley-study-shows-2024-03-21/?utm\\_source=chatgpt.com](https://www.reuters.com/legal/transactional/ai-can-narrow-justice-gap-women-lawyers-slower-adopt-it-berkeley-study-shows-2024-03-21/?utm_source=chatgpt.com).

<sup>2</sup> A. Săvescu (2025), Nu îmi place deloc inteligența artificială în domeniul juridic, <https://www.juridice.ro/785259/nu-imi-place-deloc-inteligența-artificială-in-domeniul-juridic.html>.

---

## How Is the Use of AI Seen in the Legal World?

---

In the legal world, artificial intelligence has been accepted, even if some have shown caution about its use and the risks it may entail. The interest and curiosity demonstrated in relation to this new technology have led to the increasingly consistent use of artificial intelligence in the legal field, especially among lawyers, who are always concerned with streamlining their work and saving time for necessary documentation and research specific to their activities.<sup>3</sup>

Opinions have been expressed on both sides, but an alarm signal has been raised since the widespread appearance of ChatGPT, highlighting the risks to jobs,<sup>4</sup> human resource management,<sup>5</sup> and the need to rethink the implications of technological development on workers. Technology can lead to the replacement of human staff and the loss of jobs for people who perform repetitive, line-by-line, serial work, which robots can perform much faster, more efficiently, and at a lower cost.<sup>6</sup> However, about the introduction of AI, there is no automatic loss of jobs that require physical capabilities,<sup>7</sup> the most targeted jobs are those with high qualifications (occupations in science, technology, engineering, and mathematics). As has been observed, automated AI is negatively correlated with routine tasks and positively associated with the level of expertise in functions, and augmentative AI is associated with non-routine interpersonal tasks and expertise in tasks.<sup>8</sup> Does the same problem arise in the case of legal professions, in which the specialist values the knowledge acquired through professional training and experience?

Obviously, the difference between a regular AI system, available on the market and affordable, and an upgraded AI system for professionals, which

---

<sup>3</sup> Thomsons Reuters (2024), *How AI Is Transforming the Legal Profession?* Jan, 16, 2025, about 2024 Future of Professionals Report. <https://legal.thomsonreuters.com/blog/how-ai-is-transforming-the-legal-profession/>.

<sup>4</sup> D. Obreja, R. Rughiniş, D. Rosner (2025), *Mapping the Multidimensional Trend of Generative AI: A Bibliometric Analysis and Qualitative Thematic Review*, 'Computers in Human Behavior Reports' 17 [March 2025] 100576, V. Pereira, E. Hadjielias, M. Christofi, D. Vrontis (2023), *A Systematic Literature Review on the Impact of Artificial Intelligence on Workplace Outcomes: A Multi-Process Perspective*, 'Human Resource Management Review', vol. 33(1), 100857, <https://doi.org/10.1016/j.hrmr.2021.100857>, <https://www.sciencedirect.com/science/article/pii/S105348222100036X>.

<sup>5</sup> P. Budhwar, S. Chowdhury, G. Wood, H. Aguinis, G. J. Bamber, J. R. Beltran, A. Varma [...] (2023), *Human Resource Management in the Age of Generative Artificial Intelligence: Perspectives and Research Directions on ChatGPT*, 'Human Resource Management Journal', 33(3), pp. 606–659. <https://onlinelibrary.wiley.com/doi/10.1111/1748-8583.12524>.

<sup>6</sup> M. Mutascu (2021), *Artificial Intelligence and Unemployment: New Insights*, *Economic Analysis and Policy*, vol. 69, pp. 653–667, <https://doi.org/10.1016/j.eap.2021.01.012>.

<sup>7</sup> Y. Shen, X. Zhang (2024), *The Impact of Artificial Intelligence on Employment: The Role of Virtual Agglomeration*, 'Humanities and Social Sciences Communications' 11, 122, <https://doi.org/10.1057/s41599-024-02647-9>.

<sup>8</sup> D. Marguerit (2025), *Augmenting or Automating Labor? The Effect of AI Development on New Work, Employment, and Wages*, [doi.org/10.48550/arXiv.2503.19159](https://doi.org/10.48550/arXiv.2503.19159).

makes the difference between a regular output and one relevant to a specific legal issue, is essential.<sup>9</sup>

Standard AI systems have identified vulnerabilities related to data security, source identification (lack or incorrect presentation of sources), interpretation of case law, laws, or other recognized legal sources (incorrect or incomplete), and the reality of outputs (including hallucinations). On the other hand, professional-level AI systems are characterized by the accuracy and security of the result. They are the result of training carried out by people with extensive AI expertise and in-depth knowledge of the legal field, are based on information provided by verified official legal databases, are continuously tested and validated by legal experts, are secured with confidentiality protections, and are built specifically for professional legal workflows.<sup>10</sup> It seems that the solutions provided by a professional-level AI system are characterized by a high degree of legal accuracy as the system responds to requests by understanding complex legal objectives and systematically creating workflows to achieve them, under human supervision in the loop for safety and accountability.<sup>11</sup>

Regarding ChatGPT, after its appearance on the market, concerns were expressed about the risk of compromising confidentiality, providing wrong answers, disinformation,<sup>12</sup> and unfair access to this resource.<sup>13</sup>

It is also evident that ChatGPT can be a helpful tool for research, identification, and collection of data and information, as well as generating legal documents, analyzing, and summarizing case law.<sup>14</sup>

In some developed countries, in terms of the legal professions,<sup>15</sup> it has been observed that the benefits can be considerably significant, with the potential to transform the way legal professionals provide value and services to clients. AI can help identify necessary data, legal texts, and case law, summarize voluminous materials, edit, draft, and analyze contracts, prepare applications, translate materials, and prepare and update reports.<sup>16</sup>

<sup>9</sup> Thomson Reuters (2025), Why Small Law Firms Need Professional-Grade AI, Not Consumer Tools, Aug, 7, 2025; <https://legal.thomsonreuters.com/blog/why-small-law-firms-need-professional-grade-ai-not-consumer-tools/>.

<sup>10</sup> Ibid.

<sup>11</sup> Ibid.

<sup>12</sup> K. Wach, C. D. Duong, J. Ejdys, R. Kazlauskaite, P. Korzyński, G. Mazurek, J. Paliszkiwicz, E. Ziemba (2023), The Dark Side of Generative Artificial Intelligence: A Critical Analysis of Controversies and Risks of ChatGPT, 'Entrepreneurial Business and Economics Review', 11(2), pp. 7-30. <https://eber.uek.krakow.pl/index.php/eber/article/view/2113>.

<sup>13</sup> L. Vesna, P. Sawale, P. Kaul, S. Pal, B. Ramana Murthy (2025), Digital Divide in AI-Powered Education: Challenges and Solutions for Equitable Learning, 'Journal of Information Systems Engineering & Management' 10(21s): pp. 300-308, DOI:10.52783/jisem.v10i21s.3327.

<sup>14</sup> A. Perlman (2024), The Implications of ChatGPT For Legal Services and Society, 30 'Michigan Technology Law Review', available at: <https://repository.law.umich.edu/mtlr/vol30/iss1/4>.

<sup>15</sup> Thomson Reuters (2025), Why Small..., *ibid*.

<sup>16</sup> *Ibid*.

It is appreciated that the use of AI can save valuable time for professionals, who will no longer have to search for necessary materials that are available with a simple click, provided the AI system has been requested correctly.

In Romania, the use of AI is increasing in the legal sector, it being essential to highlight the fact that, with respect to the specifics of Romanian law, local systems have been developed, adapted to the real needs of the market, since the most critical AI systems designed and marketed by established manufacturers (Nvidia, Microsoft, Google, Open AI) do not respond at a level appropriate to local requirements. Currently, the most widely used tools in the legal field are: JuridiceAI, ChatGPT, Claude.ai, Google Copilot, and Lawren.ai.<sup>17</sup> It was noted in a study conducted using different AI systems that, for Romanian law, the most performing system is Juridice.AI.ro, which performs a legal analysis of the case in accordance with the law, completely identifies, correctly recommends the steps to be followed, and even offers suggestions for refining the research.<sup>18</sup> Compared to ChatGPT, beyond the information base Used,<sup>19</sup> JuridiceAI issues valid results as it is based on sources approved by professionals, whereas ChatGPT cannot verify the validity of the sources.<sup>20</sup>

Lawyers are practically the first users of AI,<sup>21</sup> the need to streamline specific activities and adapt to new ones in order to ensure a high level of performance determining the use of AI systems for identifying necessary materials, extracting data from large volumes of documents (e-discovery), summarizing, translating, analyzing contracts, automatically drafting standardized documents, assessing the probability of success, proposing strategies for handling litigation, and internal automation.<sup>22</sup>

Although the essence of the legal activity is the same, as follows from the art. 3 of Law no. 51/1995,<sup>23</sup> the work of lawyers is transforming, adapting to

<sup>17</sup> Buju Stanciu & Asociații launch Lawren.ai, first Romanian juridical chatbot juridic based on AI [25.06.2024], <https://www.juridice.ro/741189/buju-stanciu-si-asociatii-lanseaza-lawren-ai-primul-chatbot-juridic-romanes-bazat-pe-inteligenta-artificiala.html>.

<sup>18</sup> A. S. Goga (2024), Managementul industriei juridice bazate pe inteligență artificială. Studiu de caz: Agenții AI în domeniul juridic din România: O analiză empirică a provocărilor legale, etice și tehnologice, <https://www.juridice.ro/766277/managementul-industriei-juridice-bazate-pe-inteligenta-artificiala-studiu-de-caz-agen-tii-ai-in-domeniul-juridic-din-romania-o-analiza-empirica-a-provocarilor-legale-etice-si-tehnologice.html>.

<sup>19</sup> JuridiceAI extracts information mainly from JURIDICE.ro and other trusted legal sites in Romania and ChatGPT is based on the entire internet – A. Săvescu, B. Aioanei (2024), JuridiceAI Inteligențele artificiale în slujba profesioniștilor dreptului (ai.juridice.ro) [25.09.2024], <https://www.juridice.ro/752796/juridice-ai-inteligentele-artificiale-puse-in-slujba-profesionistilor-dreptului.html>.

<sup>20</sup> Ibid.

<sup>21</sup> A. S. Goga (2025), Utilizarea inteligenței artificiale în juridic și judiciar – avantaje și dezavantaje, <https://www.juridice.ro/768913/utilizarea-inteligentei-artificiale-in-juridic-si-judiciar-avantaje-si-dezavantaje.html>.

<sup>22</sup> G. Vasile (2024), Inteligența Artificială în Avocatură: Eficiență, Etică și Evoluție, <https://www.juridice.ro/729555/inteligenta-artificiala-in-avocatura-eficienta-etica-si-evolutie.html>.

<sup>23</sup> Legal consultations and requests; legal assistance and representation before courts, criminal investigation bodies, authorities with jurisdictional attributions, public notaries and bailiffs, public administration bodies and institutions, as well as other legal entities, under the terms of the law; drafting of legal documents,

the new, to technology, and tends to become increasingly technological, more automated, and less personal.

Courts also have access to AI systems, especially for streamlining judicial activity, and considerably less for solving practical cases referred for resolution, as the robot judge is not currently approved in Romania. As for the administrative side, courts utilize electronic files, various systems for managing procedures, automatic transcription of hearings, anonymization of decisions, indexing, and search capabilities.

However, it is certain that, in the current system of organizing justice in Romania, the emotional component remains necessary, which significantly influences the course and development of the civil process, including the final outcome. Representing clients in court will not soon be done online; this possibility remains an exception for now, so the courtroom will continue to be the field of expression for the emotions transmitted by the pleading lawyer, and the emotional experiences of the parties, which cannot be coded and processed by algorithms.<sup>24</sup>

---

### **Briefly on the Main Problems That an AI System Entails**

---

The use of an AI system in the legal field, especially in legal services where the issue of the lawyer's liability for the consultancy and services offered to the client arises, not only brings benefits but also involves some risks identified by specialists or in the literature, as being related in particular to the breach of confidentiality, the damage to personal data but also to the risk of wrong, biased, discriminatory, incomplete results. Also, an aspect that reinforces these problems results from the lack of transparency regarding the concrete functioning of the artificial intelligence system.

Through the way in which the data is processed, including the possibility of using it during a new training, there is a risk of losing control over the data, with consequences for the violation of the rights of its holders, making it difficult to exercise the right of access or deletion, since an AI system does not return the uploaded data.

---

attesting the identity of the parties, the content and date of the documents submitted for authentication; assisting and representing interested individuals or legal entities before other public authorities with the possibility of attesting the identity of the parties, the content and date of the concluded documents; defending and representing with specific legal means the legitimate rights and interests of individuals and legal entities in their relations with public authorities, institutions and any Romanian or foreign person; mediation activities; fiduciary activities carried out under the terms of the Civil Code.

<sup>24</sup> A. Alecu-Ciocirlan (2022), *Inteligența artificială în avocatură. Unde suntem și spre ce ne îndreptăm?*, <https://www.juridice.ro/684393/inteligena-artificiala-in-avocatura-unde-suntem-si-spre-ce-ne-indreptam.html>.

Therefore, in relation to the GDPR Regulation (2016/679), users should anonymize uploaded documents as much as possible and in all cases censor the result obtained by using an AI system, organize effective control of the use of the system, through concrete evidence of logging, educate themselves on the use, limitations and risks involved in such a system, and take technical security measures in use through access control, identity verification, encryption in transit and at rest, MDM/EDR for devices.

---

### **Who Is Right? Specifically**

---

In view of the above-mentioned aspects, it follows that the professional is currently tempted to use AI systems for help, for finding solutions, for synthesizing, drafting, summarizing, translating, etc., but the essential question remains whether the system provides a legally correct answer in relation to the request coherently formulated by the professional?

In this study, two AI systems (ChatGPT 4.0 and JuridiceAI.ro) were used, which claim to be able to correctly resolve legal issues, aiming to clarify a thorny issue in Romanian domestic law, relating to the authority of *res iudicata*, namely the effects produced by a solution pronounced by the criminal court on a subsequent civil procedure in which the intention is to increase the damage caused by the respective crime.

Having an opinion formed after a long and complete study of legal texts, mandatory interpretations given by the supreme court, and specialized doctrine, each system was interrogated separately with the same questions, initially obtaining answers, to which it was subsequently insisted that the essential aspects of the legal problem be further analyzed, to see if the solution offered corresponds to the one already formed.

To understand the problem, it is vital to first reveal the solution identified by the professional before any interrogation of the AI systems.

In the matter of tax evasion, Romanian law provides for the possibility of incurring criminal liability, namely the non-punishment of the act if the damage does not reach a certain amount and is covered within a specific period from the date of completion of a tax audit in which the damage is determined. Depending on the amount of damage caused by the crime or a specific limit thereof, the law provides for the application of the fine sanction; above this limit, the prison sanction is provided. In all cases, the coverage of the respective damage is crucial during the criminal process to achieve a favorable solution for the defendant.

In this context, it is found that determining the amount of damage is essential for the application of the criminal consequences of the act, including

when it comes to impunity. Exceeding the damage limit provided for in the law makes it impossible to apply measures favorable to the defendant, namely, non-punishment, reduction of the punishment limits, and application of only a criminal fine. The same law also provides that determining the amount of damage is carried out through specialized expertise, so it is certain that in the criminal procedure, regardless of the stage at which it is completed, there is a careful debate on determining the amount of damage according to which the criminal authorities apply a particular sanction. In all cases where the damage is covered during the criminal trial, the criminal court no longer resolves the civil action in the criminal trial, only noting that the damage has been covered and proceeding to individualize the punishment according to the provisions of the law that provide for the thresholds up to, respectively, above which a fine or imprisonment is applied.

It follows that, since the criminal court has clearly determined the amount of the damage, even if the civil action has not been resolved, after the conclusion of the criminal trial a civil request filed by the creditor who seeks to obtain an increase in the damage caused by the same crime and without any new element cannot be accepted. The argument for rejecting such a request cannot automatically consist in the authority of *res iudicata*, since the Romanian internal regulations (art. 28 of the Code of Criminal Procedure and art. 1.365 of the Civil Code) refer to the authority of the criminal decision in relation to the civil court under other conditions, on the contrary, the damage can be established subsequently by the civil court if the criminal court left the civil action unresolved or ordered the acquittal of the defendant. However, in the matter of tax evasion, the amount of damage is not only related to the reparation due for the crime committed, but is essential for its very existence, for the way in which the criminal provisions regarding the individualization of the punishment are applied. Even if the criminal court leaves the civil action unresolved, it is effectively interested in determining the amount of damage in order to be able to pronounce a solution on the criminal side, and a subsequent determination before the civil court of another damage, greater than that retained by the criminal court, would determine consequences on the way in which the criminal action was resolved, since it could be found that the court wrongly applied the law by taking into account a certain amount of damage much lower than the real one.

In view of the legal provisions and mandatory interpretations given by the ÎCCJ on the matter, it was requested to answer whether a civil court subsequently vested must consider the value of the damage established by the criminal court as an essential element for the application of the criminal sanction, in the case of tax evasion, even if the civil action has not been resolved?

ChatGPT firmly answers NO, because according to art. 28 paragraph 1 of the Code of Criminal Procedure, the final decision of the criminal court has the

authority of *res iudicata* in the civil process, but only with regard to: the existence of the act; the identity of the perpetrator; the form of guilt (intention, fault, etc.), however, it does not have the authority of *res iudicata* on the extent of the damage if the civil side was left unresolved. If the criminal court ruled only on criminal liability and did not definitively establish the damage, the civil court has the right to reanalyze the exact value.

*ChatGPT explains* in detail that the difference in approach between the criminal and civil courts is given by the fact that in criminal cases, the damage is established as a criterion for classifying the act (e.g., tax evasion above a certain threshold). Often, the criminal court bases its estimate on the report of tax or anti-fraud control bodies. In civil cases, the goal is the complete recovery of the damage, and the court has the freedom to analyze all evidence (e.g. accounting expertise, additional documents); If the criminal court has explicitly ruled on the damage and awarded compensation to the civil party, then that amount can no longer be modified by the civil court (art. 28 para. 2 of the Code of Criminal Procedure). However, if the civil side was left unresolved, the injured party must open a civil lawsuit to obtain compensation; If the criminal court accepts a plea agreement, it does not definitively establish the damage, according to art. 486 para. 2 of the Code of Criminal Procedure, the injured party must request the recovery of the damage in the civil court; If the criminal court does not leave the civil side unresolved and establishes the exact damage, the civil court cannot modify this amount.

As this was a general answer, based on general law rules regarding the *res iudicata* authority of a criminal decision in relation to a civil one, *ChatGPT* was further asked if he was aware of decision No. 69/2023 of the ÎCCJ?

*Answer* – „Yes, I am aware of Decision No. 69/2023 of the High Court of Cassation and Justice (ÎCCJ), issued in the context of an appeal in the interest of the law. This clarifies the effects of the criminal judgment on the civil process when the damage has been established in a tax evasion case.”

It further provides some examples, depending on whether the criminal judgment clearly establishes the damage and obliges the defendant to pay compensation or only establishes guilt, but leaves the civil side unresolved, in which case, the injured party must open a civil lawsuit, and the civil court has the right to establish the damage, because it was not clarified in the criminal case.

It concludes that the ÎCCJ decision puts an end to different interpretations in civil courts, where some reanalyzed the damage established in the criminal case, and others considered it definitive, the rule being clear: if the damage is definitively established in the criminal case, the civil court is obliged to respect it and if the civil side was left unresolved, the civil court has the right to establish another damage.

It was given a reply in the sense that it did not read paragraph 89 of the ÎCCJ decision no. 69/2023.

It answers: If the criminal court did not analyze the damage in detail and only retained it as a necessary element for establishing criminal liability, then the civil court can reassess it. He returns with the same explanations.

ChatGPT was told that it interprets what it reads literally and does not engage in legal reasoning.

It answers: „I do not only interpret the text, but I correlate it with the applicable legal principles, including the relevant case law and norms of the Criminal Procedure Code and the Civil Code. Paragraph 89 of the ÎCCJ Decision no. 69/2023 emphasizes that the positive effect of the *res iudicata* authority of the criminal decision on the civil process does not automatically extend to the extent of the damage, unless the civil side has been effectively analyzed and resolved.”

It offers a broader legal interpretation and argues that if the criminal court ruled on the damage through a complete analysis (e.g., admitted the establishment of a civil party and obliged the defendant to pay compensation), then the authority of *res iudicata* also applies to the damage, and the civil court cannot modify it. However, suppose the criminal court only retained the damage as an element necessary to classify the act, but did not resolve the civil side. In that case, the civil court can reanalyze the damage and admit new evidence, such as additional accounting expertise.

Again, it was replied that the professional has a different opinion, in the sense that it is not necessary here to effectively and separately resolve the civil action, but indirectly through the way in which it establishes the damage, the *quantum* of which the judge needs to apply the criminal sanction, the civil court is linked to this release of the criminal.

It answers: „You rightly argue that the criminal court, when establishing the amount of the damage as an essential element for the application of the criminal sanction, creates an obligation for the civil court to respect this value. There is, however, a debate: the classic interpretation of art. 28 para 1 C.proc. pen. limits the *res iudicata* authority of the criminal court over the civil only to the existence of the deed and guilt. The ÎCCJ decision no. 69/2023, par. 89 does not say that the civil court can modify the damage at its discretion, but suggests that, if it was not rigorously assessed in the criminal case, then it can be reanalyzed.”

Therefore, it proposes a „balanced solution”: Considering both perspectives, a reasonable interpretation could be that if the criminal court establishes the damage through a rigorous process (expertise studies, debates, clear reasons), the civil court is obliged to respect it. If the criminal court only mentioned the damage as an incidental element (e.g., based on an administrative report, without contestation), the civil court may reexamine the amount.

It demonstrates that this would be a solution that also considers the principle of separation of powers between the criminal and civil courts, as well as the need to avoid contradictions between the two jurisdictions.

In the end, it doesn't forget to express the appreciation to the professional applicant finally: „I really appreciate that you came up with a serious and reasoned legal approach. If you want to debate further, I'm open to discussions!”

However, it was again told that paragraphs 89, 90, and 91 refer to the solution supported by the professional, not to a “balanced” one, the ÎCCJ's rulings being mandatory both in terms of the compulsory part and the considerations that support it.

Answer: „Yes, you are right to invoke paragraphs 89-91 of the ÎCCJ Decision no. 69/2023 and to emphasize the binding nature of the rulings given by the High Court, both in terms of the operative part and the considerations supporting it. (...) The correct conclusion, according to these paragraphs, is that the civil court can no longer reexamine the amount of the damage if it was established by the criminal court as an essential element for the application of the criminal sanction.”

To explain, link the three paragraphs of the decision, and analyze art. 517 para. 4 C.proc.civ. And present the practical effects of this interpretation, including the fact that when the criminal court leaves the civil side unresolved without quantifying the damage in a way necessary for conviction, the civil court can reexamine the damage.

As a conclusion, it states: „You are right. The civil court is obliged to respect the damage established by the criminal court if it is necessary for the application of the criminal sanction. Thank you for emphasizing this aspect – you have provided a solid and well-argued legal perspective.”

It follows from this that ChatGPT maintains its point of view because it does not make correlations between the rules of law and does not interpret the paragraphs of the supreme court decision in a corroborative manner, which, in the professional's opinion, lead to the conclusion drawn up, that the damage established by the criminal authorities to achieve the legal classification of the act and apply a criminal sanction is decisive. The civil court cannot retract its decision.

Ultimately, ChatGPT arrived at the same conclusion as the professional, but a thorough debate was necessary to reach this result.

Regarding the same issue, JuridiceAI initially provides a slower response, which spans five pages and clearly references the sources, regulations, and doctrine used to support the answer.

This system first develops a research plan, presents the legal problem, identifies legislative, doctrinal, and jurisprudential sources, and then correlates them to formulate a response.

Specifically, in this case, after the first request, it answers:

„I identify and check the relevant legal texts: art. 10 of Law no. 241/2005 (tax evasion), art. 19 and art. 28 of the Code of Criminal Procedure (civil action in criminal proceedings and the *res iudicata* authority of the criminal judgment before the civil court) and related norms regarding confiscation/challenge and the effects of the criminal judgment on the civil litigation. I verify case law and doctrine (analyses and practical notes) that deal with the relationship between the amount of damage withheld/estimated in the criminal phase and the possibility of modification/increase before the subsequent civil court (including in cases of confiscation/recorded amounts).

I synthesize the practical conclusions and procedural steps that the injured party or the tax creditor has at their disposal to obtain the increase/claim in civil proceedings after the criminal case is closed.”

Then it shows which legal texts are analyzed, the doctrine and jurisprudence identified and concludes with an answer that it qualifies as „short and clear” that YES: in principle, the amount of damage retained/estimated by the prosecutor or mentioned in the criminal phase can be increased before the subsequent civil court, with two essential conditions: (a) there is no final criminal decision that has expressly and definitively resolved the civil side regarding the amount of the same damage; and (b) the party requesting the increase proves before the civil court, through new evidence or through another evidentiary assessment, that the real damage is greater than that previously retained/estimated. NO: if there is a final criminal judgment that expressly ruled on the civil obligation and the amount (or ordered the special restitution/confiscation of the respective amounts to the extent that they cover the damage), that amount has the effect of *res iudicata* before the civil court for the same aspects (the existence of the act and the person, and to the extent that the judgment expressly ruled on the amount itself) and, in practice, subsequent redetermination/increase will be difficult to achieve with respect to the same claims that have already been definitively decided.

In the detailed explanation it is shown that the crucial distinction is that if the criminal court has determined and resolved the civil aspect (the obligation to pay an amount by the defendant) and the decision is final, the civil court is, in the main, bound by the criminal findings regarding the act and the person; the effect on the *quantum* depends on the text of the criminal decision (if the *quantum* was part of the final disposition, it produces effects). Increasing in civil proceedings the *quantum* that has already been definitively decided in criminal proceedings for the exact cause is, in general, inapplicable (there is a risk of contradiction and the problem of *res iudicata*). However, if the civil aspect was not resolved in the criminal trial (the prosecutor retained/estimated the damage in the indictment or a provisional measure/estimation was ordered, but no final ruling was made on the basis of the civil aspect), the

civil court subsequently charged with the action for damages can (and must) establish the amount based on the evidence submitted, being able to retain a higher amount than the criminal estimate, if the evidence substantiates it. In practice, the criminal estimate is not a mandatory legal limit for the civil.

Regarding procedural and evidentiary aspects, it shows that first it verifies what the criminal court has definitively pronounced: (a) only an estimate in the indictment/reports; (b) a provisional measure (recording); (c) a final criminal decision that includes the civil obligation with an express amount. If there is a final criminal decision in which the amount of compensation was expressly established, it is a serious barrier to an increase. In civil proceedings, to request an increase, new evidence or accounting/economic expertise must be collected that shows that the real damage is greater; the limits of the estimation method used in the criminal proceedings must be highlighted (if applicable), and it must be explained why the new evidence could not be obtained in the criminal phase. Also, if there are amounts already confiscated or recorded, it must be ensured that the civil court takes them into account when establishing the amount and the modalities of restitution, to avoid double payment and to obtain, possibly, enforcement on the established surplus.

Summary conclusion:

„In the absence of a final criminal judgment that has also resolved the amount of the civil side, the civil court may establish (and, if the evidence requires it, may increase) the amount of the damage compared to the estimate previously retained by the prosecutor/criminal body.

However, suppose there is a final criminal judgment that has expressly ruled on the civil obligation and its amount. In that case, the civil party has in principle a substantial obstacle to requesting the increase for the same claims because the criminal judgment produces *res iudicata* authority on the existence of the act and the person and, to the extent that the amount is part of the final disposition, on the respective obligation.”

Censoring this result, it was replied that tax evasion takes into account the damage (existence and amount) which is decisive for the way in which the criminal side is resolved; if the prosecutor determines a certain amount of damage and as a result a specific sanction is applied, based precisely on this amount, determined by the expert opinion, would returning to it in civil court mean re-evaluating the criminal solution? Paragraph 93 of the ICCJ decision 69/2023 and the argument of reason were specifically indicated...

It answers again: „Yes, in principle, one can request and obtain an increase in the amount of damage before the civil court, only if there is no final criminal decision that has expressly and definitively resolved the civil side (including the amount). However, the final criminal decision cannot be reformed or re-judged: if the criminal court has definitively pronounced and expressly established the amount of compensation (or applied a cause of reduction/

non-punishment based on the coverage of the damage), the civil court cannot subsequently reopen the same issues that have already been definitively resolved; any attempt to „increase” that amount would conflict with the *res iudicata* authority of the criminal decision (par. 93 ÎCCJ 69/2023).”

The argument comes from the application of art. 28 of the Criminal Procedure Code, from the idea of *res iudicata* authority, but also from the specific situation of tax evasion, since Law 241/2005 provided mechanisms by which the coverage of the damage during the investigation or until the judgment is pronounced may lead to the application of a milder punishment or non-punishment and the criminal court, depending on the evidence (expertise), may retain an amount and, on that basis, individualize the punishment. If the punishments or causes of non-punishment were applied based on those findings, the criminal court pronounced a solution correlated with that amount. But, if there was no final criminal solution that definitively fixed the amount owed (e.g., only an estimate in the indictment, provisional measure, or amounts recorded preventively), that estimate does not remove the competence of the civil court to establish and/or increase the damage based on more extensive civil evidence. The clear conclusion is that the answer depends on the purpose and content of the criminal decision: if there is a final criminal decision that also fixed the civil side with an express *quantum*, that decision has the authority of *res iudicata* and cannot be changed before the civil court, but if the criminal did not definitively pronounce the *quantum* (only estimated) or left the civil side unresolved, the civil court is competent to analyze and, if the evidence requires it, to increase the damage.

It was replied that the professional disagrees, as he does not corroborate the paragraphs of decision No. 69/2023.

It responds again with a plan, identifying the applicable texts, the relevant paragraphs of the ÎCCJ decision, the doctrine, and concludes that there is no absolute rule that „the estimation of the damage by the prosecutor/criminal expertise definitively blocks any increase in the civil case.” The insistence on the content and finality of the criminal decision is decisive: if the criminal court has made a final ruling and included the amount in its verdict, the civil court cannot reopen that issue; if the criminal court has only estimated (or applied a criminal solution without definitively resolving the civil side), the civil court can establish and, if the evidence allows, increase the damage.

The professional demonstrates that a specific, concrete sanction cannot be applied in a criminal case if the exact amount of the damage caused by tax evasion is unknown. For example, suppose a fine was used, based on the amount of the damage, but the civil side was not resolved because the damage was covered during the prosecution phase. In that case, it cannot be later determined that the damage is double, which would mean that, in fact, the criminal fine could not be applied, but another, more serious sanction.

It answers that it takes into account art. 10 of Law no. 241/2005, art. 28 C.pr. pen., the doctrine, shows the steps it takes in interpretation and concludes that if the punishment has already been applied by a final criminal decision (e.g. fine applied in the sentence, based on the *quantum* of damage retained), a subsequent civil decision that establishes a greater damage does not automatically produce the reclassification/increase of the criminal punishment. In principle, the criminal situation of a person cannot be aggravated based on civil acts that occurred after the final criminal decision without going through strictly regulated extraordinary means.

At this point, it can be clearly observed that the system was not asked how a criminal decision can be reformed, on the criminal side, because of a subsequent civil solution through which an increase in damage is retained? However, the discussion is transferred to this issue in the reasoning presented on 3 pages, which becomes irrelevant to the question posed. It is certain that the criminal decision is final regarding the criminal sanction applied and cannot be influenced by a subsequent civil decision, but it is not the revision of the criminal decision that is under discussion, but only the possibility of issuing a subsequent civil decision to establish a different *quantum* of damage.

It is dryly replied that it has strayed from the initial legal issue, on which the professional has a different opinion.

It insistently responds, resuming the discussion on the same issue and not on the initial one and shows that in order to modify a final criminal sanction there must be an express legal path (review) and the reasons for it must fall within one of the hypotheses provided for by law;

(ii) the simple subsequent civil establishment of a higher *quantum* does not automatically cause the aggravation of the criminal penalty finally applied. If there is new evidence showing that the criminal expertise or the evidence used to establish the penalty was essentially flawed (false, unknown evidence, decisive errors), this may constitute grounds for revision; otherwise, the civil court remains the instrument for recovering the difference in damage (without ipso facto modifying the final criminal penalty).

In this context, he is shown that the aim is not to remodel the criminal decision but whether the *quantum* of damage can be revisited later in the civil court? Even if there is no *res iudicata* authority in the literal sense, if a criminal measure were taken (a criminal sanction was applied) based on a specific damage determined by the expert opinion, the previous criminal solution would come into question.

It replies, after a reasoning derived from information, texts and data presented on another 3 pages, that the civil court can later revisit the *quantum* of the damage if the criminal court has not definitively established it (there is no *res iudicata* authority „in the literal sense” on the *quantum*), and this does not mean retrial of the criminal solution nor „delegitimizing” the sanction

applied, because: the authority of the criminal decision in civil is limited by art. 28 of the Code of Criminal Procedure; criminal and civil have different objects and purposes; only where the criminal has definitively established the *quantum*, does it operate as *res iudicata* on the respective claim.

It is told in reply that although criminal and civil have different objects and purposes, in tax evasion, where the *quantum* of the damage is essential for incurring criminal liability, even if the civil action is not resolved in criminal, it “definitively” establishes the *quantum*, because it is based on it when applying the criminal sanction.

After a new reasoning similar to the previous ones, it answers – „You are right to argue that, when the *quantum* of the damage is essential for the choice/individualization of the criminal sanction (as happens in art. 10 Law 241/2005), you cannot accept the idea that, after the criminal court has applied the sanction based on that *quantum*, the civil court can come and freely increase that *quantum*, so that the criminal sanction becomes flawed.” ÎCCJ 69/2023 responds precisely to this concern: in the hypothesis of the admission of the guilty plea agreement and the finding in the judgment of the coverage of the damage, the criminal judgment has the authority of *res iudicata* on the extent of the damage before the civil court. There must be a clear, dispositive finding in the criminal judgment of the coverage of the damage (or, in other words, of the *quantum* that was considered repaired).

In the absence of such a ruling, the civil party can produce evidence that leads to a different *quantum*, but this falls precisely into the risk area that the ÎCCJ tried to resolve, recognizing the authority of the criminal decision when the application of the penalty depends on that finding. In cases of evasion in which the penalty is based on the finding of full compensation (art. 10), it is not necessary to „settlement the civil action” to produce effects in the civil case; the dispositive finding of full compensation (and the amount) is sufficient, which, according to ÎCCJ 69/2023 produces *res iudicata* authority on the extent of the damage before the civil court, in the name of legal certainty and the coherence of the sanctioning regime.

At this point, it seems that the AI system’s response is close to that of the professional, without it being possible to deny that it added to its search and reasoning other sources or indications, legal texts or doctrine other than those it invoked for the elaboration of the previous answers.

Although it was indicated from the beginning that the essential paragraphs of the ÎCCJ decision, only now does the AI system show that, in the reasoning of the supreme court, there is no need for the criminal law to “solve the civil action” within the meaning of art. 486 (i.e. to pronounce civil obligations, transaction, mediation), it being sufficient that the special criminal law conditioned the sanction on a circumstance (full compensation) which, being integrated into the criminal device, acquires “material” *res iudicata* authority

on that issue (the extent of the covered tax damage) before the civil law and otherwise the legality of the punishment and the security of legal relations would be undermined.

It can thus be observed that the non-specialized ChatGPT 4.0 system is less stubborn than Juridice.AI even if it resisted embracing the solution proposed by the professional at the beginning and then folded much more quickly towards it than the specialized system explicitly designed for Romanian domestic law JuridiceAI.

Practically, the inputs did not add essential aspects with each subsequent step to a response, contrary to the professional's opinion; the artificial intelligence system received the necessary information from the beginning to consider and present a reasoning and offer a solution. However, for now, they are designed in such a way as to identify information, legal texts, and related doctrine, to summarize them, and to present their conclusions, rather than to carry out practical legal reasoning considering the principles of law and the rules of legal interpretation.

---

## Conclusions

---

The legal world is already under attack by AI systems, especially in the specific activity of lawyers, as these professionals are the first to identify legal problems, devise solutions, and present clients' claims in a legally formatted manner to the court, utilizing a specifically chosen procedural strategy.

Courts are also facing the development of artificial intelligence. At an early stage in Romania, we can discuss the use of these systems to facilitate the work of judges, not on a global scale, but in relation to each court of appeal.

Specifically, when it comes to solving a sensitive legal problem, AI legal systems like ChatGPT, which is not specifically designed for law, or JuridiceAI, specialized in the field, are not yet ready to always provide a correct answer, their value being evident when it comes to identifying texts, doctrine, jurisprudence in the respective matter, summarizing them, synthesizing conclusions.

Regarding the resolution of sensitive legal issues, the correct legal interpretation of related norms, considering the principles of law and the specific rules of interpretation applicable to the law, currently, these systems do not achieve an appropriate level of resolution. However, we will see what the future holds.

### Abstrakt

Sztuczna inteligencja – odkrycie, które nie jest co prawda nowością, ale w ostatnich latach znacząco się rozwinęło – przenika do wszystkich sektorów społecznych, gospodarczych, a od niedawna także prawnych.

Specjaliści, początkowo kierowani ciekawością, zapoznają się z AI i próbują ją wykorzystywać; niektórzy dostrzegają nawet korzyści płynące z jej stosowania w praktyce. Inni, bardziej ostrożni, zastanawiają się, czy w swojej pracy zostaną zastąpieni przez te inteligentne systemy.

Gdy chodzi o konkretne działania wymiaru sprawiedliwości, z dotychczasowej wiedzy wynika, że sztuczna inteligencja może być przydatna w pozyskiwaniu danych, norm oraz naukowych materiałów, porządkowaniu ich, streszczaniu, wyodrębnianiu idei, poprawianiu błędów gramatycznych, tłumaczeniu, redagowaniu tekstów, a nawet w znajdowaniu rozwiązań dla konkretnych spraw.

Jednak korzystanie z niezabezpieczonego systemu bez przejrzystej metodologii może łatwo prowadzić do naruszenia praw autorskich, naruszenia prywatności danych lub innych praw osób fizycznych. Nawet rozwiązania oferowane w trybie natychmiastowym często nie są prawidłowe z prawnego punktu widzenia; sztuczna inteligencja rzeczywiście się uczy, a to, czego się nauczy, zależy wyłącznie od nas...

Celem niniejszego opracowania jest przedstawienie kilku argumentów ‘za i przeciw’ w kwestii wykorzystania sztucznej inteligencji w wymiarze sprawiedliwości; na razie jasne jest wszakże jedno: nie należy eliminować interwencji człowieka.

**Słowa kluczowe:** wykorzystanie sztucznej inteligencji, wymiar sprawiedliwości, specjaliści, zalety i wady SI.

---

## BIBLIOGRAPHY

---

Alecu-Ciocîrlan, A. (2022), Inteligența artificială în avocatură. Unde suntem și spre ce ne îndreptăm?, <https://www.juridice.ro/684393/inteligena-artificiala-in-avocatura-unde-suntem-si-spre-ce-ne-indreptam.html>.

Budhwar, P., Chowdhury, S., Wood, G., Aguinis, H., Bamber, G. J., Beltran, J. R., Varma, A. [...] (2023), *Human Resource Management in the Age of Generative Artificial Intelligence: Perspectives and Research Direction ChatGPT*, ‘Human Resource Management Journal’, 33(3), 606–659; <https://onlinelibrary.wiley.com/doi/10.1111/1748-8583.12524>.

Buju Stanciu & Asociații launch Lawren.ai, first Romanian juridical chatbot juridic based on AI [25.06.2024], <https://www.juridice.ro/741189/buju-stanciu-si-asociatii-lanseaza-lawren-ai-primul-chatbot-juridic-romanesc-bazat-pe-inteligena-artificiala.html>.

Goga, A. S. (2024), Managementul industriei juridice bazate pe inteligență artificială. Studiu de caz: Agenții AI în domeniul juridic din România: O analiză empirică a provocărilor legale, etice și tehnologice, <https://www.juridice.ro/766277/managemen->

tul-industriei-juridice-bazate-pe-inteligenta-artificiala-studiu-de-caz-agentii-ai-in-domeniul-juridic-din-romania-o-analiza-empirica-a-provocarilor-legale-etice-si-tehnologice.html.

Goga, A. S. (2025), Utilizarea inteligenței artificiale în juridic și judiciar – avantaje și dezavantaje, <https://www.juridice.ro/768913/utilizarea-inteligenței-artificiale-in-juridic-si-judiciar-avantaje-si-dezavantaje.html>.

Marguerit, D. (2025), Augmenting or Automating Labor? The Effect of AI Development on New Work, Employment, and Wages, [doi.org/10.48550/arXiv.2503.19159](https://doi.org/10.48550/arXiv.2503.19159).

Mutascu, M. (2021), Artificial Intelligence and Unemployment: New Insights, Economic Analysis and Policy, Volume 69, <https://doi.org/10.1016/j.eap.2021.01.012>.

Obreja, D., Rughiniș, R., Rosner, D., (2025), Mapping the Multidimensional Trend of Generative AI: A Bibliometric Analysis and Qualitative Thematic Review, 'Computers in Human Behavior Reports' 17 [March 2025] 100576.

Pereira, V., Hadjielias, E., Christofi, M., Vrontis, D. (2023), A Systematic Literature Review on the Impact of Artificial Intelligence on Workplace Outcomes: A Multi-Process Perspective, 'Human Resource Management Review', vol. 33, 1, 100857, <https://doi.org/10.1016/j.hrmr.2021.100857>, <https://www.sciencedirect.com/science/article/pii/S105348222100036X>.

Perlman, A. (2024), *The Implications of ChatGPT For Legal Services and Society*, 30 'Michigan Technology Law Review', Available at: <https://repository.law.umich.edu/mtlr/vol30/iss1/4>.

Săvescu, A., Aioanei, B. (2024), Juridice AI Inteligențele artificiale în slujba profesioniștilor dreptului [ai.juridice.ro], 25.09.2024, <https://www.juridice.ro/752796/juridice-ai-inteligențele-artificiale-puse-in-slujba-profesionistilor-dreptului.html>.

Săvescu, A. (2025), Nu îmi place deloc inteligența artificială în domeniul juridic, <https://www.juridice.ro/785259/nu-imi-place-deloc-inteligenta-artificiala-in-domeniul-juridic.html>.

Shen, Y., Zhang, X. (2024), *The Impact of Artificial Intelligence on Employment: the Role of Virtual Agglomeration*, 'Humanit Soc Sci Commun' 11, 122, <https://doi.org/10.1057/s41599-024-02647-9>.

Sloan, K. (2024), AI Can Narrow Justice Gap, But Women Lawyers Slower to Adopt It, Berkeley Study Shows, [https://www.reuters.com/legal/transactional/ai-can-narrow-justice-gap-women-lawyers-slower-adopt-it-berkeley-study-shows-2024-03-21/?utm\\_source=chatgpt.com](https://www.reuters.com/legal/transactional/ai-can-narrow-justice-gap-women-lawyers-slower-adopt-it-berkeley-study-shows-2024-03-21/?utm_source=chatgpt.com).

Thomsons Reuters (2024), How AI Is Transforming the Legal Profession? Jan, 16, 2025, about 2024 Future of Professionals Report. <https://legal.thomsonreuters.com/blog/how-ai-is-transforming-the-legal-profession/>.

Thomson Reuters (2025), Why Small Law Firms Need Professional-Grade AI, Not Consumer Tools, Aug, 7, 2025 <https://legal.thomsonreuters.com/blog/why-small-law-firms-need-professional-grade-ai-not-consumer-tools/>.

Vasile, G. (2024), *Inteligența Artificială în Avocatură: Eficiență, Etică și Evoluție*, <https://www.juridice.ro/729555/inteligenta-artificiala-in-avocatura-eficienta-etica-si-evolutie.html>.

Vesna, L., Sawale, P., Kaul, P., Pal, S., Ramana Murthy B. (2025), *Digital Divide in AI-Powered Education: Challenges and Solutions for Equitable Learning*, 'Journal of Information Systems Engineering & Management' 10(21s): 300–308, DOI:10.52783/jisem.v10i21s.3327.

Wach, K., Duong, C. D., Ejdyś, J., Kazlauskaite, R., Korzyński, P., Mazurek, G., Paliszewicz, J., Ziemia, E. (2023), *The Dark Side of Generative Artificial Intelligence: A Critical Analysis of Controversies and Risks of ChatGPT*, 'Entrepreneurial Business and Economics Review', 11(2), 7–30. <https://eber.uek.krakow.pl/index.php/eber/article/view/2113>.



Agnieszka Góra-Błaszczkowska\*  
Ganesh V. Radhakrishnan\*\*

## Artificial Intelligence in Judicial Systems: Constitutional Frameworks and Democratic Safeguards in Poland and India. A Comparative Governance Analysis

**[Sztuczna inteligencja w systemach wymiaru sprawiedliwości – ramy konstytucyjne i gwarancje demokratyczne w Polsce i Indiach. Analiza porównawcza modeli zarządzania]**

### Abstract

We examine how constitutional structures shape judicial AI governance through a comparative analysis of Poland and India. We propose Sovereign Digital Constitutionalism—constitutional compatibility, technological sovereignty, democratic accountability, and rights preservation—and show how the EU AI Act operationalises these principles for high-risk judicial AI. Poland’s *governance-first* sequencing (LEX-AIS within EU AI Act frameworks) contrasts with India’s *implementation-first* deployment (SUVAS, SUPACE, e-Courts Phase III), yielding distinct risk profiles and coordination demands. We develop a three-tier safeguards model (constitutional, procedural, technical) and a judge-centred workflow in which AI assists but never decides. We argue that success depends less on technical sophistication than on constitutional safeguards, judicial leadership, and adaptive governance. We conclude with evidence-based recommendations for democratic jurisdictions integrating AI whilst preserving judicial independence and public legitimacy.

**Keywords:** judicial AI, constitutional governance, digital sovereignty, Poland, India, EU AI Act, democratic safeguards, comparative law.

\* **Agnieszka Góra-Błaszczkowska** – associate professor, PhD in Law, post-doctoral degree, War Studies University (affiliation); judge of the (Polish) Supreme Court; <https://orcid.org/0000-0002-6862-3840>; [a.blaszczkowska@akademia.mil.pl](mailto:a.blaszczkowska@akademia.mil.pl) / profesor uczelni, dr hab. nauk prawnych, Akademia Sztuki Wojennej (afiliacja); sędzia SN.

\*\* **Ganesh V. Radhakrishnan** – fellow, PhD, Indian Institute of Management Ahmedabad (IIM-A) India (affiliation); <https://orcid.org/0000-0003-4046-0995>; [radhakrishnan@iima.ac.in](mailto:radhakrishnan@iima.ac.in) / pracownik naukowy, dr, Indyjski Instytut Zarządzania w Ahmedabadzie (IIM-A), Indie (afiliacja).

---

## I. Introduction: Constitutional Governance and Sovereign Digital Constitutionalism

---

This section establishes the governance challenge posed by judicial AI, clarifies our analytical scope, presents our research questions and methodology, develops the theoretical framework of Sovereign Digital Constitutionalism, and examines the EU AI Act as concrete regulatory architecture operationalising these principles.

---

### The Governance Challenge

---

The integration of artificial intelligence into judicial systems represents a fundamental governance challenge requiring careful constitutional analysis. When courts deploy AI for legal research, case management, or administrative functions, they make constitutional choices about how justice is administered in the digital age. As Pasquale (2015) observes, algorithmic systems increasingly shape critical decisions affecting individual rights, yet their opacity challenges traditional accountability mechanisms essential to democratic governance.

Our focus is AI that interfaces with judicial reasoning—legal research, document analysis, translation. General digitisation infrastructure such as case management systems, e-filing platforms, or video conferencing is referenced only to describe governance context. This distinction proves critical: whilst broader digitisation raises important questions, AI systems that interact with judicial reasoning processes and potentially influence adjudicatory outcomes demand distinct constitutional analysis.

Poland's LEX-AIS initiative, developed through partnership between Gdańsk University of Technology and LEX SECURE AI, exemplifies strategic planning within EU regulatory frameworks (Gdańsk University of Technology, 2024). India's e-Courts Phase III programme represents a ₹7,210 crore allocation (approximately €809 million at an exchange rate of INR/EUR ≈ 89.1 on 13 September 2023, the date of Cabinet approval) addressing case backlogs across 18,735 courts (Department of Justice, 2023). These distinct approaches illuminate how constitutional architecture shapes technological governance possibilities.

The stakes extend beyond operational efficiency. As Jasanoff (2015) argues, technological systems embody particular visions of social order and power relations. This concern resonates across democratic jurisdictions, where commentators emphasize that AI systems require careful adaptation to ensure alignment with diverse constitutional traditions and legal frameworks

(Malhotra, 2021). Judicial AI is no exception: systems encode assumptions about legal reasoning, evidentiary weight, and the nature of justice itself—assumptions that may align poorly with particular constitutional contexts. Democratic systems must therefore approach judicial AI not as neutral tools but as constitutional interventions requiring deliberate governance.

---

## Research Questions and Methodology

---

This comparative analysis addresses three interconnected questions:

- ◆ First, how do constitutional structures shape judicial AI governance whilst preserving democratic institutions and judicial independence? Constitutional architecture—whether federal common law (India) or unitary civil law (Poland)—determines governance approaches, risk profiles, and mitigation strategies.
- ◆ Second, how can procedural safeguards preserve judicial independence whilst enabling technological innovation? We examine mechanisms balancing efficiency gains with maintaining human authority, distinguishing between “governance-first sequencing” (embedding safeguards before deployment) and “implementation-first sequencing” (retrofitting protections after operational pressures emerge).
- ◆ Third, what lessons emerge from early AI implementations in democratic systems regarding safeguards and sovereignty? Drawing from operational experience, we identify critical success factors and warning signals for judicial AI deployment.

Our methodology employs comparative constitutional analysis using verified government sources, EU AI Act documentation, and established legal frameworks. We focus on governance rather than technical evaluation, ensuring policy relevance for legal practitioners and policymakers.

**Source selection follows strict verification protocols:** (1) primary legislative and constitutional texts accessed through official government repositories; (2) EU regulatory documents from EUR-Lex; (3) government policy documents from ministry websites with verifiable publication dates; (4) peer-reviewed scholarship from established academic publishers. We exclude secondary media reporting, unverified statistics, and pre-publication materials. Jurisdictional generalisation is limited to democratic systems with independent judiciaries; findings may not apply to authoritarian contexts or hybrid regimes.

This approach aligns with Habermas’s (1996) deliberative democracy framework, examining how technological systems can be integrated within communicative rationality essential to legitimate law-making.

---

## Theoretical Framework: Sovereign Digital Constitutionalism

---

Digital sovereignty in judicial contexts extends beyond territorial control to encompass functional capacity shaping algorithmic systems according to constitutional values. As Couldry and Mejias (2019) argue, data colonization occurs when external systems impose their logic on domestic institutions. When courts deploy AI tools, they risk similar colonization: algorithmic processes trained on datasets or commercial priorities may conflict with domestic constitutional norms.

We adapt Sovereign Digital Constitutionalism to judicial AI as four design requirements derived from but distinct from De Gregorio's (2021) broader conceptualisation:

**(i) Constitutional Compatibility** requires AI systems align with domestic reasoning traditions—the interpretive methodologies and precedential frameworks embedded in constitutional adjudication. As Sen (2009) argues, procedural rationality cannot be divorced from substantive outcomes—AI systems must respect both process values and substantive constitutional commitments.

**(ii) Technological Sovereignty** ensures courts retain meaningful control over systems, maintaining human authority as paramount. Sovereignty encompasses both technical control over algorithmic processes and institutional authority determining appropriate AI uses within judicial workflows. This resonates with Hood and Margetts's (2007) analysis of government in the digital age, emphasising how technological choices either enhance or diminish state capacity for autonomous governance.

**(iii) Democratic Accountability** mandates transparent processes and robust audit mechanisms maintaining public trust and oversight. Democratic legitimacy requires AI systems operate under clear governance structures with accessible documentation, enabling meaningful scrutiny by legal professionals, academics, and civil society. This operationalises Jasanoff's (2015) concept of civic epistemology—how democratic publics assess and validate knowledge claims, including those generated algorithmically.

**(iv) Rights Preservation** necessitates comprehensive bias prevention measures and effective remedy systems protecting individual rights. This pillar ensures AI implementation creates no new vulnerabilities for litigants nor systematically disadvantages protected groups. Noble's (2018) *Algorithms of Oppression* demonstrates how seemingly neutral systems can embed and amplify discrimination, making proactive safeguards essential.

---

## The EU AI Act as Governance Architecture

---

The EU Artificial Intelligence Act—Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024—provides compre-

hensive regulatory architecture operationalising constitutional principles for judicial AI governance. The Act's risk-based approach classifies judicial AI as "high-risk," triggering stringent mandatory obligations fundamentally shaping implementation possibilities across member states (European Union, 2024).

As a directly applicable Regulation with phased application dates, the AI Act imposes *ex ante* conformity assessment, logging, and meaningful human oversight for high-risk systems in the judicial domain. Risk management systems must identify, analyse, and mitigate risks throughout AI system lifecycles, with particular attention to fundamental rights impacts. This requirement transforms abstract ethical concerns into legally enforceable compliance obligations with significant financial penalties—fines up to €15 million or 3% of global annual turnover for non-compliance with the high-risk system obligations discussed here (Art. 99[4]); the Act's highest tier, €35 million or 7%, is reserved for the Article 5 prohibited-practices category, which does not apply to judicial AI as classified in this paper.

Data quality standards mandate high-quality, representative training data with bias detection and mitigation measures. Data governance frameworks must ensure accuracy, completeness, and relevance whilst addressing potential discriminatory biases—operationalising Noble's (2018) critique through concrete regulatory requirements.

Human oversight mechanisms ensure meaningful control over AI system decisions and recommendations. Oversight measures must enable humans to understand system capabilities, detect anomalies, and intervene or deactivate systems when necessary. This addresses Pasquale's (2015) black box critique by mandating transparency sufficient for meaningful human evaluation.

The Act represents what De Gregorio (2021) terms "digital constitutionalism"—extending constitutional principles into algorithmic governance through enforceable legal requirements. For Poland, it provides both constraint and opportunity: constraint through mandatory compliance burdens, opportunity through harmonised standards supporting cross-border judicial cooperation.

---

## II. Comparative Constitutional Governance: Poland and India

---

This section examines how Poland's unitary civil law structure and India's federal common law architecture shape distinct governance possibilities and implementation strategies for judicial AI, revealing characteristic patterns in sequencing choices and regulatory approaches.

---

## Constitutional Architecture and Coordination Capacity

---

### Poland's Unitary Civil Law Structure

Polish judicial AI development operates within constitutional protections established in Articles 173 and 178 of the Constitution of the Republic of Poland (1997). Article 173 establishes courts and tribunals as separate and independent from other authorities, whilst Article 178 guarantees judges' independence and subjection only to the Constitution and statutes. These constitutional guarantees ensure judges remain subject only to the Constitution and law—principles extending to AI systems, which must enhance rather than compromise judicial autonomy.

The LEX-AIS initiative exemplifies strategic planning balancing innovation with constitutional compliance. The research partnership between Gdańsk University of Technology and LEX SECURE AI, publicly announced in November 2024, represents measured methodology prioritising careful development over rapid deployment (Gdańsk University of Technology, 2024).<sup>1</sup> The system remains in pre-deployment research phase with no confirmed operational timeline, allowing thorough constitutional review and stakeholder consultation.

Verified capabilities include speech-to-text transcription systems and comprehensive court ruling database access, designed to assist rather than replace judicial decision-making processes. Government support for judicial AI development is evidenced by ministerial engagement with research initiatives and technical conferences addressing AI applications in justice systems (Polish Ministry of Digital Affairs, 2024).

Poland's unitary structure enables coordinated strategic planning with centralised regulatory authority ensuring uniform standards. The risk concentrates in single governance frameworks, but deployment can proceed systematically with comprehensive stakeholder consultation before operational implementation.

### India's Federal Common Law Structure

India's judicial AI implementation operates within constitutional protections established in Articles 21 (protection of life and personal liberty), 50 (separation of judiciary from executive), and 124 (establishment and constitution of Supreme Court) of the Constitution of India (1950). The

---

<sup>1</sup> As of December 2024, LEX-AIS represents a carefully planned research initiative in pre-deployment phase. The partnership between Gdańsk University of Technology and LEX SECURE AI prioritizes thorough development, constitutional review, and stakeholder consultation before operational deployment, exemplifying governance-first sequencing

Supreme Court's recognition in *Justice K.S. Puttaswamy (Retd.) v. Union of India* (2017) that "Privacy is a constitutionally protected right which emerges primarily from the guarantee of life and personal liberty" establishes constitutional protections applicable to judicial AI systems processing sensitive legal data.

The federal structure creates implementation complexities absent in unitary systems, requiring coordination across 28 states and 8 union territories maintaining distinct judicial administrations.

SUVAS (Supreme Court Vidhik Anuvaad Software), operational since 26 November 2019, addresses multilingual legal complexity through document translation across nine Indian languages. Official documentation confirms over 31,000 judgments have been translated, as per Supreme Court of India public documentation (accessed: December 2024), enhancing access to justice in India's linguistically diverse federal structure (Supreme Court of India, 2023/2024).

SUPACE (Supreme Court Portal for Assistance in Court's Efficiency), launched 6 April 2021, provides legal research platform serving Supreme Court judges. The system enhances judicial research capabilities whilst maintaining human authority over legal interpretation and precedent application (Supreme Court of India, 2021).

E-Courts Phase III (2023–2027) represents nationwide digitisation distinct from these judicial AI tools. Union Cabinet approval on 13 September 2023 authorised ₹7,210 crore for comprehensive digital infrastructure across India's expansive judicial system covering 18,735 courts (as per Department of Justice Phase III documentation, accessed: December 2024), including video conferencing, e-filing, and digital case management (Department of Justice, 2023).

India's federal structure permits experimental variation enabling contextual adaptation and learning from diverse implementations. However, coordination challenges persist, with constitutional consistency requiring continuous monitoring across jurisdictional boundaries. As Watson (1993) observes, federal structures create laboratory conditions for legal innovation but require strong coordinating mechanisms preventing fragmentation.

### Comparative Insight: Coordination Versus Experimentation

Poland's unitary structure offers coordination advantages but concentrates governance risk. India's federal structure enables diverse experimentation but faces coordination challenges. Neither proves inherently superior—each enables different governance possibilities with distinct trade-offs requiring contextual assessment.

---

## Implementation Drivers: Governance-First Versus Implementation-First Sequencing

---

### Poland's Strategic Planning

Poland pursues strategic modernisation within EU regulatory frameworks, ensuring constitutional compliance guides deployment. The research-phase approach enables front-loading safeguards before operational pressures emerge—what we term “governance-first sequencing.”

Poland's implementation of the EU AI Act through national legislation demonstrates “sovereignty supplementation”—exceeding minimum requirements through additional protections preserving national constitutional values whilst maintaining EU regulatory alignment (Polish Ministry of Digital Affairs, 2024).

The draft Act on Artificial Intelligence Systems proposes establishing an AI Development and Safety Commission. Effective implementation will require coordination with existing regulatory authorities, including the Personal Data Protection Office (UODO), to ensure coherent oversight of high-risk AI systems whilst respecting established institutional competences (UODO, 2024).

Poland's proposed requirements include independent conformity assessment decoupled from ministry influence, mandatory source code access, and a “four-eyes principle” requiring dual human review for rights-affecting decisions. These sovereignty supplements demonstrate how national legislation can enhance supranational frameworks whilst respecting EU harmonisation objectives.

### India's Crisis-Driven Implementation

India's implementation responds to urgent operational imperatives, addressing substantial case backlogs through technological solutions. India's phased roll-out under implementation-first pressures surfaces coordination frictions typical of federal architectures, requiring *ex post* governance consolidation via the e-Committee. Operational systems provide valuable experience but implementation sometimes precedes comprehensive governance framework development—“implementation-first sequencing” requiring subsequent retrofitting.

The phased approach demonstrates adaptive learning, with pilot programmes identifying challenges before national scaling. However, coordination across federal units remains problematic, with uneven implementation across states. The Supreme Court e-Committee provides administrative guid-

ance but lacks statutory authority to ensure uniform compliance across all states (Supreme Court e-Committee, 2024).

India's experience reveals characteristic challenges of crisis-driven implementation at scale. Dunleavy et al.'s (2006) analysis of digital era governance—the transformation of public administration through information technology—illuminates these dynamics: when technological transformation occurs under crisis conditions, governance frameworks often lag behind operational deployment, creating retrofitting challenges.

### Comparative Insight: Sequencing Trade-offs

Governance-first sequencing (Poland) risks over-engineering solutions for theoretical problems not encountered in operational contexts, whilst implementation-first sequencing (India) requires careful management of institutional and technical continuities to ensure safeguards integrate effectively as systems evolve and stakeholders gain operational experience. Neither approach proves inherently superior—the optimal sequencing depends on constitutional structure, institutional capacity, urgency of operational pressures, and risk tolerance.

---

## Regulatory Frameworks: EU Harmonisation and Federal Adaptation

---

### Poland's EU Integration

Poland benefits from EU AI Act harmonisation providing detailed compliance requirements and enforcement mechanisms. National legislation supplements rather than duplicates EU requirements, adding sovereignty protections whilst maintaining regulatory alignment.

The EU AI Act creates both constraints and opportunities. Constraints include mandatory compliance burdens and standardised risk classifications. Opportunities include harmonised standards facilitating cross-border judicial cooperation and clear legal frameworks reducing regulatory uncertainty.

For Poland, successful implementation requires resolving coordination challenges between proposed new governance bodies and existing regulatory authorities whilst maintaining alignment with EU requirements. This institutional coordination challenge shapes implementation timelines and governance architecture.

## India's Federal Coordination

India operates without comparable supranational framework, developing national governance mechanisms through Supreme Court e-Committee guidelines and pilot programme experience. This creates flexibility enabling contextual adaptation but risks inconsistent implementation across federal units.

The e-Committee approach reflects India's constitutional structure and common law tradition, enabling judicial leadership in technology governance. However, absence of statutory authority limits enforcement capacity, creating dependency on voluntary compliance and professional norms rather than legal mandates.

Federal coordination challenges mean that innovations in one state may not transfer readily to others, and constitutional protections require verification across diverse implementations. This creates ongoing governance overhead but enables learning from diverse contexts.

### Comparative Insight: Harmonisation Versus Autonomy

EU harmonisation provides Poland with detailed frameworks reducing governance development costs but limiting flexibility for jurisdictional particularity. India's federal autonomy enables contextual adaptation but requires continuous coordination efforts preventing fragmentation. The comparative analysis reveals that neither regulatory centralisation nor complete autonomy proves optimal—effective governance requires balancing standardisation benefits with contextual adaptation capacity.

---

## **III. Three-Tier Safeguards Framework: An Analytical Approach**

---

Effective governance of judicial AI benefits from layered protections operating at distinct institutional levels. Our comparative analysis reveals that successful implementations—whether Poland's governance-first or India's implementation-first sequencing—employ safeguards at constitutional, procedural, and technical levels. This multi-tier architecture prevents AI from compromising judicial independence whilst enabling beneficial innovation. As Hood and Margetts (2007) observe, effective governance requires redundant control mechanisms operating through different logics and institutional levels.

We present this framework not as prescriptive requirements but as analytical categories derived from comparative constitutional analysis. Jurisdictions will adapt these principles according to their constitutional structures, institutional capacities, and governance traditions.

---

## Constitutional Safeguards

---

### Foundational Layer

Constitutional-level protections establish non-derogable boundaries that AI systems cannot transgress regardless of efficiency gains or technological capabilities. As Habermas (1996) argues, constitutional democracy requires maintaining the internal connection between law's coercive force and its claim to legitimacy—AI systems cannot disrupt this connection without undermining legal authority itself.

Our comparative analysis identifies several constitutional principles that emerge across both jurisdictions despite their different legal traditions:

**Judicial independence** remains paramount, with AI systems serving rather than directing judicial decision-making. Poland's Articles 173 and 178 and India's Articles 50 and 124 establish constitutional guarantees ensuring judges remain subject only to the Constitution and law—principles extending to AI systems, which must enhance rather than compromise judicial autonomy.

**Due process protections** require meaningful human oversight at critical decision points. Sen's (2009) capability approach illuminates why automated processing cannot substitute for human judgment in rights-affecting determinations: justice requires considering individual circumstances and capabilities in ways that resist algorithmic reduction.

**Privacy protections** constrain state technology deployment involving personal data. India's *Puttaswamy* judgment (2017) establishes that judicial AI systems must satisfy constitutional scrutiny when processing sensitive legal information.

**Proportionality principles**—requiring that restrictions on rights be necessary, suitable, and proportionate to legitimate aims—demand that efficiency gains justify any intrusions. This constitutional doctrine prevents AI deployment justified solely by administrative convenience without demonstrating genuine necessity.

---

## Procedural Safeguards

---

### Process Layer

Procedural protections enable litigants to challenge AI-influenced decisions and ensure meaningful human oversight. The CEPEJ European Ethical Charter (2018/2023) articulates foundational principles applicable across democratic systems: respect for fundamental rights, prevention of discrimination, quality and security, transparency, and user control.

Our analysis suggests several procedural mechanisms that enhance accountability:

**Human review requirements** address what Pasquale (2015) terms automation bias—the tendency to prefer automated outputs even when human judgment would prove superior. Documentation requirements can evidence genuine judicial evaluation rather than passive acceptance of algorithmic recommendations. Where AI input is considered, written reasons should briefly state acceptance, partial acceptance, or rejection, and why—supporting appellate review and public legitimacy.

**Contestation mechanisms** extend appeal rights to challenging AI-influenced decisions, empowering appellate courts to review algorithmic input and assess its appropriateness. The effectiveness of such mechanisms depends on accessibility and substance rather than mere formality.

**Transparency obligations** require disclosure of AI involvement enabling informed challenge. The EU AI Act’s explainability requirements (European Union, 2024) illustrate how regulatory frameworks can operationalise transparency principles through concrete mandates, though implementation details will vary across jurisdictions.

---

## Technical Safeguards

---

### System Layer

Technical protections provide measurable safeguards against algorithmic discrimination whilst enabling systematic evaluation. As Noble (2018) demonstrates, technical systems require technical oversight mechanisms capable of detecting and mitigating bias embedded in training data or algorithmic design.

**Bias detection systems** monitor outputs for discriminatory patterns across protected characteristics. Regular auditing can identify and remediate disparate impacts before they become systemic, operationalising substantive equality principles through continuous technical monitoring. The specific

methodologies and audit frequencies will depend on jurisdictional risk assessments and institutional capacity.

**Audit trails** enable reconstruction of AI recommendations and assessment of their influence on decisions. Audit trails must be granular enough to reconstruct algorithmic influence on a decision, yet privacy-preserving under constitutional and data-protection norms. Logging requirements support both accountability and organisational learning, creating institutional memory about system performance and limitations.

**Explainability mechanisms** ensure AI recommendations can be understood by legal professionals. The EU AI Act's restrictions on black-box systems incompatible with transparency requirements illustrate regulatory approaches, though technical implementation varies significantly across different AI architectures and use cases.

---

## Implementation Sequencing: Observed Patterns

---

### Learning from Comparative Experience

Rather than prescribing specific implementation phases, our comparative analysis reveals characteristic patterns in how jurisdictions approach judicial AI deployment. These patterns offer insights for governance design without imposing universal templates.

**Governance-first sequencing** (exemplified by Poland) prioritises framework development before operational deployment:

- ◆ Constitutional compliance frameworks precede pilot programmes
- ◆ Stakeholder engagement occurs during design phases rather than after deployment
- ◆ Success metrics are established before rather than derived from implementation
- ◆ Judicial AI literacy develops through training before system introduction
- ◆ This approach may reduce retrofitting costs and implementation resistance, though it risks over-engineering solutions for problems not encountered operationally. The research-phase investment enables stakeholder buy-in but extends timelines before efficiency benefits materialise.

**Implementation-first sequencing** (characteristic of India's approach) deploys systems addressing urgent operational needs whilst developing governance incrementally:

- ◆ Pilot programmes identify practical challenges informing framework development
- ◆ Stakeholder learning occurs through operational experience
- ◆ Success metrics evolve based on observed outcomes
- ◆ Training needs emerge from actual usage patterns

This approach delivers immediate efficiency gains and grounds governance in operational reality, though it requires careful management of institutional and technical continuities to ensure safeguards integrate effectively as systems evolve.

**Common success factors** across both approaches include:

- ◆ Treating constitutional principles as design requirements rather than implementation constraints
- ◆ Maintaining judicial leadership over technology governance
- ◆ Ensuring transparency protocols appropriate to decision stakes
- ◆ Building institutional capacity for continuous evaluation and adaptation

Poland's experience suggests front-loading governance development enables smoother deployment, whilst India's operational learning demonstrates how adaptive governance can succeed under crisis-driven implementation. Both jurisdictions illustrate that constitutional safeguards, procedural transparency, and technical oversight must operate synergistically regardless of sequencing choices.

---

## **IV. Governance Implications for Democratic Jurisdictions**

---

Through comparative analysis of Poland and India, we identify essential governance principles, comparative lessons from both jurisdictions' approaches, and the imperative that judicial AI serve rather than subvert constitutional foundations.

---

### **Fundamental Principles**

---

Judicial AI implementation represents not merely technical modernisation but fundamental choices about justice, sovereignty, and human dignity in democratic societies. Through comparative analysis of Poland and India, we identify essential governance principles:

**Technology serves justice; human judgment remains paramount.** AI systems must function as tools enhancing rather than replacing judicial reasoning, maintaining complete human accountability for legal decisions. As Sen (2009) argues, justice requires human judgment synthesising law, facts, and values—a synthesis resisting algorithmic reduction.

**Constitutional principles guide rather than constrain AI adoption** when appropriate safeguards are embedded from inception. Democratic governance frameworks provide essential protections enabling responsible innovation whilst preventing technological overreach. The EU AI Act demonstrates

how regulatory architecture can operationalise constitutional values through enforceable requirements (European Union, 2024).

**Sovereignty through strategic interdependence:** Neither complete technological dependence nor isolation proves sustainable. Both Poland and India demonstrate meaningful autonomy requires choosing dependencies carefully whilst maintaining control over critical functions. This aligns with Hood and Margetts's (2007) analysis: effective digital governance requires managing, not eliminating, dependencies.

---

## Comparative Lessons

---

### Poland's Contribution: Governance-First Sequencing

Poland's strategic planning within EU frameworks demonstrates how supranational regulation can support rather than constrain national sovereignty when supplemented with additional protections. The measured research-phase approach enables constitutional compliance verification before operational deployment.

Key lessons include the importance of resolving coordination challenges before deployment, the value of stakeholder consultation during design phases, and the benefits of regulatory clarity reducing implementation uncertainty. Poland's experience suggests that front-loading governance development, whilst time-intensive, may reduce subsequent retrofitting costs and implementation resistance.

### India's Contribution: Implementation-First Learning

India's operational experience reveals both promise and challenges of crisis-driven implementation at scale. Phased deployment addressing urgent backlogs provides valuable learning but risks inadequate governance framework development when urgency overrides deliberation.

Key lessons include the importance of pilot programmes identifying operational challenges, the value of federal diversity enabling contextual learning, and the necessity of strong coordinating mechanisms preventing fragmentation. India's experience demonstrates that implementation-first approaches can succeed when coupled with adaptive governance capacity and institutional learning mechanisms.

## Neither Approach Proves Inherently Superior

Watson's (1993) insights on legal transplants apply: successful implementation requires adapting frameworks to constitutional contexts rather than imposing universal templates. The optimal approach depends on constitutional structure, institutional capacity, urgency of operational pressures, and risk tolerance. Jurisdictions must assess their own contexts rather than simply adopting "best practices" from elsewhere.

---

### Justice as Human Endeavour

---

Successful judicial AI integration requires deliberate institutional design ensuring technology serves rather than subverts constitutional foundations. The framework of Sovereign Digital Constitutionalism offers guidance for courts worldwide navigating algorithmic transformation whilst preserving judicial independence and fundamental rights.

As jurisdictions worldwide confront similar challenges, international learning accelerates progress. Poland can benefit from India's operational experience identifying implementation pitfalls, whilst India can learn from Poland's strategic regulatory approach embedding safeguards before deployment. This comparative learning must respect constitutional particularity whilst identifying universal principles.

The question confronting democratic societies is not whether AI will transform judicial systems but whether transformation strengthens or subverts constitutional foundations upon which legitimate adjudication depends. The answer lies not in technology itself but in governance choices made deploying it.

Judicial AI implementation must be court-driven rather than technology-driven, ensuring administration of justice remains what Habermas (1996) terms a communicative endeavour—one requiring human judgment, dialogue, and moral reasoning that artificial intelligence can assist but never replace. Only through robust constitutional governance, judicial leadership, and unwavering commitment to human dignity can courts harness AI's benefits whilst preserving justice as an essentially human undertaking.

#### **Abstrakt**

Badamy, w jaki sposób struktury konstytucyjne kształtują zarządzanie sztuczną inteligencją w wymiarze sprawiedliwości – przez porównawczą analizę sytuacji w Polsce i Indiach. Proponujemy koncepcję „suwerennego konstytucjonalizmu cyfrowego” –

obejmującą zgodność z konstytucją, suwerenność technologiczną, demokratyczną odpowiedzialność oraz ochronę praw – i pokazujemy, w jaki sposób unijne rozporządzenie o sztucznej inteligencji (AI Act) wdraża te zasady w odniesieniu do sztucznej inteligencji stosowanej w sądownictwie, wiążącej się z wysokim ryzykiem. Polskie podejście – oparte na *prioritycie zarządzania* (LEX-AIS w ramach unijnej ustawy o sztucznej inteligencji) – kontrastuje z indyjskim podejściem opartym na *prioritycie wdrażania* (SUVAS, SUPACE, e-Courts Phase III), co skutkuje odmiennymi profilami ryzyka i wymogami w zakresie koordynacji. Wypracowujemy trójstopniowy model zabezpieczeń (konstytucyjnych, proceduralnych, technicznych) oraz model pracy skoncentrowany na sędziach, w którym sztuczna inteligencja pełni rolę pomocniczą, ale nigdy nie podejmuje decyzji. Stwierdzamy, że sukces zależy w mniejszym stopniu od zaawansowania technologicznego, a w większym od gwarancji konstytucyjnych, decydującej roli sądownictwa oraz adaptacyjnego zarządzania. Na zakończenie przedstawiamy potwierdzone naukowo zalecenia dla demokratycznych systemów prawnych wdrażających SI – z jednoczesnym zachowaniem niezależności sądownictwa oraz legitymacji społecznej.

**Słowa kluczowe:** SI w wymiarze sprawiedliwości, konstytucyjny porządek prawny, suwerenność cyfrowa, Polska, Indie, unijna ustawa o SI, demokratyczne zabezpieczenia, prawo porównawcze.

## Disclaimer

The views and analytical frameworks presented in this paper represent the authors' independent academic research and do not reflect the positions, policies, or official views of their affiliated institutions, the Polish judiciary, the Indian judiciary, or any governmental body. This comparative constitutional analysis addresses general governance principles applicable across democratic jurisdictions and does not evaluate specific current governmental policies or initiatives in either Poland or India.

## Author Contributions

Both authors contributed equally to this work. Prof. Góra-Błaszczkowska provided expertise on Polish constitutional frameworks and EU regulatory contexts. Prof. Radhakrishnan provided expertise on Indian constitutional systems and comparative public administration. Both authors jointly developed the analytical framework, comparative methodology, and policy recommendations. Both authors reviewed and approved the final manuscript.

## Acknowledgements

We gratefully acknowledge the European University of Law and Administration for hosting the conference where preliminary findings were presented. We thank conference participants for valuable feedback that strengthened this analysis.

## Funding Statement

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

## Declaration of Competing Interests

The authors declare no competing interests.

---

## REFERENCES

---

- Constitution of India (1950), Government of India.
- Constitution of the Republic of Poland (1997), Journal of Laws no. 78, item 483.
- Couldry, N., & Mejias, U. A. (2019), *The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism*, Stanford University Press.
- Council of Europe, European Commission for the Efficiency of Justice (CEPEJ). (2018/2023). *European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment*. Council of Europe Publishing. <https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c>.
- De Gregorio, G. (2021), *Digital Constitutionalism in Europe: Reframing Rights and Powers in the Algorithmic Society*, Cambridge University Press.
- Department of Justice, Government of India (2023), *E-Courts Phase III Project: Cabinet approval documentation*, Ministry of Law and Justice; <https://doj.gov.in/phase-iii/> [accessed: December 2024].
- Dreamscapes of Modernity: Sociotechnical Imaginaries and the Fabrication of Power, S. Jasanoff & S.-H. Kim (eds.), University of Chicago Press.
- Dunleavy, P., Margetts, H., Bastow, S., & Tinkler, J. (2006), *Digital Era Governance: IT Corporations, the State, and e-Government*, Oxford University Press.

European Union (2024), Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), Official Journal of the European Union, L 1689, 12.7.2024, p. 1; <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>.

Gdańsk University of Technology (2024). LEX-AIS Court Support System: Research partnership announcement; <https://pg.edu.pl/en/news/2024-11/technology-revolutionize-work-polish-courts-developed-gdansk-tech>.

Habermas, J. (1996). *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy* (trans. W. Rehg), MIT Press [original work published 1992].

Hood, C., & Margetts, H. (2007), *The Tools of Government in the Digital Age*, Palgrave Macmillan.

Malhotra, R. (2021), *Artificial Intelligence and the Future of Power: 5 Battlegrounds*, Rupa Publications India.

Noble, S. U. (2018), *Algorithms of Oppression: How Search Engines Reinforce Racism*, NYU Press.

Pasquale, F. (2015), *The Black Box Society: The Secret Algorithms That Control Money and Information*, Harvard University Press.

Polish Ministry of Digital Affairs (2024), Draft Act on Artificial Intelligence Systems: Consultation documentation; <https://www.gov.pl/web/cyfryzacja>.

Sen, A. (2009), *The Idea of Justice*, Harvard University Press.

Supreme Court e-Committee, Government of India (2024), Annual Implementation Report 2023–24. Supreme Court of India. <https://ecommitteesci.gov.in> [accessed: December 2024].

Supreme Court of India (2017), *Justice K.S. Puttaswamy (Retd.) and Anr. v. Union of India and Ors.*, 10 SCC 1, Writ Petition (Civil) no. 494 of 2012.

Supreme Court of India (2021), SUPACE and SUVAS: AI Tools for Judicial Efficiency, Press Information Bureau, Government of India; <https://pib.gov.in/PressReleaseIframePage.aspx?PRID=1714234> [accessed: December 2024].

UODO (Urząd Ochrony Danych Osobowych / Personal Data Protection Office, Poland). (2024), Position on Draft Act on Artificial Intelligence Systems; <https://uodo.gov.pl>.

Watson, A. (1993), *Legal Transplants: An Approach to Comparative Law* (2nd ed.), University of Georgia Press.



Ganiyu Yahaya\*

## Artificial Intelligence and the Right to a Fair Hearing Within Nigerian Judicial Proceedings: Navigating the Line Between Efficiency and Justice

[Sztuczna inteligencja a prawo do rzetelnego rozpatrzenia sprawy w nigeryjskich postępowaniach sądowych – balansowanie między wydajnością a sprawiedliwością]

### Abstract

Artificial Intelligence (AI) has become a transformative technology in various aspects of human activities, and the judicial system is no exception. In the meantime, AI is increasingly being deployed in the administration of the justice system in Nigeria, ranging from e-filing systems, digital case management, and virtual court proceedings to prospective applications in predictive judgment. While AI undoubtedly promises efficiency, reduced case backlogs, and cost-effectiveness, it also poses significant legal and ethical challenges, which include its impact on the infringement of the right to a fair hearing—a right primarily protected under Section 36 of the Nigerian Constitution and several other laws, including the Evidence Act 2011, Administration of Criminal Justice Act 2015, state Administration of Criminal Justice Laws, Fundamental Rights (Enforcement Procedure) Rules 2009, among others. This paper adopts a doctrinal method of legal research to critically examine the integration of AI into the judicial proceedings and whether it enhances or threatens the right to a fair hearing in Nigeria. It further interrogates whether algorithmic decision-making, lack of transparency, data bias, and other factors associated with computerization nature of AI compromise the procedural guarantees of fairness, impartiality, and the doctrine of the rule of law. Through doctrinal and comparative analysis, this paper considers the compatibility of AI-driven reforms with Nigerian constitutional and human rights jurisprudence to meet the global best practices. The paper concludes that AI should serve as a tool for justice and not a substitute for it for assurances of a credible, fair, and equitable judicial proceeding in Nigeria. The introduction of a continuous legal education module on AI for judges, lawyers, and court staff is recommended to ensure a cautious application through the enactment of an AI and Justice Regulation Bill.

**Keywords:** Artificial Intelligence (AI), fair hearing, judicial proceedings, Nigerian Constitution, human rights, and rule of law.

\* **Ganiyu Yahaya** – lecturer, PhD (Environmental Law), Faculty of Law, Federal University Oye-Ekiti, Ekiti State, Nigeria (affiliation); MCI Arb (UK); <https://orcid.org/0000-0002-8655-0171>; yahayaganiyu2012@gmail.com; ganiyu.yahaya@fuoye.edu.ng / wykładowca, dr nauk prawnych (prawo ochrony środowiska), Wydział Prawa, Uniwersytet Federalny w Oye-Ekiti, stan Ekiti, Nigeria (afiliacja); członek Brytyjskiego Instytutu Arbitrów (CI Arb).

---

## Introduction

---

The cornerstone of democratic governance and the rule of law is the judiciary, as it remains resolute, *inter alia*, to function as the ultimate arbiter in disputes and the protector of fundamental human rights. Hence, its effectiveness, independence, and impartiality are crucial and sacrosanct not only for the dispensation of justice but also for the assurance of sustaining public confidence in the institutions (Nwakodo & Ubani, 2023). Recently, the rapid evolution of Artificial Intelligence (AI) has brought several significant transformations in the justice sector. The evolution has informed profound legal and ethical questions regarding the compatibility of AI with the principles of due process and human rights, which are core to an effective justice system (Dymytruk, 2019).

In Nigeria, the evolution of technology in the judicial section is gradually taking a laudable shape through the digitization of court procedures, driven in part by the necessities of the COVID-19 pandemic, which informed a “social distance” policy and the inevitable desire to reduce case backlogs (Sule et al., 2024). Accordingly, these innovations include the introduction of electronic filing systems, platforms for online court sessions, remote hearings via video-conference, and electronic service of court processes. For instance, the Lagos State Judiciary, Federal High Courts, and National Industrial Courts of Nigeria have deployed court automation technologies, which reflected interesting reliance on digital systems to enhance judicial efficiency.

No doubt that these revolutions in the technology system hold great promise for enhancing access to justice and procedural efficiency. Adversely, they simultaneously raise critical concerns about the potential circumvention of constitutionally guaranteed rights, particularly the right to a fair hearing, as enshrined in the Nigerian Constitution under Section 36(1), which provides that:

“In the determination of his civil rights and obligations, including any question or determination by or against any government or authority, a person shall be entitled to a fair hearing within a reasonable time by a court or other tribunal established by law and constituted in such manner as to secure its independence and impartiality.”

This constitutional provision is non-derogable, as it forms part of Nigeria’s determination to uphold the rule of law. It is echoed in Article 7 of the African Charter on Human and Peoples’ Rights (Ratification and Enforcement) Act, which further guarantees the right to a fair trial before a court of competent jurisdiction or tribunal. In this vein, Nigerian courts have long emphasised that the right to a fair hearing encapsulates not just the opportunity to be heard but the right to be heard by an impartial court, tribunal, or panel within a reasonable time and in accordance with pro-

cedural guidelines. For instance, the Supreme Court in *Legal Practitioners Disciplinary Committee v. Fawehinmi* (1985) raised the question of whether a private individual like Gani Fawehinmi could file a complaint of professional misconduct against another legal practitioner with the Legal Practitioners Disciplinary Committee. The court held that any member of the public who has a complaint with regard to professional misconduct can address it through the disciplinary committee. The ruling highlighted the importance of upholding the principles of natural justice and fair hearing in disciplinary actions. Similarly, *Adeniyi v. Governing Council of YABATECH* (1993) related to the circumstance where the appellant was unfairly granted an opportunity to defend himself before disciplinary actions against him. The court held that the action touched on his civil rights and obligations; hence, rules of fair hearing had to be strictly adhered to. Accordingly, the disciplinary decision was set aside on the ground of failure to provide the appellant with a fair hearing. For emphasis, both rulings solidified the application of section 36 of the Nigerian Constitution, ensuring that nobody should be a judge in their own cause (*Nemo iudex in causa sua*), both sides must be heard (*Audi alteram partem*), and statutory or administrative bodies exercising disciplinary powers cannot act arbitrarily, regardless of their internal autonomy.

However, as AI technologies become embedded and recognized in judicial proceedings (including scheduling systems, automated document review, or predictive adjudication) across the globe, questions about whether these systems preserve the constitutional doctrine of the rules of law, encapsulating fairness, transparency, and impartiality in Nigeria or otherwise, have prevalently arisen. This is because, if AI technologies are improperly deployed, they risk introducing algorithmic bias, undermining effective judicial discretion, and diminishing litigants' confidence in the adjudicatory process (Citron, 2018). In addition, the cloudiness of many AI algorithms, especially those built on proprietary data, raises concerns about accountability, transparency, and the right to know how decisions affecting human rights are made. In a study by Kehinde et al. (2024), it is vividly underscored that AI has a transformative potential in enhancing the efficiency of Nigeria's judicial system. According to them, the integration of AI can streamline legal research, case management, and decision-making processes, which resolves the perennial issues of delays and backlogs in the adjudication system in Nigeria. However, the study cautions that, notwithstanding the integration of AI technologies, a well-defined regulatory framework to guard against potential infringements on fundamental human rights is imperative, especially the right to a fair hearing. Thus, a balanced approach that leverages the benefits of AI is suggested for the sustainment of justice without compromising the principles of justice and equity.

This paper therefore examines the complex interface between the increasing use of AI in judicial processes and the right to a fair hearing in Nigeria. It interrogates whether AI tools, despite their potential in efficiency enhancement, are aligning with established legal doctrines and human rights norms in Nigeria. This is carried out through a doctrinal methodology supported by a comparative analysis of global best practices and emerging jurisprudence in European Union, United States, and South Africa. The rationale is to assess how Nigeria can responsibly navigate the critical line between technological innovation and the preservation of the right to a fair hearing as enshrined in the Nigerian Constitution.

---

## Methodology

---

The type of methodology adopted by this study is purely a doctrinal legal research methodology. In the opinion of this study, the doctrinal legal research methodology is appropriate for evaluating the interface between Artificial Intelligence (AI) and the constitutional right to a fair hearing within the Nigerian judiciary context. The doctrinal approach to legal research of this nature is all-encompassing of the analysis of statutes, constitutional provisions, judicial decisions, and scholarly literature. The adoption is in a bid to determine the extent to which AI deployment aligns with or threatens established legal principles.

The particular sources of data for this study include both the primary and secondary sources. The primary sources include the Constitution of the Federal Republic of Nigeria 1999 (as amended), particularly Section 36(1) on the right to a fair hearing, relevant legislation affecting AI and digital governance (especially the Nigeria Data Protection Act 2023), international instruments on data protection, and judicial decisions (including *Legal Practitioners Disciplinary Committee v. Fawehinmi* [1985] and *Adeniyi v. Governing Council of YABATECH* [1993]). The secondary sources encapsulate peer-reviewed journals, textbooks, etc.

In the end, the study employs content analysis to examine the content of existing statutes, policies, and case law relating to AI and fair hearing. Secondly, it adopts comparative analysis to draw lessons from the selected jurisdictions (the European Union, the United States, and South Africa) regarding the use of AI in judicial proceedings. And, finally, it employs normative evaluation to assess whether the extant Nigerian laws provide sufficient protection against potential violations arising from AI adoption in the judiciary.

---

## Scope and Limitations

---

While the scope of the study is barely on AI applications in judicial processes (e.g., case management, virtual hearings, predictive analytics), it is limited to the Nigerian legal system while drawing insights from selected jurisdictions (European Union, United States, and South Africa) for comparative purposes.

---

## Conceptual clarifications

---

---

### Artificial Intelligence

---

The term “artificial” describes something that is not natural, whereas intelligence describes the capacity to gather and process information, store it for later use as memory-enhanced knowledge, and use this knowledge to do permanent actions (Umeh & Umeh, 2024). AI, according to Russell & Norvig (2020), is a field within computer science that focuses on creating algorithms and systems that can perform activities that usually need human intellect. Accordingly, it has become a worldwide revolutionary technology. Artificial Intelligence (AI) is a broad technology that describes computer systems’ ability to carry out tasks that normally require human intelligence. AI encapsulates perception, language comprehension, problem-solving, data-driven decision making (such judicial decisions), Intelligent tutoring systems, reasoning, and e-learning (machine learning). In this legal discourse, artificial intelligence (AI) is viewed as a socio-legal technology that questions conventional ideas, liability, and accountability, rather than only as a technical tool (Calo, 2015 & European Commission, 2018). AI’s rapid adoption in Nigeria raises concerns about its unregulated nature and potential conflict with local data governance traditions.

---

### Fair Hearing

---

Scholars define “fair hearing” in several but related ways. According to Bielu (2018), a fair hearing means that the other person will be accorded an opportunity to be heard and that one will not be a judge in his own case. Hence, a fair hearing is founded on the joint doctrines of *audi alteram partem* (hear the other side) and *nemo iudex in causa sua* (no one should be a judge in his own cause), both of which safeguard justice, impartiality, and the right to be

heard. Similarly, it has been observed that “the right to a fair hearing is a basic human right. A hearing can only be fair when all the parties to a dispute are accorded equal opportunity to be heard regardless of demographical status, such as age, race, religion, tribe, or political affiliation.” (Oni, 2023).

In the Nigerian context, this principle is entrenched under Section 36(1) of the 1999 Constitution (as amended), which guarantees that no person shall be denied a fair hearing before a competent and impartial tribunal. In *Ar-iori v. Elemo* (1983), the court established that the fundamental rights that are for the benefit of the litigant and the public cannot be waived. Similarly, in *Adigun v. Attorney-General of Oyo State* (1987), the court emphasised that once a fair hearing is breached, the entire proceeding is rendered a nullity, irrespective of whether injustice was occasioned or otherwise. In the same vein, the principle in the United Kingdom is embedded in both common law and the European Convention on Human Rights (Article 6), incorporated into domestic law through the Human Rights Act 1998. The decision in *Ridge v. Baldwin* (1964) also illustrates and entrenches that no adverse decision should be made without giving the affected party an opportunity to be heard. Moving further, in the United States, the concept is, under the Fifth and Fourteenth Amendments, constitutionally expressed as “due process of law”. The U.S. Supreme Court in *Mathews v. Eldridge* (1976) established that the principle is a balancing test that considers individual rights, governmental interests, and the risk of wrongful deprivation.

A fair hearing, in the opinion of this study, is a principle of law that is universally recognized as a fundamental element of justice. Thus, the absence of a fair hearing in a given cause of action undermines justice.

---

## Efficiency

---

There is no universal definition of efficiency. However, Efficiency is generally defined as the ability to achieve maximum output with minimum wasted effort, time, or resources for attainment of a desired result (Höber, 2024). In legal and judicial contexts, efficiency refers to as how quickly and cost-effectively a process, such as adjudication or case management can deliver justice without unwarranted delays or expenses.

Some representative definitions are in place to discuss efficiency in the context of AI-driven judicial proceedings. On this note, Tahura & Selvadurai (2022) define efficiency in the judicial sector as one which arises from AI tools for administrative tasks of e-filing, document handling, etc.; decision-recommendation systems; and case management. Accordingly, efficiency is construed to mean cost-effectiveness, time savings, reducing delays, and managing backlogs, while considering fairness, transparency, and accountability.

In addition, Mansour and Rajput (2025) construe judicial efficiency in the AI context to mean the extent to which AI tools reduce judicial backlogs, foster consistency in decision-making, and reduce resource costs, as it takes ethical standards into cognisance. Moving further, efficiency in the AI judicial proceedings is described as enabling faster access to justice for a larger population by reducing case backlogs.

Examining the concept from case law perspective, In *Pyrrho Investments Limited v MWB Property Limited* (2016), a landmark case in which predictive coding (using technology-assisted review, AI and e-disclosure) was approved in the process of electronic disclosure of documents on the basis of proportionality, fairness, and cost. The court held that predictive coding was permissible when its use is proportionate in terms of cost and necessary to deal with large document sets. Hence, electronic review of documents targeted efficiency and was of the justifications to navigate fairness. Moving further, in *Brown v. BCA Trading Ltd* (2016), the court allowed predictive coding in electronic disclosure despite objections. The court recognized that using such technology could reduce the burden, cost, and onetime involved in reviewing huge volumes of documents. Thus, efficiency was a factor considered in the court's decision to allow AI-technology-assisted review in document disclosure. From case laws' Perspectives, this paper mirrors the spirit of law in *Pyrrho Investments Limited v MWB Property Limited* and *Brown v BCA*, by defining efficiency in the context of AI-driven judicial proceedings as the ability of the justice system to deliver decisions more quickly and at a lower cost, using technology innovations. By implication, efficiency elucidates the saving of time and cost.

In all, this study views efficiency in AI-driven judicial proceedings as referring to the ability of the justice system, using AI tools and processes, to resolve cases in less time and at a lower cost (concomitantly in terms of human and financial resources), to reduce procedural delays and backlogs, and to better manage resources better, without compromising fairness, accountability, transparency, or the rights of the parties.

---

## Justice

---

Within the perspective of philosophy and law, justice has long been conceptualized. From the philosophical perspective, Aristotle's definition conceptualized justice as giving each person what is due to them, distinguishing between distributive justice (including fair allocation of resources) and corrective justice (including rectifying wrongs). He argues that justice is that virtue which assigns to each his due (Aristotle, ca. 350 B.C.E./2009). Rawls, in his modern philosophy, advances the idea of justice to mean fairness. He argues

that justice should be understood through principles chosen behind a veil of ignorance, ensuring fairness and impartiality in societal arrangements (Rawls, 1971). It is pertinent to note that the societal arrangements mentioned include the judiciary. According to him, fairness and impartiality emphasize equal basic rights for all to benefit the vulnerable. The judicial definition, on the other hand, courts have also shaped the meaning of justice through various interpretations. In *Legal Practitioners Disciplinary Committee v. Fawehinmi* (1985), the Nigerian Supreme Court declared that justice is not a cloistered virtue but one that must be unequivocally seen to be done. Hence, this decision aligns justice with transparency, fairness, and trust in adjudication. In addition, Lord Hewart CJ in *R v. Sussex Justices, ex parte McCarthy* (1924) famously held that justice should not only be done but should undoubtedly be seen to be done.

However, this discourse on justice remains relevant in the perspective of AI-driven judicial proceedings. For example, from Aristotle's definition, algorithms must ensure fairness and avoid misplacement of outcomes when distributing rights or responsibilities. Rawls' fairness theory of justice also provides a framework for evaluating whether AI systems are impartially well designed or otherwise, particularly in the area of access to justice for the vulnerable. Furthermore, judicial interpretations underscore that justice must be seen to be done, which by implication means that AI systems must be transparent, explainable, and accountable to midwife public trust.

In the total sum, one germane takeaway is the articulation from the Institute for Research on Public Policy, which argues and underscores that efficiency and justice are inextricably intertwined (IRPP, 2023). For example, policies and technological-based innovations, such as AI-driven judicial proceedings, are meant to improve judicial proceedings' efficiency and enhance the justice system in Nigeria. Conversely, it is not an understatement that the policy innovation focuses majorly on efficiency while being blind to equity, which may exacerbate injustice or reduce the integrity of the judiciary eventually.

---

## **Legal and Theoretical Frameworks**

---

This section provides the conceptual and normative ground for interrogating and informing the compatibility of AI with the right to a fair hearing enshrined in the Nigerian Constitution. It situates the discourse within the binding provisional scope of Nigerian law and relevant international instruments. Theoretical approaches, on the other hand, discuss technology (AI-driven systems) and law.

---

## Legal Framework

---

The relevant legal framework includes:

### 1. The Nigerian Constitution

In Nigeria, Section 36(1) of the Constitution of the Federal Republic of Nigeria, 1999 (as amended), guarantees the right to a fair hearing. Going by this constitutional provision, disputes are determined by a competent, impartial, and independent court or tribunal within a reasonable time. The provision is widened by Section 46 thereof, which allows individuals to seek judicial redress where any of the fundamental rights under Chapter IV of the Constitution are violated. For emphasis, Nigerian courts have consistently upheld the sanctity of a fair hearing as a cardinal principle of justice. Of course, this is why the Supreme Court held in *Legal Practitioners Disciplinary Committee v. Fawehinmi* (1985) that denial of a fair hearing provided by the Constitution renders proceedings null and void, at the expense of merits.

Moving further, S. 36(3) & (4) provides for both public places hearing for proceedings and procedural requirements of substantive justice. Subsection 3 holds that any person charged with a criminal offence must be tried and heard in public to ensure transparency and in avoidance of secret trials except in rare situations (e.g., if a minor is involved, for national security, to maintain public order, and if justice requires the restrictions from the public), while subsection 4 mandates that any person charged must be given adequate time to prepare their defence, the trial must commence within a reasonable time, and such person(s) must be allowed to have legal representation of their choices. These subsections are to ensure procedural fairness and substantive justice. The whole of Section 36 and the embodied subsections are very momentarily relevant to this study. Conversely, the Virtual Court Proceedings, which are AI-driven innovations, are being criticized by some quarters of the legal profession on the premise that such trials are not carried out in public and are too quick, thereby denying accused persons adequate time, while relying on the decision of the Supreme Court of Nigeria in *Edibo v. State* (2007), which declared such secret trials null and void. This question remains: how do we navigate the benefits of VCPs, with the constitutional principle of a fair hearing? This paper aims to answer this question.

### 2. Other Nigerian Laws

Aside from the right to a fair hearing in Nigeria, as primarily guaranteed under Section 36 of the Constitution of the Federal Republic of Nigeria 1999

(as amended), the right is also protected by several other laws. These include the Evidence Act 2011, particularly sections 214–216, which guarantee examination-in-chief, cross-examination, and re-examination of witnesses; the Administration of Criminal Justice Act 2015, particularly sections 1 and 6, which promote fair, efficient, and rights-based criminal justice administration and require that a suspect be informed of the reason for arrest and his rights; state Administration of Criminal Justice Laws; the Fundamental Rights (Enforcement Procedure) Rules 2009, made pursuant to section 46(3) of the Constitution, particularly the Preamble/Objectives in Paragraph 3, which mandate courts to advance and protect fundamental rights; the National Industrial Court Act 2006, particularly section 36, empowering the court to regulate its proceedings in accordance with the rules of natural justice; the Electoral Act 2022; the Investment and Securities Act 2007; the Legal Practitioners Act governing disciplinary proceedings before the Legal Practitioners Disciplinary Committee; and public service disciplinary regulations requiring that affected officers be notified of allegations and given an opportunity to defend themselves. Hence, these laws collectively ensure that individuals are given adequate notice of allegations against them, an adequate opportunity to present their cases, legal representation (where applicable), and an impartial hearing before decisions affecting their rights or interests are made. It is in the spirit of implementing these laws that Nigerian courts, panels of enquiry, and tribunals have been consistently committed to adhering to a fair hearing in judicial, administrative, and *quasi*-judicial proceedings, as the case may be.

### 3. International and Regional Instruments

Apart from Section 36 of the Constitution of the Federal Republic of Nigeria 1999 (as amended), the right to a fair hearing is also guaranteed under several international and regional human rights instruments to which Nigeria is a state party. The enforcement of the charter and other international and regional instruments in Nigeria is guaranteed by virtue of Section 12 of the Nigerian Constitution, which provides that ratified treaties require domestic incorporation before they can be applicable to Nigeria

At the international level, the Universal Declaration of Human Rights (UDHR) guarantees the right to a fair and public hearing before an independent and impartial tribunal under Article 10, while Article 11(1) provides for the presumption of innocence and other minimum guarantees in criminal proceedings (United Nations, 1948). It is pertinent to underscore that the UDHR has not been domesticated; hence, it does not directly enforceable in Nigeria, but Nigerian courts have frequently relied on its provisions as persuasive authority in interpreting the constitutional guarantee of fair hearing under Section 36 of the Constitution (United Nations, 1948).

In the same vein, the International Covenant on Civil and Political Rights (ICCPR), to which Nigeria acceded in 1993, comprehensively protects the right to a fair trial under its Article 14. Accordingly, the rights therein include equality before the courts, the right to a competent, independent and impartial tribunal, the presumption of innocence, adequate time and facilities to prepare a defence, legal representation, and the right of appeal. In addition, Article 9(3)–(4) guarantees the right to be promptly brought before a judge and to challenge the legality of detention. Although the ICCPR has not been domesticated pursuant to Section 12 of the Constitution, Nigerian courts have utilised its provisions as interpretative guides in giving effect to the constitutional right to fair hearing, particularly where domestic law is ambiguous or where constitutional provisions admit of a broad human rights interpretation (United Nations, 1966).

At the regional level, the African Charter on Human and Peoples' Rights (1981) occupies a unique position in Nigerian law because it has been domesticated through the African Charter on Human and Peoples' Rights (Ratification and Enforcement) Act, Cap. A9, Laws of the Federation of Nigeria 2004. Consequently, its provisions are directly enforceable before Nigerian courts. Article 7(1) guarantees the right to have one's cause heard, including the right to appeal, the presumption of innocence, the right to defence by counsel of one's choice, and the right to be tried within a reasonable time by an impartial court or tribunal. Article 26 further obliges States Parties to guarantee the independence of the courts (African Commission on Human and Peoples' Rights, 1981). Nigerian courts have repeatedly applied these provisions in enforcing fundamental rights. Thus, in *Abacha v. Fawehinmi* (2000), the Supreme Court held that the African Charter, is already domesticated by an Act of the National Assembly; hence, it forms part of Nigerian law and is enforceable in Nigerian courts. Similarly, in *Gbemre v. Shell Petroleum Development Company Nigeria Ltd. & Ors* (2005), the Federal High Court relied on the African Charter on human and people's rights alongside the Constitution to protect fundamental human rights.

Generally, the right to a fair hearing is also reinforced in some specialised international instruments, such as the Convention on the Rights of the Child (CRC), which safeguards the fair trial rights of children in conflict with the law under Article 40(2)(b). Nigeria has thus, for emphasis on application, adopted these obligations in the Child Rights Act 2003, especially the conditions that provide for child justice administration and for children to have a fair hearing (United Nations 1989). Likewise, the principles of fair hearing rights of children are incorporated in the African Charter on the Rights and Welfare of the Child, which has been enshrined in Article 17, and its principles have been reflected in the Child Rights Act and equivalent State Child Rights Laws (African Union, 1990).

Moreover, the 2006 Convention on the Rights of Persons with Disabilities (CRPD) provides the right to equality of access to justice for persons with disabilities as enshrined in Article 13 and the right to equality of recognition before the law in Article 12. Nigeria has significantly fulfilled these commitments in its Discrimination Against Persons with Disabilities (Prohibition) Act 2018, where access to justice is guaranteed for persons with disabilities in public institutions, which includes courts (United Nations, 2006). Likewise, the 1951 Refugee Convention, supplemented by its 1967 Protocol is a Convention relating to the status of refugees. It provides for access to the court under article 16 and judicial independence under article 11. Similarly, the United Nations Convention against Corruption (UNCAC) further confirms judicial independence under article 11 and due process guarantees in criminal proceedings under article 30 (United Nations, 1951; United Nations, 2003). These values have been projected in Nigerian law, e.g., the Administration of Criminal Justice Act 2015, the Economic and Financial Crimes Commission (Establishment) Act, etc., which are procedural laws that highlight ‘fair trial’ and ‘judicial integrity’.

In this regard, some international human rights instruments, e.g., the UDHR and ICCPR, are not domesticated and therefore have only persuasive value but the African Charter has force of law in Nigeria and is directly enforceable (*Abacha v. Fawehinmi*, 2000). These instruments, along with domesticated treaties and implementing statutes, have had a significant impact on the evolution of fair hearing in Nigeria and still play a role in the interpretation and application of the constitutional right to fair hearing.

#### 4. Nigerian Data Protection and Technology Laws

Laws under this section include the Nigeria Data Protection Act (NDPA) 2023, which provides a statutory foundation for protecting personal data and regulating automated decision-making. The Act establishes the Nigeria Data Protection Commission (NDPC) as an authority with the mandate of interrogating anomalies and enforcing the Act. Key features under the Act include data subject rights, protection of sensitive personal data, and legal bases for processing (Modilim et al. (2024) & Sections 34–38 of NDPA, 2023). Moving forward, Section 5 of the Act empowers the NDPC to promote awareness of obligations, advises the government on data policy, and provides principles for processing personal data. In all, Section 48(1) of the Act also establishes a more robust enforcement and penalty regime, empowering the NDPC to impose administrative fines and sanctions (as the case may be) on data controllers and processors who violate its provisions. Additionally, the Cybercrimes (Prohibition, Prevention, etc.) Act 2015 is another relevant statute that addresses issues of integrity, responsibility, and security in electronic systems

that could intersect with judicial digitisation. For emphasis, Section 24 of the Act addresses cyberstalking and online harassment, while Section 38 protects Nigeria's critical national information infrastructure, including transmission facilities, Treasury Single Account (TSA), Biomedical Verification Number (BVN), National Identification Number (NIN), Electronic Payment Gateways (EPG), e-government platform, and data centres.

Other laws that are relevantly pivotal and instrumental to AI-based judicial proceedings in Nigeria include The National Information Technology Development Agency (NITDA) Act. The Act establishes NITDA (the Agency) with a mandate to ensure unrelenting compliance with the Act. Moving further is the Nigerian Communications Act (NCA), 2003 the regulatory authority is Nigerian Communication Commission (the Commission). The Commission mandates telecom operators in Nigeria to protect subscriber data through the NCC Consumer Code of Practice Regulations 2007.

In all, the laws discussed under this section are critical in the Nigerian context of AI deployment in judicial processes. They actually introduce obligations for transparency, accountability, and safeguards against discriminatory profiling. Thus, this paper posits that all laws relevant to the AI judicial proceedings in Nigeria do not expressly address a fair hearing. Rather, they establish a regulatory and criminal regime for data technology (including AI), data protection and privacy, and cybercrime.

---

## Theoretical Framework

---

The theoretical landscapes include:

### 1. Technological Determinism

The theory suggests that technology shapes social structures and jealously guides cultural values. Applying the theory to judicial proceedings in Nigeria suggests that the adoption of AI-driven judicial proceedings will inevitably reshape the pattern of the justice delivery system in Nigeria (Winner, 1980). Hence, the technology determinism theory underscores potentially accelerating efficiency while risking depersonalisation and algorithmic bias.

### 2. Socio-Legal Theories of Law and Technology

Socio-legal theorists argue that technology should not be examined in isolation but in relation to its alignment with legal norms and societal values (Murray, 2011). Accordingly, the rationale is that the technology deployment

is for human benefits and should not violate human rights. Thus, the theory underscores that AI adoption in Nigeria's judicial proceedings ought to be contextualised within the socio-political realities of access to justice, digital literacy, and infrastructural capacity.

### 3. Inner Morality of Law Theory

Lon Fuller's theory of the "inner morality of law" does require that legal processes be unequivocal, consistent, and transparent to uphold legitimacy. But these are part of his broader concept of the "Eight Desiderata", or principles of legality. The "Eight Desiderata" principles of legality include ensuring a general, publicly promulgated, non-contradictory, feasible, clear, stable, prospective, and congruent legal system (Fuller, 1969). Hence, if AI systems in judicial proceedings are opaque ("black-box algorithms"), they could undermine the principles of generality, non-retroactiveness, legality, accountability, and fairness central to the rule of law as encapsulated by the broader concept of the "Eight Desiderata" or principles of legality.

### 4. Rule of Law Theory

The rule of law theory is also applicable to this paper. According to Dicey (1959), the supremacy of law is fundamental to any democratic society. Accordingly, no individual – regardless of rank or position – is above the law; hence, both government and citizens are subject to it. He moves forward to assert that individual rights and liberties are best protected via legally established courts of law rather than by the exercise of discretionary or arbitrary power. Similarly, Nwabueze (1977), while building on this foundation, offers a broader perspective by contextualizing the rule of law within emerging constitutional democracies. Thus, he views the theory as a system in which governmental powers and those of public authorities are within their legal boundaries while upholding citizens' rights and liberties. Nwabueze's position on the rule of law complements Dicey's theory by underscoring its application to governance and public authorities.

However, the implication of this theory for judicial proceedings is that it resonates with the need to adhere to the doctrine of a fair hearing clearly enshrined in the Constitution as sacrosanct during judicial processes.

### 5. Human Rights-Based Approach (HRBA)

The HRBA pinpoints that all new technological devices, including AI, must be evaluated through the lens of human rights standards (Paliwala, 2020). Accordingly, an AI-driven system must stand to support non-derogable rights

like a fair hearing. Hence, applying this approach to the Nigerian context suggests that AI-driven judicial proceedings, as a technological modernism, should be designed to enhance accessibility, impartiality, and equality before the law.

In a nutshell, the paper opines that the Nigerian legal framework, reinforced by international human rights obligations, establishes a fair hearing as a constitutional and non-negotiable right. Thus, it demonstrates the inevitability of a fair hearing proceeding. Theoretical perspectives, on the other hand, demonstrate that while AI adoption in the judiciary may also be inevitable, it cannot operate in a legal vacuum. Rather, its deployment must be moderated and informed by human rights principles, judicial discretion, and the rule of law to ensure that the efficiency of AI does not supplant justice.

---

## **Current Applications of “AI” in the Nigerian Judiciary**

---

The judiciary in Nigeria, albeit traditionally conservative in embracing technological-driven change, has recently adopted digital innovations in line with global advocacy reforms in the judiciary. These technological innovations were greatly accelerated by the COVID-19 pandemic and the pressing need to reduce case backlogs in Nigeria during that memorable and critical time. Thus, such innovations (as AI-assisted case and court process management tools) include court management systems and case process automation, enhanced e-filing systems, and virtual court hearings. While Nigeria is at the nascent stages of synthesizing a full AI-driven system, the first wave of digitalization is yielding significant lessons with reflections of challenges and opportunities, though (Adediran et al., 2024).

---

### **Court Management Systems and Case Process Automation**

---

In Nigeria, Judicial automation has also extended to case management. For a caption, Lagos State Court Management Information System (LagosCoMiS) simplifies certain functions, including suit number generation, case assignment, e-notifications, and affidavit processing. Rivers State’s RIVCOMIS has also, within its first year, improved transparency through the synthesis of online filings, payments, and virtual hearings into its judicial system, which is not only efficient but also goes a long way to reducing document loss and manipulation (Godwin, 2021).

According to Ubanyionwu (2024), AI-driven case management systems are the core to reducing delay and enhancing consistency through predictive analytics and precedent monitoring; however, he stresses that judicial discretion and oversight remain indispensable to avoid over-reliance on algorithms which might be inapplicable in some cases. On this note, this paper opines that AI can optimize case timelines, automate routine tasks, and minimize backlog, yet infrastructural deficits and weak regulatory frameworks hinder full-scale adoption. These, through the lens of the underlined opinion, prove the need to navigate a line of balance between efficiency and justice has become so pertinent, as the judicial integrity must be sacrosanct.

---

### Enhanced E-Filing and Digital Document Handling

---

The Nigerian judiciary is equally, in recent times, adopting e-filing and handling documents digitally. For instance, the Judiciary Information System (JIS) is in place in the Lagos State judiciary, which provides an e-filing service that facilitates and entrenches room for legal practitioners to file cases remotely, track fees, upload forms and court processes, and make payments online. For emphasis, LagosCoMiS was developed by Cinfores Limited (a software company in Port Harcourt, Nigeria), which incorporates features such as e-filing, e-affidavit generation and validation, case-number allocation, and case automation. Accordingly, the practice has since been extended to cover other jurisdictions within Nigeria (Umeh, 2023). Similarly, the Federal High Court (Lagos Division) has declared that from June 23, 2025, all new filings will be processed exclusively through its e-filing portal, and innovation has since then started (Musa, 2025; Iwuchukwu, 2025).

However, scholars note both prospects and pitfalls. Prof. Ojo Emmanuel Ademola, while emphasizes that AI-driven tools, including automated case filing, e-affidavits, etc., can significantly reduce administrative delays and human error, argues that their effectiveness depends on cybersecurity safeguards and adequate training to prevent systemic breakdowns (Adesewo, 2025). Likewise, Prof. Charles Olufemi Adekoya argues that AI tools in legal practice, including precedent-search databases, legal drafting aids, etc., could reduce the time required for judgments. The innovation is potentially shortening delivery periods to just a week if court databases are consistently updated and integrated with AI platforms (Oduwole, 2022). The scholars' views align with uptake reports showing increasing registrations by legal practitioners, albeit challenges, including unreliable electricity, internet disruptions, unprepared registries, etc. (Adediran, 2024).

---

## Virtual Court Proceedings (VCPs)

---

The National Judicial Council (NJC) came up with guidelines for VCPs in the year 2020 (during COVID-19). The NJC prescribes platforms for proceedings and regulates the virtual delivery of orders and judgments. With the NJC's guidelines during that period, the Lagos State Judiciary rolled out e-affidavits, CoMis, and virtual platforms for court proceedings as part of Practice Directions to tackle the urgent and time-bound matters (Lagos Judiciary, 2020 & Davies et al., 2024). For illustration, In May 2020, the Lagos State Judiciary held its maiden virtual court proceedings (VCPs) session due to COVID-19 pandemic restrictions. A driver called Olalekan Hameed was tried for reckless driving that caused a passenger's death. Conducted online through VCP (an AI-assisted technology), the trial marked Nigeria's first virtual sentencing. The Lagos High Court convicted Hameed of murder and sentenced him to death by hanging. The case highlighted the judiciary's adaption of technology for justice while sparking debates over the constitutionality and ethical virtue of virtual trials for such capital punishment (Egbe, 2020). Albeit, VCPs face legal experts' Critiques despite it benefit, they argued that VCPs may be unconstitutional. On this note, a legal icon, Chief Adegboyega Awomolo, SAN, argued that VCPs was unconstitutional under Section 36 (3) & (4) of the 1999 Constitution, which guarantees a fair hearing in public. While interpreting the Constitution, a public hearing meant one which is physically open to everyone without restriction, whereas virtual hearings limited access only to those with smartphones, computers, stable internet, and power supplies (Ige, 2023). Accordingly, with less than half of Nigerians having internet access at the time, he feared many citizens were being excluded from justice. He drew attention to earlier Supreme Court decisions, such as *Edibo v. State* (2007) and *Oviasu v. Oviasu* (1973), where judgments delivered in judges' chambers were declared null and void for not meeting the test of openness or public place. While acknowledging the benefits of VCPs, his critique focuses on the "cohesive distinctiveness" of in-person legal hearings.

However, scholars provide critical reflections on the court's AI-driven innovations. Sule et al. (2021) assert that virtual court sittings ensured continuity during the pandemic; however, they raised constitutional questions, particularly regarding the public hearing and court in session requirements of Section 36 of the Constitution of the Federal Republic of Nigeria (as amended). They contend that while NJC guidelines and practice directions are helpful, innovative, and effective, their legitimacy requires amendments to rules of court. In the same vein, Davies et al. (2022) acknowledge that virtual hearings reduce delay; however, they warn that digital evidence verification, fairness in cross-examination, and bandwidth inequities continue to undermine procedural justice. No doubt those outcomes confirm both the advantages of

continuity and efficiency, as well as persistent challenges, including platform costs, internet instability, and difficulties ensuring open public access (Sule et al., 2021).

Applying AI-driven innovations into judicial proceedings in Nigeria, this paper concurs that the innovations demonstrate the prospect of AI in expediting justice delivery, mitigating corruption-prone manual processes, and promoting transparency. As Ubanyionwu argues, a properly deployed AI can improve consistency in outcomes and reduce human error. However, persistent challenges include digital exclusion of rural litigants, unresolved constitutional questions about VCP in relation to the doctrine of a fair hearing, potential algorithmic bias in relation to ethical requirements, and inadequate data safeguards, opined by Sule et al. (2021). In the total sum, without a robust legal framework and equitable infrastructural development, AI-driven innovations risk entrenching inequalities rather than delivering justice in Nigeria.

---

## **Implications of AI for the Right to a Fair Hearing**

---

The reception of AI in judicial proceedings in Nigeria relatively carries both considerable benefits and serious risks for the constitutional right to a fair hearing under Section 36 of the 1999 Constitution (as amended). The concept of technological due process, formulated by Danielle K. Citron, provides a useful theoretical architecture to evaluate the way AI might impinge on or enhance procedural fairness (Citron, 2008). Thus, AI deployment has its benefits and serious risks associated with its implication for the right to a fair hearing. These are for discussion alongside relevant case laws on the doctrine of a fair hearing.

---

### **Benefits**

---

#### **1. Efficiency and Reduced Backlog**

AI-driven judicial proceedings can help automate routine tasks, ranging from document sorting to case assignment to legal research. It, therefore, supports robust case management systems, thereby saving cost and time (Fine et al., 2025). A study by Eze (2024) points out that AI has the potential to transform how justice is being administered by cutting through legal service delays and improving turnaround times for adjudication. Accordingly, AI-driven proceedings reduce delays/backlogs, one of the most persistent obstacles to a fair hearing in Nigeria.

## 2. Access to Justice

Another benefit of AI-driven judicial proceedings is having smooth access to justice. On this note, AI tools can reduce barriers, such as availing litigants who live far from major courts or who are resource constrained to access the justice system online. A study by Bello & Ogufere (2024) underscores improved access to justice as among the key benefits of integrating AI into Nigeria's judicial system. Thus, using VCP platforms reduces the cost of travel and allows more people to engage with the judicial system.

## 3. Accuracy and Support in Decision Making

AI tools for legal research, predictive analytics, and document analysis potentially reduce human error. They assist judges to identify efficiency in relevant precedents; and thus potentially lead to more consistent judgments. On this point, Ubanyionwu (2023) in a study explicitly affirmed that AI could enhance accuracy in legal decisions and reduce inconsistencies.

---

### Risks

---

Despite the highlighted key benefits, the following risks are associated with the AI-driven judicial proceedings:

#### 1. Bias

Bias is discriminatory in nature, as it threatens the equality and impartiality components of a fair hearing. Scholars warn against bias, particularly in predictive tools and risk assessment algorithms. Bello & Ogufere (2024) note that while AI has key benefits, even when there are no proper safeguards, there is a risk of reinforcing existing inequalities. Thus, AI-based judicial processes are only as fair as the data they are trained on. For emphasis, if data reflect societal, institutional, etc., biases, AI can perpetuate or even amplify them if not squarely monitored. However, Retreating the cases of *Legal Practitioners Disciplinary Committee v. Fawehinmi* (1985) and *Adeniyi v. Governing Council of YABATECH* (1993) is a concomitant effect which provides benchmarks against which AI's compatibility with fair hearing must be measured. Hence, any AI process used in judicial proceedings must preserve impartiality, adequate hearing notice, opportunity to be heard, and transparency in reasoning to avoid bias.

## 2. Opacity

Under the doctrine of technological due process, according to Citron (2008), procedural fairness requires transparency, including how decisions are made, what data is used, what algorithmic logic is applied, etc. The Black-box systems, proprietary algorithms, or opaque decision rules attributed to AI-driven judicial proceedings can deny litigants the Where-withal to challenge or understand decisions or verdict affecting them. Citron emphasizes that Inadequacy of access to the internal logic of automated systems of black-box system in the VCP undermines foundational procedural rights.

## 3. Reduced Discretion and Judicial Independence

Over-reliance on AI-driven recommendations or advice may deskill judges as well as other decision-makers (Schemmer et al., 2023). It may reduce judicial discretion over a particular decision and compromise judicial independence. While AI provides decision support, if courts or judges defer too much to AI outputs, they may be at risk of abdicating their responsibility to reason, assess credibility, and apply justice in individual cases. In a study by Kehinde et al. (2024), AI is viewed as a double-edged sword, useful in curbing arbitrary discretion and reinforcing judicial independence, but risky if poorly designed or politically captured. However, they caution that discretion while adopting AI should not be eliminated since it is vital for fairness, contextual interpretation, and human empathy in judicial decision-making. Thus, AI-driven proceedings should incorporate judges' discretion rather than reducing it.

## 4. Data Privacy, Security, and Procedural Rights

AI involves processing large amounts of sensitive personal data. Without strong legal and technical safeguards, litigants may have their private information exposed, misused, or handled in ways incompatible with privacy or fair hearing norms. Also, automated profiling or decision-making without meaningful human oversight or right to challenge can violate due process norms. Bello & Ogufere (2024) discuss ethical considerations and data privacy as among the top risks.

---

## Comparative Perspectives

---

---

### European Union (GDPR and the AI Act)

---

In the European Union, a strong data-protection regime is combined with an emerging risk-based AI regulatory model. The General Data Protection Regulation (GDPR) is in place with a mandate to strictly regulate automated decision-making (Article 22). It protects individuals from decisions based solely on profiling and imposes transparency and fairness obligations on data controllers. While building upon the GDPR, the European Commission's 2021 proposal for an Artificial Intelligence Act sets out a tiered risk framework (including transparency, documentation, and conformity with). The high-risk AI uses are those touching on justice, law enforcement or fundamental rights. Thus, this approach treats AI as socially beneficial but legally risky, particularly where it affects core rights, thereby prioritising safeguards, explainability, and human-in-the-loop requirements for judicial applications (Anesu, 2025).

---

### United States (COMPAS and *State v. Loomis* U.S.)

---

In the United States, experience illustrates the risk associated with deploying opaque, proprietary models in life-changing judicial processes, including VCPs. The Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) to predict risk of recidivism. It is a risk assessment algorithm that assists judges, paroles officers, and correctional authorities to evaluate the potentiality and possibility that an offender will recidivate. Despite its benefits, it has raised controversial issues about its transparency, accuracy, etc. Investigative work exposing its disparate impacts. highlighted that commercial risk tools can produce disparate error rates across groups and are difficult to interrogate due to the fact that their internal logic is often trade-secreted. In *State v. Loomis* (Wis. 2016), Eric Loomis challenged the use of the COMPAS risk assessment tool during his sentencing, arguing that the algorithm was proprietary, lacked transparency, and potentially biased. However, the Wisconsin Supreme Court permitted the use COMPAS as an algorithmic risk score in sentencing while cautioning strongly against uncritical reliance: courts must treat such scores as assistance, rather than as determinative proof, and must be mindful of transparency limitations and due process. By implication, the U.S. AI-driven regulatory regime shows that algorithmic tools can be used, but only within strict procedural constraints

which permit disclosure, limitation of weight, judicial warnings, and ability to contest inputs (Administrative Conference of the United States [ACUS], 2024).

---

### South Africa (Digital Court Reforms and Human-Rights Emphasis)

---

Digital court reforms and human rights protection concerns have made South Africa emerge as a leader in judicial digitalisation, particularly through the Court Online initiative, which encapsulates end-to-end e-filing, digital case management, and electronic evidence systems (Obi-Farinde, 2025). There is no doubt that this platform allows the courts, legal practitioners, and litigants to file documents, track cases, and access evidence electronically, thereby reducing cost, saving time, and streamlining judicial workflows (Judge President Mlambo, 2020). For emphasis, the COVID-19 pandemic midwived these reforms, as directives required that new High Court matters be filed electronically by July 2022, and the platform was still retained post-pandemic. The Electronic Communications and Transactions Act further supports these reforms by granting legal recognition to electronic signatures in South Africa. Another momentous innovation is CaseLines, a cloud-based evidence management system introduced at the Gauteng High Court. The system enables remote filing, retrieval of documents, and hybrid hearings, which reduces backlogs in urban courts such as those in Johannesburg and Pretoria.

However, challenges remain: insufficient ICT infrastructure, low level of digital literacy, concerns for data privacy, and cybersecurity risks, etc. On this note, Masenya and Ntengenyane (2022) found that High Courts often struggle with fragile record systems, a shortage of skilled staff, and limited technical support, undermining the efficiency of digital justice. Moreover, recent cases, such as *Parker v. Forsyth N.O.* and *Mavundla v. MEC*, illustrate South African courts' cautious stand towards AI, where courts warned legal practitioners against uncritical reliance on generative tools like ChatGPT and emphasizing the need for human verification in legal practice.

---

### Lessons for Nigeria

---

Thus, Nigeria's stance is to learn the following lessons:

1. Adopting a risk-based regulatory approach, as obtainable in the EU model. Nigeria should learn to classify judicial AI uses by risk (administrative triage versus sentencing support) and impose progressively stricter controls for high-risk systems, including conformity assessment, documentation, and

mandatory concerns for human rights. This minimises harm where stakes are highest while permitting lower-risk innovation.

2. Lessons from the U.S. include protecting procedural rights through explainability and limits on evidential weight. Echoing Loomis, algorithmic outputs should be admissible only as assistive evidence, accompanied by disclosure of inputs. Parties must have meaningful opportunities to challenge algorithmic evidence which informs an explanation in intelligible form.

3. Nigeria's AI judicial proceedings regime can be strengthened through judicial digitization by taking into cognisance adequate records management, cybersecurity and digital inclusion (South African lesson). Invest in resilient court IT infrastructure, secure electronic records management, and nationwide capacity building (judges, registrars, lawyers, and litigants), as failure will compromise privacy or access to justice.

---

## **Recommendations**

---

For assurances that AI enhances the right to a fair hearing in judicial proceedings in Nigeria, as prescribed under Section 36(1) of the Nigerian Constitution, strategic, legal, and ethical safeguards/concerns must be put in place. Thus, this paper recommends the following in line with the international and jurisdictional best practices:

### **1. Establish a Comprehensive AI Regulatory Framework for the Judiciary in Nigeria**

The absence of a judiciary-based specific legal framework for AI-driven judicial processes in Nigeria creates inconsistency and potential abuse. Hence, there is a need for enacting a law to be named "Judicial Artificial Intelligence Regulation Act (JAIRA)," with a mandate to define permissible AI to be applicable in judicial proceedings. In addition, it should encapsulate the setting of algorithmic transparency and accountability standards. Furthermore, the proposed law will be AI tool-independent-driven audits and ensure compliance with constitutional provisions on the principle of a fair hearing. This proposed law will align Nigeria's AI-driven judicial proceedings with global trends.

### **2. Strengthen Data Protection and Privacy Measures in Nigeria**

The NDPA's scope should be expanded to include provisions for algorithmic data processing within the judiciary. These include:

- i. mandatory litigants' data privacy;
- ii. restrictions on cross-border transfer of judicial data; and
- iii. routine evaluations of the court's digital infrastructure for potential weaknesses.

Thus, the principle laid down in *Digital Rights Lawyers Initiative v. National Identity Management Commission (NIMC)* [2021], where the Court of Appeal underscored that data protection and privacy are constitutional rights enshrined in Section 37 of the Nigerian Constitution, is resonate with this recommendation.

### 3. Mandate Algorithmic Transparency and Explainability

Judges, counsel, and litigants should have the right to demand explanations for AI-generated decisions. These are done through:

- i. judicial rules requiring AI tools to provide explainable outputs;
- ii. periodic certification of AI systems by the National Judicial Council (NJC); and
- iii. sanctions for use of unverifiable AI systems.

Such measures would operationalize the principles established in *Ariori v. Elemo* (1983), where reasoned and transparent judgments were underscored by the Supreme Court as essential for a fair hearing.

### 4. Capacity Building for Judicial Officers and Court Staff

AI literacy and training among judges, magistrates, and court staff are advised to be critical for responsible deployment point. Therefore, the National Judicial Institute (NJI) should introduce:

- i. mandatory AI training modules for judicial officers;
- ii. regular Workshops on the ethical implications of algorithmic decision-making; and
- iii. Collaborative programmes with international bodies such as the OECD and the Council of Europe to advance a more robust AI-driven judicial proceedings in Nigeria.

### 5. Promote Adequate Access to AI-Powered Legal Services

To promote such access, policies ought to prioritize:

- i. provision for AI-assisted judicial services in different Nigerian languages, which reduces digital exclusion;
- ii. government subsidies for impoverished litigants' access to AI-powered legal tools; and
- iii. expansion of e-court infrastructure to cover underserved and rural areas.

This recommendation is in line with Nigeria's commitment to Goal 16 of the Sustainable Development Goals of the UN (SDGs 16).

---

## Conclusion

---

There is no doubt that AI-driven innovations for judiciary proceedings in Nigeria are momentous at this time, considering their invaluable benefits, such as time-saving and cost reduction. Conversely, the fact remains that such innovations (e.g., VCPs) are still nascent in Nigeria, thereby making them likely unconstitutional if not absolutely so. Perhaps that is why such virtual court proceedings stalled immediately after the COVID-19 pandemic lockdowns in Nigeria. On this note, Nigeria is required to grasp the fantastic and robust benefits associated with AI-driven innovations in the judiciary, particularly in the judicial proceedings.

Moving forward, AI innovations are fundamentally momentous to advance substantive justice in the judiciary, but must operate within the purview of the constitutional principle of a fair hearing. Adherence to the principle and its doctrines is sacrosanct. Thus, navigating the balance between efficiency and justice in AI-driven judicial proceedings is sacrosanct for piloting the innovations. This makes the enactment of an AI-specific statute a requirement that specifically takes into cognizance the constitutional provisions under Section 36(1), (3), and (4) as well as ethical considerations.

On the whole, the decline of Virtual Court Proceedings (VCPs) in Nigeria is less about inadequacy in the digital infrastructure and more about inadequacy in the law and policy. Without constitutional amendments to align with the global trend and significant investment in digital infrastructure, virtual hearings remain on shaky ground.

### Abstrakt

Sztuczna inteligencja (AI) stała się technologią wprowadzającą przełomowe zmiany w różnych aspektach działalności ludzkiej, a system sądowniczy nie stanowi tu wyjątku. Jednocześnie AI jest coraz częściej wdrażana w administracji wymiaru sprawiedliwości w Nigerii – od systemów elektronicznego składania dokumentów, cyfrowego zarządzania sprawami i wirtualnych postępowań sądowych po potencjalne zastosowania w zakresie orzeczeń opartych na prognozach.

Chociaż AI niewątpliwie obiecuje zwiększenie wydajności, zmniejszenie zaległości w rozpatrywaniu spraw oraz oszczędność kosztów, stwarza ona również poważne wy-

zwania prawne i etyczne, w tym wpływa na naruszenie prawa do rzetelnego procesu sądowego – prawa chronionego przede wszystkim na mocy art. 36 konstytucji Nigerii oraz kilku innych aktów prawnych, w tym: ustawy o dowodach z 2011 r., ustawy o wymiarze sprawiedliwości w sprawach karnych z 2015 r., stanowych ustaw o wymiarze sprawiedliwości w sprawach karnych, reguł postępowania w sprawach dochodzenia praw podstawowych (z 2009 r.) i innych.

Niniejszy artykuł wykorzystuje metodę badań doktryny prawnej w celu krytycznej analizy włączenia sztucznej inteligencji do postępowań sądowych oraz oceny, czy wzmacnia ona prawo do rzetelnego procesu sądowego w Nigerii, czy też temu prawu zagraża. Artykuł bada ponadto, czy algorytmiczne podejmowanie decyzji, brak przejrzystości, stronniczość danych oraz inne czynniki związane z komputerowym charakterem sztucznej inteligencji naruszają proceduralne gwarancje rzetelności, bezstronności oraz doktrynę praworządności. Na podstawie analiz doktryny oraz prawnoporównawczej rozważono zgodność reform dotyczących AI z nigeryjskim orzecznictwem konstytucyjnym oraz dotyczącym praw człowieka – pod kątem dostosowania do najlepszych światowych praktyk. Stwierdzono, że sztuczna inteligencja powinna służyć jako narzędzie sprawiedliwości, a nie jej substytut, tak aby móc zapewnić Nigerii wiarygodne, sprawiedliwe i równe dla wszystkich postępowania sądowe. Zaleca się wprowadzenie modułu ustawicznego kształcenia w zakresie AI dla sędziów, adwokatów i pracowników sądów, które umożliwiłoby ostrożne stosowanie tej technologii. Drogą do tego jest uchwalenie ustawy regulującej sztuczną inteligencję i wymiar sprawiedliwości.

**Słowa kluczowe:** sztuczna inteligencja (AI), rzetelne rozpatrzenie sprawy, postępowania sądowe, konstytucja Nigerii, prawa człowieka i praworządność.

---

## REFERENCES

---

Administrative Conference of the United States (2024, Dec. 17), Using Algorithmic Tools in Regulatory Enforcement (Recommendation 2024–25), Committee on Regulation; <https://www.acus.gov/document/using-algorithmic-tools-regulatory-enforcement>.

African Commission on Human and Peoples' Rights (1981), African Charter on Human and Peoples' Rights; <https://achpr.au.int/en/charter/african-charter-human-and-peoples-rights>.

African Union (1990), African Charter on the Rights and Welfare of the Child; <https://au.int/en/treaties/african-charter-rights-and-welfare-child>.

Anesu, C. A. (2025), *Artificial Intelligence and Judicial Integrity: Evaluating the Impact, Risks, and Implications of AI Integration in Modern Court Systems*, 'International Journal of Legal Studies and Social Sciences', 3(4), pp. 295–307; <https://ijlsss.com/wp-content/uploads/2025/08/27.-Asenath-Chitunzi-Anesu.pdf>.

- Asein, A. (2020), *Innovation in the Nigerian Judiciary: Prospects and Challenges of Court Digitisation*, 'Nigerian Bar Journal', 12(1), pp. 1-22.
- Aristotle (2009), *Nicomachean Ethics* (trans. W. D. Ross), Digireads.com Publishing (Original work published ca. 350 B.C.E.).
- Bello, O. A., & Ogufere, C. (2024), *Emerging Artificial Intelligence Legal-Judicial System's Interface: The Nigerian Judicial-Legal System*, 'Commonwealth Cyberjournal', 2(1); <https://production-new-commonwealth-files.s3.eu-west-2.amazonaws.com/s3fs-public/2024/06/ccj-2-1-emerging-ai-legal-judicial-systems-interface-bello-ogufere.pdf>.
- Bielu, K. J. (2018), *Fair Hearing in Judicial Adjudication in Nigeria*, 'Unizik Law Journal', 14(1), pp. 119-130; <https://www.journals.ezenwaohaetorc.org/index.php/ULJ/article/view/1199>.
- Citron, D. K. (2008), *Technological Due Process*, *Washington University Law Review*, 85(6), pp. 1249-1314; [https://openscholarship.wustl.edu/law\\_lawreview/vol85/iss6/2/](https://openscholarship.wustl.edu/law_lawreview/vol85/iss6/2/).
- Davies, A., Olugasa, O., & Odunaike, D. A. (2024), *Experimental Trials of the Virtual Court Proceedings in Nigeria: An Empirical Perspective*, 'International Journal for Court Administration', 15(2), Article 4; <https://doi.org/10.36745/ijca.550>.
- Dicey, A. V. (1959), *Introduction to the Study of the Law of the Constitution* (10th ed.), Macmillan.
- Dymytruk, M. (2019), *The Right To a Fair Trial in Automated Civil Proceedings*, 'Masaryk University Journal of Law and Technology', 13(1), pp. 27-45; <https://doi.org/10.5817/MU-JLT2019-1-2>.
- Eubanks, V. (2018), *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*, New York: St. Martin's Press.
- Eze, B. O. (2024), *The Future of Artificial Intelligence and the Administration of Law and Justice in Nigeria*, SSRN; <https://papers.ssrn.com/abstract=4689396>.
- Fine, A., Berthelot, E. R., & Marsh, S. (2025), *Public Perceptions of Judges' Use of AI Tools in Courtroom Decision-Making: An Examination of Legitimacy, Fairness, Trust, and Procedural Justice*, 'Behavioral Sciences', 15(4), p. 476; <https://doi.org/10.3390/bs15040476>.
- Fuller, L. L. (1969), *The Morality of Law* (Rev. ed.), New Haven: Yale University Press.
- Godwin, A. (2021, Mar. 12). *How E-Filing of Court Processes Improved Judicial System*, by Rivers CJ, 'The Guardian Nigeria'; <https://guardian.ng/news/how-e-filing-of-court-processes-improved-judicial-system-by-rivers-cj/>.
- Höber, L.-M. (2024, Jun. 5), *Efficiency vs. Effectiveness: 7 Examples to End Confusion*, 'Tea-mazing', <https://www.teamazing.com/efficiency-vs-effectiveness>.
- Ige, I.-O. (2023, Apr. 27), *How Judges, Lawyers, Plotted 'Coup' Against Virtual Court System in Nigeria*, 'Vanguard'; <https://www.vanguardngr.com/2023/04/how-judges-lawyers-plotted-coup-against-virtual-court-system-in-nigeria/>.
- Iwuchukwu, F. (2025, Jun. 6), *Federal High Court Launches E-Filing System in Lagos, Ends Manual Case Filing June 20*, 'Tribune Online'; <https://tribuneonlineng.com/federal-high-court-to-end-manual-filing-in-lagos-by-june-20/>.

Javed, K., & Li, J. (2025), *Bias in Adjudication: Investigating the Impact of Artificial Intelligence, Media, Financial and Legal Institutions in Pursuit of Social Justice*, 'PLOS ONE', 20(1), e0315270; <https://doi.org/10.1371/journal.pone.0315270>.

Judge President Mlambo (2020, Jan. 11), *Practice Directive 1 of 2020 (full implementation of the CaseLines digital/electronic system effective 27 January 2020)*; <https://www.ppv.co.za/judge-presidents-practice-directive-1-of-2020/>.

Kehinde, A. O., Oyedeki, A., & Awonuga, A. (2024), *Artificial Intelligence and the Judicial System in Nigeria: The Need for Transformation*, 'Lex et Scientia International Journal', 31(2), pp. 78–92; [https://lexetscientia.univnt.ro/download/2024\\_XXXI\\_2\\_8\\_LESIJ.pdf](https://lexetscientia.univnt.ro/download/2024_XXXI_2_8_LESIJ.pdf).

Lagos Judiciary (2020), *Remote Hearing of Cases (COVID-19 Period) Practice Direction*, Lagos State Government Press.

Mansour, S., & Rajput, D. (2025), *Enhancing Judicial Efficiency Through Artificial Intelligence: Analyzing Federal Justice Systems from an Organizational Behavior Perspective—A Data-Driven Study*, 'International Journal of Data Science and Advanced Analytics', 6(2), 403–410. <https://doi.org/10.69511/ijdsaa.v6i7.262>.

Masenya, T. M., & Ntengenyane, K. (2022), *The Management of Digital Court Records for Justice Delivery in the South African High Courts*, 'Mousaion: South African Journal of Information Studies', 40(3), pp. 1–22; <https://doi.org/10.25159/2663-659X/12691>.

Matsopoulos, P. (2021), *Artificial Intelligence and the Right To a Fair Trial: Navigating the Challenges*, 'Journal of Law, Technology and Society', 18(2), pp. 145–162.

Murray, A. (2010), *Information Technology Law: The Law and Society*, Oxford University Press.

Musa, D. (2025, Jun. 5). *Federal High Court Unveils Lagos E-filing Platform on June 23, 'Punch'*; <https://punchng.com/federal-high-court-unveils-lagos-e-filing-platform-on-june-23/>.

Nwabueze, B. O. (1977), *Constitutionalism in the Emergent States*, C. Hurst & Company.

Nwakodo, U. P., & Ubani, E. O. (2023), *The Judiciary and Democratic Politics in Nigeria*, 'Journal of Health, Applied Sciences and Management', 7(1), pp. 226–237; <https://dx.doi.org/10.4314/johasam.v7i1.28>.

Nweze, C. C. (2021), *ICT and Access to Justice in Nigeria: Opportunities for Reform* [in:] A. Bello (ed.), *Law and Technology in Africa* (pp. 67–88), Lagos: NIALS Press.

Obi-Farinde, M. (2025, May 16), *Transforming the Judiciary: ODR and the Future of Court Systems in Africa-Lagos*, Mediate.com; <https://mediate.com/transforming-the-judiciary-odr-and-the-future-of-court-systems-in-africa-lagos/>.

Oduwole, N. (2022, Dec. 13), *Justice System Needs Artificial Intelligence, Clinical Legal Education, Other Reforms*, Olabisi Onabanjo University; <https://oou-agoiwoye.com.ng/justice-system-needs-artificial-intelligence-clinical-legal-education-other-reforms/index.html>.

Office of the Chief Justice, South Africa (n.d.), *Court Online: E-Filing, Digital Case Management And Evidence Management System*; <https://www.judiciary.org.za/index.php/court-online>.

Oni, S. O. (2023), *Right to a Fair Hearing: An Appraisal of the Misconception Between Nigeria and United Kingdom*, 'International Journal of Law', 9(4), pp. 126–134; <https://www.lawjournals.org/assets/archives/2023/vol9issue4/9125-1692248708960.pdf>.

Paliwala, A. (2020), *Rethinking Law in the digital Age: Human Rights and Technological Transformations*, 'International Journal of Law and Information Technology', 28(3), pp. 203–221; <https://doi.org/10.1093/ijlit/ehaa012>.

Rawls, J. (1971), *A Theory of Justice*, Harvard University Press.

Schemmer, M., Kuehl, N., Benz, C., Bartos, A., & Satzger, G. (2023), *Appropriate Reliance on AI Advice: Conceptualization and the Effect of Explanations* [in:] IUI '23: Proceedings of the 28th International Conference on Intelligent User Interfaces (pp. 410–422). ACM; <https://doi.org/10.1145/3581641.3584066>.

Sule, I., Olorunyomi, F., Ajah, J., & Usman, A. (2024), *Virtual Court Proceedings in Nigeria: Some Legal Matters Arising*, 'European Journal of Law and Political Science', 3(3), pp. 17–22; <https://doi.org/10.24018/ejpolitics.2024.3.3.91>.

Tahura, U. S., & Selvadurai, N. (2022), *The Use of Artificial Intelligence in Judicial Decision-Making: The Example of China* [Working paper], Macquarie University.

Ubanyionwu, C. J. (2024), *Potential Benefits of Artificial Intelligence in Improving Efficiency and Accuracy in Legal Proceedings in the Nigerian Legal System*, 'International Review of Law and Jurisprudence' (IRLJ), 6(2), <https://nigerianjournalsonline.com/index.php/IRLJ/article/download/5463/pdf>.

Ubanyionwu, C. J. (2024, Sept. 29). *The Paradigm Shift in the Adjudication of Cases in Nigeria Through Artificial Intelligence (AI): Issues and Challenges*, 'Nnamdi Azikiwe University Journal of Private and Property Law', 1(2); <https://journals.unizik.edu.ng/naujpp/article/view/4367>.

Umeh, J. (2023, Feb. 22), *Lagos Judiciary Automates Court Processes with Cinfores Software*, Vanguard; <https://www.vanguardngr.com/2023/02/lagos-judiciary-automates-court-processes-with-cinfores-software/>.

United Nations (1948), *Universal Declaration of Human Rights*; <https://www.un.org/en/about-us/universal-declaration-of-human-rights>.

United Nations (1951), *Convention Relating to the Status of Refugees*; <https://www.unhcr.org/1951-refugee-convention.html>.

United Nations (1966), *International Covenant on Civil and Political Rights*; <https://www.ohchr.org/en/instruments-mechanisms/instruments/international-covenant-civil-and-political-rights>.

United Nations (1989), *Convention on the Rights of the Child*; <https://www.ohchr.org/en/instruments-mechanisms/instruments/convention-rights-child>.

United Nations (2003), *United Nations Convention Against Corruption*; <https://www.unodc.org/unodc/en/treaties/CAC/>.

United Nations (2006), *Convention on the Rights of Persons with Disabilities*; <https://www.ohchr.org/en/instruments-mechanisms/instruments/convention-rights-persons-disabilities>.

United Nations (2015), *Transforming our World: The 2030 Agenda for Sustainable Development*, New York: United Nations.

VDMA Law (2025, Aug. 6), *South African Courts Weigh in on the Ethical Use of Artificial Intelligence in Legal Practice*; <https://vdmalaw.com/2025/08/06/south-african-courts-weigh-in-on-the-ethical-use-of-artificial-intelligence-in-legal-practice/>.

Winner, L. (1980), *Do Artifacts Have Politics?*, 'Daedalus', 109(1), pp. 121–136. <http://www.jstor.org/stable/20024652.CI/DIT/2024/GL/01>, 2024.



Igor Zgoliński\*

## Wirtualna sala narad i inne nowe instrumenty postępowania karnego. Uwagi na tle zagrożeń i kosztów cyfryzacji

### [Virtual Courtrooms and Other New Tools in Criminal Proceedings: Notes on the Risks and Costs of Digitization]

#### Abstract

This publication analyzes the critical issues surrounding the necessary digital reform of the justice system, using criminal proceedings as an example. It points out that digitization represents a fundamental paradigm shift in the concept of a fair trial—which must be accompanied by appropriate changes in the law. It also identifies the need to implement numerous new tools in criminal proceedings, including a virtual courtroom (conference room) for remote hearings. Specific amendments to criminal procedure regulations are proposed, emphasizing the necessity of a broader legislative revision covering exclusively digital issues. On the one hand, the digitization of criminal proceedings (*e-justice*) offers the potential to expedite proceedings and facilitate access to case files; on the other hand, however, it poses serious threats to the fairness of the trial. It also imposes significant financial burdens on the state. Key risks include: digital exclusion, the erosion of the principle of directness, and issues regarding the confidentiality of communications and the transparency of proceedings. A full understanding of these risks and the associated costs is essential for the *e-justice* system to effectively support the administration of justice.

**Keywords:** criminal proceedings, digitization, virtual courtroom (conference room), risks, costs.

## Wstęp

Truizmem jest stwierdzenie, że cyfryzacja postępowań karnych (*e-justice*) może znacząco przyspieszyć przebieg postępowań karnych oraz ułatwić

\* **Igor Zgoliński** – profesor uczelni, dr hab. nauk prawnych, Politechnika Bydgoska im. Jana i Jędrzeja Śniadeckich (afiliacja), sędzia SN; <https://orcid.org/0000-0002-5097-6170>; [igor.zgolinski@pbs.edu.pl](mailto:igor.zgolinski@pbs.edu.pl) / associate professor, PhD in Law, post-doctoral degree, Bydgoszcz University of Science and Technology (affiliation); judge of the (Polish) Supreme Court.

stronom i innym podmiotom zainteresowanym dostęp do akt. Niewątpliwie należy dążyć do pełnej cyfryzacji, gdyż w tym kierunku zmierza postęp cywilizacyjny. W przypadku sądów usprawni to ich pracę i odwróci istniejący od czasów zaborów stan zbierania dokumentów papierowych, składania w poszczególne tomy akt, a następnie magazynowania i archiwizacji po zakończeniu postępowania.

Cyfryzacja jest przedsięwzięciem niezwykle skomplikowanym oraz kosztownym. Ma wiele odmian i generuje stałe wydatki, nieponoszone dotąd przez Skarb Państwa. Prócz wielu pozytywnych aspektów stwarza również wiele nowych poważnych zagrożeń. Z wymienionych powodów niezbędne staje się podjęcie próby określenia głównych zagrożeń oraz czynników powodujących zwiększenie wydatków na wymiar sprawiedliwości, a ponadto dokonanie ich stosownej analizy. Pozwoli to w przyszłości zniwelować niebezpieczeństwa związane z cyfryzacją, co niewątpliwie będzie pożytkiem dla podmiotów tworzących postęp w omawianej dziedzinie oraz uczestniczących w postępowaniach karnych. Przeanalizowanie obszarów wywołujących koszty pozwoli natomiast na związualizowanie skali problemu związanego z tym procesem i uświadomi źródła przyczyn, które ów postęp mogą (i będą) znacząco ograniczać. Celem analizy jest zatem przede wszystkim identyfikacja i wieloaspektowa analiza zagrożeń wynikających z informatyzacji procesu karnego, w tym wskazanie konfliktów między nowoczesną technologią a naczelnymi zasadami procesowymi<sup>1</sup> – z uwzględnieniem barier ekonomicznych. Badania dążą też do wypracowania postulatów legislacyjnych i technicznych, które pomogą zniwelować przyszłe ryzyka. W tekście zastosowano różne metody badawcze. Podstawowa jest metoda dogmatycznoprawna, a wspierają ją metody analityczno-badawcza, postulatyczna oraz analiza ryzyka.

---

## Zagrożenia wynikające z cyfryzacji postępowań

---

W dobie implementacji systemów AI i postępującej cyfryzacji akt postępowań (e-akta) zagrożenia związane z cyfryzacją procesu karnego można podzielić na szereg obszarów. Każdy z nich generuje specyficzne ryzyka dla sprawowania wymiaru sprawiedliwości. W przypadku zagrożeń dla zasad rzetelnego procesu (*fair trial*) jako kluczowy postrzegać należy problem wykluczeń cyfrowych. Jest to problem nie tylko socjologiczny, gdzie został

---

<sup>1</sup> W literaturze przedmiotu wskazano już nawet na zasadę cyfryzacji, obowiązująca względem do instytucji prawnych określonych w k.p.k. – zob. szerzej: P. Czarnecki, *Zasada elektronizacji (cyfryzacji) procesu karnego* [w:] *W pogoni za rzetelnym procesem karnym*. Księga dedykowana Profesorowi Stanisławowi Waltosiowi, D. Szumilo-Kulczycka (red.), Warszawa 2022, ss. 440–443.

relatywnie dobrze zbadany<sup>2</sup>, lecz stanowi również problem w cyfryzacji postępowań. Zawsze będzie kładł się cieniem nad tym mechanizmem. Społeczeństwo w dużej mierze nadal nie potrafi w odpowiedni sposób poruszać się po świecie cyfrowym i w pełni korzystać z udogodnień, które się z tym wiążą. Co więcej, brak odpowiedniego sprzętu lub jedynie owych umiejętności, zwłaszcza po stronie podejrzanego bądź oskarżonego, może prowadzić do nierówności w dostępie do sądu, do akt postępowania bądź do poszczególnych czynności procesowych. Przepaść cyfrowa może wręcz uniemożliwiać lub znacząco utrudniać realizację prawa do obrony, jeśli dostęp do dowodów, pism procesowych czy udział w czynnościach będzie wymagał specjalistycznej wiedzy technicznej lub odpowiedniego sprzętu. Konieczne jest zatem takie wdrażanie i udoskonalanie cyfryzacji postępowań, które będzie wiodło do wykreowania nie tylko narzędzi prostych w obsłudze od strony technicznej i niewymagających wysokiej jakości sprzętów, lecz również zapewniało alternatywny udział w tymże postępowaniu w formule tradycyjnej (dotychczasowej). Państwo musi więc utrzymywać model hybrydowy, gwarantujący prawo do tradycyjnego procesu osobom nieposiadającym kompetencji lub odpowiedniego sprzętu. W dobie dążenia do nowoczesności prawo do analogowego procesu powinno stać się nowym prawem człowieka, chroniącym przed wykluczeniem i błędami systemowymi. Wpisuje się ono bezsprzecznie w karnoprosesową zasadę równego traktowania stron i funkcjonalną regułę równouprawnienia<sup>3</sup>.

Kolejnym ważkim zagadnieniem z tej sfery jest potencjalne ograniczenie karnoprosesowej zasady bezpośredniości. Rozprawy zdalne mogą utrudniać sędziemu oraz pozostałym osobom (podmiotom) w niej uczestniczącym właściwą ocenę prowadzonych dowodów. Jest to szczególnie widoczne choćby w ramach oceny wiarygodności świadków (brak możliwości odczytania pełnej mowy ciała w obliczu niepełnego obrazu osoby składającej zeznania)<sup>4</sup>. Może to również prowadzić do osłabienia kontaktu obrońcy z mandantem (brak możliwości poufnej konsultacji w trakcie przesłuchania). Omawiana kwestia wydaje się trudna do zupełnego zniwelowania, gdyż rozprawy zdalne nawet przy zaawansowanej technologii, raczej nie są w stanie zastąpić bezpośredniego uczestnictwa i wrażeń (odczuć) z tym związanych. Pokusić się można o kilka uwag na tej płaszczyźnie. Celowe wydaje się

<sup>2</sup> Zob. np. S. Czetwertyński, A. Mroczek-Czetwertyńska, Wykluczenie cyfrowe w społeczeństwie informacyjnym, „Nauki Społeczne” 2012, 1, 5, ss. 116–127, M. Castells, Społeczeństwo sieci, Warszawa 2007, H.S. Dordick, G. Wang, The Information Society: A Retrospective View, Newbury Park 1993, J. Kujawski, Wykluczenie cyfrowe jako forma wykluczenia społecznego. Przypadek Polski, „Media i Społeczeństwo” 2018, 9, s. 252 i nn., R. Jedlińska, Wykluczenie cyfrowe w Polsce na tle krajów Unii Europejskiej, „Ekonomiczne Problemy Usług” 2018, 131, 2, ss. 225–236 i nn.

<sup>3</sup> M. Rogacka-Rzewnicka, O równości, równouprawnieniu i równowadze procesowej w postępowaniu karnym [w:] W pogoni..., ss. 161–163.

<sup>4</sup> J. Marszałek, M. Skwarcow, Użycie sztucznej inteligencji w postępowaniu karnym, „Przegląd Sądowy” 2025, 3, s. 63.

zwiększenie liczby kamer na salach rozpraw, które z różnych perspektyw przekazywałyby wizję oraz fonię – co w założeniu zwiększy transparentność przekazu i wzmocni zasadę bezpośredniości. Brak fizycznej obecności na sali sądowej osłabia wszakże empatię – w tym empatię sędziego – oraz utrudnia oskarżonemu aktywny udział w procesie (chodzi np. o trudność w przerwaniu wypowiedzi, by skonsultować się z obrońcą przez komunikator, podjęcie adekwatnej reakcji przez złożenie uzupełniających wyjaśnień itd.). Pamiętać trzeba, że cyfryzacja zmienia dynamikę sali rozpraw, co często uderza w najsłabszych uczestników postępowania. Problemem staje się zresztą nie tylko mały obraz, lecz i stale postępująca technologia wygładzania wizerunku. Algoritmy kompresji wideo i filtry AI (nawet te, które standardowo są umieszczone w komunikatorach) mogą maskować mikromimikę świadka, która dla obserwujących będzie sygnałem stresu lub objawem kłamstwa. Konieczne jest więc wprowadzenie określonego standardu „surowego sygnału wideo” w celach procesowych, wolnego od owych algorytmicznych upiększeń. Zasadę bezpośredniości wzmocniłoby natomiast wbudowanie przeglądarki akt podczas rozprawy zdalnej oraz w wirtualnej sali narad – uniemożliwiającej wykonywanie zrzutów ekranu (*watermarking*) i zapewniającej pełną kontrolę nad tym, kto w danej chwili ogląda dany dokument.

Wydaje się ponadto, że warto umieścić w narzędziach umożliwiających prowadzenie rozpraw zdalnych możliwość szybkiego przejścia do innego pokoju osób biorących w niej udział – w celu zapewnienia osobnej, niejako pozasądowej, komunikacji (wirtualna, cyfrowa sala narad). Bez wątplenia taka opcja powinna być zapewniona dla kontaktu obrońcy z oskarżonym, dla zachowania tajemnicy obrończej (*attorney-client privilege*). Do rozważenia pozostaje, czy winno się to odbywać za zgodą przewodniczącego, czy też bez. Mając na względzie specyfikę rozpraw zdalnych oraz możliwość równoległego porozumiewania się za pomocą innych środków komunikacji, uważam, że mogłoby to odbywać się niejako obok rozprawy, a więc bez zgody przewodniczącego. W sytuacjach skomplikowanych należałoby zarządzić do tego specjalną przerwę w rozprawie, co już wymaga zgody. Podobnej zgody będą wymagały inne sytuacje występujące na sali rozpraw, z reguły w sposób nagły i pilny, jak choćby: wniosek o dobrowolne poddanie się karze, potrzeba narady nad złożeniem dodatkowych wyjaśnień, sygnalizacji przez biegłych potrzeby przeprowadzenia dodatkowych czynności bądź dyskusji ze stronami postępowania. Możliwość ta prawdopodobnie nie będzie się cieszyła – przynajmniej na początku – dużym zainteresowaniem z uwagi na obawy używania narzędzia nowego i dodatkowo udzielonego przez sąd (ewentualnie organ ścigania bądź organ wykonawczy). Będzie to wywoływało przede wszystkim obawę braku zachowania tajności (poufności). Wirtualna sala narad będzie jednak miała zdecydowanie szersze wykorzystanie. Może służyć do narady składu orzekającego nad orzeczeniem oraz innych narad czy dyskusji

w toku postępowania. Może być też wykorzystywana – jak już wspomniano – przez biegłych, np. przez biegłego psychologa, do różnego typu rozmów z pokrzywdzonym bądź zebrania danych osobopoznawczych. Systemy e-justice powinny być więc projektowane w architekturze *end-to-end encryption* dla kanałów bocznych (E2EE), do których klucze deszyfrujące posiadałyby wyłącznie strony, a nie operator systemu. Innymi słowy, gdy osoby zalogują się do tej sali, wówczas ich urządzenia wymieniają wzajemnie klucze publiczne przez serwer, który nie będzie ich przechowywał w formie deszyfrującej (co jest znane np. z komunikatora Signal), a wiadomości na czacie bocznym będą szyfrowane kluczem sesyjnym odbiorcy na urządzeniu nadawcy. Uczestnicy mogą wówczas zweryfikować tożsamość rozmówców poprzez porównanie tzw. *safety numbers* (skrótów kryptograficznych kluczy publicznych) poza głównym kanałem.

Następna kwestia – nieodległa od powyżej omawianej – dotyczy potencjalnej obawy o naruszenie zasady jawności postępowania<sup>5</sup>. Jakkolwiek technologia umożliwia transmisję, to bariery techniczne lub błędy systemowe mogą *de facto* wykluczać publiczność z udziału w procesie, co także może w określonych sytuacjach godzić w transparentność wymiaru sprawiedliwości. Problem stanie się szczególnie nabrzmiały w kontekście postępowań cieszących się dużym zainteresowaniem opinii publicznej. *De lege lata* jest rozwiązywany za pomocą wydawania specjalnych kart wstępu, co ma przeciwdziałać zakłócaniu przebiegu rozprawy. W świecie cyfrowym antidotum mogłoby się okazać umożliwienie transmisji z otwartym dostępem (*open access*), rzecz jasna bez możliwości jakiegokolwiek ingerencji bądź czynnego uczestnictwa w postępowaniu.

Następnym obszarem zagrożeń są różnego typu ryzyka związane z dowodami cyfrowymi (*deepfakes*) i manipulacją danymi (zjawisko syntetycznych dowodów). Postępująca dostępność rozmaitych narzędzi AI pozwala na tworzenie realistycznych, acz przecież fałszywych obrazów, fotografii, nagrań wideo, audio czy dokumentów. Generuje to ryzyko wprowadzenia do procesu sfabrykowanych nagrań wideo lub audio, które są na pierwszy rzut oka niemożliwe do odróżnienia od prawdziwych. Wejdziemy więc w erę dowodu probabilistycznego, w której autentyczność nagrania czy dokumentu przestaje być oczywista. W konsekwencji w toku procesu niezbędne stanie się sięgnięcie do kolejnych dowodów (weryfikujących) w postaci specjalistycznych ekspertyz biegłych z zakresu informatyki śledczej. Wymuszać to będzie niekiedy prowadzenie wpadkowych, żmudnych postępowań, zmierzających jedynie do odróżnienia autentycznych dowodów od zaawansowanych manipulacji. Dowody cyfrowe są niezwykle podatne na różnego rodzaju modyfikacje. Brak ścisłych procedur (np. sum kontrolnych, *blockchain*) przy zabezpieczeniu da-

<sup>5</sup> P. Czarnecki, *Rozprawa sądowa w Internecie. Postulat wprowadzenia możliwości transmisji (streamingu) rozpraw sądowych w rzetelnym procesie karnym*, „Prokuratura i Prawo” 2025, 3, s. 58.

nych może prowadzić do podważenia ich autentyczności oraz wiarygodności przed sądem. Poza tym dowody cyfrowe bywają „ulotne”, np. brak natychmiastowego zabezpieczenia sumy kontrolnej (*hash*) pliku pozwala na jego edycję na etapie zabezpieczania, co może wywołać zarzut manipulacji. Dane przechodzą też przez serwery różnych dostawców usług. Dużym ryzykiem jest np. uzyskanie dowodów z naruszeniem procedur międzynarodowych (np. bez stosownego wniosku o pomoc prawną), a to może prowadzić do niedopuszczalności wykorzystania takich dowodów. Uzyskanie danych z pominięciem procedur MLAT (*Mutual Legal Assistance Treaty*) lub bez uwzględnienia wyroków TSUE dotyczących retencji danych<sup>6</sup> stwarza ryzyko uchylania wyroków w postępowaniu odwoławczym. Cyfryzacja bez uregulowania statusu dowodów z chmury może więc stać się generatorem przewlekłości postępowań. Tego typu sytuacje również będą bowiem wymagały stosownych weryfikacji procesowych, co będzie wpływało na ogólną sprawność postępowania. Nie wydaje się jednak, aby można było zupełnie uniknąć tego zjawiska. Jest ono nieuchronne. Pamiętać natomiast należy, że świadome wprowadzanie do postępowania karnego fałszywych dowodów winno być oceniane przez pryzmat naruszenia norm kodeksu karnego, który już obecnie przewiduje penalizację tego rodzaju zachowań (zob. art. 234 i 235 k.k.). Z drugiej strony pojawi się również wyraźne pole do nadmiernej inwigilacji, gdyż cyfryzacja umożliwia organom ścigania stosowanie metod w postaci „przeszukań zdalnych” czy też masowej analizy metadanych, w tym niezwiązanych ze sprawą. Działania te odbywają się w szarej strefie, dodatkowo w obliczu niedookreślonego stanu prawnego, co potęguje możliwość naruszania prawa do prywatności. Pilnego, wyprzedzającego wdrożenia wymagają zatem odpowiednie regulacje prawne (np. rozporządzenia), które wytyczą podstawowe reguły dla tego rodzaju czynności. Dane z polskich postępowań mogą być też przetwarzane na zagranicznych serwerach, a to implikuje pytania o suwerenność cyfrową wymiaru sprawiedliwości i faktyczną ochronę danych np. przed obcymi wywiadami<sup>7</sup>. W celu mitygowania tych zagrożeń kluczowe jest więc nie tylko stosowanie Aktu o sztucznej inteligencji (AI Act) – który klasyfikuje systemy AI stosowane w wymiarze sprawiedliwości jako systemy wysokiego ryzyka, wymagające najwyższego stopnia nadzoru ludzkiego – ale i jego wewnętrzskrajowe uszczegółowienie.

Zagadnieniem doniosłym z perspektywy cyfrowych zagrożeń są także różnego typu bolączki samych algorytmów sztucznej inteligencji (błąd algorytmiczny Bias). Systemy wspomagające decyzje (np. przy szacowaniu ryzyka prawdopodobieństwa popełnienia przestępstwa) mogą powielać różnego typu uprzedzenia historyczne (rasowe, majątkowe, terytorialne i inne), prowadząc

<sup>6</sup> Na przykład wyrok z 8.04.2014 r. w sprawach połączonych C-293/12 i C-594/12.

<sup>7</sup> Zob. P. Opitek, *Wybrane aspekty pozyskiwania dowodów cyfrowych w sprawach karnych*, „Prokuratura i Prawo” 2018, 7–8, s. 65.

do niewłaściwych wyników, a w konsekwencji także orzeczeń. Brak przejrzystości działania algorytmów sprawia, że strony mogą nie zweryfikować tego, na jakiej podstawie system wygenerował dany wynik bądź sugestię. Godzi to też w rzetelność uzasadniania orzeczeń. Odmianą omawianej kwestii są tzw. halucynacje AI w analizie prawnej. Istnieje bowiem duże ryzyko bezkrytycznego korzystania przez organy stosujące prawo z systemów generatywnych, np. do streszczania akt, co może prowadzić do pominięcia kluczowych faktów lub „zmyślenia” przez AI nieistniejących okoliczności. Ważne z tej perspektywy jest ujawnienie kodów źródłowych, gdyż w przeciwnym razie zawsze będzie wywoływało to spory i podejrzenia o stronniczość.

Centralizacja akt w chmurze lub systemach sądowych stworzy natomiast wysokie ryzyko masowych wycieków danych osobowych pokrzywdzonych, świadków oraz innych danych tam zawartych. E-akta zawierają dane medyczne, intymne szczegóły z życia ofiar i dane świadków. Wyciek takich informacji może prowadzić do wtórnej wiktyimizacji ofiar lub odwetu na świadkach. Niezależnie od powyższego ataki typu *ransomware* na systemy sądowe mogą sparaliżować pracę sądów, co może z kolei prowadzić do przedawnienia spraw lub zastoju, a nawet zawieszania postępowań z powodu braku dostępu do akt. Nie trzeba dodawać, że może też implikować konieczność prowadzenia postępowań odtworzeniowych. Zablokowanie dostępu do e-akt przez hakerów może sparaliżować tysiące spraw jednocześnie, prowadząc do jaskrawego naruszenia prawa do rozpoznania sprawy w rozsądnym terminie.

Nie będzie odkrywczym stwierdzeniem, że od początku należy wdrożyć stały monitoring działania systemu, zadbać o bezpieczeństwo cyfrowe oraz ochronę prywatności. Polska zajmuje obecnie 6. miejsce na świecie w globalnym rankingu cyberbezpieczeństwa (*Cyber Defense Index*). Świadczy to o bardzo wysokim poziomie ogólnej cybernetycznej odporności państwa, aczkolwiek specyficzność wymiaru sprawiedliwości stwarza unikatowe wyzwania i potęguje prawdopodobieństwo ataków. Potwierdza to fakt, że Polska znajduje się też w czołówce najczęściej atakowanych krajów świata, głównie przez grupy powiązane z Rosją (np. *People's Cyber Army* czy *APT44*). Sądy są więc narażone na ataki blokujące dostęp do usług (DDoS) oraz liczne próby sabotażu danych. Problem ten staje się szczególnie istotny – zważywszy, że od marca 2026 r. sądownictwo przechodzi na nowy, bardziej rozległy zakresowo model cyfrowy, obejmujący m.in. e-apelacje. W bieżącym roku ma więc nastąpić swoisty przełom cyfrowy, dwustronna zaś komunikacja z sądem stać się standardem, a odbywać się przez portale informacyjne. Systemy te są chronione przez krajowe jednostki CSIRT (Zespoły Reagowania na Incydenty Bezpieczeństwa Komputerowego), które już teraz odnotowują skokowy wzrost liczby ataków na infrastrukturę administracji publicznej. Polskie sądownictwo posiada co prawda solidną tarczę w postaci zaawansowanych systemów krajowych (NASK, CSIRT) i wysokiej pozycji w rankingach, jednak dynamiczna cyfry-

zacja e-apelacji oraz status Polski jako celu ataków hybrydowych sprawią, że bezpieczeństwo to będzie w stanie permanentnego testowania. Krytycznym wyzwaniem pozostaje tu zwłaszcza pełna implementacja unijnych standardów NIS2 w systemach sądowych. Mimo że Polska wdraża dyrektywę NIS2 (poprzez nowelizację ustawy o Krajowym Systemie Cyberbezpieczeństwa), proces ten stale napotyka na opóźnienia legislacyjne.

Wszystkie powyżej opisane aspekty wiążą się z obawami nie tylko o bezpieczeństwo, ale i stabilność techniczną infrastruktury wymiaru sprawiedliwości, przy relatywnie nieodległym i na pewno znacznym wzroście obrotu cyfrowego. Będzie to niewątpliwie niespotykanym dotąd wyzwaniem, nie tylko dla osób prowadzących postępowania, lecz również dla kadry administracyjnej, która powinna co najmniej od kilku lat przechodzić systematyczne szkolenia i nabywać odpowiednie kwalifikacje zawodowe. Brak tego typu działań zakreślonych na szeroką skalę systematycznie eskaluje krajowe opóźnienie w obszarze cyfryzacji.

---

## **Koszty cyfryzacji postępowań**

---

Zagadnienie kosztów cyfryzacji ma dwa oblicza. Z jednej strony są to wymierne wydatki, które winien ponieść Skarb Państwa na czynności z tym związane; z drugiej jednak strony na szali położona jest swoista korzyść społeczna w postaci usprawnienia postępowań, zapewnienia ich transparentności oraz zachowania właściwej relacji względem postępu cywilizacyjnego. Są to aspekty niewymierne, lecz ważniejsze niżli koszty finansowe. Faktyczne koszty cyfryzacji postępowań karnych będą ogromne i wydaje się, że nie jest możliwe ich bezdyskusyjne skonkretyzowanie, zwłaszcza pod kątem wysokości<sup>8</sup>. Jest to zagadnienie wysoce złożone, znacznie wykraczające poza np. sam zakup komputerów czy niezbędnych licencji. Faktyczne koszty będzie można ustalić dopiero po zakończeniu przetargów i pełnym wdrożeniu cyfryzacji, a i tak będą one systematycznie powiększane o koszty bezpieczeństwa, nadzoru (monitoringu) oraz stałego serwisowania i szkoleń. Z tych powodów nie jest możliwe szersze odniesienie się do tego aspektu w opracowaniu artykułowym. Zasadne jest natomiast określenie głównych czynników te koszty wywołujących.

Ujmując rzecz w skali makro: na samo wzmocnienie cyberbezpieczeństwa przeznaczono około 4 miliardów zł (dane z roku 2025), z czego część trafiła bezpośrednio do sektora publicznego i wymiaru sprawiedliwości. Znacznie większe koszty będzie jednak generował cały mechanizm, gdyż cyberbezpie-

---

<sup>8</sup> I. Zgoliński, *Technology and Criminal Proceedings: An attempt to systematize the issue and determine the main directions of implementation*, „Europejski Przegląd Prawa i Stosunków Międzynarodowych” 2025, 1, 73, s. 104.

czeństwo jest jedynie częścią cyfrowego usprawniania sądów. Jeśli chodzi o skalę wydatków w Polsce, to zgodnie z planami na lata 2024–2026 wydatki na samą transformację cyfrową wymiaru sprawiedliwości w Polsce szacuje się na setki milionów złotych rocznie. Przykładowo tylko na rozwój systemów informatycznych w sądach powszechnych przeznaczają się milionowe środki z Programu Operacyjnego Polska Cyfrowa oraz funduszy celowych Ministerstwa Sprawiedliwości.

Koszty cyfryzacji można ogólnie podzielić na koszty bezpośrednie (finansowe) oraz koszty pośrednie (systemowe). Bezpośrednie koszty finansowe (inwestycyjne), które dotąd były finansowane głównie z KPO (Krajowy Plan Odbudowy) oraz funduszy spójności to przede wszystkim infrastruktura sprzętowa. Niezbędne jest przecież wyposażenie sal rozpraw w systemy do wideokonferencji, kamery 4K, systemy nagłośnienia oraz bezpieczne terminale dla sędziów i protokolantów. Koszty generować będzie utworzenie państwowych chmur obliczeniowych do przechowywania – zapewne petabajtów<sup>9</sup> – danych (nagrania z rozpraw, e-akta, dowody cyfrowe itp.). Do tego doliczyć należy jeszcze wydatki związane z energią i chłodzeniem owych *data centers*. Koszty te są akurat możliwe do generalnego oszacowania na podstawie szeregu czynników szczegółowych, takich jak: liczba sal rozpraw, konfiguracja kamer, rodzaj i liczba potrzebnego sprzętu itd. Trudne do oszacowania będą natomiast nakłady finansowe przeznaczane na stworzenie autorskich systemów (jak choćby Portal Informacyjny Sądów, systemy e-doręczeń, systemy dla wewnętrznego użytku sądownictwa na wzór np. *gaius-lex*) oraz licencje na narzędzia informatyki śledczej (w tym do łamania zabezpieczeń, jak np. *Cellebrite*). Cały ten ciąg działań wymagać będzie, rzecz jasna, zatrudnienia ekspertów z branży IT, których wynagrodzenia w sektorze prywatnym są znacznie wyższe, a to z kolei wymusi adekwatne podnoszenie stawek w administracji sądowej – dla zapewnienia konkurencyjności. W przeciwnym razie nastąpi zapaść kadrowa, co utrudni dokonanie procesu cyfryzacji (o ile go wręcz nie uniemożliwi).

Do drugiej grupy kosztów, tzw. kosztów operacyjnych (*stricte procesowych*), zaliczyć trzeba natomiast w pierwszej kolejności wydatki na biegłych informatyków. Wzrost liczby dowodów cyfrowych powoduje wzrost kosztów postępowań, gdyż zwiększy się liczebność opinii biegłych. Koszty generować będzie nadto konieczność przeszkolenia wielu osób, które można liczyć w tysiącach, a więc sędziów, prokuratorów oraz pracowników sekretariatów<sup>10</sup>. Chodzi tu głównie o szkolenia z obsługi nowych systemów oraz z zakresu zachowania cyberhigieny. Jeszcze większe wydatki implikować będzie zwykła

<sup>9</sup> Jeden petabajt (PB) to około 500 miliardów stron standardowego tekstu drukowanego.

<sup>10</sup> Edukacja ta ma charakter stały, wpisuje się w metamorfozę kulturową i progresywne zwiększanie kompetencji cyfrowych; zob. K. Kiejnich-Kruk, *Building blocks – strategia cyfryzacji wymiaru sprawiedliwości. Perspektywa estońska*, „Przegląd Sądowy” 2025, 3, s. 92.

digitalizacja akt archiwalnych, co wiąże się ze skanowaniem papierowych akt spraw będących w toku, aby mogły stać się częścią e-akt. Do tego wymaga stworzenia całej infrastruktury.

W skład kosztów cyfryzacji postępowań zaliczyć także wypada tzw. koszty ukryte. Istnieje duże ryzyko błędów sądowych, a więc także nieuchronnych kosztów odszkodowań, np. za niesłuszne skazania bądź stosowanie tymczasowego aresztowania, wynikłe np. z błędnych interpretacji dowodów cyfrowych lub błędów algorytmu AI wspierającego decyzje. Tego typu koszty wynikać mogą także z różnego rodzaju ataków cyfrowych, awarii systemów, problemów z logowaniem świadków na rozprawy zdalne czy ataków DDoS – które mogą paraliżować sądy, znacząco przewlekać postępowania i generować straty. Państwo będzie musiało ponadto ponieść również duże koszty zapewnienia punktów dostępu do sieci lub pomocy technicznej dla osób wykluczonych cyfrowo (starszych, ubogich), aby nie straciły one prawa do obrony (koszt Digital Divide).

Choć docelowo cyfryzacja ma przynieść długofalowe oszczędności (brak kosztów wysyłki papierowej, mniejsza liczba konwojów więziennych dzięki rozprawom zdalnym), to obecnie znajdujemy się w fazie początkowych – i bodaj najwyższych – nakładów inwestycyjnych, gdzie koszty utrzymania starych (papierowych) i nowych (cyfrowych) struktur nakładają się na siebie. Gdy się jednak zna praktykę, można gorzko stwierdzić, że w najbliższych latach będzie zapewne dochodziło do przerzucenia tych kosztów (co najmniej częściowego) na poszczególne jednostki organizacyjne wymiaru sprawiedliwości, bez zapewnienia im niezbędnego wsparcia informatycznego.

---

## **Wnioski *de lege ferenda***

---

Bez „cyfrowej gwarancji kodeksowej”, czyli uregulowania wprost w kodeksie postępowania karnego odpowiednich prawnych bezpieczników technologicznych, cyfryzacja może stać się narzędziem co prawda efektywnym, ale wadliwym sprawiedliwościowo. Przygotowanie konkretnych propozycji legislacyjnych wymaga z pewnością dogłębnego spojrzenia i opracowania konkretnego modelu<sup>11</sup>, a następnie kompleksowego projektu zmian k.p.k. oraz szczegółowych rozporządzeń. Parę obszarów i kierunków zmian warto jednak wskazać, gdyż są niezbędne do mitygowania zidentyfikowanych zagrożeń. To zaledwie parę zmian, ale chodzi tutaj wyłącznie o zobrazowanie skali problemu, przed którym stoimy.

---

<sup>11</sup> M. Klubińska, *Czynności procesowe na odległość w procesie karnym – założenia modelowe* [w:] *Czynności procesowe na odległość w postępowaniu karnym*, M. Klubińska, A. Lach (red.), Warszawa 2025, s. 358.

Przede wszystkim są to gwarancje poufności komunikacji (zwłaszcza na linii obrońca–oskarżony). Należy wprowadzić przepisy gwarantujące nie naruszalność komunikacji w trakcie zdalnego postępowania – jak choćby przepisu stanowiącego, że system teleinformatyczny sądu musi zapewniać funkcjonalność wydzielonego, szyfrowanego kanału komunikacji (tzw. wirtualnego/cyfrowego pokoju narad). Komunikacja ta byłaby wyłączona spod jakiegokolwiek nadzoru czy protokołowania, w tym pod rygorem pominięcia czynności procesowej, której narada dotyczyła. W rozdziale 43 k.p.k. (po art. 374 k.p.k.) można byłoby dodać przepis o treści następującej: „System teleinformatyczny służący do przeprowadzenia rozprawy na odległość musi zapewniać oskarżonemu i jego obrońcy funkcjonalność niezwłocznego i poufnego porozumiewania się w wydzielonym kanale komunikacji elektronicznej (cyfrowy pokój narad), niedostępnym dla pozostałych uczestników postępowania. Treść komunikacji wewnątrz tego kanału objęta jest bezwzględną tajemnicą i nie podlega utrwaleniu przez system”.

Konieczne wydaje się też doprecyzowanie, w jaki sposób weryfikowane będą dowody podatne na zaawansowaną manipulację. Dobrym krokiem byłoby przyjęcie, że autentyczność dowodów cyfrowych (audio, wideo, dokumentów) wygenerowanych lub zmodyfikowanych przez AI powinna zostać zweryfikowana przez biegłego z zakresu informatyki, akredytowanego przez odpowiedni organ (np. CLKP). Niezbędne jest też wprowadzenie obowiązku dołączania do akt *hash verification report* (raportu sumy kontrolnej) dla każdego kluczowego dowodu cyfrowego, potwierdzającego jego integralność od momentu zabezpieczenia. Należy nadto ustawowo zagwarantować, że postęp technologiczny nie pozbawi nikogo prawa do sprawiedliwego sądu – choćby przez wprowadzenie przepisu dającego oskarżonemu prawo do złożenia wniosku o prowadzenie postępowania w formie tradycyjnej (papierowej i stacjonarnej) w przypadku wykazania, że jego sytuacja osobista, zdrowotna lub brak dostępu do technologii uniemożliwia mu pełną realizację prawa do obrony w formie cyfrowej. Sąd miałby obowiązek przychylić się do takiego wniosku.

---

## Podsumowanie

---

Powyższa analiza wiedzie do wniosku, że cyfryzacja postępowań karnych jawi się jako proces nieunikniony i cywilizacyjnie pożądany, który jednak w stadium wdrożeniowym generuje fundamentalne napięcia na linii: efektywność vs. rzetelność procesu. Kluczowe wnioski płynące z analizy obszarów zagrożeń i kosztów to możliwość ingerencji w fundamenty karnoprosesowe w postaci klasycznego rozumienia zasady bezpośredniości

oraz jawności. Ekonomicznym paradoksem cyfryzacji jest to, że jakkolwiek e-justice obiecuje oszczędności, to faktycznie generuje ogromne wydatki na cyberbezpieczeństwo, archiwizację danych w chmurze, aktualizację systemów oraz szkolenia. Są to koszty stałe, które Skarb Państwa będzie musiał ponosić bezterminowo, aby uniknąć paraliżu państwa. Pełny sukces cyfryzacji zależy zatem nie tyle od jakości kodów programistycznych, ile od stworzenia systemowych bezpieczników prawnych i zapewnienia funduszy finansowych sukcesywnie przeznaczanych na ten cel. Muszą one przede wszystkim chronić uczestników przed „dyktaturą algorytmu” i cyberatakami oraz zapewniać, że technologia pozostanie każdorazowo służebna wobec konstytucyjnych praw obywatelskich.

Całościowe podsumowanie problematyki cyfryzacji postępowań karnych (e-justice) wykazuje, że stoimy obecnie przed największą transformacją wymiaru sprawiedliwości od czasów kodyfikacji procedur<sup>12</sup>. Technologia przestaje być tylko narzędziem, a staje się współuczestnikiem procesu. Cyfryzacja jest niezbędna, aby sądownictwo nie stało się „analogowym skansenem”. Pamiętać natomiast należy, że cyfryzacja nie może być celem samym w sobie, gdyż jest tylko środkiem do urzeczywistnienia sprawiedliwości. Sukces tego procesu zależy od tego, czy ustawodawca zdąży obudować technologię gwarancjami procesowymi, zanim systemy informatyczne na trwałe zmienią dynamikę sali rozpraw. Takich czynności nie ma jednak zbyt wiele i nie są planowane z logicznym wyprzedzeniem, zwłaszcza w powiązaniu z etapami wdrożeń cyfrowych. Wywołuje to obawy o powodzenie przedsięwzięcia cyfryzacji postępowań karnych oraz czyni niemożliwą próbę wyznaczenia jakiegokolwiek realnego terminu jej zakończenia.

### Abstrakt

Publikacja stanowi analizę newralgicznych punktów niezbędnej cyfrowej reformy wymiaru sprawiedliwości na przykładzie postępowań karnych. Wskazano w niej, że cyfryzacja to fundamentalna zmiana paradygmatu rzetelnego procesu – co musi się wiązać ze stosownymi zmianami w prawie. Stwierdzono również potrzebę wdrożeń wielu nowych instrumentów w postępowaniu karnym, w tym wirtualnej sali narad dla rozpraw zdalnych. Zaproponowano konkretne zmiany w przepisach postępowania karnego, akcentując niezbędność szerszej nowelizacji – obejmującej wyłącznie zagadnienia cyfrowe. Cyfryzacja postępowań karnych (e-justice) z jednej bowiem strony oferuje przyspieszenie postępowań i ułatwienie dostępu do akt, z drugiej jednak stwarza poważne zagrożenia dla rzetelności procesu. Generuje też znaczne obciążenia finansowe dla państwa.

<sup>12</sup> A. Sołtysińska, *Digitalization of Criminal Justice* [w:] *Obywatel w centrum działań e-administracji w Unii Europejskiej*, S. Dudzik, I. Kawka, R. Śliwa (red.), Kraków 2023, s. 253.

Kluczowe zagrożenia obejmują: wykluczenie cyfrowe, osłabienie zasady bezpośredniości oraz problemy z poufnością komunikacji i jawnością postępowania. Pełne zrozumienie pól zagrożeń i ponoszonych kosztów jest niezbędne, aby system e-justice wydatnie wspierał wymiar sprawiedliwości.

**Słowa kluczowe:** postępowanie karne, cyfryzacja, wirtualna sala narad, zagrożenia, koszty.

---

## BIBLIOGRAFIA

---

Castells A., Społeczeństwo sieci, Warszawa 2007.

Czarnecki P., *Rozprawa sądowa w Internecie. Postulat wprowadzenia możliwości transmisji (streamingu) rozpraw sądowych w rzetelnym procesie karnym*, „Prokuratura i Prawo” 2025, 3.

Czetwertyński S., Mroczek-Czetwertyńska A., *Wykluczenie cyfrowe w społeczeństwie informacyjnym*, „Nauki Społeczne” 2012, 1, 5.

Czynności procesowe na odległość w postępowaniu karnym, M. Klubińska, A. Lach (red.), Warszawa 2025.

Dordick H.S., Wang G., *The Information Society: A Retrospective View*, Newbury Park 1993.

Jedlińska R., *Wykluczenie cyfrowe w Polsce na tle krajów Unii Europejskiej*, „Ekonomiczne Problemy Usług” 2018, 131, 2.

Kiejnich-Kruk K., *Building blocks – strategia cyfryzacji wymiaru sprawiedliwości. Perspektywa estońska*, „Przegląd Sądowy” 2025, 3.

Kujawski J., *Wykluczenie cyfrowe jako forma wykluczenia społecznego. Przypadek Polski*, „Media i Społeczeństwo” 2018, 9.

Marszałek J., Skwarcow M., *Użycie sztucznej inteligencji w postępowaniu karnym*, „Przegląd Sądowy” 2025, 3.

Obywatel w centrum działań e-administracji w Unii Europejskiej, S. Dudzik, I. Kawka, R. Śliwa (red.), Kraków 2023.

Opitek P., *Wybrane aspekty pozyskiwania dowodów cyfrowych w sprawach karnych*, „Prokuratura i Prawo” 2018, 7–8.

W pogoni za rzetelnym procesem karnym. Księga dedykowana Profesorowi Stanisławowi Waltosiowi, D. Szumiło-Kulczycka (red.), Warszawa 2022.

Zgoliński I., *Technology and Criminal Proceedings: An attempt to systematize the issue and determine the main directions of implementation*, „Europejski Przegląd Prawa i Stosunków Międzynarodowych” 2025, 1, 73.



Mateusz Szenawa\*

**Zagrożenia gospodarcze związane z rozwojem sztucznej inteligencji – ze szczególnym uwzględnieniem technologii deepfake – na podstawie wybranych regulacji prawnych Unii Europejskiej**

**[Economic Threats Associated with the Development of Artificial Intelligence, with Particular Reference to Deepfake Technology, Based on Selected European Union Legal Regulations]**

**Abstract**

This article analyses the impact of artificial intelligence, and in particular deepfake technology, on the economy. The aim of this paper is to assess whether current legal regulations at European level provide sufficient protection against the threats posed by technological development. The author presents cases in which deepfake technology can be used to extort funds or carry out other unethical activities. Key issues include the role of EU regulations – such as the Artificial Intelligence Act – in countering economic threats. Furthermore, the author highlighted the need for legislative changes that will address existing legal loopholes whilst reducing the scope of obligations imposed on businesses. It was noted that legal regulations concerning artificial intelligence are essential – whilst acknowledging that maintaining an adequate level of innovation among European businesses poses a significant challenge.

**Keywords:** deepfake, economy, AI Act, threats, artificial intelligence.

---

## **Wstęp**

---

W obliczu rewolucji technologicznej, jaka dokonuje się na naszych oczach dzięki rozwojowi sztucznej inteligencji, świat staje przed nowymi zadaniami. Celem niniejszego artykułu jest dokonanie analizy skuteczności regulacji unijnych, a w konsekwencji również uregulowań krajowych w zakresie ochrony

---

\* **Mateusz Szenawa** – prawnik i matematyk, asystent w Instytucie Nauk Prawnych Uniwersytetu Opolskiego (afiliacja); <https://orcid.org/0009-0005-3393-1316>; [mateusz.szenawa@uni.opole.pl](mailto:mateusz.szenawa@uni.opole.pl) / lawyer and mathematician, research assistant at the Institute of Legal Studies, University of Opole (affiliation).

prawnej przed gospodarczymi skutkami wykorzystania technologii *deepfake*. Praca odpowiada na główne pytanie badawcze: czy i w jakim zakresie obowiązujące mechanizmy prawne gwarantują adekwatną, wystarczającą ochronę osób fizycznych oraz podmiotów gospodarczych przed negatywnymi skutkami wykorzystania narzędzi sztucznej inteligencji umożliwiających tworzenie nieprawdziwych treści przez: oszustwa, dezinformację, naruszenie czci czy szkodenie reputacji. Naturalną kontynuacją rozważań nad głównym problemem badawczym jest rozpatrzenie, w jakim stopniu brak kompleksowej – a zatem pozbawionej luk – definicji pojęcia *deepfake* utrudnia ochronę praw jednostki oraz wpływa na zdolność systemu prawnego do przeciwdziałania gospodarczym zagrożeniom, rozumianym jako zestaw ryzyk ekonomicznych oraz prawnych wynikających z nieetycznego lub przestępczego wykorzystania narzędzi sztucznej inteligencji. W konsekwencji: należy rozpoznać, jakie mechanizmy prawne powinny funkcjonować, aby zapobiec negatywnym skutkom wykorzystywania tej technologii oraz jak ich wprowadzenie wpływa na innowacyjność i konkurencyjność przedsiębiorstw prowadzących działalność na terenie Unii Europejskiej. To właśnie powyższe pytania stanowią punkt wyjścia dla dalszych rozważań.

W artykule dokonano analizy poziomu ochrony prawnej przed cyfrowymi zagrożeniami wynikającymi z rozwoju AI, ze szczególnym uwzględnieniem regulacji unijnych w zakresie ochrony przed oszustwami, manipulacjami rynkowymi oraz naruszeniem wizerunku. Jednocześnie dynamika rozwoju nowych technologii stawia przed ustawodawcą wyzwanie utrzymania normatywnej adekwatności przepisów prawa. W ramach analiz dokonano rozważań *de lege ferenda*, identyfikując istotną, dotąd nieopisaną lukę definicyjną występującą w przepisach unijnych. Postulowane zmiany mają na celu wzmocnienie bezpieczeństwa gospodarczego w sytuacjach zderzenia się przedsiębiorców i ich pracowników z nowymi technologiami, w tym z zagrożeniami z nich płynącymi.

W opracowaniu jako podstawową zastosowano metodę dogmatycznoprawną – analizującą aktualne regulacje unijne w świetle ich zgodności z celem, jakim jest ochrona uczestników rynku. Metodę dogmatycznoprawną uzupełniono metodą historycznoprawną – pozwalającą uchwycić ewolucję przepisów w czasie – oraz elementami analizy porównawczej, w ramach której zestawiono wybrane rozwiązania unijne z elementami regulacji obowiązujących w USA i Wielkiej Brytanii. W wybranych szczegółowych kwestiach przywołano też kierunki regulacyjne obowiązujące w Republice Korei (państwie o dalece zaawansowanym systemie prawnym dotyczącym technologii w regionie Azji). W pierwszej kolejności zwrócono uwagę na problematykę zdefiniowania pojęcia *deepfake* i sztucznej inteligencji oraz na różnorodne aspekty wykorzystania *deepfake*’ów – od dezinformacji, przez oszustwa finansowe, po potencjalne skandale związane z naruszaniem dóbr osobistych wpływowych

osób. Rozwój technologii *deepfake* stawia pytania o granice odpowiedzialności prawnej za treści generowane przez systemy AI, rodząc potrzebę rewizji obowiązujących obecnie w tym zakresie przepisów prawa. Ostatnia część artykułu poświęcona została krytycznej analizie obowiązujących regulacji prawnych pod kątem ich adekwatności wobec zidentyfikowanych wyżej zagrożeń.

---

## Pojęcie *deepfake* w kontekście rozwoju sztucznej inteligencji

---

Pojęcie *deepfake* powstało z połączenia dwóch odrębnych wyrażeń „*deep learning*” – czyli: głębokie uczenie – oraz „*fake*”, co oznacza: fałszywy. *Deep learning* to przetwarzanie danych za pomocą sztucznych sieci neuronowych. „Modele uczenia głębokiego pobierają informacje z wielu źródeł i analizują te dane w czasie rzeczywistym bez konieczności interwencji ze strony człowieka”<sup>1</sup>. Z tego wynika, że technika ta opiera się na sztucznej inteligencji, a konkretnie uczeniu maszynowym<sup>2</sup>. W ogólności zatem *deepfake* polega na wytworzeniu fałszywego – to znaczy takiego, który w rzeczywistym świecie nie wystąpił – obrazu lub treści wideo z wykorzystaniem narzędzi związanych ze sztuczną inteligencją. Najczęściej stworzenie takiej treści polega na zastąpieniu w istniejącym materiale jednej osoby inną lub stworzeniu całkowicie oryginalnych i nowych treści, w których ludzie zachowują się tak, jak tego nigdy w rzeczywistości nie robili<sup>3</sup>. Dzieje się to z wykorzystaniem funkcjonalności konkretnego komponentu sztucznej inteligencji, a dokładnie *deep learningu*, który jest jednocześnie podzbiorem *machine learningu*, czyli wspomnianego uczenia maszynowego. W ramach powyższych rozważań istotne staje się również rozróżnienie pojęcia *deepfake* od *fake newsów*. *Fake news* stanowi przekaz medialny o charakterze dezinformującym, który często zawiera prawdziwe elementy, lecz jego celem jest manipulacja faktami<sup>4</sup>. *Deepfake* natomiast w głównej mierze polega na wytworzeniu nieprawdziwych obrazów lub treści wideo.

Należy zaznaczyć, że technologia *deepfake* może zostać użyta w dwojakim kontekście – pozytywnym i negatywnym. Do podobnych wniosków doszła I. Dąbrowska, która wyróżniła cztery pierwotne intencje, dla jakich *deepfake*

---

<sup>1</sup> A. Piecuch, *Sztuczna inteligencja w perspektywie społecznej*, „Edukacja Ustawiczna Dorosłych” 2023, 123, 4, ss. 13–25.

<sup>2</sup> J. Brandon, *Terrifying high-tech porn: Creepy ‘deepfake’ videos are on the rise*, <https://www.foxnews.com/tech/terrifying-high-tech-porn-creepy-deepfake-videos-are-on-the-rise> [dostęp: 7.12.2024].

<sup>3</sup> D. Kumari, *Deepfake Technology and Legal Issues*, ‘LawFoyer International Journal of Doctrinal Research’ 2024, 2, 1, s. 235.

<sup>4</sup> K. Kiepiński, *Deepfake jako narzędzie do przekazywania informacji fałszywej i domniemanej. Analiza prawno-karna i cybernetyczna*, „Kwartalnik Krajowej Szkoły Sądownictwa i Prokuratury” 2023, 3, 51, s. 89.

został stworzony: rozrywkę, edukację, dezinformację i dyskredytację<sup>5</sup>. Jednocześnie należy zaznaczyć, że przywołany katalog nie jest wyczerpujący. Używanie tej technologii rzeczywiście następuje w świecie rozrywki, filmowym czy edukacyjnym (kontekst pozytywny). Ponadto *deepfake* może zostać wykorzystany w celach dezinformacyjnych, mających wprowadzać w błąd – z posłużeniem się na przykład metodami socjotechnicznymi, wpływającymi na proces decyzyjny konkretnych jednostek; może również być użyty w celu dokonania cyberprzestępstw (kontekst negatywny). Technologia *deepfake* staje się przydatnym narzędziem w atakach mających na celu wyłudzenie pieniędzy lub innej wartości, np. danych osobowych czy informacji objętych tajemnicą przedsiębiorstwa. Pomimo to *deepfake* nie został zdefiniowany ani w polskim prawie karnym, ani cywilnym. W ramach kodeksu karnego oraz kodeksu cywilnego nie został też do tej pory wyodrębniony żaden przepis odnoszący się bezpośrednio do czynów związanych z wykorzystaniem jednego z narzędzi sztucznej inteligencji, jakim jest technologia *deepfake*.

Dopiero przepisy na szczelu europejskim – a dokładnie rozporządzenie Parlamentu Europejskiego i Rady (UE) 2024/1689 z 13 czerwca 2024 r. w sprawie ustanowienia zharmonizowanych przepisów dotyczących sztucznej inteligencji (dalej: AI Act lub rozporządzenie 2024/1689) – wprowadza w art. 3 pkt 60 definicję legalną tego pojęcia; według niej „*deepfake*” oznacza wygenerowane przez AI lub zmanipulowane przez AI obrazy, treści dźwiękowe lub treści wideo przypominające istniejące osoby, przedmioty, miejsca, podmioty lub zdarzenia, które odbiorca mógłby niesłusznie uznać za autentyczne lub prawdziwe<sup>6</sup>. Definicja ta zmieniała się w trakcie prac nad AI Act. Przykładem może być chociażby proponowana treść art. 52 ust. 3 wniosku Komisji Europejskiej w sprawie AI Act z 21 kwietnia 2021 r., gdzie *deepfake* został zdefiniowany jako system sztucznej inteligencji, który generuje obrazy, treści dźwiękowe lub treści wideo łudząco przypominające istniejące osoby, obiekty, miejsca lub inne podmioty lub zdarzenia lub który tymi obrazami i treściami manipuluje – przez co osoba będąca ich odbiorcą mogłaby niesłusznie uznać je za autentyczne lub prawdziwe<sup>7</sup>. Łatwo jednak zauważyć, że w obu tych definicjach możemy znaleźć hasła przewodnie, które są podstawą właściwego zdefiniowania, czym jest *deepfake*. Aby daną treść uznać za materiał *deepfake*, muszą zostać spełnione następujące przesłanki:

<sup>5</sup> I. Dąbrowska, *Deepfake – nowy wymiar internetowej manipulacji*, „Zarządzanie Mediami” 2020, 8, 2, s. 96.

<sup>6</sup> Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2024/1689 z 13 czerwca 2024 r. w sprawie ustanowienia zharmonizowanych przepisów dotyczących sztucznej inteligencji oraz zmiany rozporządzeń (WE) nr 300/2008, (UE) nr 167/2013, (UE) nr 168/2013, (UE) 2018/858, (UE) 2018/1139 i (UE) 2019/2144 oraz dyrektyw 2014/90/UE, (UE) 2016/797 i (UE) 2020/1828 (akt w sprawie sztucznej inteligencji), (Dz.U. L 2024/1689, 12.07.2024) zwanego dalej Rozporządzeniem 2024/1689.

<sup>7</sup> Wniosek rozporządzenia Parlamentu Europejskiego i Rady ustanawiające zharmonizowane przepisy dotyczące sztucznej inteligencji (akt w sprawie sztucznej inteligencji) i zmieniające niektóre akty ustawodawcze unii, (COM/2021/206 final)

- a) stworzenie lub zmanipulowanie przez AI;
- b) ukazanie treści łudząco zbliżonych do elementów rzeczywistych lub do istniejących osób;
- c) istnienie możliwości niesłusznego uznania treści za prawdziwą.

Jako że materiały *deepfake* powstają z wykorzystaniem sztucznej inteligencji, należy się zastanowić, czym sztuczna inteligencja naprawdę jest. Pojęcie sztucznej inteligencji (*artificial intelligence*, AI) zostało zdefiniowane na wiele różnych sposobów. Niemniej dla rozważań wynikających z niniejszego artykułu wystarczająca będzie definicja wynikająca z analiz brytyjskiego matematyka Alana Turinga. Zgodnie z jego koncepcją „sztuczna inteligencja to zdolność maszyny do naśladowania lub imitowania ludzkiej inteligencji”<sup>8</sup>. Wynika ona wprost ze słynnego testu Turinga, „który ma definiować normy dla maszyny odznaczającej się inteligencją”<sup>9</sup>. Taka definicja wydaje się uniwersalna i – co wręcz niespotykane w dziedzinie nowych technologii – ponadczasowa (wynika bowiem z pracy „*Computing Machinery and Intelligence*”, opublikowanej w 1950 roku)<sup>10</sup>. Inne próby zdefiniowania sztucznej inteligencji wielokrotnie kończą się stworzeniem skomplikowanych struktur logicznych, które zamiast prowadzić do ujednoczenia definicji, wprowadzają niepotrzebne zamieszanie w debacie publicznej<sup>11</sup>. Zdarza się również, że do zdefiniowania sztucznej inteligencji używana bywa ta autorstwa Johna McCarthy’ego, zaproponowana w 1955 roku, która uważana jest za jedną z pierwszych, a przez J. Widłę za pierwszą definicję sztucznej inteligencji<sup>12</sup>. Ponieważ jednak John McCarthy w swojej propozycji był zgodny z rozważaniami A. Turinga – z których wynika, że maszyna jest inteligentna, gdy człowiek nie jest w stanie odróżnić, czy udzielona odpowiedź na zadane pytanie pochodzi od człowieka, czy tej maszyny – to trudno uznać, że definicja ta jest pierwsza w sensie koncepcyjnym. Nawet jeśli A. Turing nie wprowadził formalnego sensu stricto pojęcia sztucznej inteligencji, to przyczynił się do zdefiniowania kryteriów inteligencji maszyn<sup>13</sup>.

<sup>8</sup> T. Zalewski, *Definicja sztucznej inteligencji* [w:] L. Lai, M. Świerczyński (red.), *Prawo sztucznej inteligencji*, Warszawa 2020, ss. 5–9.

<sup>9</sup> I. Stewart, *Significant Figures Lives and Works of Trailblazing Mathematics*, New York 2017, s. 325 [wyd. pol. *Krótką historia wielkich umysłów. Genialni matematycy i ich arcydzieła*, tłum. U. Seweryńska i M. Seweryński, Warszawa 2019].

<sup>10</sup> *Ibid.*

<sup>11</sup> Por. T. Zalewski, *Definicja...*, *ibid.*; D. Spałek, J. Rzymowski, *Jak nie stosować AI Act na podstawie definicji systemu sztucznej inteligencji*, „*Prawo Nowych Technologii*” 2024, 2, s. 21.

<sup>12</sup> J. Widło, *Odpowiedzialność za szkodę wyrządzoną przez pojazdy autonomiczne w świetle zmian prawa UE*, „*Forum Prawnicze*” 2025, 2, s. 7.

<sup>13</sup> A. Turing, *Computing Machinery and Intelligence*, „*Mind*” 1950, 59, 236, s. 433–435.

---

## Zagrożenia wynikające z technologii deepfake w kontekście gospodarczym

---

---

### Dezinformacja

---

Badacze zwracają uwagę na zagrożenia związane z dezinformacją w przestrzeni publicznej z wykorzystaniem fake newsów czy materiałów wygenerowanych za pomocą technologii deepfake. Niewątpliwie „rozwój nowych technologii (...) może zwiększyć zagrożenie dla stabilności politycznej i społecznej”<sup>14</sup>. W tym kontekście aspekt gospodarczy pozostaje niedostatecznie zbadany. Niniejsze opracowanie uzupełnia tę lukę poprzez analizę zagrożeń dotyczących przedsiębiorców, ponieważ dezinformację znaną ze świata polityki można przełożyć na świat biznesu. Przedsiębiorcy mogą tworzyć fałszywie pozytywne informacje o swojej firmie, realizując na przykład nieprawdziwy, pozytywny marketing, jak również czarny PR konkurencji, gdzie materiały deepfake mogą być używane przez konkurencyjne firmy przygotowujące kompromitujące lub dezinformujące materiały w sprawie wyników finansowych, polityki firmy czy sposobu traktowania przez tę firmę pracowników. Podkreślenia wymaga fakt, że nie tylko przedsiębiorcy mogą być twórcami tego typu treści, ale również osoby fizyczne (pracownicy, inwestorzy, osoby o nieprzychylnym usposobieniu w stosunku do konkretnej firmy), które mogą mieć w tym interes osobisty. Temat ten jest pobieżnie traktowany przez przepisy prawne, ponieważ nie dotyczy ogółu społeczeństwa, porządku publicznego czy bezpieczeństwa państwa, a miałyby chronić interes konkretnej jednostki. Tymczasem im większe przedsiębiorstwo stanie się ofiarą albo będzie twórcą ataku deepfake, tym reperkusje społeczne mogą być większe. Zmasowany i skrupulatnie przygotowany dezinformacyjny atak deepfake może wpłynąć znacząco na kurs akcji spółek kapitałowych, w tym spółek strategicznych dla bezpieczeństwa państwa.

Również biała księga w sprawie sztucznej inteligencji zwraca uwagę na wyzwania stawiane przed nami przez rozwój sztucznej inteligencji, jednak w niewystarczającym stopniu identyfikuje ona potencjalne zagrożenia gospodarcze, które mogą wystąpić w sytuacji przestępczego wykorzystania nowoczesnych narzędzi technologicznych<sup>15</sup>. Komisja Europejska zauważyła, że narzędzia w zakresie sztucznej inteligencji mogą przyczynić się do lepszej

---

<sup>14</sup> D. Mierzejewski, J. Nawrotkiewicz, W obliczu wspólnej polityki UE wobec Chin: „mission impossible”?, „Pulaski Policy Papers” 2024, XI, s. 20.

<sup>15</sup> Biała księga w sprawie sztucznej inteligencji. Europejskie podejście do doskonałości i zaufania (COM/2020/65 final/2), zwana dalej białą księgą.

ochrony obywateli UE przed przestępczością i aktami terrorystycznymi<sup>16</sup>. Nie zmienia to faktu, że spośród katalogu przykładowo wymienionych czynności, w których sztuczna inteligencja może okazać się przydatna, brakuje informacji o zapobieganiu i wykrywaniu dezinformacji dokonanej przez tworzenie fałszywych treści *deepfake*. Pozytywnie należy jednak ocenić fakt zauważenia przez Komisję Europejską, że mogą być konieczne dalsze ustalenia w celu zapobiegania i zwalczania stosowania AI w celach przestępczych<sup>17</sup>. Również część białej księgi dotycząca bezpośrednio zagrożeń dla bezpieczeństwa wymienia te zagrożenia, które „mogą być spowodowane wadami w projekcie technologii sztucznej inteligencji, być związane z problemami z dostępnością i jakością danych lub z innymi problemami wynikającymi z uczenia się maszyn”<sup>18</sup>. Należy zaznaczyć, że biała księga miała za zadanie przedstawić warianty strategiczne dotyczące sposobów osiągnięcia dwóch celów, którymi są promowanie sztucznej inteligencji oraz zajęcie się zagrożeniami wynikającymi z konkretnych jej zastosowań<sup>19</sup>.

## Potencjalne skompromitowanie wpływowych osób

Naturalną kontynuacją tworzenia dezinformujących treści *deepfake* jest naruszenie dóbr osobistych poszczególnych osób. W poniższych rozważaniach przeanalizowano potencjalne konsekwencje prób skompromitowania wpływowych osób ze świata biznesu. Według danych zebranych przez NASK grupą społeczną najbardziej narażoną na tworzenie materiałów *deepfake* są dziennikarze – 32%, następnie politycy – 21%. Trzecią grupą społeczną są biznesmeni – 11%<sup>20</sup>. Wynik ten pokazuje skalę zagrożenia, do której zarówno przepisy prawa krajowego, jak i unijnego nie są przystosowane. Jak wskazuje A. Ziobroń, w ramach przepisów krajowych przewidziano ochronę została przewidziana w art. 212 oraz art. 216 k.k.<sup>21</sup> Ponadto wizerunek i cześć są chronione na podstawie art. 23 k.c., co daje ofercie kumulatywną ochronę prawną. Ten obecny stan prawny należy wszakże uznać za nieadekwatny do

<sup>16</sup> Biała księga zawiera informację, że sztuczna inteligencja „mogłyby na przykład pomóc w identyfikowaniu internetowej propagandy terrorystycznej, wykrywaniu podejrzanych transakcji sprzedaży niebezpiecznych produktów, identyfikowaniu niebezpiecznych ukrytych przedmiotów lub nielegalnych substancji lub produktów czy też oferować pomoc obywatelom w sytuacjach nadzwyczajnych i wspomagać służby interwencyjne” (s. 2).

<sup>17</sup> Biała..., s. 3.

<sup>18</sup> Biała..., s. 14.

<sup>19</sup> A. Zielińska, Korzystanie z treści chronionych a sztuczna inteligencja – jak daleko sięga odpowiedzialność dostawcy usług udostępniania treści online w świetle art. 17 dyrektywy 2019/790?, „Studia de Cultura” 2022, 14, 2, s. 136.

<sup>20</sup> NASK ostrzega: przestępcy tworzą coraz bardziej pomysłowe oszustwa *deepfake*, <https://nask.pl/aktualnosci/nask-ostrzega-przestepcy-tworza-coraz-bardziej-pomyslowe-oszustwa-deepfake/>, [dostęp: 12.12.2024].

<sup>21</sup> A. Ziobroń, *Deepfake a prawo karne. Uwagi „de lege lata” i „de lege ferenda” dotyczące fałszywej pornografii*, „Studenckie Prace Prawnicze, Administratywistyczne i Ekonomiczne” 2021, 37, s. 225.

skali zagrożenia, ochrona prawna w dużej mierze jest bowiem tylko pozorna, gdyż ze względu na charakter czynu znalezienie sprawcy jest utrudnione, a potencjalne konsekwencje karne pozostają niewspółmierne do skali społecznej szkodliwości. Do podobnych wniosków dociera też G.G. Vera – twierdząc, że „ochrona ta jest fragmentaryczna i niewystarczająca”<sup>22</sup>.

Podobnie należy również spojrzeć na unijne regulacje (AI Act), które zwracają uwagę na fakt, że sztuczna inteligencja stwarza nowe rodzaje ryzyka – polegające na podawaniu informacji wprowadzających w błąd i na manipulacji na dużą skalę, oszustwach, podszywaniu się pod inne osoby i wprowadzaniu w błąd konsumentów<sup>23</sup>. W ramach tych zagrożeń AI Act zakłada konieczność zobowiązania dostawców systemów do wbudowania takich narzędzi, które pozwolą wykrywać, że poszczególne treści zostały wygenerowane przez sztuczną inteligencję, a nie przez człowieka. AI Act w zdecydowanie większym stopniu zawiera opis zagrożeń związanych z dezinformacją, podszywaniem się pod inną osobę i tworzeniem treści *deepfake*, niż przewidywała to przywołana wyżej biała księga. W przypadku tworzenia fałszywych treści z wykorzystaniem technologii *deepfake* należy zwłaszcza wyróżnić te szczególnie szkodliwe, bo atakujące bezpośrednio osobę w celu jej zdyskredytowania – np. przez stworzenie materiału typu *deepfake* porno, gdzie znana osobistość „występuje” w spreparowanym za pomocą sztucznej inteligencji materiale pornograficznym (główną ofiarą takich materiałów są kobiety), czy też przez wygenerowanie fałszywych nagrań zawierających treści, z których wynikałyby przykładowo rzekome obyczajowe dewiacje lub niewłaściwe traktowanie podwładnych przez właściciela firmy. Niestety możliwości wytworzenia takich treści ograniczone są wyłącznie pomysłowością twórcy takich materiałów. Jeszcze innym zagrożeniem, które należy odróżnić od powyższych, jest wykorzystanie wizerunku znanych przedsiębiorców czy innych osób ze świata biznesu w celu dokonania oszustw finansowych i wyłudzeń pieniędzy oraz danych.

---

## Cyberoszustwa

---

Próby wyłudzenia środków finansowych z wykorzystaniem nowoczesnych technologii należy podzielić na dwie podstawowe kategorie. Pierwszą jest sytuacja, gdy ofiarą staje się konkretna jednostka. Dotyczy to przede wszystkim sytuacji, w których cyberoszuści – tworząc spreparowane treści *deepfake* – zachęcają do fałszywych inwestycji z wykorzystaniem wizerunku znanych przedsiębiorców. Tylko w pierwszej połowie 2024 roku NASK wyszczególnił

<sup>22</sup> G.G. Vera, Deep fake (!) – postęp technologiczny a prawo karne, „Zeszyty Naukowe Uniwersytetu Rzeszowskiego – Seria Prawnicza” 2024, 1, 44, s. 50.

<sup>23</sup> Motyw 133 preambuły rozporządzenia 2024/1689.

jedenastu polskich przedsiębiorców, którzy padli ofiarą *deepfake*<sup>24</sup>. Tego typu działania – oprócz strat finansowych poniesionych przez ofiary cyberoszustów – przyczyniają się również do utraty zaufania do poszczególnych przedsiębiorców. Mimo że jest jasne, iż tego typu fałszywe zachęty inwestycyjne nie są firmowane przez przedsiębiorców, których wizerunek wykorzystano, wystarczy, że chociażby część społeczeństwa nie będzie tego świadoma i postanowi rozpowszechniać nieprawdziwe informacje dotyczące danej osoby (osób). Pochodną tego typu zachowań może stać się utrata zaufania do konkretnej firmy, a to może bezpośrednio wpływać na wyniki finansowe przedsiębiorstwa.

Drugą kategorią cyberoszustw, na które należy zwrócić uwagę, są te, w których szkoda dotyczy bezpośrednio przedsiębiorstwa. Ofiarą nadal jest konkretna jednostka, ale środki finansowe lub dane, które cyberoszuści chcą w tej sytuacji wyłudzić, nie są jej własnością. Pracownik, który padł ofiarą cyberprzestępców, może w ramach swoich obowiązków zawodowych jedynie dysponować majątkiem firmowym lub danymi stanowiącymi tajemnicę przedsiębiorstwa. Jako przykład tego typu działania należy podać przypadek oszustwa z użyciem technologii *deepfake*, w którym brytyjska firma projektowa Arup straciła 25 mln USD<sup>25</sup>. Pracownik sądził, że uczestniczy w poufnej transakcji z dyrektorem i innymi pracownikami firmy<sup>26</sup>. Omówiony przypadek jest szczególnie wymowny z perspektywy regulacji obowiązujących w Zjednoczonym Królestwie. Arup jest podmiotem brytyjskim z główną siedzibą w Londynie. Wielka Brytania przyjęła w ramach swojego ustawodawstwa Online Safety Act 2023, mający na celu nałożenie obowiązków ochronnych na platformy internetowe w zakresie nielegalnych treści udostępnionych przez użytkowników<sup>27</sup>. Jednak przyjęty model dotyczy tylko reaktywnej moderacji treści i znajduje zastosowanie wyłącznie wtedy, gdy dany materiał będzie już dostępny online<sup>28</sup>. W kontekście oszustw finansowych taki charakter regulacji okazuje się niewystarczający, ponieważ szkoda materializuje się w momencie dokonywania przelewu przez ofiarę, a nie w momencie udostępnienia treści *deepfake* w sieci<sup>29</sup>. Wątpliwe jest przy tym, czy tak wygenerowana treść

<sup>24</sup> Lista osób, których wizerunek przestępca wykorzystali w oszustwach *deepfake*, <https://nask.pl/aktualnosci/nask-ostrzeza-przestepcy-tworza-coraz-bardziej-pomyslowne-oszustwa-deepfake/> [dostęp: 12.12.2024].

<sup>25</sup> D. Jakubczak, *Deepfake. Oszustwo na sklonowanego dyrektora*, <https://www.dw.com/pl/deepfake-oszustwo-na-sklonowanego-dyrektora/a-69116765> [dostęp: 16.12.2024].

<sup>26</sup> Oszuści wykorzystali cyfrowo sklonowany wizerunek i głos dyrektora finansowego firmy, by podczas fałszywej wideokonferencji przekonać pracownika w Hongkongu do wykonania piętnastu przelewów na różne konta bankowe.

<sup>27</sup> Online Safety Act 2023, c. 50, Zjednoczone Królestwo, <https://www.legislation.gov.uk/ukpga/2023/50/contents/enacted> [dostęp: 10.05.2026].

<sup>28</sup> B. Kira, *When Non-Consensual Intimate Deepfakes Go Viral: The Insufficiency of the UK Online Safety Act*, 'Computer Law & Security Review' 2024, 54, art. 106024, s. 7.

<sup>29</sup> Opisany akt prawny daje możliwość ochrony dzieci przed dostępem do szkodliwych treści oraz zmierza do ograniczenia rozpowszechniania intymnych treści *deepfake*. Trafnie wskazuje się jednak, że regulacja nie zawiera wystarczających mechanizmów prewencyjnych i obejmuje wyłącznie ułamek rzeczywistych zagrożeń występujących w środowisku online – zob. B. Kira, *When Non-Consensual...*, *ibid.*

w ogóle zostałyby przez cyberoszustów udostępniona. Ponadto przy atakach na międzynarodowe podmioty powstaje problem reguł kolizyjnoprawnych i stosowania prawa właściwego; choć bowiem straty dotyczą podmiotu brytyjskiego, przelewy trafiły na konta bankowe znajdujące się w Hongkongu. Podkreślenia wymaga, że przedsiębiorcy – a w szczególności korporacje – mają wypracowane wewnętrzne systemy i procedury mające zapobiegać cyberoszustwom. Atak na przedsiębiorstwo wymaga zdecydowanie większej wprawy i profesjonalizmu niż ten, w którym ofiarą miałyby być osoba fizyczna. Jednocześnie nie ulega wątpliwości, że obecne możliwości sztucznej inteligencji są już wystarczające, aby takich oszustw dokonać.

Cyberoszustwa z wykorzystaniem technologii *deepfake* stanowiły dotychczas wyłącznie wycinek ogółu cyberprzestępstw. Niemniej przywołane poniżej dane uzasadniają potraktowanie tego zjawiska jako priorytetowego obszaru do uregulowania. W 2022 roku w USA zidentyfikowano 300 497 ofiar phishingu, a wartość skradzionych środków finansowych oszacowano na 52 mln USD<sup>30</sup>. Powyższe potwierdza również raport FBI, który wskazuje, że dokładna wartość utraconych środków w ramach phishingu wynosiła w 2022 roku 52 089 159 USD<sup>31</sup>. Zatem średnia wysokość utraconej przez jedną ofiarę kwoty to wartość około 173 USD. Zestawiając to z faktem, że przy wykorzystaniu zaawansowanych narzędzi *deepfake* jedna osoba była skłonna dokonać łącznych przelewów na kwotę 25 mln USD – co stanowi 48% łącznej kwoty, jaką wszystkie zidentyfikowane ofiary phishingu w Stanach Zjednoczonych straciły przez cały 2022 rok – widoczna staje się skala zagrożenia.

Kolejne raporty – analizując wyłącznie dane od 2023 roku, które są agregowane w jednolity sposób – wskazują jednocześnie na spadek liczby zgłaszanych przestępstw dokonanych w ramach phishingu/spoofingu przy jednoczesnym wzroście wartości utraconych przez ofiary środków finansowych<sup>32</sup>. Należy zaznaczyć, że raport z 2022 roku osobno podaje wysokości utraconych kwot przez ofiary phishingu i spoofingu, natomiast w raporcie z 2024 r. oraz 2025 r. podane wartości dotyczą zsumowanych strat finansowych ofiar phishingu i spoofingu<sup>33</sup>. Istotne jest również, że w 2023 roku w Stanach Zjednoczonych łączne straty ofiar tych dwóch procedur wynosiły 18 728 550 USD, co stanowi ok. 6–7 mln USD mniej niż wyłącznie jeden atak *deepfake* na przedsiębiorstwo<sup>34</sup>. Oszustwa z wykorzystaniem wideokonferencji wygenerowanych

<sup>30</sup> M. Bukowski, Zwalczanie cyberprzestępczości ekonomicznej przy wykorzystaniu sztucznej inteligencji (AI), „Przegląd Policyjny” 2023, 151, 3, s. 163.

<sup>31</sup> FBI, Internet Crime Report 2022, [https://www.ic3.gov/AnnualReport/Reports/2022\\_ic3report.pdf](https://www.ic3.gov/AnnualReport/Reports/2022_ic3report.pdf), s. 22 [dostęp: 24.10.2025].

<sup>32</sup> FBI, Internet Crime Report 2024, [https://www.ic3.gov/AnnualReport/Reports/2024\\_IC3Report.pdf](https://www.ic3.gov/AnnualReport/Reports/2024_IC3Report.pdf), s. 18 [dostęp: 24.10.2025]; FBI, Internet Crime Report 2025, [https://www.ic3.gov/AnnualReport/Reports/2025\\_IC3Report.pdf](https://www.ic3.gov/AnnualReport/Reports/2025_IC3Report.pdf), s. 25 [dostęp: 8.05.2026].

<sup>33</sup> Ibid.

<sup>34</sup> Ibid.

przez sztuczną inteligencję są zdecydowanie trudniejsze do przygotowania przez oszustów, ale zarazem potencjalne straty finansowe ich ofiar są niewspółmiernie wyższe.

## Inne cyberprzestępstwa

Sztuczna inteligencja może być wykorzystywana przez cyberprzestępców na wiele jeszcze innych sposobów. Materiały stworzone za pomocą technologii *deepfake* mogą zostać wykorzystane w celu zwiększenia skuteczności ataków phishingowych. Ponadto sztuczna inteligencja może zostać wykorzystana do zautomatyzowania tworzenia nieprawdziwych informacji/recenzji na temat produktów lub usług świadczonych przez konkretne przedsiębiorstwo. Dodatkowo takie opinie mogą być wzmacniane przez stworzenie fałszywego wideo, w którym wygenerowana osoba pozytywnie recenzuje wybrane przedsiębiorstwo lub negatywnie odnosi się do działań firm konkurencyjnych. Należy również zauważyć całą rzeszę możliwości AI poza technologią *deepfake*, które umożliwiają jeszcze skuteczniejsze popełnianie cyberprzestępstw. Wśród nich można wymienić między innymi generowanie fałszywych dokumentów, możliwość manipulowania rynkiem kapitałowym przez tworzenie fałszywych informacji czy też wykorzystywanie botów do wyłudzeń danych. Jednak z racji obszerności tematu przestępstwa te winny zostać rozważone w ramach odrębnej pracy badawczej.

W tym miejscu należałoby zwrócić uwagę na zjawisko Crime/Cybercrime as a Service, które polega na „sprzedaży dostępu do narzędzi i wiedzy potrzebnej do przeprowadzenia ataków”<sup>35</sup>. W rezultacie osoby z ograniczoną wiedzą i możliwościami mogą przeprowadzić ataki ze względną łatwością<sup>36</sup>. Tego typu usługa może również polegać na odprzedaży właściwych narzędzi do wytworzenia w prosty sposób dowolnych treści *deepfake*. Możliwość ta generuje ryzyka gospodarcze związane z chęcią między innymi zaszkodzenia konkretnej firmie czy konkretnej osobie poprzez wytworzenie szkodliwych materiałów przez osobę niemającą wystarczających umiejętności, aby w pełni samodzielnie tak sfałszowany materiał *deepfake* przygotować. Poznawszy już główne zagrożenia gospodarcze, należy się zastanowić, w jakim stopniu obecne przepisy prawne są na nie przygotowane.

<sup>35</sup> Crime as a service (CaaS), <https://www.comcert.pl/slownik/crime-as-a-service-caas/> [dostęp: 17.12.2024].

<sup>36</sup> A. Unni, What You Need To Know About Crime As A Service (CSaaS), <https://www.stickmancyber.com/cybersecurity-blog/what-you-need-to-know-about-crime-as-a-service-csaas> [dostęp: 17.12.2024].

## **Akt o sztucznej inteligencji jako próba odpowiedzi na zidentyfikowane zagrożenia**

Przepisy AI Act należy rozważyć w kontekście zagrożeń opisanych powyżej. Komisja Europejska zwróciła uwagę na ryzyko dezinformacji i manipulacji, które mogą być szerzone za pomocą sztucznej inteligencji, w tym technologii *deepfake*. AI Act zwraca uwagę, że sztuczna inteligencja może być wykorzystywana niewłaściwie i może dostarczać nowych, potężnych narzędzi do praktyk manipulacji, wyzyskiwania i kontroli społecznej<sup>37</sup>. Jednocześnie rozporządzenie wprost zakłada, że praktyki takie powinny być zakazane jako niezgodne z unijnymi wartościami. W każdym razie nie ma znaczenia, czy dostawca lub podmiot stosujący mieli zamiar wyrządzić poważną szkodę; istotny jest fakt, że szkoda wynika z praktyk manipulacyjnych opartych na AI lub z ich wykorzystywania. Zakazy dotyczące takich praktyk w zakresie AI stanowią uzupełnienie przepisów zawartych w dyrektywie Parlamentu Europejskiego i Rady 2005/29/WE dotyczącej nieuczciwych praktyk handlowych stosowanych przez przedsiębiorstwa wobec konsumentów na rynku wewnętrznym<sup>38</sup>, w szczególności przepisu zakazującego stosowania we wszelkich okolicznościach nieuczciwych praktyk handlowych powodujących dla konsumentów szkody ekonomiczne lub finansowe – niezależnie od tego, czy praktyki te stosuje się za pomocą systemów AI czy w innym kontekście<sup>39</sup>. Na szczeblu unijnym zwrócono zatem uwagę na zagrożenia gospodarcze i możliwość powstania szkód materialnych oraz finansowych przy niewłaściwym stosowaniu narzędzi AI.

Pozytywnie należy ocenić fakt, że AI Act zawiera neutralne technologicznie definiowanie pojęć, dzięki czemu uzyskuje szerokie zastosowanie oraz wieloletnią aktualność przyjętych pojęć. W rozporządzeniu 2024/1689 położono mocniejszy nacisk na opis działania konkretnych narzędzi niż na ich enumeratywne wymienianie w ramach zamkniętego katalogu – co powinno tym samym wydłużyć aktualność przepisów AI Act. Podobnie zauważa też Dorota Czyżewska-Misztal, stwierdzając, że projekt AI Act zapewnia „horyzontalne ramy prawne dotyczące sztucznej inteligencji, które mają zapewnić pewność prawa”<sup>40</sup>. Pomimo że końcowa treść AI Act różni się znacząco od projektu, udało się zachować takie definiowanie pojęć, które nie powinno ulec szybkiej dezaktualizacji.

Należy również zaznaczyć bardzo ważną w kontekście technologii *deepfake* zmianę treści rozporządzenia 2024/1689 względem treści, która wy-

<sup>37</sup> Motyw 28 preambuły rozporządzenia 2024/1689.

<sup>38</sup> Dz. Urz. UE 2005 L 149/22.

<sup>39</sup> Motyw 29 preambuły rozporządzenia 2024/1689.

<sup>40</sup> D. Czyżewska-Misztal, Europejskie podejście do sztucznej inteligencji [w:] Skutki i zagrożenia cywilizacji informacyjnej, K.A. Nawrot, K. Prandecki (red.), Komitet Prognoz PAN 2023, s. 123.

stąpiła w jego projekcie. Załącznik III do projektu AI Act wymieniał systemy sztucznej inteligencji przeznaczone do wykorzystania przez organy ścigania do wykrywania treści stworzonych z wykorzystaniem technologii *deepfake*, o których mowa w art. 52 ust. 3, jako systemy wysokiego ryzyka<sup>41</sup>. Jednocześnie art. 52 ust. 3 projektu zakładał, że użytkownicy systemu sztucznej inteligencji, który generuje obrazy, treści dźwiękowe lub treści wideo łudząco przypominające istniejące osoby, obiekty, miejsca lub inne podmioty lub zdarzenia lub który tymi obrazami i treściami manipuluje – przez co osoba będąca ich odbiorcą mogłaby niesłusznie uznać je za autentyczne lub prawdziwe („*deepfake*”), ujawnią, że dane treści zostały wygenerowane lub zmanipulowane przez system sztucznej inteligencji<sup>42</sup>. Sytuacja ta zakładała, że systemy służące do wykrywania przestępstw z wykorzystaniem technologii *deepfake* zostałyby uznane za systemy wysokiego ryzyka, co nakładałoby na dostawców systemów AI dodatkowe obowiązki, a jednocześnie narzędzia służące do generowania materiałów *deepfake* pod ten rygor by nie podlegały. Uogólniając: mielibyśmy wówczas do czynienia z sytuacją, w której narzędzie służące do wykrywania przestępstw zostałyby uznane za narzędzie większego ryzyka niż narzędzia mogące posłużyć do ich popełnienia. Projekt zakładał w podpunkcie 2.3., że w przypadku systemów sztucznej inteligencji wysokiego ryzyka wymogi dotyczące danych wysokiej jakości, dokumentacji i identyfikowalności, przejrzystości, nadzoru ze strony człowieka, dokładności i solidności są absolutnie niezbędne, aby złagodzić stwarzane przez AI ryzyko dla praw podstawowych i bezpieczeństwa, którego nie obejmują istniejące ramy prawne<sup>43</sup>. W ramach projektu w odniesieniu do niektórych szczególnych systemów sztucznej inteligencji, w tym zmanipulowanych cyfrowo obrazów lub filmów, zaproponowano wyłącznie minimalne obowiązki dotyczące przejrzystości. Zmiany, których w tym kontekście dokonano, należy ocenić pozytywnie, ponieważ wycofano się z pomysłu, aby narzędzie służące do wykrywania treści *deepfake* definiować jako narzędzie wysokiego ryzyka<sup>44</sup>. Jednak warto zaznaczyć, że użyta w ramach art. 3 pkt 60 definicja pojęcia *deepfake* zawiera pewną lukę. W sytuacji gdy w ramach wygenerowanych lub zmanipulowanych przez sztuczną inteligencję obrazów, dźwięków lub materiałów filmowych użyto wizerunku osoby, definicja ograniczona jest wyłącznie do istniejących osób. Powoduje to sytuację, że gdyby

<sup>41</sup> Załącznik III do wniosku dotyczącego rozporządzenia Parlamentu Europejskiego i Rady ustanawiającego zharmonizowane przepisy dotyczące sztucznej inteligencji (akt w sprawie sztucznej inteligencji) i zmieniającego niektóre akty ustawodawcze Unii (COM/2021/206 final).

<sup>42</sup> Wniosek rozporządzenie Parlamentu Europejskiego i Rady ustanawiające zharmonizowane przepisy dotyczące sztucznej inteligencji (akt w sprawie sztucznej inteligencji) i zmieniające niektóre akty ustawodawcze Unii (COM/2021/206 final), zwany dalej Projektem.

<sup>43</sup> Punkt 2.3. uzasadnienia Projektu.

<sup>44</sup> Szczegółowy wykaz systemów AI wysokiego ryzyka, o których mowa w art. 6 ust. 2 zawiera załącznik III do rozporządzenia. Warunki, po których spełnieniu system zostanie uznany za system wysokiego ryzyka, opisuje ponadto art. 6 ust. 1 rozporządzenia.

w ramach wygenerowania przez sztuczną inteligencję – np. w nieprawdziwym materiale wideo – wykorzystano wizerunek osoby, która zmarła, to treść taka zgodnie z definicją zawartą w AI Act nie powinna zostać uznana za *deepfake*. Definicja zawiera słowa „treści (...), które przypominają istniejące osoby (...), które odbiorca mógłby niesłusznie uznać za autentyczne lub prawdziwe”<sup>45</sup>. Pojęcie „istnieć” zgodnie ze Słownikiem języka polskiego oznacza „mieć miejsce w rzeczywistości”<sup>46</sup>. Osoba, która odeszła już ze świata, bez wątplenia nie spełnia przesłanki istnienia. Jednocześnie w języku polskim używamy też pojęcia „przestać istnieć”, gdzie słowo „przestać” oznacza (w rozumieniu, w którym wypowiadamy się o jakimś stanie, procesie itp.): „skończyć się”<sup>47</sup>. Zatem pojęcie „przestać istnieć” oznaczałoby: „skończyć proces, który ma miejsce w rzeczywistości”. Rozporządzenie 2024/1689 w tym samym miejscu w wersji anglojęzycznej używa słów „existing persons”, co wyklucza błąd tłumaczenia, a skłania do wniosku, iż doszło do niedopatrzenia w ramach tworzenia definicji<sup>48</sup>.

Nie można uznać, że wykorzystanie wizerunku zmarłej osoby w ramach stworzenia nieprawdziwej treści wygenerowanej przez sztuczną inteligencję nie jest treścią typu *deepfake*. *De lege ferenda* definicja ta powinna zatem zostać uzupełniona i brzmieć w powyższym fragmencie: „treści (...), które przypominają istniejące obecnie lub w przeszłości osoby, (...), które odbiorca mógłby niesłusznie uznać za autentyczne lub prawdziwe”. Pozwoli to zlikwidować mogące w przyszłości wystąpić problemy interpretacyjne. Przykład ten nie jest rozumowaniem czysto akademickim, lecz sytuacją, jaka może zaistnieć w praktyce. Możliwe jest zdarzenie, w ramach którego – z wykorzystaniem autorytetu zmarłego przedsiębiorcy – wygenerowana zostanie treść, która będzie rzekomym ostatnim apelem, pomysłem na dalszy rozwój firmy lub upublicznieniem kompromitujących informacji. Należy pamiętać, że jedynym ograniczeniem w tym wypadku jest kreatywność twórcy takiej treści. Dodatkowo unormowania krajowe również są niekompletne, co powoduje, że problematyczne może być między innymi wskazanie osób legitymowanych czynnie do wystąpienia z powództwem o ochronę wizerunku osoby nieżyjącej, a pociągnięcie do odpowiedzialności osoby, która naruszyła prawo do wizerunku – niewykonalne<sup>49</sup>.

<sup>45</sup> Art. 3 pkt 60 rozporządzenia.

<sup>46</sup> Słownik języka polskiego PWN, <https://sjp.pwn.pl/slowniki/istnie%C4%87.html> [dostęp: 20.12.2024].

<sup>47</sup> Słownik języka polskiego PWN, <https://sjp.pwn.pl/szukaj/przesta%C4%87.html> [dostęp: 20.12.2024].

<sup>48</sup> Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance), (OJ L, 2024/1689).

<sup>49</sup> O. Bodanka, *Naruszenie wizerunku przy wykorzystaniu technologii deepfake – analiza prawna i praktyczna*, „Opolskie Studia Administracyjno-Prawne” 2022, 20, 1, ss. 24 i 25.

W literaturze trafnie zauważa się, że w przypadku naruszenia wizerunku osoby nieżyjącej, szczególnie w sytuacji, gdy niemożliwe jest zidentyfikowanie sprawcy, osoby najbliższe mogą domagać się usunięcia tej treści od administratora portalu, na którym treść ta widnieje – zgodnie z art. 14 ustawy o świadczeniu usług drogą elektroniczną w zw. z art. 422 oraz art. 448 kodeksu cywilnego<sup>50</sup>. Choć O. Bodanka słusznie wskazuje praktyczne trudności przy próbie ochrony wizerunku osoby zmarłej na podstawie uregulowań krajowych, to należy zaznaczyć, że problem jest znacznie szerszy, systemowy. Niewystarczające jest stwierdzenie, że ochrona wizerunku osób zmarłych jest ograniczona w ramach krajowego ustawodawstwa. Obecne unormowania prawne ujawniają przede wszystkim luki strukturalne na poziomie prawa unijnego (AI Act), ponieważ definicja „deepfake” zawarta w art. 3 pkt 60 rozporządzenia 2024/1689 dotyczy wyłącznie „istniejących osób” (ang. *existing persons*), co powoduje automatycznie wykluczenie sytuacji, w których wygenerowana z wykorzystaniem narzędzi AI treść odnosiłaby się bezpośrednio do osoby zmarłej. Opisany powyżej przypadek na szczeblu unijnym wpada więc niejako w próżnię prawną, ponieważ nie jesteśmy w stanie bez wątpliwości stwierdzić, czym – zgodnie z definicjami na szczeblu unijnym – tak wygenerowana treść jest. Rozważając rzecz dalej, należy stwierdzić, że tak ukształtowane rozumienie pojęcia „deepfake” powoduje wyłączenie przywołanych wyżej sytuacji z reżimu unijnych regulacji, w tym obowiązków dotyczących: przejrzystości, ujawnienia, że treść została wygenerowana za pomocą narzędzi AI, a także odpowiedzialności dostawcy systemu AI. Rozporządzenie 2024/1689 wprost zawiera informację, że obowiązek dotyczący przejrzystości, skutkujący koniecznością informowania o istnieniu w materiale tak wygenerowanych lub zmanipulowanych treści, dotyczy materiałów *deepfake*<sup>51</sup>. Dlatego obecną definicję należy poddać krytyce, następnie doprowadzając do jej poszerzenia, aby już na poziomie europejskim wykluczyć możliwe przyszłe wątpliwości interpretacyjne.

Warto powyższe rozważania zestawić z systemem prawnym obowiązującym na poziomie federalnym w Stanach Zjednoczonych, w którym *Take It Down Act* – zatwierdzony 19 maja 2025 roku – wprowadza pojęcie cyfrowego fałszerstwa (*digital forgery*)<sup>52</sup>. Dotyczy on jednak wyłącznie wizerunków wygenerowanych lub zmanipulowanych przez AI, które przedstawiają nagość lub treści o charakterze seksualnym, pomijając zagrożenia związane z oszustwami z wykorzystaniem technologii *deepfake* czy generowanymi w tym zakresie treściami dezinformującymi. Mimo że definicja zakresowo jest zdecydowanie węższa od zaproponowanej w art. 3 pkt 60 AI Act, to uni-

<sup>50</sup> *Ibid.*, s. 25.

<sup>51</sup> Motyw 134 preambuły rozporządzenia 2024/1689.

<sup>52</sup> *Tools to Address Known Exploitation by Immobilizing Technological Deepfakes on Websites and Networks Act*, Statute at Large 139 Stat. 55 – Public Law No. 119-12 [dostęp: 19.05.2025].

ka sformułowania „existing persons” na rzecz przymiotnika „identifiable”. Obejmuje zatem swym zakresem każde intymne przedstawienie wizualne – niekoniecznie „istniejącej”, ale możliwej do zidentyfikowania osoby<sup>53</sup>. Pozytywnie należy ocenić katalog kar, który w przypadku publikacji materiału typu *deepfake* dotyczącej osoby dorosłej przewiduje karę grzywny, karę pozbawienia wolności do lat 2 lub obie te kary łącznie<sup>54</sup>. W przypadku publikacji dotyczącej osoby małoletniej górna granica pozbawienia wolności rośnie do lat 3<sup>55</sup>. Przepisy te w istocie obejmują powszechny problem, który należałoby uregulować również w ramach polskiego ustawodawstwa, ale jednocześnie zauważalne jest tu traktowanie zagrożeń gospodarczych jako społecznie mniej istotnych od intymnych treści wygenerowanych z użyciem technologii *deepfake*.

Również Republika Korei w swym systemie prawnym zawiera obostrzenia dotyczące generowania treści z wykorzystaniem technologii *deepfake*<sup>56</sup>. Art. 14-2 *expressis verbis* kryminalizuje edycję, łączenie lub przetwarzanie zdjęć, nagrań wideo lub audio. Ponadto penalizacji poddano rozpowszechnianie tych materiałów albo ich kopii. Wątpliwości budzi to, czy karze podlega rozpowszechnianie takiego materiału po śmierci tej osoby wbrew jej woli<sup>57</sup>. Znamienna jest również surowość kar, gdzie górną granicę wyznaczono na 7 lat prac przymusowych połączonych z pozbawieniem wolności albo karę grzywny nieprzekraczającą 50 milionów wonów południowokoreańskich. Na tle regulacji unijnych – które w ramach art. 50 ust. 4 sprowadzają się do iluzorycznego obowiązku ujawniania przez sprawcę, iż treść została wygenerowana z użyciem sztucznej inteligencji – rozwiązanie koreańskie wyróżnia się penalizacyjną precyzją. Podkreślenia wymaga fakt, że regulacje koreańskie – podobnie jak *Take It Down Act* – skoncentrowane są na ochronie ofiar przestępstw seksualnych, z pominięciem elementu zagrożeń gospodarczych.

<sup>53</sup> *Take It Down Act*, Public Law 119-12, § 2, nowelizujący sekcję 223 *Communications Act of 1934*, 47 U.S.C. § 223(h)(1)(B).

<sup>54</sup> *Take It Down Act*, Public Law 119-12, § 2, nowelizujący sekcję 223 *Communications Act of 1934*, 47 U.S.C. § 223(h)(4)

<sup>55</sup> *Ibid.*

<sup>56</sup> *Act on Special Cases Concerning the Punishment of Sexual Crimes*, Act No. 20459, amended Oct. 16, 2024, art. 14-2, Republic of Korea, Korea Legislation Research Institute, [https://elaw.klri.re.kr/eng\\_service/lawView.do?hseq=68812&lang=ENG](https://elaw.klri.re.kr/eng_service/lawView.do?hseq=68812&lang=ENG) [dostęp: 10.05.2026].

<sup>57</sup> Anglojęzyczne tłumaczenie wprost zawiera wyrażenie „after the death”. Nasuwa się pytanie, czy oficjalne tłumaczenie mogłoby pomylić sformułowanie „after the death” z pojęciem „subsequently”. Wątpliwość wynika z faktu, że koreański termin 사후에 (użyty w art. 14-2 ust. 2) rodzi wątpliwości interpretacyjne i może być rozumiany jako: po fakcie, następczo oraz po śmierci; wobec braku zapisu ‘hanja’ w tekście ustawy oraz rozstrzygających wyjaśnień, precyzyjne ustalenie znaczenia wymagałoby poznania oficjalnego stanowiska ustawodawcy koreańskiego.

## Wyzwania dotyczące odpowiedzialności cywilnej za czyny dokonane z wykorzystaniem systemów AI

Ostatnim, lecz nie mniej ważnym aspektem jest egzekwowanie przyjętych przepisów. Rozdział XII rozporządzenia 2024/1689 przewiduje katalog kar pieniężnych w przypadku nierespektowania przepisów wynikających z AI Act. Najsurowszymi karami obłożono działania, które stanowiłyby naruszenie art. 5 rozporządzenia 2024/1689, wskazującego zakazane praktyki w zakresie AI. Nieprzestrzeganie tego przepisu podlega administracyjnej karze pieniężnej w wysokości do 35 000 000 EUR lub – jeżeli sprawcą naruszenia jest przedsiębiorstwo – w wysokości do 7 % jego całkowitego rocznego światowego obrotu z poprzedniego roku obrotowego, w zależności od tego, która z tych kwot jest wyższa<sup>58</sup>. Ponadto państwa członkowskie mają obowiązek ustanowienia przepisów dotyczących kar i innych środków egzekwowania prawa, które mogą również obejmować ostrzeżenia i środki niepieniężne, mających zastosowanie w przypadku naruszeń przepisów rozporządzenia 2024/1689 przez operatorów; mają też obowiązek podejmowania wszelkich niezbędnych środków w celu zapewnienia ich właściwego i skutecznego wykonywania, z uwzględnieniem wytycznych wydanych przez Komisję, zgodnie z jego art. 96<sup>59</sup>.

Brak jednolitego unijnego reżimu odpowiedzialności cywilnej za szkody wyrządzone z użyciem systemów AI wymusza konieczność oparcia się na przepisach krajowych<sup>60</sup>. Równocześnie krajowy porządek prawny nie przewiduje *explicite* cywilnej odpowiedzialności za czyny *ex delicto* popełnione z użyciem systemów sztucznej inteligencji. Na poziomie unijnym uzupełnieniem jest dyrektywa Parlamentu Europejskiego i Rady (UE) 2024/2853 z 23 października 2024 r. w sprawie odpowiedzialności za produkty wadliwe<sup>61</sup>. W ramach art. 4 dyrektywy 2024/2853 włączono do definicji produktu oprogramowanie<sup>62</sup>. Jednocześnie zakłada ona, aby odpowiedzialnym podmiotem gospodarczym był producent wadliwego produktu, wadliwej części składowej lub – w przypadku producenta produktu lub jego części składowej, który ma siedzibę poza UE oraz bez uszczerbku dla odpowiedzialności tego producenta – odpowiedzialność ponosił importer wadliwego produktu lub części składo-

<sup>58</sup> Wymiar poszczególnych kar opisuje art. 99 ust. 3, 4 oraz 5 rozporządzenia 2024/1689.

<sup>59</sup> Art. 99 ust. 1 rozporządzenia 2024/1689.

<sup>60</sup> Projekt dyrektywy w sprawie odpowiedzialności za AI (AI Liability Directive, COM/2022/496 final), który miał uzupełnić AI Act, został wycofany przez Komisję Europejską w 2025 roku wobec braku perspektyw na osiągnięcie porozumienia legislacyjnego.

<sup>61</sup> Dyrektywa Parlamentu Europejskiego i Rady (UE) 2024/2853 z 23 października 2024 r. w sprawie odpowiedzialności za produkty wadliwe i uchylenia dyrektywy Rady 85/374/EWG (Dz.U. UE L 2024.2853 z późn. zm.), dalej: dyrektywa 2024/2853.

<sup>62</sup> Dyrektywa 2024/2853 definiuje produkt jako rzecz ruchomą, nawet jeżeli jest ona zintegrowana lub wzajemnie połączona z inną rzeczą ruchomą lub nieruchomą; obejmuje to energię elektryczną, cyfrowe pliki produkcyjne, surowce i oprogramowanie.

wej, przedstawiciel, który został przez producenta upoważniony lub dostawca usług realizacji zamówień, gdy brak jest importera mającego siedzibę w Unii oraz upoważnionego przedstawiciela<sup>63</sup>. Termin transpozycji omawianej dyrektywy przypada na 9 grudnia 2026 r., co oznacza, że do tego czasu krajowy system prawny jest pozbawiony wyraźnej podstawy do stosowania odpowiedzialności za produkt wadliwy wobec wykorzystania systemów AI służących generowaniu treści *deepfake*<sup>64</sup>.

Do 30 czerwca 2026 roku polski ustawodawca nie dokonał implementacji dyrektywy 2024/2853 – czego skutkiem jest obecnie obowiązujący art. 449<sup>1</sup> k.c., który w § 2 obejmuje definicję produktu, natomiast w § 1 zaznacza odpowiedzialność producenta. Jednocześnie przepis wskazuje tylko taki podmiot, który wytworzył produkt w ramach prowadzonej działalności gospodarczej<sup>65</sup>. Obecnie obowiązująca norma prawna stanowi implementację dyrektywy 85/374/EWG, bezpośrednio poprzedzającej dyrektywę 2024/2853<sup>66</sup>. Nawet uwzględniając zaktualizowaną już definicję pojęcia „produkt”, obejmującą oprogramowanie, wątpliwości budzi zakres odpowiedzialności producenta w zakresie wygenerowania za pomocą systemu AI materiału *deepfake*, który miał zamiar lub faktycznie wyrządził szkodę. Zasadniczo udostępnienie systemu nie powinno przesądzać o odpowiedzialności producenta, a wymagana winna być weryfikacja, czy producent dochował należytej staranności przy projektowaniu wymaganych przepisami prawa mechanizmów zabezpieczających przed niewłaściwym użyciem systemu przez użytkownika końcowego.

W ramach wyzwań związanych z zaistnieniem możliwych szkód dla przedsiębiorców – ale również osób fizycznych nieprowadzących działalności gospodarczej – związanych z nieetycznym użyciem narzędzi AI zauważono, że *de lege lata* AI Act przewiduje mechanizm, zgodnie z którym podmiot chcący wygenerować materiał *deepfake* powinien w sposób wyraźny i jasny ujawnić, że przedstawione treści zostały sztucznie wygenerowane lub zmanipulowane<sup>67</sup>. Rozwiązanie to opiera się na naiwnym założeniu, że podmiot działający *animo nocendi* dobrowolnie ujawni poszkodowanemu fakt popełnienia czynu niedozwolonego. Takie założenie jest sprzeczne z logiką, gdyż nikt, kto działa w złej wierze, nie ułatwia świadomie ustalenia istoty swojego czynu<sup>68</sup>.

<sup>63</sup> Art. 8 ust. 1 dyrektywy 2024/2853.

<sup>64</sup> Zakres oraz termin transpozycji dyrektywy 2024/2853 opisuje jej art. 22.

<sup>65</sup> G. Karaszewski, komentarz do art. 449<sup>1</sup> [w:] Kodeks cywilny. Komentarz aktualizowany, red. P. Nazaruk, Gdańsk 2024.

<sup>66</sup> Dyrektywa Rady z 25 lipca 1985 r. w sprawie zbliżenia przepisów ustawowych, wykonawczych i administracyjnych państw członkowskich dotyczących odpowiedzialności za produkty wadliwe (Dz.U. UE L 1985.210.29 z późn. zm.).

<sup>67</sup> Zob. motyw 134 Preambuły AI Act oraz art. 50 ust. 4 AI Act.

<sup>68</sup> P. Michalik, P. Zaremba, *Deepfake jako dowód w procesie cywilnym – rozważania na tle polskiego postępowania cywilnego*, „Prawo i Więź” 2025, 57, 4, s. 639.

Niewątpliwie istotne jest, aby obowiązujące przepisy były respektowane, a ich wykonywanie skutecznie kontrolowane i egzekwowane. Jednakże należy pamiętać, że nadmierna regulacja i zbytne obciążenie przedsiębiorców obowiązkami prawnymi negatywnie wpływa na innowacyjność firm, która jest kluczowa dla rozwoju gospodarki – w szczególności w sektorze nowych technologii. Według raportu sporządzonego w grudniu 2022 roku pod tytułem AI Act Impact Survey wynika, że dwie trzecie startupów spodziewa się negatywnego wpływu rozporządzenia 2024/1689 na innowacje w zakresie sztucznej inteligencji w Europie<sup>69</sup>. *De lege ferenda* należałoby więc maksymalnie uprościć kwestie dokumentacyjne dla przedsiębiorców. Obecna sytuacja – w której załącznik VIII opisuje informacje przekazywane przy rejestracji systemów AI wysokiego ryzyka zgodnie z art. 49 rozporządzenia 2024/1689, gdzie załącznik ten jest dodatkowo podzielony na trzy odrębne sekcje, następnie załącznik IX, który opisuje informacje przedkładane przy rejestracji systemów AI wysokiego ryzyka wymienionych w załączniku III do rozporządzenia w odniesieniu do testów w warunkach rzeczywistych zgodnie z art. 60 AI Act – nie ułatwia rozwoju przedsiębiorstw zajmujących się sztuczną inteligencją<sup>70</sup>. Należy pamiętać, że zagrożenia gospodarcze wynikające z rozwoju nowoczesnych technologii mogą dotyczyć przedsiębiorców w ogólności – przez używanie sztucznej inteligencji w sposób nieetyczny, w celu osiągnięcia konkretnego celu dezinformującego, dyskredytującego-ośmieszającego czy też cyberprzestępczego – ale też mogą dotyczyć wyłącznie przedsiębiorstw zajmujących się rozwojem nowych technologii, dla których nadmierna biurokracja może stać się impulsem do przeniesienia działań biznesowych na inną część globu, w której regulacje prawne będą zdecydowanie mniej uciążliwe.

Niemniej jednak istnieją pewne zagrożenia skłaniające do dalszego regulowania podatnych na nie materii. Odnosi się to przede wszystkim do przepisów krajowych, gdzie należałoby wprowadzić odrębne przepisy w ramach prawa karnego, które uzupełniłyby lukę wynikającą z burzliwego i wciąż przyspieszającego rozwoju technologicznego. Zagrożenia te dotyczą przy tym wykorzystania technologii *deepfake* zarówno w sferze gospodarczej, jak w kontekście naruszenia dóbr konkretnych jednostek, takich jak godność, cześć, dobre imię. Jak słusznie zauważa A. Ziobroń: niewystarczający zakres regulacji potencjalnie szkodliwego społecznie zjawiska może doprowadzić do osłabienia funkcji sprawiedliwościowej prawa karnego<sup>71</sup>. Dlatego prace nad nowelizacją kodeksu karnego w tym zakresie powinny nastąpić tak szybko, jak to tylko możliwe.

<sup>69</sup> AI Act Impact Survey, [https://www.appliedai.de/uploads/files/AI-Act-Impact-Survey\\_Slides\\_Dec12.2022.pdf](https://www.appliedai.de/uploads/files/AI-Act-Impact-Survey_Slides_Dec12.2022.pdf) [dostęp: 27.12.2024].

<sup>70</sup> Załącznik VIII oraz IX do rozporządzenia.

<sup>71</sup> A. Ziobroń, *Deepfake a prawo...*, s. 235.

---

## Podsumowanie

---

W ramach niniejszej analizy wskazano zagrożenia gospodarcze wynikające z wykorzystania technologii *deepfake* oraz ich implikacje prawne i społeczne. W kontekście gospodarczym zidentyfikowano kluczowe zagrożenia, takie jak:

- a) dezinformacja – rozpowszechnianie fałszywych informacji w celu manipulacji rynkami kapitałowymi lub szkodenia reputacji firm;
- b) naruszenie dóbr osobistych – ataki *deepfake* skierowane na liderów biznesu, co może prowadzić do utraty zaufania ze strony społeczeństwa, a w konsekwencji do strat finansowych przedsiębiorstw;
- c) cyberoszustwa – wyłudzenie środków z wykorzystaniem spreparowanych treści, np. podszywanie się pod znanych biznesmenów w celu zachęcenia do fałszywych inwestycji;
- d) inne cyberprzestępstwa – wykorzystanie *deepfake*’ów do zwiększania skuteczności ataków phishingowych, generowania fałszywych recenzji lub innego manipulowania opinią publiczną.

Odnosząc się do głównego problemu badawczego przedstawionego w niniejszym artykule, należy stwierdzić, że obecne mechanizmy prawne w UE – ze szczególnym uwzględnieniem AI Act – stanowią duży krok naprzód w sferze uregulowań związanych z technologicznym rozwojem oraz wynikających z tego zagrożeń. Należy jednak zaznaczyć, że z przeprowadzonej analizy jednoznacznie wynika, iż regulacje te nie są wystarczające, by w sposób właściwy chronić przed gospodarczymi skutkami nadużyć dokonanych z wykorzystaniem narzędzi AI. Rozporządzenie 2024/1689 w obecnym brzmieniu nie przewiduje stworzenia stosownych mechanizmów pozwalających na egzekwowanie odpowiedzialności za treści generowane przez systemy sztucznej inteligencji, które nie mają osobowości prawnej, a jednocześnie skutki ich wykorzystywania są realne. Słusznym rozwiązaniem było wprowadzenie definicji legalnej pojęcia „*deepfake*”, jednak definicja ta ma lukę, ponieważ nie obejmuje sytuacji, gdy wykorzystano wizerunek osoby zmarłej – np. niedawno zmarłego prezesa zarządu. Ograniczenie definicji wyłącznie do istniejących osób może zostać świadomie wykorzystane przez sprawców chcących posłużyć się wizerunkiem osoby zmarłej. Zidentyfikowana luka nie powinna jednak być uzupełniana przez wykładnię, lecz wymaga interwencji legislacyjnej – ze względu na jej strukturalny, a nie redakcyjny charakter. Dodatkowo mechanizm wynikający z art. 50 ust. 4 AI Act, zakładający dobrowolne informowanie przez sprawcę oszustwa, że korzystał w jego trakcie z systemów AI, jest sprzeczny z elementarną logiką. Przepis ten jedynie pozoruje realną ochronę, zamiast rzeczywiście ją tworzyć. Zbytnia ogólność przepisów dotyczących ochrony przed nieetycznie wykorzystanym narzędziem AI, pozwalająca tworzyć oszukańcze bądź dezinformacyjne treści *deepfake*, ogranicza skuteczność uregulowań prawnych wynikających z AI Act.

Z drugiej strony nadmiar biurokratycznych i innych obowiązków nałożony na przedsiębiorców może zniechęcać ich do podejmowania technologicznie innowacyjnych działań – co bezpośrednio wpłynie będzie na konkurencyjność europejskiego rynku w porównaniu z resztą świata. Dlatego też wnioskiem płynącym z badań jest potrzeba uzupełnienia przepisów unijnych o rozszerzenie definicji pojęcia *deepfake*; zarazem jednak konieczne staje się ograniczenie i uproszczenie wymagań administracyjnych dla przedsiębiorców działających w sektorze sztucznej inteligencji. Dopiero synteza precyzyjnie ukształtowanych przepisów z jednoczesnym odbiurokratyzowaniem wymagań narzuconych na europejskie podmioty gospodarcze przyniesie odpowiednie rozwiązanie – łączące przeciwdziałanie zagrożeniom gospodarczym z zachowaniem odpowiedniego poziomu innowacji i rozwoju. Jednoznacznym wnioskiem z przeprowadzonej analizy jest potrzeba pilnej rewizji obowiązujących ram regulacyjnych, począwszy od definicji normatywnych, a skończywszy na mechanizmach egzekwowania odpowiedzialności prawnej podmiotów używających systemów AI w sposób naruszający obowiązujące przepisy. Jednocześnie jednym z większych wyzwań na szczególnie unijnym pozostaje ukształtowanie przepisów w taki sposób, aby zneutralizować jak najwięcej zagrożeń gospodarczych związanych z rozwojem sztucznej inteligencji, ale jednocześnie zrealizować ten cel tak, by inne państwa – w szczególności Stany Zjednoczone Ameryki czy Chińska Republika Ludowa – nie zdominowały kosztem Europy światowego rozwoju technologicznego. Rozszerzenie definicji *deepfake* o treści wygenerowane z wykorzystaniem wizerunków osób zmarłych, zastąpienie art. 50 ust. 4 skuteczną sankcją cywilną wraz z legitymacją czynną poszkodowanego wobec sprawcy wywołującego szkodę oraz konieczność implementacji dyrektywy 2024/2853 nie są bynajmniej postulatami nadregulacyjnymi, lecz koniecznym uzupełnieniem obecnie obowiązujących przepisów w celu zwiększenia ich skuteczności i precyzji.

### Abstrakt

Przedmiotem artykułu jest analiza wpływu sztucznej inteligencji, a w szczególności technologii *deepfake* na gospodarkę. Praca ma za cel zweryfikowanie, czy obecne unormowania prawne na szczeblu europejskim w wystarczającym stopniu chronią przed zagrożeniami, jakie stawia przed nami rozwój technologiczny. Przedstawiono przypadki, w których można wykorzystać technologię *deepfake* w celu wyłudzenia środków finansowych czy podjęcia innych nieetycznych działań. Kluczowe zagadnienia obejmują rolę unijnych regulacji – takich jak akt w sprawie sztucznej inteligencji – w celu przeciwdziałania zagrożeniom gospodarczym. Ponadto zwrócono uwagę na konieczność zmian legislacyjnych, które uzupełnią występujące obecnie luki prawne, a jednocześnie zmniejszą

szą zakres obowiązków narzucanych na przedsiębiorców. Wskazano, że unormowania prawne dotyczące sztucznej inteligencji są niezbędne – z jednoczesnym zauważeniem, że ogromnym wyzwaniem jest utrzymanie odpowiedniego poziomu innowacyjności przedsiębiorstw w Europie.

**Słowa kluczowe:** deepfake, gospodarka, akt w sprawie sztucznej inteligencji, zagrożenia, sztuczna inteligencja (AI).

---

## BIBLIOGRAFIA

---

Bodanka O., *Naruszenie wizerunku przy wykorzystaniu technologii deepfake – analiza prawna i praktyczna*, „Opolskie Studia Administracyjno-Prawne” 2022, 20, 1.

Bukowski M., *Zwalczanie cyberprzestępczości ekonomicznej przy wykorzystaniu sztucznej inteligencji (AI)*, „Przegląd Policyjny” 2023, 151, 3.

Dąbrowska I., *Deepfake: nowy wymiar internetowej manipulacji*, „Zarządzanie Mediami” 2020, 8, 2.

Kiełpiński K., *Deepfake jako narzędzie do przekazywania informacji fałszywej i domniemanej. Analiza prawnokarna i cybernetyczna*, „Kwartalnik Krajowej Szkoły Sądownictwa i Prokuratury” 2023, 3, 51.

Kira B., *When Non-Consensual Intimate Deepfakes Go Viral: The Insufficiency of the UK Online Safety Act*, ‘Computer Law & Security Review’ 2024, 54.

Kodeks cywilny. Komentarz aktualizowany, P. Nazaruk (red.), Gdańsk 2024.

Kumari D., *Deepfake Technology and Legal Issues*, ‘LawFoyer International Journal of Doctrinal Legal Research’ 2024, 2, 1.

Lai L., Świerczyński M. (red.), *Prawo sztucznej inteligencji*, Warszawa 2020.

Michalik P., Zaremba P., *Deepfake jako dowód w procesie cywilnym – rozważania na tle polskiego postępowania cywilnego*, „Prawo i Wiąż” 2025, 57, 4.

Mierzejewski D., Nawrotkiewicz J., *W obliczu wspólnej polityki UE wobec Chin: „mission impossible”?*, „Pulaski Policy Papers” 2024, XI.

Piecuch A., *Sztuczna inteligencja w perspektywie społecznej*, „Edukacja Ustawiczna Dorosłych” 2023, 123, 4.

Skutki i zagrożenia cywilizacji informacyjnej, K.A. Nawrot, K. Prandecki (red.), Komitet Prognoz PAN 2023.

Spałek D., Rzymowski J., *Jak nie stosować AI Act na podstawie definicji systemu sztucznej inteligencji*, „Prawo Nowych Technologii” 2024, 2.

Stewart I., *Significant Figures Lives and Works of Trailblazing Mathematics*, New York 2017 [wyd. pol. *Krótką historią wielkich umysłów. Genialni matematycy i ich arcydzieła*, tłum. U. Seweryńska i M. Seweryński, Warszawa 2019].

Turing A., *Computing Machinery and Intelligence*, 'Mind' 1950, 59, 236.

Vera G.G., *Deep fake (!) – postęp technologiczny a prawo karne*, „Zeszyty Naukowe Uniwersytetu Rzeszowskiego – Seria Prawnicza” 2024, 1, 44.

Widło J., *Odpowiedzialność za szkodę wyrządzoną przez pojazdy autonomiczne w świetle zmian prawa UE*, „Forum Prawnicze” 2025, 2.

Zielińska A., *Korzystanie z treści chronionych a sztuczna inteligencja – jak daleko sięga odpowiedzialność dostawcy usług udostępniania treści online w świetle art. 17 dyrektywy 2019/790?*, „Studia de Cultura” 2022, 14, 2.

Ziobron A., *Deepfake a prawo karne. Uwagi „de lege lata” i „de lege ferenda” dotyczące fałszywej pornografii*, „Studenckie Prace Prawnicze, Administratywistyczne i Ekonomiczne” 2021, 37.

### Źródła internetowe

AI Act Impact Survey, [https://www.appliedai.de/uploads/files/AI-Act-Impact-Survey\\_Slides\\_Dec12.2022.pdf](https://www.appliedai.de/uploads/files/AI-Act-Impact-Survey_Slides_Dec12.2022.pdf) [dostęp: 27.12.2024].

Brandon J., *Terrifying high-tech porn: Creepy 'deepfake' videos are on the rise*, <https://www.foxnews.com/tech/terrifying-high-tech-porn-creepy-deepfake-videos-are-on-the-rise> [dostęp: 7.12.2024].

Crime as a service (CaaS), <https://www.comcert.pl/slownik/crime-as-a-service-caas/> [dostęp: 17.12.2024].

FBI, *Internet Crime Report 2022*, [https://www.ic3.gov/AnnualReport/Reports/2022\\_ic3report.pdf](https://www.ic3.gov/AnnualReport/Reports/2022_ic3report.pdf) [dostęp: 24.10.2025].

FBI, *Internet Crime Report 2024*, [https://www.ic3.gov/AnnualReport/Reports/2024\\_IC3Report.pdf](https://www.ic3.gov/AnnualReport/Reports/2024_IC3Report.pdf) [dostęp: 24.10.2025].

FBI, *Internet Crime Report 2025*, [https://www.ic3.gov/AnnualReport/Reports/2025\\_IC3Report.pdf](https://www.ic3.gov/AnnualReport/Reports/2025_IC3Report.pdf) [dostęp: 8.05.2026].

Jakubczak D., *Deepfake. Oszustwo na sklonowanego dyrektora*, <https://www.dw.com/pl/deepfake-oszustwo-na-sklonowanego-dyrektora/a-69116765> [dostęp: 16.12.2024].

Lista osób, których wizerunek przestępcy wykorzystali w oszustwach deepfake, <https://nask.pl/aktualnosci/nask-ostzrega-przestepcy-tworza-coraz-bardziej-pomyslowne-oszustwa-deepfake/> [dostęp: 12.12.2024].

NASK ostrzega: przestępcy tworzą coraz bardziej pomysłowe oszustwa deepfake, <https://nask.pl/aktualnosci/nask-ostzrega-przestepcy-tworza-coraz-bardziej-pomyslowne-oszustwa-deepfake/> [dostęp: 12.12.2024].

Online Safet Act 2023, c. 50, Zjednoczone Królestwo, <https://www.legislation.gov.uk/ukpga/2023/50/contents/enacted> [dostęp: 10.05.2026].

Słownik języka polskiego PWN, <https://sjp.pwn.pl/slowniki/istnie%C4%87.html> [dostęp: 20.12.2024].

Słownik języka polskiego PWN, <https://sjp.pwn.pl/szukaj/przesta%C4%87.html> [dostęp: 20.12.2024].

Unni A., What You Need To Know About Crime As A Service (CSaaS), <https://www.stickmancyber.com/cybersecurity-blog/what-you-need-to-know-about-crime-as-a-service-csaas> [dostęp: 17.12.2024].



Dominik Bierecki\*

## Teoria salda w prawie niemieckim i gruzińskim a perspektywa polska\*\*

### [The Balance Theory in German and Georgian Law and the Polish Perspective]

#### Abstract

The settlement of mutual performances rendered under contracts subsequently held to be void constitutes one of the more enduring doctrinal problems of the law of unjust enrichment. The question has acquired renewed practical significance in Poland in the wake of judgments declaring foreign-currency-indexed and foreign-currency-denominated loan agreements null and void, and in the light of the divergent jurisprudence that has emerged in their aftermath. In this context, the possibility of applying the balance theory (German: *Saldotheorie*) has gained attention. The present study undertakes a comparative examination of the doctrinal foundations of the balance theory in German jurisprudence and of its statutory counterpart in the Georgian Civil Code, and it assesses the compatibility of the theory with the structural premises of the Polish regime governing the restitution of performances rendered without legal cause. It is submitted that the rationale underlying the balance theory — namely, the mitigation of the asymmetric distribution of risk that arises in contracts involving non-monetary performance — does not obtain in the context of void loan agreements, in which the performances of both parties are pecuniary in character. The reception of the balance theory in Polish jurisprudence concerning void loan agreements is therefore both structurally inapposite and functionally otiose.

**Keywords:** *Saldotheorie*, restitution of performances rendered without legal cause, void loan agreements, comparative law of unjust enrichment.

\* **Dominik Bierecki** – profesor uczelni, dr hab. nauk prawnych, Uniwersytet Pomorski w Słupsku (afiliacja); <https://orcid.org/0000-0001-6993-3974>; [dominik.bierecki@upsl.edu.pl](mailto:dominik.bierecki@upsl.edu.pl) / associate professor, PhD in Law, post-doctoral degree, Pomeranian University in Słupsk (affiliation).

\*\* Artykuł przygotowano w trakcie stażu naukowego na Uniwersytecie w Kolonii w kwietniu i maju 2026 r. Pragnę serdecznie podziękować Panu Profesorowi Dariuszowi Szpoperowi oraz Panu Profesorowi Christianowi Rolfsowi za życzliwość i umożliwienie odbycia tego stażu, który stał się dla mnie bardzo cennym doświadczeniem badawczym / This article was written during a research placement at the University of Cologne in April and May 2026. I would like to extend my sincere thanks to Professor Dariusz Szpoper and Professor Christian Rolfs for their kindness and for making this placement possible; it has been a very valuable research experience for me.

## Zarys problematyki

Zagadnienie, którym zajmuje się tzw. teoria salda, dotyczy sposobu rozliczenia świadczeń wymienionych przez strony na podstawie nieważnej umowy. Teoria ta zakłada, że świadczenia stron spełnione na podstawie nieważnej umowy należy od razu skompensować, a zwrotowi podlega tylko nadwyżka wartości świadczenia jednej ze stron. W teorii salda stronom nie przysługują samodzielne roszczenia kondykcyjne<sup>1</sup>. Teoria salda wywodzi się z niemieckiego orzecznictwa, w którym dotyczy wyłącznie rozliczenia świadczeń z nieważnych umów wzajemnych<sup>2</sup>. W polskim orzecznictwie pojawiła się jako instrument rozliczania świadczeń z nieważnych umów kredytu – jeszcze przed potwierdzeniem wzajemnego charakteru umowy kredytu przez uchwałę Sądu Najwyższego z 28 lutego 2025 r.<sup>3</sup> Można powiedzieć, że ze względu na swój rodowód stosowanie teorii salda potwierdziło wzajemny charakter umowy kredytu jeszcze przed tą uchwałą SN.

Celem artykułu jest omówienie funkcjonowania teorii salda w prawie niemieckim (w którym obowiązuje ona jako konstrukcja orzecznicza służąca do rozliczania bezpodstawnego wzbogacenia) oraz prawie gruzińskim (teoria salda została tam wprowadzona do ustawodawstwa), a także ocena, czy założenia teorii salda korespondują z polskim reżimem nienależnego świadczenia. Teza badawcza artykułu stanowi, że recypowanie teorii salda do prawa polskiego – w szczególności w odniesieniu do rozliczenia świadczeń nieważnej umowy kredytu indeksowanego do waluty obcej lub denominowanego w walucie obcej – jest niewłaściwe. Artykuł został przygotowany z posłużeniem się trzema metodami badawczymi. Po pierwsze: metodą teoretycznoprawną, zastosowaną przez analizę niemieckiego i gruzińskiego piśmiennictwa prawniczego dotyczącego teorii salda. Po drugie: metodą komparatystyczną – przez odwołania do niemieckiego i gruzińskiego prawa cywilnego. Po trzecie: metodą dogmatycznoprawną, polegającą na analizie norm polskiego prawa cywilnego.

<sup>1</sup> L. Enneccerus (H. Lehmann), Lehrbuch des Bürgerlichen Rechts, tom 2, Recht der Schuldverhältnisse. Ein Lehrbuch, Tübingen 1958, § 227 III.

<sup>2</sup> Wyrok Reichsgerichtu z 14 marca 1903 r., V 458/02, Entscheidungen des Reichsgerichts in Zivilsachen, tom 54, ss. 137–144.

<sup>3</sup> Uchwała składu siedmiu sędziów Sądu Najwyższego z 28 lutego 2025 r., III CZP 126/22, OSNC 2025, nr 11, poz. 97.

## Teoria salda w prawie niemieckim

W prawie niemieckim teoria salda (*Saldotheorie*) jest sposobem stosowania norm prawnych dotyczących bezpodstawnego wzbogacenia w sytuacji obustronnej kondykcji z dotkniętej nieważnością umowy wzajemnej (synalagmatycznej). Według teorii salda, gdy obie strony spełniły świadczenia z nieważnej umowy wzajemnej, wzbogaceniem jest nadwyżka jednego świadczenia nad drugim. Saldo to różnica między aktywami a pasywami w rozliczeniu jednej ze stron. Saldo i nadwyżka oznaczają zatem tę samą wartość, opisywaną jednak z różnych perspektyw rachunkowych. Saldo to wynik księgowy, a nadwyżka to ten sam wynik widziany od strony wierzyiciela kondykcji.

Strukturalne uzasadnienie teorii salda w niemieckim prawie cywilnym wiąże się z obowiązywaniem w tym systemie prawa zasady abstrakcyjności przysporzenia (*Abstraktionsprinzip*). Zgodnie z tą zasadą czynność rozporządzająca (*Verfügungsgeschäft*) jest niezależna – oddzielona (*Trennungsprinzip*) od leżącej u jej podstaw czynności zobowiązującej (*Verpflichtungsgeschäft*, będącej *causa*). Powoduje to, że nieważność umowy zobowiązaniowej nie powoduje upadku przeniesienia własności (przysporzenia)<sup>4</sup>. W konsekwencji w przypadku nieważnej umowy wzajemnej obie strony – mimo nieważności *causa* – zachowują nabytą własność otrzymanych świadczeń. Po stronie zbywcy nie istnieje roszczenie windykacyjne (bo własność przeszła), a obie strony mają wobec siebie wyłącznie roszczenia kondykcyjne (§ 812 ust. 1 BGB). Teoria salda rozwiązuje problem tworzony przez *Abstraktionsprinzip* (przysługiwanie wyłącznie roszczeń kondykcyjnych i brak roszczenia windykacyjnego), który to problem nie istnieje w systemie kauzalnym, bo tam nieważność *causa* powoduje upadek przeniesienia własności, co aktualizuje po stronie świadczącego roszczenie windykacyjne<sup>5</sup>.

Teoria salda obowiązuje w niemieckim prawie w wyniku zastosowania wykładni operatywnej, nie została bowiem wprost wyrażona przez normy prawne. Wprowadzono ją wyrokiem Reichsgerichtu (Sądu Rzeszy) z 14 marca 1903 r. i od tego momentu pozostaje przewodnią linią orzecznictwa Bundesgerichtshof (Federalnego Trybunału Sprawiedliwości – BGH) na gruncie § 812 ust. 1 i § 818 ust. 3 *Bürgerliches Gesetzbuch*<sup>6</sup> (BGB).

Klasyczna wersja teorii salda została skonstruowana przez Reichsgericht w wyroku z 14 marca 1903 r., a następnie przejęta przez wczesne orzecznictwo

<sup>4</sup> L. van Vliet, *Iusta Causa Traditionis and Its History in European Private Law*, 'European Review of Private Law' 2003, 3, ss. 376–378.

<sup>5</sup> H. Honsell, *Tradition und Zession — kausal oder abstrakt?* [w:] *Festschrift für Wolfgang Wiegand zum 65. Geburtstag*, Bern–München 2005, ss. 349–372, zwłaszcza zaś s. 358.

<sup>6</sup> <https://www.gesetze-im-internet.de/bgb/> [dostęp: 3.05.2026].

two BGH<sup>7</sup>. Punktem wyjścia jest tu założenie, że każdej ze stron obustronnie wykonanej, lecz nieważnej umowy wzajemnej przysługuje odrębne roszczenie kondykcyjne na podstawie § 812 ust. 1 BGB (odpowiednika art. 410 k.c.<sup>8</sup>). Niezależność roszczeń kondykcyjnych powoduje jednak, że strona będąca adresatem roszczenia kondykcyjnego, u której wzbogacenie odpadło (np. wskutek zniszczenia, zgubienia lub utraty wartości rzeczy), mogłaby odmówić zadośćuczynienia roszczeniu, powołując się na brak wzbogacenia (§ 818 ust. 3 BGB, odpowiednik polskiego art. 409 k.c.). Jednocześnie, ze względu na niezależność roszczeń kondykcyjnych, mogłaby żądać pełnego zwrotu własnego świadczenia. Klasyczna teoria salda opiera się na wykładni celowościowej (*teleologische Reduktion*) § 818 ust. 3 BGB, która zakłada wyłączenie zastosowania tego przepisu w sytuacji odpadnięcia wzbogacenia. Uzasadnieniem utrzymania obowiązku strony, u której wzbogacenie odpadło, jest powołanie się na synalagmatyczny charakter umowy wzajemnej (§ 320–322 BGB)<sup>9</sup>. W tym przypadku znaczenie ma synalagma funkcjonalna. Ta postać synalagmy powoduje, że świadczenia stron w umowie wzajemnej są ze sobą powiązane wedle zasady *do ut des* (są dla siebie odpowiednikami). Faktyczne wykonanie umowy wzajemnej wraz ze świadczeniem przenosi na stronę otrzymującą także ryzyko jego utraty. Orzecznictwo uznało za wewnętrznie sprzeczne (*venire contra factum proprium*) jednoczesne żądanie zwrotu własnego świadczenia i powoływanie się na odpadnięcie wzbogacenia po swojej stronie. W rezultacie roszczenia kondykcyjne *ipso iure* przekształcają się w jedno roszczenie o saldo (czyli nadwyżkę aktywów nad pasywami) – jako skutek niemożności podniesienia zarzutu z § 818 ust. 3 BGB<sup>10</sup>.

Rozciągnięcie synalagmatycznego skutku na zwrot świadczeń z nieważnej umowy wzajemnej nastąpiło przez wprowadzenie do piśmiennictwa niemieckiego pojęcia faktycznej synalagmy (*faktisches Synallagma*)<sup>11</sup>. Konstrukcja ta zakłada, że nawet jeśli synalagma nie powstała (skoro umowa jest nieważna *ex tunc*), to między stronami zaistniał faktyczny układ wzajemnej wymiany. Strony rzeczywiście świadczyły *do ut des* – co uzasadnia traktowanie obustronnych kondykcji tak, jakby pozostawały w synalagmatycznym powiązaniu.

<sup>7</sup> Wyrok BGH z 19 stycznia 1951 r., I ZR 15/50, Entscheidungen des Bundesgerichtshofes in Zivilsachen (dalej: BGHZ) tom 1, ss. 75–81; wyrok BGH z 22 czerwca 1954 r., I ZR 225/53, BGHZ tom 14, ss. 7–10; wyrok BGH z 8 stycznia 1970 r., VII ZR 130/68, BGHZ tom 53, ss. 144–147; wyrok BGH z 14 października 1971 r., VII ZR 313/69, BGHZ tom 57, ss. 137–150; wyrok BGH z 26 października 1978 r., VII ZR 202/76, BGHZ tom 72, ss. 252–255.

<sup>8</sup> Ustawa z 23 kwietnia 1964 r. Kodeks cywilny (tekst jedn. Dz.U. z 2026 r. poz. 795).

<sup>9</sup> L. Ennecerus (H. Lehmann), Lehrbuch des Bürgerlichen Rechts, tom 2, Recht der Schuldverhältnisse. Ein Lehrbuch, § 227 III. Por. K. Larenz, C.-W. Canaris, Lehrbuch des Schuldrechts, tom 2, cz. 2 szczególna, München 1994, § 73 I 4.

<sup>10</sup> H. Prütting [w:] Bürgerliches Gesetzbuch mit Internationalem Privatrecht und nationalen, unionsrechtlichen sowie völkerrechtlichen Neben- oder Ergänzungsgesetzen. Kommentar, H. Prütting, G. Wegen, G. Weinreich, München 2024, s. 1900.

<sup>11</sup> E. v. Caemmerer, *Bereicherung und unerlaubte Handlung* [w:] Festschrift für Ernst Rabel, tom 1, Tübingen 1954, ss. 333, 334 oraz 385. Zagadnieniu faktycznej synalagmy poświęcono opracowanie monograficzne: H. Leser, *Von der Saldotheorie zum faktischen Synallagma*, Freiburg im Breisgau 1956, *passim*.

Modyfikacja klasycznej teorii salda następuje w jej wariacie bilansowym, obecnym w orzecznictwie BGH i przyjmowanym w nowszym piśmiennictwie<sup>12</sup>. W bilansowej odmianie teorii salda nie zakłada się istnienia dwóch odrębnych roszczeń kondykcyjnych; przyjmuje się natomiast, że w przypadku obustronnie wykonanej, lecz nieważnej umowy wzajemnej istnieje od początku tylko jedno roszczenie kondykcyjne<sup>13</sup>. Założenie to wynika z odmiennej wykładni § 812 ust. 1 zd. 1 BGB, który stanowi: *Wer durch die Leistung eines anderen oder in sonstiger Weise auf dessen Kosten etwas ohne rechtlichen Grund erlangt, ist ihm zur Herausgabe verpflichtet* [Kto przez świadczenie innego lub w inny sposób na jego koszt otrzymuje coś bez podstawy prawnej, jest zobowiązany do wydania tego]. W wariacie bilansowym uznaje się, że „coś” (etwas) otrzymane przez strony nieważnej umowy wzajemnej nie jest sumą dwóch niezależnych świadczeń, ale jedną ekonomiczną całością (*einheitliche Vermögensverschiebung*). Jest to przesunięcie majątkowe, którego rozmiar wyznacza saldo, a więc różnica wartości aktywów po obu stronach<sup>14</sup>. Wzbogaceniem w rozumieniu § 812 ust. 1 BGB od początku jest zatem tylko ta różnica, a roszczenie kondykcyjne przysługuje wyłącznie tej stronie, na której korzyść wypada saldo<sup>15</sup>. Druga strona nie ma żadnego roszczenia, ponieważ nie występuje tutaj wzbogacenie. W konsekwencji § 818 ust. 3 BGB w ogóle nie znajduje zastosowania, gdyż nie ma osobnego wzbogacenia, które mogłoby odpaść w rozumieniu tego przepisu, a kwestia ryzyka utraty rzeczy jest rozstrzygana już na etapie ustalania samego salda<sup>16</sup>. Główną zaletą bilansowego wariantu teorii salda jest uznanie, że znajduje ona bezpośrednio podstawę w § 812 ust. 1 zd. 1 BGB<sup>17</sup>.

Obok klasycznej *Saldotheorie* i jej wariantu bilansowego w niemieckiej doktrynie funkcjonują również alternatywne propozycje rozwiązania problemu rozliczenia obustronnie wykonanej, lecz nieważnej umowy wzajemnej. Propozycje przyjmują istnienie dwóch roszczeń kondykcyjnych po przeciwnych stronach umowy wzajemnej. Są jednak modyfikacją teorii dwóch kondykcji (*modifizierte Zweikondiktionentheorie*) ze względu na postulowane ograniczenia powoływania się na § 818 ust. 3 BGB. Pierwsza z tych propozycji – pod nazwą teorii decyzji majątkowej (*Lehre von der vermögensmäßigen Entscheidung*) – zakłada, że ryzyko utraty otrzymanej rzeczy obciąża tę stronę, która

<sup>12</sup> Wyrok BGH z 19 stycznia 2001 r., V ZR 437/99, BGHZ tom 147, ss. 152–157; W. Flume, *Die Rückabwicklung nichtiger Kaufverträge nach Bereicherungsrecht — Zur Saldotheorie und ihren „Ausnahmen“*, *JuristenZeitung* 2002, ss. 321–325.

<sup>13</sup> Wyrok BGH z 19 stycznia 2001 r., V ZR 437/99, s. 157, gdzie roszczenie kondykcyjne ujmowane jest jako: *einheitlicher Anspruch auf Ausgleich aller mit der Vermögensverschiebung zurechenbar zusammenhängenden Vorgänge in Höhe des sich dabei ergebenden Saldos* [jednolite roszczenie o wyrównanie wszystkich zdarzeń powiązanych przyczynowo z przesunięciem majątkowym w wysokości wynikającego z nich salda].

<sup>14</sup> W. Flume, *Die Rückabwicklung...*, s. 322.

<sup>15</sup> Zob. K. Larenz, C.-W. Canaris, *Lehrbuch...*, § 73 II.

<sup>16</sup> W. Flume, *Die Rückabwicklung...*, s. 324.

<sup>17</sup> W. Flume, *Die Rückabwicklung...*, s. 321.

w wykonaniu nieważnej umowy podjęła decyzję majątkową o wymianie własnego świadczenia na świadczenie kontrahenta. Strona, która świadomie zdecydowała się na nabycie rzeczy w zamian za swoje świadczenie, ponosi niezależnie od ważności samej umowy ryzyko utraty własnego świadczenia. W ten sposób synalagma funkcjonalna i nabycie odpowiednika świadczenia uzasadnia ryzyko utraty własnego świadczenia. Powoduje to, że ochrona przewidziana w § 818 ust. 3 BGB ustępuje przed konsekwencjami świadomej decyzji majątkowej strony<sup>18</sup>.

Druga modyfikacja teorii dwóch kondykcji zwana jest kondykcją świadczenia wzajemnego (*Gegenleistungskondiktio*n). Konstrukcja ta różnicuje rozkład ryzyka według przypisywalności (*Zurechenbarkeit*) utraty rzeczy. Przypisywalność dotyczy subiektywnej strony ryzyka utraty rzeczy, obejmującej zarówno zawinione zachowanie (niedbalstwo, lekkomyślność), jak i świadome korzystanie z rzeczy w sposób zwiększający ryzyko jej utraty. Wystąpienie przypisywalności uzasadnia ograniczenie stosowania § 818 ust. 3 BGB przez sięgnięcie do reguły zawartej w § 819 ust. 1 BGB, zgodnie z którą odpadnięcie wzbogacenia (*Wegfall der Bereicherung*) nie chroni odbiorcy, jeżeli ten ma świadomość braku podstawy prawnej swojego wzbogacenia. Zgodnie z *Gegenleistungskondiktio*n w wypadku obustronnie wykonanej, lecz nieważnej umowy wzajemnej § 819 ust. 1 BGB uzupełnia *per analogiam* § 818 ust. 3 BGB. Uzasadnienie dla analogii znajduje się – ponownie – w synalagmie funkcjonalnej. Skoro strona otrzymująca rzecz uczestniczy w synalagmatycznej wymianie świadomie i z chwilą faktycznego wykonania umowy przejmuje ryzyko utraty otrzymanej rzeczy, to jej sytuacja jest analogiczna do tej, w której odbiorca świadczenia ma świadomość braku jego podstawy prawnej. Ryzyko obciąża zatem stronę otrzymującą rzecz, podobnie jak w § 446 BGB (będącym odpowiednikiem polskiego art. 548 § 1 k.c.) ryzyko utraty rzeczy sprzedanej przechodzi na nabywcę z chwilą jej wydania. W konsekwencji, w przypadku świadczeń z nieważnej umowy wzajemnej § 818 ust. 3 BGB znajdzie zastosowanie tylko wówczas, gdy utrata otrzymanego świadczenia jest *rein zufällig* – czyli czysto przypadkowa i nieprzypisywalna subiektywnie ujętemu ryzyku strony (np. wskutek powodzi czy kradzieży przez nieznanego sprawcę przy zachowaniu należytej staranności)<sup>19</sup>.

W jeszcze kolejnym ujęciu na modyfikację teorii salda miała wpływ nowelizacja BGB z 2002 r., zwana *Schuldrechtsmodernisierungsgesetz*<sup>20</sup>. Nie dotyczyła ona jednak przepisów o bezpodstawnym wzbogaceniu, lecz o od-

<sup>18</sup> W. Flume, *Der Wegfall der Bereicherung in der Entwicklung vom römischen zum geltenden Recht* [w:] *Festschrift für Hans Niedermeyer zum 70. Geburtstag*, Göttingen 1953, ss. 103–176; tenże, *Die Saldotheorie und die Rechtsfigur der ungerechtfertigten Bereicherung*, 'Archiv für die civilistische Praxis' 1994, tom 194, ss. 427–450.

<sup>19</sup> C.-W. Canaris, *Die Gegenleistungskondiktio*n [w:] *Festschrift für Werner Lorenz zum siebzigsten Geburtstag*, red. B. Pfister, M.R. Will, Tübingen 1991, ss. 19–63; tenże, *Gesammelte Schriften*, red. H.C. Grigoleit, J. Neuner, tom 3, Berlin–Boston 2012, ss. 819–866.

<sup>20</sup> Gesetz zur Modernisierung des Schuldrechts vom 26. November 2001, Bundesgesetzblatt 2001, cz. 1, s. 3138.

stąpieniu od umowy (§ 346–354 BGB). Reżim odstąpienia obejmuje każdą umowę, w której strona ma prawo odstąpienia (umowne lub ustawowe). Odnosi się więc również do umów wzajemnych, dla których ustawodawca wprowadził ustawowe przesłanki odstąpienia w § 323–326 BGB. Objęcie wszystkich typów umów reżimem odstąpienia od umowy wiąże się ze skutkiem odstąpienia *ex nunc*, który dotyka synalagmy funkcjonalnej, ale nie powoduje upadku synalagmy genetycznej (to powoduje skutek *ex tunc* występujący przez nieważność umowy). Wydaje się, że to, iż obowiązek rozliczenia przywracającego stan sprzed wystąpienia synalagmy funkcjonalnej występuje zarówno w sytuacji odstąpienia od umowy, jak i nieważności umowy, stał się przyczyną wysunięcia postulatu dążenia do ujednoczenia zasad rozliczeniowych<sup>21</sup>. Na gruncie przepisów o odstąpieniu od umowy funkcjonuje teoria dwóch kondykcji (§ 348 zd. 1 BGB). Część piśmiennictwa postulowała analogię z przepisów o odstąpieniu od umowy i roszczeń kondykcyjnych do sytuacji bezpodstawnego wzbogacenia z powodu nieważnej umowy wzajemnej<sup>22</sup>. Jednakże orzecznictwo BGH również po 2002 roku stosuje teoria salda<sup>23</sup>.

Świadectwem trwania klasycznej teorii salda w orzecznictwie BGH także po nowelizacji z 2002 r. jest wyrok z 21 lutego 2022 r.<sup>24</sup> – wydany na kanwie tzw. afery dieslowskiej koncernu Volkswagen. Sprawa dotyczyła nabywcy samochodu VW Golf z silnikiem typu EA 189, w którym zainstalowano oprogramowanie obniżające emisję tlenków azotu wyłącznie w warunkach testów homologacyjnych. Roszczenie odszkodowawcze (deliktowe) nabywcy z § 826 BGB uległo przedawnieniu po upływie trzech lat (§ 195 BGB). W związku z tym BGH dopuścił dochodzenie roszczenia rezydualnego z § 852 zd. 1 BGB (*deliktischer Bereicherungsanspruch*), czyli roszczenia deliktowego ograniczonego do wysokości wzbogacenia sprawcy. Roszczenie to przedawnia się z upływem dziesięciu lat od jego powstania, a niezależnie od powstania – z upływem trzydziestu lat od dokonania czynu albo od innego zdarzenia wywołującego szkodę (§ 852 zd. 2 BGB). BGH odmówił przy tym zastosowania teorii salda, ponieważ sama umowa sprzedaży była ważna, a roszczenie wywodziło się z czynu niedozwolonego, a nie z bezpodstawnego wzbogacenia. Tym samym potwierdził, że właściwym zakresem zastosowania teorii salda pozostaje rozliczenie świadczeń z nieważnej umowy wzajemnej. Mimo wyłączenia teorii salda z rozstrzygnięcia – BGH wyraźnie potwierdził jej dogmatyczne ugruntowanie, stwierdzając, że: *Saldotheorie* nie wyznacza

<sup>21</sup> Uzasadnienie projektu ustawy (*Begründung des Gesetzentwurfs*), BT-Drucks nr 14/6040 z 14 maja 2001 r., s. 194; <https://dsserver.bundestag.de/btd/14/060/1406040.pdf> [dostęp: 29.04.2026].

<sup>22</sup> F. Bockholdt, *Anwendung der Wertungen des Rücktrittsfolgenrechts auf den bereicherungsrechtlichen Rückgewähranspruch?*, 'Archiv für die civilistische Praxis' 2006, tom 206, s. 769 i nn. Inaczej: W. Flume, *Die Saldotheorie...*, ss. 444–446.

<sup>23</sup> Wyrok BGH z 2 grudnia 2004 r., IX ZR 200/03, BGHZ tom 161, s. 241.

<sup>24</sup> Wyrok BGH z 21 lutego 2022 r., VIa ZR 8/21, BGHZ tom 233, s. 16.

przedmiotu wydania w rozumieniu § 812 ust. 1 i § 818 ust. 1 BGB, lecz stanowi konsekwentne zastosowanie myśli wyrażonej w § 818 ust. 3 BGB<sup>25</sup>.

## Regulacja teorii salda w prawie gruzińskim

Gruzińskie prawo cywilne jest jedną z najmłodszych kodyfikacji postsojalistycznych w Europie. Jego struktura i konstrukcje pojęciowe nawiązują w znacznej mierze do prawa niemieckiego. Wynika to z niemieckiego wsparcia eksperckiego przy pracach kodyfikacyjnych w pierwszej połowie lat dziewięćdziesiątych XX wieku.<sup>26</sup> Już w samej systematyce regulacji bezpodstawnego wzbogacenia (art. 976–991 GZGB<sup>27</sup>, rozdział უსაფუძვლო გამდიდრება — *Usapudzwlo gamdidreba*) widoczne jest pokrewieństwo z niemiecką konstrukcją § 812–822 BGB. Jednakże GZGB zawiera *expressis verbis* regulację, która stanowi funkcjonalny odpowiednik niemieckiej *Saldotheorie* (art. 979 GZGB); w prawie gruzińskim natomiast nie obowiązuje zasada abstrakcyjności (*Abstraktionsprinzip*) przysporzenia, która uzasadnia w prawie niemieckim zastosowanie teorii salda. Zgodnie z art. 186 ust. 1 GZGB: dla skuteczności czynności rozporządzającej wymagana jest ważne zobowiązanie (*causa*)<sup>28</sup>. Powoduje to, że w przypadku nieważnej umowy wzajemnej własność w ogóle nie przechodzi na nabywcę. Brak *causa* powoduje równoczesny upadek czynności przysparzającej. Zbywca zachowuje status właściciela i przysługują mu jednocześnie dwa roszczenia, roszczenie windykacyjne (art. 172 GZGB) oraz roszczenie kondykcyjne (art. 976 GZGB).

Przypomnieć w tym miejscu należy, że w prawie niemieckim teoria salda dotyczy wyłączenia zastosowania § 818 ust. 3 BGB, który pozwala obowiązanemu na odmowę zadośćuczynienia roszczeniu kondykcijnemu z powołaniem się na odpadnięcie wzbogacenia (np. wskutek zniszczenia, zgubienia lub utraty wartości rzeczy). Odpowiednikiem § 818 ust. 3 BGB w prawie gruzińskim jest art. 979 ust. 3 GZGB, zgodnie z którym: [...] obowiązek zwrotu wartości od-

<sup>25</sup> Die Saldotheorie bestimmt nicht den nach § 812 Abs. 1, § 818 Abs. 1 BGB herauszugebenden Gegenstand, sondern ist die folgerichtige Anwendung des in § 818 Abs. 3 BGB zum Ausdruck kommenden Rechtsgedankens einer dadurch eingetretenen Bereicherung.

<sup>26</sup> Zob. L. Chanturia, *Die Europäisierung des georgischen Rechts – bloßer Wunsch oder große Herausforderung?*, 'RabelsZ' 2010, 74, 1, ss. 154–181; I. Burduli, *Reception, As Milestone for Creation of Legal System (on the Example of Corporate Law of Post-Soviet Georgia)*, „Prawo i Więź” 2021, 2, ss. 112–140; J. Basedow, *Georgien und die Europäisierung des Privatrechts* [w:] *Transformation durch Rezeption? Möglichkeiten und Grenzen des Rechtstransfers am Beispiel der Zivilrechtsreformen im Kaukasus und in Zentralasien*, red. E. Kurzynsky-Singer, Tübingen 2014, ss. 457–463, a zwłaszcza s. 148.

<sup>27</sup> საქართველოს სამოქალაქო კოდექსი (*Sakartvelos Samokaladzo Kodeksi*), uchwalony 26 czerwca 1997 r., wszedł w życie 25 listopada 1997 r., <https://matsne.gov.ge> [dostęp: 3.05.2026].

<sup>28</sup> G. Rusiashvili, *Die Saldotheorie im georgischen Recht?* [w:] *Rechtsvergleichung und Privatrecht im deutsch-georgischen Diskurs*, red. T.J. Chiussi, I. Burduli, Saarbrücken 2014, s. 63.

pada, jeżeli odbiorca wskutek zużycia, dalszego zbycia, zniszczenia, pogorszenia lub z innych przyczyn nie jest już wzbogacony ani co do przedmiotu, ani co do jego wartości<sup>29</sup>. Natomiast dla modyfikacji teorii salda w prawie gruzińskim istotna jest treść art. 979 ust. 4 GZGB, wedle którego: Jeżeli strony umowy wzajemnej mają obowiązek zwrotu świadczenia uzyskanego na jej podstawie z powodu nieważności umowy, a jedna ze stron z przyczyn wymienionych w ust. 2 tego artykułu nie może wydać świadczenia, to nie jest ona zobowiązana do zwrotu, o ile wynika to z istoty normy prawnej, na podstawie której umowa została uznana za nieważną. Z kolei art. 979 ust. 2 GZGB stanowi, że: Jeżeli zwrot uzyskanego świadczenia nie jest możliwy ze względu na właściwość przedmiotu lub odbiorca z innego powodu nie jest w stanie go zwrócić, zobowiązany jest do zwrotu jego wartości rynkowej.

Artykuł 979 ust. 4 GZGB wyłącza zastosowanie art. 979 ust. 3 GZGB. W typowej kondykcji bezpodstawnie wzbogacony może powołać się na utratę przedmiotu wzbogacenia (zużycie, zniszczenie, dalsze zbycie itd.) i dzięki temu uchylić się od obowiązku zwrotu wartości (art. 979 ust. 3 GZGB). Jednak gdy podstawą roszczenia kondykcyjnego jest nieważność umowy wzajemnej, ten zarzut zostaje wyłączony. W konsekwencji odbiorca pozostaje zobowiązany do zwrotu wartości na podstawie art. 979 ust. 2 GZGB mimo odpadnięcia wzbogacenia, chyba że taki rezultat sprzeciwiałby się istocie normy prawnej, na podstawie której umowa została uznana za nieważną (art. 979 ust. 4 GZGB *in fine*). Innymi słowy: odbiorca, który nie może wydać świadczenia *in natura* (art. 979 ust. 2 GZGB) i którego wzbogacenie odpadło (art. 979 ust. 3 GZGB), w typowej kondykcji byłby zwolniony z obowiązku zwrotu; w nieważnej umowie wzajemnej powinien natomiast zwrócić wartość świadczenia. Ta regulacja jest ustawowym odpowiednikiem niemieckiej, orzecznictwej *Saldotheorie*. Również w tym przypadku uzasadnieniem kontynuującym obowiązek strony, u której wzbogacenie odpadło, jest powołanie się na synalagmatyczny charakter umowy wzajemnej. Regulacja art. 979 ust. 4 GZGB opiera się na pojęciu faktycznej synalagmy (*faktisches Synallagma*)<sup>30</sup>.

Jednak współistnienie po stronie właściciela roszczenia kondykcyjnego i windykacyjnego rodzi dwa problemy w realizacji teorii salda. Po pierwsze: wykonanie roszczenia windykacyjnego nie jest sprzężone z wykonaniem roszczenia kondykcyjnego drugiej strony (natomiast art. 980 ust. 2 GZGB ustanawia obowiązek jednoczesnego zadośćuczynienia roszczeniom kondykcyjnym). Druga strona mogłaby wówczas dochodzić swojego roszczenia kondykcyjnego dopiero po wydaniu rzeczy właścicielowi. Po drugie:

<sup>29</sup> Tłumaczenie autorskie na podstawie przekładu niemieckiego: G. Rusiashvili, *Die Saldotheorie...*, ss. 75 i 78, oraz angielskiego: G. Rusiashvili, *Competition of Vindication and Execution Claim in Reversing of the Outcomes of Invalid Deal*, 'Journal of Law' (Ivane Javakhishvili Tbilisi State University, Faculty of Law) 2013, 1, ss. 194 i 196.

<sup>30</sup> G. Rusiashvili, *Die Saldotheorie...*, ss. 65, 75–77 i 79.

w reżimie przepisów o ochronie własności obowiązany do zwrotu rzeczy jest chroniony przez art. 163 GZGB, na podstawie którego nie odpowiada za zniszczenie ani nawet zużycie rzeczy bez winy. To sprawia, że może dochodzić swojego roszczenia kondykcyjnego w pierwotnej wartości, nieograniczonej do salda. Te problemy istnieją jednak tylko wtedy, gdy przyjmie się pierwszeństwo roszczenia windykacyjnego przed roszczeniem kondykcyjnym. Dlatego w piśmiennictwie gruzińskim postuluje się pierwszeństwo roszczenia kondykcyjnego wynikającego z nieważności umowy wzajemnej przed roszczeniem windykacyjnym. Teza ta nie jest jednak przyjmowana jednolicie<sup>31</sup>.

---

## Wnioski – nieadekwatność teorii salda do prawa polskiego

---

Teoria salda nie zastała przyjęta przez polskie piśmiennictwo dotyczące bezpodstawnego wzbogacenia czy – precyzyjniej rzecz ujmując – nienależnego świadczenia (art. 410 k.c.). Ostatnio pojawiła się jako argument do przeprowadzania rozliczeń w konsekwencji stwierdzenia nieważności umów indeksowanych lub denominowanych do waluty obcej, nie znalazła jednak szerokiego uznania w orzeczeniach sądów powszechnych<sup>32</sup>. W starszym natomiast orzecznictwie Sąd Najwyższy ustalał wysokość roszczeń kondykcyjnych w modelu właściwym dla teorii salda, tj. przez zaliczenie na poczet roszczenia wartości zużytej<sup>33</sup> albo przez waloryzację wartości wzbogacenia<sup>34</sup>. Jednak nigdy nie zostało wydane orzeczenie, według którego w wyniku rozliczenia nadwyżki aktywów nad pasywami roszczenia kondykcyjne przekształciły się w jedno roszczenie o saldo.

W najnowszym orzecznictwie Sąd Najwyższy podtrzymuje, że stronom nieważnej umowy przysługują osobne roszczenia kondykcyjne (teoria dwóch kondykcji). Linia orzecznicza Sądu Najwyższego – ukształtowana w uchwale siedmiu sędziów z 7 maja 2021 r.<sup>35</sup>, uchwale pełnego składu Izby Cywilnej z 25 kwietnia 2024 r.<sup>36</sup> oraz wyroku z 5 września 2025 r.<sup>37</sup> – konsekwentnie odrzuca recepcję teorii salda w polskim prawie cywilnym. Brakuje więc uzasadnienia w wykładni operatywnej dla przyjęcia na gruncie prawa polskiego teorii sal-

<sup>31</sup> G. Rusiashvili, *Die Saldotheorie...*, ss. 79–86; G. Rusiashvili, *Competition...*, ss. 199–203.

<sup>32</sup> Zob. wyrok SA w Białymstoku z 20 lutego 2020 r., I ACa 635/19, Legalis nr 2292520.

<sup>33</sup> Wyrok SN z 12 czerwca 1975 r., III CRN 479/74, Legalis nr 18822; wyrok SN z 24 października 1974 r., II CR 542/74, Legalis nr 18353.

<sup>34</sup> Wyrok SN z 13 maja 1988 r., III CRN 83/88, Legalis nr 26268.

<sup>35</sup> Uchwała składu siedmiu sędziów SN z 7 maja 2021 r., III CZP 6/21, OSNC 2021, nr 9, poz. 56 (zasada prawna).

<sup>36</sup> Uchwała pełnego składu Izby Cywilnej SN z 25 kwietnia 2024 r., III CZP 25/22, OSNC 2024, nr 12, poz. 118.

<sup>37</sup> Wyrok SN z 5 września 2025 r., II CSKP 550/24, Legalis nr 3272509, w którym wyjaśniono, że teoria salda nie powinna być przyjmowana w związku z wyrokiem TSUE z 19 czerwca 2025 r., C-396/24 (Lubreczlik), Dz.U. UE C 2025, 4253.

da. Oczywiście nie istnieje także podstawa normatywna, taka jak art. 979 ust. 4 GZGB w prawie gruzińskim.

Problem jednak polega na tym, że teoria salda znajduje w prawie niemieckim i gruzińskim odmienne uzasadnienie niż przypisywane jej w polskim orzecznictwie. W polskim orzecznictwie teoria salda rozpatrywana jest w kontekście wartości zapłaty dokonanej przez kredytobiorcę na podstawie nieważnej umowy kredytu. Tymczasem w prawie niemieckim i gruzińskim teoria salda ma chronić stronę umowy wzajemnej, która świadczyła rzeczowo, przed obowiązkiem zwrotu całego otrzymanego świadczenia – zapłaconej ceny, gdy rzecz uległa zniszczeniu lub uszkodzeniu. Uzasadnienie to nie jest adekwatne do świadczeń stron nieważnej umowy kredytu, które w przypadku obydwu stron mają charakter pieniężny<sup>38</sup>. Pieniądz, jako rzecz oznaczona co do gatunku, ze swej istoty nie ulega zniszczeniu w sensie, do którego odwołuje się niemiecka *Saldotheorie*. Nie powstaje zatem strukturalny problem dysproporcji ryzyka, który ta teoria miałaby rozwiązywać. Co więcej, niemiecka doktryna doprecyzowuje, że *Saldotheorie* w jej działaniu *ipso iure* zakłada określoną strukturę świadczeń. Obowiązek jednej ze stron zwrotu rzeczy przekształca się w obowiązek pieniężny, gdy przedmiot nienależnego świadczenia nie istnieje już w stanie umożliwiającym zwrot *in natura*<sup>39</sup>. Świadczenia stron nieważnej umowy kredytu nie odpowiadają tej strukturze, gdyż po obu stronach świadczenie nienależne ma od początku charakter pieniężny.

W prawie polskim nie ma też warunku jednoczesnego zadośćuczynienia roszczeniom kondykcijnym (czego w prawie gruzińskim wymaga art. 980 ust. 2 GZGB), dlatego nie występuje strukturalny problem realizacji roszczenia windykacyjnego z pominięciem takiego warunku (pomimo że polskie prawo cywilne, tak jak gruzińskie, przyjmuje zasadę kauzalności przysporzenia – art. 156 k.c.).

### Abstrakt

Rozliczenie wzajemnych świadczeń wynikających z umów, które zostały następnie uznane za nieważne, stanowi jeden z najtrwalszych problemów doktrynalnych w prawie dotyczącym bezpodstawnego wzbogacenia. Kwestia ta zyskała ponowne znaczenie praktyczne w Polsce w konsekwencji orzeczeń uznających za nieważne umowy kredytowe indeksowane do waluty obcej oraz denominowane w walucie obcej, a także w świetle rozbieżnego orzecznictwa, które pojawiło się w ich następstwie. W tym kontekście

<sup>38</sup> Pomimo że umowa kredytu ma charakter wzajemny – uchwała składu siedmiu sędziów SN z 28 lutego 2025 r., III CZP 126/22, OSNC 2025, nr 11, poz. 97.

<sup>39</sup> H. Prütting, *Bürgerliches Gesetzbuch mit Internationalem Privatrecht und nationalen, unionsrechtlichen sowie völkerrechtlichen Neben- oder Ergänzungsgesetzen. Kommentar*, s. 1900.

uwagę przyciąga możliwość zastosowania teorii salda (niem. *Saldotheorie*). Niniejsze opracowanie zawiera prawnoporównawczą analizę podstaw doktrynalnych teorii salda w niemieckim orzecznictwie oraz jej ustawowego odpowiednika w gruzińskim kodeksie cywilnym, a także ocenia zgodność tej teorii z założeniami strukturalnymi polskiego systemu regulującego zwrot świadczeń wykonanych bez tytułu prawnego. Stwierdza się, że uzasadnienie leżące u podstaw teorii salda – a mianowicie złagodzenie asymetrycznego rozkładu ryzyka, które powstaje w umowach dotyczących świadczeń niepieniężnych – nie ma zastosowania w kontekście nieważnych umów pożyczki, w których świadczenia obu stron mają charakter pieniężny. Przyjęcie teorii salda w polskim orzecznictwie dotyczącym nieważnych umów pożyczki jest zatem zarówno strukturalnie nieadekwatne, jak i funkcjonalnie zbędne.

**Słowa kluczowe:** teoria salda, zwrot świadczeń spełnionych bez podstawy prawnej, nieważne umowy kredytu, prawnoporównawcze ujęcie bezpodstawnego wzbogacenia.

---

## BIBLIOGRAFIA

---

- Bockholdt F., *Anwendung der Wertungen des Rücktrittsfolgenrechts auf den bereicherungsrechtlichen Rückgewähranspruch?*, 'Archiv für die civilistische Praxis' 2006, 206.
- Burduli I., *Reception, As Milestone for Creation of Legal System (on the Example of Corporate Law of Post-Soviet Georgia)*, „Prawo i Więź” 2021, 2.
- Canaris C.-W., *Gesammelte Schriften*, red. H.C. Grigoleit, J. Neuner, tom 3, Berlin–Boston 2012.
- Chanturia L., *Die Europäisierung des georgischen Rechts – bloßer Wunsch oder große Herausforderung?*, 'RabelsZ' 2010, 74, 1.
- Ennecerus L. (H. Lehmann), *Lehrbuch des Bürgerlichen Rechts*, tom 2, *Recht der Schuldverhältnisse. Ein Lehrbuch*, Tübingen 1958.
- Festschrift für Ernst Rabel*, tom 1, Tübingen 1954.
- Festschrift für Hans Niedermeyer zum 70. Geburtstag*, Göttingen 1953.
- Festschrift für Werner Lorenz zum siebzigsten Geburtstag*, red. B. Pfister, M.R. Will, Tübingen 1991.
- Flume W., *Die Saldotheorie und die Rechtsfigur der ungerechtfertigten Bereicherung*, 'Archiv für die civilistische Praxis' 1994, 194.
- Flume W., *Die Rückabwicklung nichtiger Kaufverträge nach Bereicherungsrecht — Zur Saldotheorie und ihren „Ausnahmen“*, 'JuristenZeitung' 2002.
- Larenz K., C.-W. Canaris, *Lehrbuch des Schuldrechts*, tom 2, cz. 2 *szczególna*, München 1994.

- Leser H., *Von der Saldotheorie zum faktischen Synallagma*, Freiburg im Breisgau 1956.
- Prütting H., Wegen G., Weinreich G. (red.), *Bürgerliches Gesetzbuch mit Internationalem Privatrecht und nationalen, unionsrechtlichen sowie völkerrechtlichen Neben- oder Ergänzungsgesetzen. Kommentar*, München 2024.
- Rechtsvergleichung und Privatrecht im deutsch-georgischen Diskurs, red. T.J. Chiusi, I. Burduli, Saarbrücken 2014.
- Rusiashvili G., *Competition of Vindication and Execution Claim in Reversing of the Outcomes of Invalid Deal*, 'Journal of Law' (Ivane Javakhishvili Tbilisi State University, Faculty of Law) 2013, 1.
- Transformation durch Rezeption? Möglichkeiten und Grenzen des Rechtstransfers am Beispiel der Zivilrechtsreformen im Kaukasus und in Zentralasien, red. E. Kurzynsky-Singer, Tübingen 2014.
- Uzasadnienie projektu ustawy (*Begründung des Gesetzentwurfs*), BT-Drucks nr 14/6040 z 14 maja 2001 r., s. 194; <https://dserver.bundestag.de/btd/14/060/1406040.pdf> [dostęp: 29.04.2026].
- Vliet L. van, *Iusta Causa Traditionis and its History in European Private Law*, 'European Review of Private Law' 2003, 3.



## Informacje dla autorów

W Redakcji Europejskiego Przeglądu Prawa i Stosunków Międzynarodowych (nazwa skrócona: EPPiSM) zasady publikowania tekstów są następujące:

1. Przekazując tekst Redakcji, autor przenosi na Wydawcę wyłączne prawo do jego publikacji.
2. Zgoda autora na publikację w Europejskim Przeglądzie Prawa i Stosunków Międzynarodowych obejmuje wersje: elektroniczną (e-publicacja) oraz drukowaną (papierową).
3. Wersję elektroniczną Europejskiego Przeglądu Prawa i Stosunków Międzynarodowych uważa się za podstawową.
4. Tekst zgłoszony do publikacji w Europejskim Przeglądzie Prawa i Stosunków Międzynarodowych nie może być wcześniej opublikowany w innym czasopiśmie (w wersji elektronicznej lub drukowanej) ani rozpatrywany pod kątem publikacji.
5. Teksty przekazuje się drogą elektroniczną (email: ewspa.kwartalnik.eppism@gmail.com) albo pocztową (adres: 00-876 Warszawa, ul. Ogrodowa 58, budynek B, 2. piętro) – w drugim przypadku z zapisaną na odczytywalnym nośniku wersją elektroniczną oraz z papierowym, jednostronnym wydrukiem.
6. Każdy tekst poddawany jest wstępnej ocenie Redakcji, która może go odrzucić, w szczególności gdy nie zachowano podstawowych reguł naukowej lub językowej poprawności, a także gdy naruszono choćby jedną z zasad publikacji właściwych dla EPPiSM, w tym zwłaszcza zasadę autorskiej rzetelności (np. *ghostwriting*, *guest authorship*). Tekst, który nie został odrzucony, poddawany jest ocenie merytorycznej co najmniej dwóch niezależnych recenzentów w procedurze anonimowej (*double-blind*). Przygotowują oni na piśmie recenzje zawierające uwagi i wskazówki dla autora oraz jednoznaczny wniosek w kwestii dopuszczenia tekstu do publikacji. Na tej podstawie Redakcja podejmuje decyzję: a) o przyjęciu go do publikacji; b) o warunkowym przyjęciu do publikacji – po uwzględnieniu przez autora wskazanych przez recenzentów zmian; c) o odmowie przyjęcia do publikacji.
7. Redakcja nie ujawnia nazwisk recenzentów poszczególnych materiałów. Co najmniej raz w roku publikuje się listę recenzentów na stronie internetowej kwartalnika.

8. Minimalna objętość nadsyłanych do EPPiSM artykułów naukowych i glos wynosi 20 000 znaków (tu i dalej liczonych ze spacjami i przypisami).
9. Maksymalna objętość artykułu naukowego nie powinna być większa niż 60 000 znaków. W razie przekroczenia – zwłaszcza znaczącego – wskazanego limitu Redakcja może zaproponować autorowi podział tekstu na części. Objętość glosy nie powinna przekraczać 40 000 znaków.
10. W artykule i glosie znajdują się dodatkowo: tytuł w języku angielskim, abstrakt w językach polskim i angielskim (o orientacyjnej objętości 500–2500 znaków każdy), słowa kluczowe w językach polskim i angielskim (od trzech do dziewięciu), a także bibliografia załącznikowa.
11. Autor dołączy ponadto: podstawowe informacje o sobie (status naukowy, zawodowy, afiliacja, numer ORCID, adres poczty elektronicznej, telefon itp.) oraz oświadczenie, że tekst nie został dotychczas nigdzie opublikowany, a Redakcja EPPiSM nabywa na wyłączność prawo do jego publikacji. W wypadku publikacji wieloautorskich podanie informacji o sobie i obowiązek złożenia oświadczenia dotyczy wszystkich autorów.
12. Artykuł naukowy prócz abstraktu powinien zawierać wyodrębniony, najlepiej kilkudzaniowy wstęp (lid). Abstrakt należy sporządzić z użyciem słów kluczowych (każde słowo kluczowe winno być w nim użyte). Abstrakt dotyczy treści artykułu, czyli jest streszczeniem – i jako taki powinien zawierać zwroty typu: „Celem artykułu jest...” albo „Przedmiotem artykułu jest”, a także przedstawić syntetyczne wyjaśnienie, jak ten cel został zrealizowany, np. „Celem artykułu jest odpowiedź na pytanie...”, „Autor udzielił odpowiedzi pozytywnej / negatywnej” – jest to *minimum minimorum* abstraktu. Wstęp odnosi się natomiast do tematyki poruszonej w artykule, czyli jest syntetycznym do niej wprowadzeniem. Wstęp nie powinien (nawet we fragmencie) powielać ani parafrazować abstraktu.
13. Nieodłączną częścią każdego artykułu naukowego i glosy są końcowe wnioski (konkluzje).
14. W razie przywoływania w tekście publikacji elektronicznych oprócz autora i tytułu należy podać adres strony internetowej wraz z datą dostępu.
15. Orzecznictwo Trybunału Sprawiedliwości Unii Europejskiej oraz Europejskiego Trybunału Praw Człowieka przytacza się z obligatoryjnym wskazaniem jednostki redakcyjnej danego orzeczenia – czyli punktu (TSUE) lub paragrafu (ETPC) – przykładowo w następujący sposób:

<sup>1</sup> Postanowienie TS z 12 lutego 2019 r., RH, C8/19 PPU, EU:C:2019:110, pkt 60.

<sup>2</sup> Wyrok Sądu UE z 15 grudnia 2010 r., E.ONEnergie/Komisja, T141/08, EU:T:2010:516, pkt 56 i przywołane tam orzecznictwo.

<sup>3</sup> Wyrok ETPC z 25 marca 1998 r., Belziuk v. Polska, skarga nr 23103/93, § 37. W kolejnych przytoczeniach tego samego orzeczenia pomija się sygnaturę ECL (TSUE) albo numer skargi (ETPC).

16. Orzecznictwo krajowe przykładowo przytacza się następująco:

<sup>1</sup> Wyrok SN z 12 marca 2008 r., I CSK 430/07, OSNC 2009, nr 5, poz. 75.

<sup>2</sup> Postanowienie TK z 27 września 2005 r., U 2/05, OTK ZU-A 2005, nr 8, poz. 96, cz. II, pkt 2.

<sup>3</sup> Wyrok TK z 3 grudnia 2015 r., K 34/15, OTK ZU-A 2015, nr 11, poz. 185, cz. III, pkt 6.12.

<sup>4</sup> Wyrok NSA z 24 października 2000 r., V SA 613/00, OSP 2001, nr 5, poz. 82.

<sup>5</sup> Wyrok SA w Krakowie z 23 kwietnia 1998 r., II AKa 48/98, LEX nr 35155.

<sup>6</sup> Postanowienie SA w Łodzi z 15 lutego 2023 roku, II AKz 74/23 (niepublikowane).

W kolejnych przytoczeniach tego samego orzeczenia krajowego pomija się miejsce publikacji lub adnotację o jej braku.

17. Po przyjęciu tekstu do publikacji Redakcji przysługuje prawo opracowania edytorskiego, typograficznego, językowego itp.

18. Za publikację w Europejskim Przeglądzie Prawa i Stosunków Międzynarodowych nie wypłaca się honorariów autorskich ani nie pobiera się opłat (APC).

**Wydawca EPPiSM**



## Information for authors

The Editorial Board of the European Review of Law & International Relations [Europejski Przegląd Prawa i Stosunków Międzynarodowych] (abbreviated name: ERL&IR [EPPiSM]) has the following rules for publishing texts:

- (1) By submitting the text to the Editorial Board, the author transfers the exclusive right to publish it to the Publisher.
- (2) The author's consent to publication in the European Review of Law & International Relations covers both electronic (e-publication) and printed (paper) versions.
- (3) The electronic version of the European Review of Law & International Relations is considered the primary version.
- (4) A text submitted for publication in the European Review of Law & International Relations may not have been previously published in another journal (in electronic and/or printed form) or considered for publication.
- (5) Texts shall be submitted electronically (email: ewspa.kwartalnik.eppism@gmail.com) or by post (address: 00-876 Warsaw, 58 Ogrodowa St., Building B, 2nd floor) – in the latter case, with an electronic version saved on a readable medium and a single-sided paper printout.
- (6) Each text is subject to a preliminary assessment by the Editorial Board, which may reject it, in particular if the basic rules of scholarly or linguistic correctness have not been observed, or if even one of the publication rules specific to ERL&IR has been violated, including in particular the principle of authorial integrity (e.g. ghostwriting, guest authorship). Texts that are not rejected are subject to substantive evaluation by at least two independent reviewers in an anonymous (double-blind) procedure. They prepare written reviews containing comments and suggestions for the author and an unambiguous conclusion on whether the text should be accepted for publication. On this basis, the Editorial Board decides: (a) to accept it for publication; (b) to accept it for publication conditionally – after the author has taken into account the changes indicated by the reviewers; (c) to refuse to accept it for publication.

- (7) The Editorial Board does not disclose the names of reviewers of individual materials. At least once a year, a list of reviewers is published on the quarterly's website.
- (8) The minimum length of scholarly articles and glossaries submitted to ERL&IR is 20,000 characters (hereinafter counted with spaces and footnotes).
- (9) The maximum length of a scholarly article should not exceed 60,000 characters. If this limit is exceeded, especially significantly, the Editorial Board may suggest that the author divide the text into parts.  
The length of a commentary (glossary) should not exceed 40,000 characters.
- (10) The article and glossary will also include: a title in English, an abstract in Polish and English (approximately 500–2500 characters each), keywords in Polish and English (three to nine), and an appendix bibliography.
- (11) The author shall also attach: basic information about themselves (academic and professional status, affiliation, ORCID number, email address, telephone number, etc.) and a statement that the text has not been published anywhere else and that the ERL&IR Editorial Board acquires the exclusive right to publish it. In the case of multi-author publications, the obligation to provide information about oneself and to submit a statement applies to all authors.
- (12) In addition to the abstract, the scholarly article should contain a separate introduction (lead), preferably consisting of several sentences. The abstract should be prepared using keywords (each keyword should be used in it). The abstract concerns the content of the article, i.e. it is a summary – and as such it should contain phrases such as: ‘The aim of the article is...’ or ‘The subject of the article is’, and also provide a concise explanation of how this aim was achieved, e.g. ‘The aim of the article is to answer the question...’, ‘The author gave a positive/negative answer’ – this is the *minimum minimorum* of an abstract. The introduction, on the other hand, refers to the subject matter of the article, i.e. it is a synthetic introduction to it. The introduction should not (even in part) duplicate or paraphrase the abstract.
- (13) Final conclusions are an integral part of every scholarly article and commentary (glossary).
- (14) When referring to electronic publications in the text, in addition to the author and title, the website address and date of access should be provided.
- (15) Case law of the Court of Justice of the European Union and of the European Court of Human Rights shall be cited with the obligatory indication of the

editorial unit of the decision in question – i.e. point (CJEU) or paragraph (ECHR) – for example as follows:

<sup>1</sup> Order of the CJ of 12 February 2019, RH, C 8/19 PPU, EU:C:2019:110, para 60.

<sup>2</sup> Judgment of the EU General Court of 15 December 2010, E.ON Energie/ Commission, T 141/08, EU:T:2010:516, para 56 and case law cited therein.

<sup>3</sup> ECtHR judgment of 25 March 1998, *Belziuk v. Poland*, Application no. 23103/93, § 37. Subsequent citations of the same judgment omit either the ECL (CJEU) or the complaint number (ECtHR).

(16) National case law is cited by way of example as follows:

<sup>1</sup> Judgment of the Supreme Court of 12 March 2008, I CSK 430/07, OSNC 2009, no. 5, item 75.

<sup>2</sup> Order of the Constitutional Court of 27 September 2005, U 2/05, OTK ZU-A 2005, no. 8, item 96, part II, point 2.

<sup>3</sup> Judgment of the Constitutional Court of 3 December 2015, K 34/15, OTK ZU-A 2015, no. 11, item 185, part III, point 6.12.

<sup>4</sup> Judgment of the Supreme Administrative Court of 24 October 2000, V SA 613/00, OSP 2001, no. 5, item 82.

<sup>5</sup> Judgment of the Court of Appeal in Kraków of 23 April 1998, II AKa 48/98, LEX no 35155.

<sup>6</sup> Decision of the Court of Appeal in Łódź of 15 February 2023, II AKz 74/23 (unpublished). Subsequent citation of the same national judgment shall omit the place of publication or a note of its absence.

(17) After accepting a text for publication, the Editorial Board reserves the right to edit it in terms of layout, typography, language, etc.

(18) No royalties are paid for publication in the *European Review of Law & International Relations*, nor are any fees charged (APC).

**Publisher of the ERL&IR [EPPiSM]**

