European Review of Digital Administration & Law - Erdal
2024, Volume 5, Issue 1, pp. 51-77
ISSN 2724-5969 – ISBN 979-12-218-1507-8 – DOI 10.53136/9791221815078 05

*Justice and Generative AI: The Constitutional Challenges**

Erik Longo
(Full professor of Public Law, University of Florence, Department of Legal Studies)

ABSTRACT This article investigates some of the multiple intersections between artificial intelligence (AI) and law by focusing on the implications of using the latest forms of generative AI in judicial processes. After introducing the sub-category of Large Language Models, the author examines the amateur use of ChatGPT in some recent cases, highlighting the potential benefits, such as improved efficiency and effectiveness of decisions, and the risks concerning transparency, personal-data protection, judicial independence, and the right to a fair trial. The article then explores some institutional documents that govern the development of digital justice in different jurisdictions (Court of Justice of the European Union, European Court of Human Rights, England and Wales Courts) and addresses AI-assisted human decision-making. The last part discusses whether we are at the dawn of a new form of justice, emphasising balancing technological advances with protecting fundamental rights.
"The law is not a machine and the judges not machine-tenders" (J. Frank, Law and the modern mind, London, Transaction Publishers, 1930).

KEYWORDS: Digital justice – AI Act – Generative AI – Predictive Justice – Constitutionalism

TABLE OF CONTENTS: 1. AI and justice. – 2. Applications of AI in the legal field. – 2.1. LLMs. – 2.2. The uses of LLMs in the field of law, with a first impression on them. – 3. Technical problems arising from the use of LLMs in the legal field. – 4. From the Ethical Charter of the CEPEJ to Documents and Guidelines on the Use of GenAI in Judicial Activities. – 5. The AI Act. – 6. Constitutional profiles relating to the use of LLM in the judicial process.

1. *AI and justice*

A significant transformation is underway: the digital revolution is profoundly reshaping our species.[1] Digital devices, particularly smartphones, have first taken over our leisure time and then progressively occupied all the crucial areas of our lives. The recent pandemic has sped up the transition to digital technologies, often driven by the need for greater efficiency. This shift has also led to a widespread integration of Artificial Intelligence (AI) in our daily lives.

We inhabit a world where the propensity to delegate human functions and decision-making processes to algorithmic machines is increasingly prevalent.[2] The governance of big data, artificial intelligence, and algorithms has become an established reality worldwide. From major shifts in macroeconomic policies to machine-generated medical diagnoses, from marketing to transportation, our world has a pronounced tendency to entrust an increasing range of management tasks to machines.[3]

This transformation has not spared even the legal domain. It has been a long time since scholars have started to study the interplay between digital technologies and the ample field of law.[4] This topic has been examined either in a "futuristic" way – e.g. the imagination of a world where robot lawyers

[*] Article submitted to double-blind peer review.

[1] Y.N. Harari, *21 Lessons for the 21st Century*, First edition, New York, Spiegel & Grau, 2018.

[2] The almost religious cult of Big Data and algorithms has even produced a philosophy, "dataism", that proposes a purely digital rationality, inevitably destined to supplant today's philosophies and rationalities, including human justice and political democracy itself. The first to speak about is C. Anderson, *The End Of Theory: The Data Deluge Makes The Scientific Method Obsolete*, in *Wired*, 23 June 2008. Retrieved from: https://www.wired.com/2008/06/pb-theory/. This idea has been propagated by many authors, like J. Van Dijck, *Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology*, in *Surveillance & society*, vol. 12, 2014; Y.N. Harari, *Homo Deus: A brief history of tomorrow*, New York, Random House, 2015.

[3] L. Floridi, *The fourth revolution: How the infosphere is reshaping human reality*, Oxford, Oxford University Press, 2014; M.A. Boden, *Artificial Intelligence: A Very Short Introduction*, Oxford, Oxford University Press, 2018; S. Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, New York, PublicAffairs, 2019; M. Benasayag and R.g. Meyran, *The Tyranny of Algorithms: A Conversation with Régis Meyran*, London, Europa Editions, 2021.

[4] See, L. Loevinger, *Jurimetrics--The Next Step Forward*, in *Minn. L. Rev.*, vol. 33, 1948; J. Weizenbaum, *Computer power and human reason: From judgment to calculation*, New York and San Francisco, W.H. Freeman and Company, 1976; G. Sartor and L.K. Branting (eds.), *Judicial applications of artificial intelligence*, The Netherlands, Kluwer, 1998. For interesting new, but controversial, perspectives see S. Deakin and C. Markou, *From Rule of Law to Legal Singularity*, in *Is Law Computable?: Critical Perspectives on Law and Artificial Intelligence*, S. Deakin and C. Markou (eds.), New York, Bloomsbury Publishing, 2020.

will replace human lawyers – or as a mere technical issue – how to automate the logic beneath the interpretation of legal provisions – until recently.[5] It is only in the last decade that more concrete and mundane analyses have been developed. Scholars and practitioners overcome this subject's mere speculative hypothesis or theoretical development by investigating both the sectoral implications of digital technologies and the consequences (also legal) of the social and economic uses of such tools in the many fields of law. Articles, books, and reports have taken a step forward, considering not just the judicial process but also the activities of the legal profession as suitable areas for the application of AI.[6]

Generally, two views of the intersection between AI systems and the legal field can be identified.[7] The first looks at AI's application to legal activities, such as drafting statutes, interpreting legal texts, applying law to cases, drafting and enforcing contracts, etc. The second examines the legal problems AI systems produce when used in human activities.

Despite the two fields being separate, they feed off each other.[8] Thus, understanding what can be done with AI by a lawyer or a judge depends on the rules concerning the consequences of using AI, and conversely, understanding what legal issues AI produces also depends on clearly defining the meaning of employing AI systems in law. Indeed, before creating technical issues and possibilities of "legal singularity",[9] the use of advanced digital technologies in the legal field raises questions regarding protecting fundamental rights and the legitimacy of technological solutions that tend to capture the decision-makers and transform their activities into mere mathematical or statistical processes.[10]

Within this broad horizon, even the perspective of constitutional law can be helpful.[11] This discipline has the dual purpose of analysing and supporting the protection of rights and defining legitimate limits to the powers of public institutions.[12] In recent years, constitutional law has contributed to the definition of the general contours of the relationship between law and technology, specifically in relation to the constitutional profiles of using Artificial Intelligence.[13] This achievement has helped to understand that the interplay between AI and law is not just a technical matter yet something to be handled with the tools of every legal expert, such as the scholars that study the impact of automated decision-making on freedom and

[5] G. Sartor, M. Araszkiewicz, K. Atkinson, F. Bex, T. Van Engers, E. Francesconi, H. Prakken, G. Sileno, F. Schilder, A. Wyner and T. Bench-Capon, *Thirty years of Artificial Intelligence and Law: the second decade*, in *Artificial Intelligence and Law*, vol. 30, 2022; A. Simoncini, *La dimensione costituzionale della giustizia predittiva. Riflessioni su intelligenza artificiale e processo*, in *Riv. dir. proc.*, vol. 79, 2024.

[6] The scholarship on these issues is abundant and too numerous to mention. However, for a sampling, see T. Sourdin, *Judges, Technology and Artificial Intelligence*, Cheltenham, Edward Elgar, 2021; R.E. Susskind, *Online Courts and the Future of Justice*, Oxford, Oxford University Press, 2019; O. Conetta, *AI in the Legal Profession*, in *The Law of Artificial Intelligence*, M. Hervey and M. Lavy (eds.), London, Sweet & Maxwell, 2021; L.A. DiMatteo, A. Janssen, P. Ortolani, F. de Elizalde, M. Cannarsa and M. Durovic (eds.), *The Cambridge Handbook of Lawyering in the Digital Age*, Cambridge, Cambridge University Press, 2021; D. Szostek and M. Załucki (eds.), *Legal Tech*, Baden-Baden, Nomos, 2021; C. Ng, *AI in the Legal Profession*, in *The Cambridge handbook of artificial intelligence: global perspectives on law and ethics*, L.A. DiMatteo, C. Poncibò and M. Cannarsa (eds.), Cambridge, Cambridge University Press, 2022.

[7] See on this issues the ample reconstruction made by G. Sartor and A. Santosuosso, *La giustizia predittiva: una visione realistica*, in *Giur. it.*, vol. 2022.

[8] For a similar understanding, see F. Donati, *The use of artificial intelligence in judicial systems: Ethics and efficiency*, in *Italian J. Pub. L.*, vol. 16, 2024.

[9] J. Cobbe, *Legal Singularity and the Reflexivity of Law*, in *Is Law Computable? Critical Perspectives on Law and Artificial Intelligence*, S. Deakin and C. Markou (eds.), London, Hart, 2020, 107.

[10] A. Simoncini, *Il linguaggio dell'Intelligenza Artificiale e la tutela costituzionale dei diritti*, in *Rivista AIC*, vol. 14, 2023.

[11] A. Simoncini and E. Longo, *Fundamental Rights and the Rule of Law in the Algorithmic Society*, in *Constitutional Challenges in the Algorithmic Society*, H.-W. Micklitz, O. Pollicino, A. Reichman, G. Sartor, A. Simoncini and G. De Gregorio (eds.), Cambridge, Cambridge University Press, 2021, 27.

[12] On this issue, see, in general, C.H. McIlwain, *Constitutionalism: Ancient and Modern*, Revised edition, Ithaca, NY, Cornell University Press, 2005; M. Fioravanti, *Constitutionalism*, in *A Treatise of Legal Philosophy and General Jurisprudence*, D. Canale, P. Grossi and B. Hofmann (eds.), Springer, 2009; N.W. Barber, *The Principles of Constitutionalism*, Oxford, Oxford University Press, 2018.

[13] See C. Casonato, *AI and Constitutionalism: The Challenges Ahead*, in *Reflections on Artificial Intelligence for Humanity*, B. Braunschweig and M. Ghallab (eds.), Cham, Springer, 2021; C. Colapietro, *The Impact of the Information and Technology Revolution on the Constitutional Rights with Particular Attention to Personal Data Protection Issues*, in *The IT Revolution and its Impact on State, Constitutionalism and Public Law*, M. Belov (ed.), London, Bloomsbury Publishing, 2021.

liberty.[14]

Constitutional scholars have extensively analysed strategies to make automated decisions more consistent and fair, reducing the risk of human error, discrimination, and abuse of power.[15] However, they emphasise that the law cannot eliminate algorithmic biases. The deployment of AI poses significant risks to individual rights and freedoms, mainly when existing training-data biases are replicated and crude predictions are made. These practices can lead to undesirable and unjustified outcomes, such as discrimination and infringements on individual freedom of choice, ultimately denying individuals access to goods or services.[16]

The underlying concern of these issues extends to the future dynamic between humans and machines.[17] The imminent danger is not relinquishing control to machines and technology but rather the emergence of a new type of authority that could potentially marginalise individuals and their rights.

In light of these problems, my task is to investigate a specific area of the constitutional law of technology that pertains to the profiles connected with the principles applicable to the use of AI in the sphere of "justice." In this study, I will refer to the uses of AI systems in the legal professions and the legal process, analysing them through the lenses of some of the most important constitutional principles regarding justice. The article will proceed as follows.

Firstly, it investigates the application of AI to the legal field, starting from the older versions of AI and moving to the latest digital technology available in this field, Generative AI (also GenAI). Secondly, it analyses some limitations and technical problems of Generative AI's use in the legal field. Thirdly, it comments on the first institutional documents released by some high-level European courts and international organisations on using Generative AI in justice-related activities. The article concludes with a summary of the most crucial constitutional issues emerging from using GenAI in the judiciary.

This article primarily focuses on European jurisdictions, although the issues depicted are important for every country and every judge, lawyer, or legal expert who uses or faces AI in their activity.

2. *Applications of AI in the legal field*

The use of AI in justice is usually subsumed within "digital justice",[18] a multifaceted expression that captures the activities that exploit the potentiality of digitalisation and big data[19] within the legal domain.[20] AI can exercise two functions in the justice field: it can be used as an "auxiliary" tool, i.e., providing support to the judge and

[14] G. Malgieri, *Automated Decision-making and Data Protection in Europe*, in *Research Handbook on Privacy and Data Protection Law*, G. González Fuster (ed.), London, Edward Elgar Publishing, 2022, 433.

[15] C.R. Sunstein, *Algorithms, correcting biases*, in *Social Research: An International Quarterly*, vol. 86, 2019, 499.

[16] A. Simoncini and S. Suweis, *Il cambio di paradigma nell'intelligenza artificiale e il suo impatto sul diritto costituzionale*, in *Riv. fil. dir.*, vol. 8, 2019.

[17] S.E. Biber, *Between Humans and Machines: Judicial Interpretation of the Automated Decision-Making Practices in the EU*, Draft Chapter submitted to the OUP Book on ADM systems and the Rule of Law, University of Luxembourg Law Research Paper. Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4662152.

[18] For a broad definition of "digital justice" see E. Longo, *Giustizia digitale e Costituzione: riflessioni sulla trasformazione tecnica della funzione giurisdizionale*, Milano, FrancoAngeli, 2023. A narrow definition of digital justice has been used mostly to cover the case of online dispute resolution systems. See S.P. de Souza and M. Spohr (eds.), *Technology, Innovation and Access to Justice*, Edinburgh, Edinburgh University Press, 2022; W. Bernhardt and A.S. Maralbaeva, *Platform Solutions for the Digital Justice: Models of Regulation*, in *Springer Nature Singapore*, 2022; M.E. Katsh and O. Rabinovich-Einy, *Digital Justice: Technology and the Internet of Disputes*, Oxford, Oxford University Press, 2017.

[19] When we talk about the AI revolution, we must start from the kind of technological advancement that allows Machine Learning (ML) algorithms to accumulate more and more data (big data), make inferences and produce outputs from which predictions can be made. Even the evolution of machine learning called "deep" (Y. LeCun, Y. Bengio and G. Hinton, *Deep learning*, in *Nature*, vol. 521, 2015) machine learning, based on neural networks similar to the human brain, can rework algorithms autonomously from the designers. This mechanism has two problems: the first and most obvious is that the algorithm does not "know" what it is doing, the operation it performs is pure mathematics; the second is related to the human being's knowledge of the mechanism, which operates like a black box whose input and output are known but whose inner workings are not. F. Pasquale, *The Black Box Society: The Secret Algorithms that Control Money and Information*, Cambridge, MA, Harvard University Press, 2015.

[20] For the definition of AI, see now OECD, *Explanatory memorandum on the updated OECD definition of an AI system*, Paris, OECD Artificial Intelligence Papers, 2024. Retrieved from: https://doi.org/10.1787/623da898-en.

other human legal practitioners to better perform specific legal functions, or it can perform tasks as an "autonomous" tool replacing judges in their related activities.[21]

At present, the uses of artificial intelligence cover mainly auxiliary functions, such as (1) the simplest and most mundane use of computers for routine operations in the field of justice; (2) intelligent databases; (3) intelligent legal information systems;[22] (4) computer-assisted legal drafting; (5) large database-management systems; (6) expert systems; (7) systems that help judicial decision-making by predicting the likelihood of a chosen event, such as the repetition of the offence.[23]

In the last two years, another bunch of functions has been swallowed by AI thanks to the diffusion of GenAI models. These systems can perform many judicial activities, from helping judges "predict" a case's decision to writing a decision and its motivation. Again, as during the old discussions regarding "prediction",[24] the magic is not in the legal capacity of the machines but solely in their computational performance. However, due to the computational capacity, AI is now so powerful and efficient in analysing and creating texts that the substitution of judges seems closer than in the past.

If big-data-analysis tools were intended to support prosecution, sentencing, or compensation policies or help anticipate the effects of a judgement, and some tools enabled sophisticated analysis of case law to inform judges whether their peers had decided on one direction or another,[25] today, AI can outperform these activities and arrive at settling the case entirely, also providing motivation.

The concept of AI being able to fully replace human decision-making has long been considered theoretical. However, with each advancement in AI, the notion of entrusting machines with the entire decision-making process becomes increasingly plausible. Although it may not have been technically feasible in the past, numerous examples from a legal perspective demonstrate that the concept of a "process outside the process" is more relevant than ever.

Today, using AI in judicial processes can leverage "Large Language Models" (LLM), which harness the capabilities of advanced deep learning models. This article predominantly focuses on this technology and aims to explore the constitutional implications arising from the use of these new AI systems in judicial operations.

Within the "Generative AI" family,[26]

---

[21] A similar distinction is made by M. Barberis, *Separazione dei poteri e giustizia digitale*, Milano, Mimesis, 2023, 93. Other definitions are made by: A. Simoncini, *La dimensione costituzionale della giustizia predittiva. Riflessioni su intelligenza artificiale e processo*; G. Sartor and A. Santosuosso, *La giustizia predittiva: una visione realistica*; M. Luciani, *La decisione giudiziaria robotica*, in *Rivista AIC*, vol. 2018. For more insight on this issue see also T. Sourdin, *Judges, Technology and Artificial Intelligence*, 2; M. Zalnieriute and F. Bell, *Technology and the Judicial Role*, in G. Appleby and A. Lynch (eds.), *The Judge, the Judiciary and the Court. Individual, Collegial and Institutional Judicial Dynamics in Australia*, Cambridge, Cambridge University Press, 2021, 116.

[22] The first three functions, for example, correspond in Italy to the digital tools of the "Processo Civile Telematico" (PCT).

[23] See: R.E. Susskind, *Expert systems in law: A jurisprudential approach to artificial intelligence and legal reasoning*, in *Mod. L. Rev.*, vol. 49, 1986; C. Tata, *The Application of Judicial Intelligence and Rules to Systems Supporting Discretionary Judicial Decision-Making*, in *Artif. Intell. Law*, vol. 6, 1998; K.D. Ashley and S. Brüninghaus, *Automatically classifying case texts and predicting outcomes*, in *Artificial Intelligence and Law*, vol. 17, 2009; D. Reiling, *Technology for justice: How information technology can support judicial reform*, Leiden, Leiden University Press, 2009; G. Lupo, *Regulating (Artificial) Intelligence in Justice: How Normative Frameworks Protect Citizens from the Risks Related to AI Use in the Judiciary*, in *European Quarterly of Political Attitudes and Mentalities*, vol. 8, 2019; G. Contissa and G. Lasagni, *When it is (also) Algorithms and AI that decide on Criminal Matters: In Search of an Effective Remedy*, in *European Journal of Crime, Criminal Law and Criminal Justice*, vol. 28, 2020, 280; S. Quattrocolo, *Artificial Intelligence, Computational Modelling and Criminal Proceedings*, Cham, Springer, 2020; D. Reiling, *Digital Justice: Nice to Have but Hard to Achieve*, in *Technology, Innovation and Access to Justice*, S.P. de Souza and M. Spohr (eds.), Edinburgh, Edinburgh University Press, 2022, 131.

[24] G. Zaccaria, *Mutazioni del diritto: innovazione tecnologica e applicazioni predittive*, in *Ars interpretandi*, vol. 26, 2021.

[25] CEPEJ, *Guidelines on how to drive change towards Cyberjustice*, Strasbourg, 2016.

[26] "Generative AI (GenAI) is a type of Artificial Intelligence that can create a wide variety of data, such as images, videos, audio, text, and 3D models. It does this by learning patterns from existing data and then using this knowledge to generate new and unique outputs. GenAI can produce highly realistic and complex content that mimics human creativity, making it a valuable tool for many industries, such as gaming, entertainment, and product design. Recent breakthroughs in the field, such as GPT (Generative Pre-trained Transformer) and Midjourney, have significantly advanced the capabilities of GenAI. These advancements have opened up new possibilities for using GenAI to solve complex problems, create art, and even assist in scientific research"

LLMs are machine learning models trained to generate text similar to human language. They are called "large" because they are trained on a large dataset and are able to generate highly realistic and coherent text that is different at every prompt.[27]

As it is well known, the best-known AI system using an LLM is OpenAI's ChatGPT, a chat-based AI created by the company OpenAI,[28] that offers the possibility to converse, answer questions and emulate the way humans write. The decisive success factor of LLMs such as ChatGpt is that they are trained to interact with human users using natural language and, therefore, produce responses in natural language, i.e., written texts, meaningful expressions, content, and arguments.

Since the release of ChatGPT by OpenAI,[29] language models (LLMs) and similar tools have become essential in our daily lives. From November 2022, these tools have increased and significantly enhanced their capabilities.[30]

One area where it was immediately obvious that these new technologies could have a major impact is the legal sector.[31] Given that the vast majority of legal documents, such as judgments, are written, and even in the (limited) cases in which they can be rendered orally, a written record must still be kept in the minutes of the hearing, the arrival of these new applications of GenAI based on texts and their statistical correlations greatly amplifies the applications to the legal world, at least to those that take a written form.

Everyone may remember that as early as March 2023, a Colombian judge handed down a ruling using ChatGPT.[32] After that case, many others popped up.[33] Many were

(see, https://generativeai.net/). See also K. Busch, *Generative artificial intelligence and data privacy: A primer*, Congressional Research Service, 2023. Retrieved from: https://crsreports.congress.gov/product/pdf/R/R47569.

[27] Generative AI appear to be good at generating coherent (but not necessarily true) text, images and sound, automating summary text, translating texts, semantic analysis and option detection, text mining and content access.

[28] OpenAI is a company founded in 2015 as a non-profit funded by many of America's leading 'big tech' companies to study the positive applications of AI. It became a commercial company in 2019. In 2023, it was acquired by Microsoft for around USD 10 billion.

[29] ChatGPT has been released in November 2022 by OpenAI. See, OPENAI, *ChatGPT: Optimizing Language Models for Dialogue*, OpenAI, 2023. Retrieved from: https://openai.com/blog/chatgpt.

[30] P.P. Ray, *ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope*, in *Internet of Things and Cyber-Physical Systems*, vol. 3, 2023.

[31] J.H. Choi, A. Monahan and D. Schwarcz, *Lawyering in the age of artificial intelligence*, 109 Minnesota Law Review (Forthcoming 2024), Minnesota Legal Studies Research Paper No. 23-31. Available at SSRN: https://ssrn.com/abstract=4626276.

[32] In January 2023, Colombia witnessed a controversy when a judge disclosed that he had used ChatGPT to determine if an autistic child's insurance should cover all expenses related to medical treatment. Ten days later, still in Colombia, a magistrate issued a court order with the assistance of ChatGPT to decide on conducting a trial in the metaverse. Furthermore, in late March 2023, a judge in Peru and a magistrate in Mexico admitted to using OpenAI's ChatGPT to support a second-instance decision and to illustrate arguments in a court hearing. Following the Mata vs. Avianca Airlines, Inc. case, where an attorney submitted falsified citations and cases generated by ChatGPT to a U.S. court, the importance of responsible use guidelines has become even more pronounced today, as we will see below (*Mata v. Avianca*, Inc., 1:22-cv-01461, available at: https://www.courtlistener.com/docket/63107798/mata-v-avianca-inc/). In general, on these cases, see R. Schworer, *Ethical Implications of Utilizing Generative Artificial Intelligence in the Legal Profession: A Cautionary Note on Potential Model Rules of Professional Conduct Violations*, in *Lincoln Memorial University Law Review Archive*, vol. 11, 2024; R. Perona and Y.C. de la Rosa, *Unveiling AI in the courtroom: exploring ChatGPT's impact on judicial decision-making through a pilot Colombian case study*, in *AI & SOCIETY*, vol. 2024. The story and the problem of the first known judge using ChatGPT are depicted by J.D. Gutiérrez, *ChatGPT in Colombian Courts: Why we need to have a conversation about the digital literacy of the judiciary*, in *VerfBlog*, 23/2/2023. Retrieved from: https://verfassungsblog.de/colombian-chatgpt/. While we were drafting this case, a decision of the Constitutional Court of Colombia was issued (T-323/204). The Second Chamber has decided to uphold the decision of the Court of Appeal with a lengthy judgment that thoroughly analyses the many constitutional issues related to using LLMs in judicial activity (https://www.corteconstitucional.gov.co/relatoria/2024/T-323-24.htm). In the specific case, there was no violation of the right to due process because the AI tool was not used to replace the role of the judge in the administration of justice but was used after the decision had been taken and justified. Therefore, the validity of the decision was not called into question. The Court considered that it was necessary to establish criteria and guidelines for using AI tools by judges in the administration of justice to respect fundamental rights and safeguard the constitutional guarantees of users of the system. Justices and clerks should apply the principles of transparency, accountability, privacy, non-substitution of human rationality, seriousness and verification, risk prevention, equality and equity, human control, ethical regulation, adaptation to best practices and collective standards, continuous monitoring and adaptation, and suitability. Finally, the Constitutional Court has asked the Superior Council of the Judiciary to prepare a guide for the implementation of generative AI in the judiciary.

[33] R. Schworer, *Ethical Implications of Utilizing Generative Artificial Intelligence in the Legal*

examples of the amatorial use of an AI system in justice. A recent survey conducted by UNESCO has uncovered an alarming trend: 44% of judicial operators, encompassing judges, prosecutors, and lawyers, are incorporating artificial-intelligence tools such as ChatGPT into their professional activities. However, a mere 9% of these professionals have undergone appropriate institutional training or have been provided with formal guidelines on using these technologies.[34]

Early cases, and the subsequent ones that we now ignore, are at the edge of two possible approaches legal professionals can take towards GenAI. On the one hand, there is the possibility of using technologies such as ChatGPT as auxiliary tools to predict decisions. For example, judges or lawyers can ask an algorithm to identify relevant precedents and select the applicable rules and case law. On the other side, some judges started to use these tools to prompt the facts of cases into a chatbot and ask the algorithm, trained from previous decisions and applicable rules, to generate the text of the sentence or the motivation.

While it is impossible to affirm that ChatGPT "decides" the case in this situation, since GenAI models lack the "understanding" of the text produced and the cognitive will necessary to decide, the system produces a text that the judge uses as a decision. Indeed, when judges ask ChatGPT to generate the text of a ruling (facts and interpretation) with the solution of a case, the model generates a text that is created in accordance with the prompt and the (potentially almost infinite) number of texts on which the (underneath) "model" has been trained.

These cases do not reveal a simple new technological trend or a strategy to reduce judicial spending. We are on the cusp of a new era in which the long-standing idea of the judiciary, ingrained in our culture and legal tradition, may be undermined.

2.1. *LLMs*

Before proceeding into a legal analysis of the fear now depicted, it is helpful to consider the advances in LLM technology within the broader field of AI and explain what has happened in the last two years.

Within the AI landscape, scholars associate LLM with the subfield known as 'Natural Language Processing' (NLP).[35] NLP is the research area that focuses on creating AI systems capable of understanding and generating human text and language. Until 2020, NLP systems could not interpret written text at a level even remotely comparable to a human person's. NLP systems of that period mainly analysed human communication using simple patterns of words and data. Therefore, although these (old) NLP systems could reliably identify keywords and text patterns within documents such as e-mails, they struggled to truly understand the content of any natural-language documents.

Although these older NLP systems were helpful in heavily language-based areas such as law, their application was mainly limited to the statistical categorisation of legal documents rather than understanding the underlying meaning or significance of the words written on a document such as a contract, motion or lease.[36] These limitations persisted for decades, leading many researchers to view the potential for computers to effectively comprehend human language as a relatively distant research objective.

By the end of 2021, the most advanced NLP models were referred to as nothing more than 'stochastic parrots'[37] because they gave

Profession: A Cautionary Note on Potential Model Rules of Professional Conduct Violations; H. Surden, *ChatGPT, Artificial Intelligence (AI) Large Language Models, and Law*, in *Fordham Law Review*, vol. 92, 2024.

[34] UNESCO, *Global Judges' Initiative: Survey on the use of AI systems by judicial operators*, Paris, 2024.

[35] The term "natural language" holds a specific significance in computer science, referring to the everyday languages people use to communicate, including English, French, and Spanish. This term is juxtaposed with computer science's idea of "formal languages," which are precise, mathematically structured "technical languages" like Python and JavaScript, used for programming computers among other tasks. Formal languages are designed for unambiguous mechanical interpretation, resulting in limited expressiveness, but reliable processing by computers. See H. Surden, *ChatGPT, Artificial Intelligence (AI) Large Language Models, and Law*.

[36] See, for example, the very famous article of N. Aletras, D. Tsarapatsanis, D. Preoţiuc-Pietro and V. Lampos, *Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective*, in *PeerJ Computer Science*, vol. 2, 2016.

[37] This expression would mean that such types of AI would only be capable of repeating behaviour displayed during their training, found on the network, or already elaborated, defined and published by others by human intelligent beings. The expression first appeared in the title of a much-cited article by a group of experts on the

the illusion of understanding natural language when, in fact, most of the seemingly coherent text generated by the systems of the time was done by remixing versions of existing sentences that people had previously written on the Internet. Thus, the lack of "understanding" of the NLP systems of that era quickly became apparent when they were asked unusual or common-sense questions.[38]

This changed abruptly in the Autumn of 2022,[39] when we discovered that ChatGPT could respond to (almost) any question we threw at it. Users interact with ChatGPT by "prompting," i.e. the text input that the user writes to interrogate the model. A prompt can take various forms, ranging from specific instructions, such as "Write a poem about dogs," "Summarize this document," or "Draft a merger agreement," to questions like "What is the capital of France?" or "What are the factors of copyright fair use?" These prompts are crucial as they provide ChatGPT with the necessary context to generate coherent and relevant responses.

The interaction with ChatGPT through prompts highlights several significant aspects of how an LLM works and its potential applications. Firstly, prompts serve as the interface through which users can leverage the model's vast computational and linguistic capabilities.[40] Secondly, the versatility of prompts underscores the adaptability of ChatGPT across different domains.[41] Moreover, the interaction through prompts also emphasises the collaborative nature of human-AI engagement.[42]

Since its answers constantly responded to what was asked (although not always completely accurately), it gave the user the impression that it was one of the first artificial-intelligence systems to "understand" the questions or instructions it was provided. ChatGPT appeared as a tool capable of simulating human "reasoning" and solving problems.[43]

Impressed by these advancements, scholars coined the expression "general-purpose AI".[44] These AI showed unexpected emergent properties that most researchers thought would not be possible for many years with old NLP, such as reasoning, problem-solving and producing relevant answers to arbitrary questions or instructions in natural language.

Somewhat counterintuitively, when ChatGPT appears to answer a question or produce lines of text (such as a poem or a legal motion), it uses its internal AI system to predict the most appropriate word (token) to generate later, one word at a time. Even when it generates long documents or comprehensive answers, it generates the text incrementally, predicting one word at a time based on what has been asked or instructed. This is accomplished through a GPT (Generative Pretrained Transformer) algorithm[45] that can read the words embedded (Word Vector) also thanks to "Instruction Finetuning" and "Reinforcement Learning".[46]

However, from the moment ChatGPT appeared, a few problems immediately emerged. On the one hand, it has been clear

ethics of artificial intelligence. See: E.M. Bender, T. Gebru, A. McMillan-Major and S. Shmitchell, *On the Dangers of Stochastic Parrots*, presented at Proceedings of the 2021 ACM Conference on "Fairness, Accountability, and Transparency", Canada, Canada.

[38] S. Bird, E. Klein and E. Loper, *Natural language processing with Python: analyzing text with the natural language toolkit*, Newton, MA, O'Reilly Media, 2009.

[39] OpenAI, *Introducing ChatGPT: Optimizing Language Models for Dialogue*, OPEN AI: BLOG (Nov. 30, 2022), https://openai.com/blog/chatgpt [https://perma.cc/8QWZ-7NKY].

[40] By formulating precise and clear prompts, users can effectively harness ChatGPT's ability to process and generate text, making it a powerful tool for a wide array of tasks.

[41] Whether it is creative writing, legal drafting, educational queries, or technical problem-solving, the model's ability to understand and respond appropriately to varied prompts illustrates its broad applicability. This adaptability is one of the key features that set LLMs apart from earlier generations of AI, which were typically limited to specific, narrow tasks.

[42] The quality of the output generated by ChatGPT is directly influenced by the clarity and specificity of the input provided by the user. This dynamic interaction requires users to develop skills in prompt engineering – crafting inputs that effectively guide the model towards the desired outcome. As such, the role of the user is not passive; instead, users actively shape the responses they receive through their prompts.

[43] ChatGPT passed the Turing test by fooling people to believe that its responses were from humans rather than machines/AI. See F. Fui-Hoon Nah, R. Zheng, J. Cai, K. Siau and L. Chen, *Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration*, in *Journal of Information Technology Case and Application Research*, vol. 25, 2023.

[44] I. Triguero, D. Molina, J. Poyatos, J. Del Ser and F. Herrera, *General Purpose Artificial Intelligence Systems (GPAIS): Properties, definition, taxonomy, societal implications and responsible governance*, in vol. 103, 2024.

[45] GPTs are pre-trained on a large amount of data to capture the nuances of language, images, and sound and generate coherent, context-aware text, images, or sound.

[46] S. Feuerriegel, J. Hartmann, C. Janiesch and P. Zschech, *Generative AI*, in *Business & Information Systems Engineering*, vol. 66, 2024.

that the chatbot makes mistakes in its answers and occasionally invents untrue but plausible information (the model is said to "self-hallucinate").[47] On the other hand, there are many issues related to the data used for the training, such as questions related to personal-data protection, copyright, and cybersecurity.

Therefore, concepts such as "understanding" and "intelligence" should be regarded as "functional metaphors";[48] it would be misleading to characterise these systems as "intelligent" in the same sense as we do with cognitively active humans. Fundamentally, these large language models (LLMs) remain highly advanced statistical pattern-matching machines despite their impressive outputs.[49]

Notwithstanding, the larger story of late 2022 and early 2023 was not centred on ChatGPT's accuracy limitations or the debate over its predicated "intelligence". Instead, it was about the remarkable leap in the usefulness and capabilities of LLMs, a development that only a few anticipated would happen so rapidly,[50] and with a wide array of tasks previously thought to be beyond narrow AI's reach.

ChatGPT's performance levels were not just marginal improvements, but quantum leaps that expanded the horizons of what AI could achieve. In sum, the year 2022 highlighted the transition from theoretical AI capabilities to practical, real-world applications that significantly impact various sectors.

Today, GPT-4o is one of the most advanced LLM systems on the market. Several other companies are working on similar models, and many others are developing models capable of creating images, videos, and sounds.[51] Bard, Copilot, Dall-E, AnthropicAI, and other text, photo, video, sound, and computer-code-generating text, images, video, and sound are already accessible.

2.2. *The Uses of LLMs in the field of law, with a first impression on them*

Although they appear as mainly "shiny objects," LLMs such as ChatGPT are poised to become a colossal industry,[52] with very attractive implications, especially for the legal professions. This innovative digital technology has immense potential to significantly increase the efficiency and productivity of many sectors by streamlining specific processes and decisions, such as facilitating the processing of notes or helping to improve teaching.[53]

Today, as mentioned above, LLM systems, such as ChatGPT, can perform new tasks in the legal field: (1) producing reasonable and quality legal documents, such as contracts, appeals and judgments; (2) performing basic legal analysis; and (3) answering questions about legal documents, such as laws, legal opinions or contracts.[54]

LLM can identify arguments or summarise very long documents and generate texts, including arguments and legal research,[55] by predicting the appropriate text to follow a given input using patterns learned from large datasets.[56]

These tools could be game-changers in many areas of law. They are much more revolutionary than the web-search engines that

[47] S. Wachter, B. Mittelstadt and C. Russell, *Do large language models have a legal duty to tell the truth?*, in *Royal Society Open Science*, vol. 11, 2024, 1.

[48] H. Surden, *ChatGPT, Artificial Intelligence (AI) Large Language Models, and Law*.

[49] R.T. McCoy, S. Yao, D. Friedman, M. Hardy and T.L. Griffiths, *Embers of Autoregression: Understanding Large Language Models Through the Problem They are Trained to Solve*, arXiv: 2309.13638, 2023. Retrieved from: https://arxiv.org/abs/2309.13638.

[50] The invention of this kid of A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, Aidan, L. Kaiser and I. Polosukhin, *Attention Is All You Need*, arXiv pre-print server, 2017. Retrieved from: https://arxiv.org/abs/1706.03762.

[51] Currently, OpenAI is creating the next-generation version within its GPT family of LLM systems, GPT-5.

[52] Editorial, *Large, creative AI models will transform lives and labour markets*, in *The Economist*, 22/4/2023. Retrieved from: https://www.economist.com/interactive/science-and-technology/2023/04/22/large-creative-ai-models-will-transform-how-we-live-and-work.

[53] J. Nelson, *The Other 'LLM': Large Language Models and the Future of Legal Education*, in *European Journal of Legal Education*, vol. 5, 2024.

[54] D. Schwarcz and J.H. Choi, *AI Tools for Lawyers: A Practical Guide*, in *Minnesota Law Review Headnotes*, vol. 108, 2023.

[55] While some LLM chatbots operate within a closed universe of information, others are open and have wider access to data, e.g. via web plug-ins or Internet connections. It is already possible to exploit LLMs on specific datasets or bodies of text and focus them on narrow tasks. Through these applications, one can create personal encyclopaedias and even real databases to train one's chatbot.

[56] For example, GPT-4 can (albeit sometimes imperfectly) engage in legal reasoning about law and facts, analyse or generate contracts, summarize legal cases, draft patents, write motions, and answer questions about legal opinions or documents. See H. Surden, *ChatGPT, Artificial Intelligence (AI) Large Language Models, and Law*.

appeared thirty years ago–although these latter have revolutionised the dissemination and usability of laws and judgments as never before. The new chatbots offer a lot more, as they can take over the roles of those people who used to draft documents, not only personal assistants or legal professionals.

The introduction of LLMs into the legal field was almost coincidentally announced. As said, the first documented instance occurred in Colombia, where a labour judge in Cartagena, Juan Manuel Padilla, prompted ChatGPT with a question regarding the constitutional protection of a social right, subsequently adhering closely to the solution provided by the AI chatbot.

Although the judgment suggests that the judge utilised the text generated by the AI chatbot merely to reinforce conclusions he had independently reached, the case quickly garnered significant media attention, indicating a potentially significant evolution in judicial practices and some important issues experienced in the same use of chatbot AI. While Justice Padilla appeared to have been playing a video game, he has opened the door to officially delegating decision-making to ChatGPT. It is not by chance that today, the first Constitutional decision on the use of new AI tools in the judicial process comes from Columbia.

3. *Technical problems arising from the use of LLMs in the legal field*

In order to grasp the constitutional implications of delegating decision-making to an AI, it is important to first delve into the technical concerns and potential risks associated with the use of LLMs. These risks entail the possibility of AI models generating faulty outputs, known as "hallucinations", as well as introducing or perpetuating biases, and further exacerbating inequalities within the judicial system.

The swift integration of GenAI across diverse economic sectors has underscored its vast potential. Nonetheless, GenAI's extensive capabilities introduce a multifaceted array of novel risks while exacerbating existing threats. In response, institutions have promptly instituted regulations and policies to mitigate these risks, such as the AI Act,[57] while academic researchers have developed and proposed various AI-safety benchmarks and comprehensive taxonomies of AI-related risks.

Scholars have highlighted at least seven categories of risks that can arise from using GenAI, such as undermining equality,[58] infringement of intellectual property rights, information leakage, malicious use, security threats, model illusion, and environmental/social and regulatory risks. Despite ongoing efforts, no unified categorisation of AI risks comprehensively covers all possible risks while considering the perspective of industry and government.[59]

In the specific legal sphere, applications of generative Al require attention, especially with regard to the following four significant risks: security risk (data security), confidentiality risk, risk of unethical use and infringement of intellectual property, risk of non-controllability of legal decisions, and risk of fraud by training with unreliable or untrue information.

These risks are the product of several limitations that underscore the performances of GenAI: 1) illusion and fiction (the "hallucination"); 2) algorithmic discrimination and bias; 3) lack of reliability; 4) non-interpretability; 5) lack of empathy; 6) complex retrieval of resources; 7) incapacity to look at the future (capture by training data); 8) security and confidentiality limitations (data security). As we have already highlighted, LLMs such as ChatGPT are prone to "hallucinations",[60] a particular error

[57] The first comprehensive regulation taking AI's risks as the main element for regulating AI is the EU Regulation "laying down harmonised rules on artificial intelligence", the so-called AI Act (Regulation (EU) No. 2024/1689). See also the Executive Order issued by the President of the United States of America: J. Biden, *Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence*, Washington, 2023. Retrieved from: https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/.

[58] Issues of discrimination and bias are well known since long time, as depicted by J. Angwin, J. Larson, S. Mattu and L. Kirchner, *Machine Bias. There's software used across the country to predict future criminals. And it's biased against blacks*, in *ProPublica*, 23-05-2016. Retrieved from: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

[59] Y. Zeng, K. Klyman, A. Zhou, Y. Yang, M. Pan, R. Jia, D. Song, P. Liang and B. Li, *AI Risk Categorization Decoded (AIR 2024): From Government Regulations to Corporate Policies*, in *arXiv preprint arXiv:2406.17864*, vol. 2024.

[60] "Hallucination" refers to generating false, nonsensical, or inaccurate information by large language models (LLMs) or other generative AI systems. For more, see

by which the output generated (e.g. text) appears coherent on the surface but may be incorrect or wholly invented.

The most significant limitation, typical of GenAI, involves the potential for some output to be fabricated or fictitious. This is partly attributable to the algorithm's configuration and the risk of erroneous correlations within the data. Furthermore, the lack of transparency in the algorithm's programming and data linkage contributes to the overall obscurity and complicates validating the provided response. Most GenAI models cannot list and credit the texts used to create the output. Additionally, this could lead to copyright infringements, which may vary from country to country. Considering this when using these tools is crucial, as the generated content may be considered plagiarism.

A typical example of ChatGPT "illusions and fictions" occurring in the legal profession happened in the case *Mata v. Avianca* in 2023 before a New York court. A lawyer submitted a pleading containing excerpts and quotes from fake cases. It turned out that the pleading was created using ChatGPT. Unaware that ChatGPT can hallucinate, perhaps trusting an inexperienced colleague, the lawyers did not check if the mentioned cases actually existed. The consequences were disastrous. Once the error was discovered, the court dismissed their client's case, sanctioned the lawyers for acting in bad faith, fined them and their firm, and exposed their actions to public ridicule.

Like any "family" of algorithms based on machine learning GenAI can perpetuate problems, errors and biases already present in the training databases. The data used by the algorithms for training and fine-tuning, as well as the construction of the mechanism itself, could be based on preconceptions or biases. It has been shown that even GenAI, in some cases, suffers from their creators' cultural settings and biases and the limitation of datasets.[61] A problem that may condition the algorithm's setting through, for instance, the inclusion or exclusion of elements that identify or refer to a protected category at risk of discrimination without the possibility of appeal. For this reason, systems based on algorithms should ensure transparency, ex-ante, at the moment of their realisation. They should always be verifiable, ex-post, at the moment when the result emerges. Yet, we know that this is not always the case as the debate over explanation shows.

This limitation brings to mind the Weberian idea of "legal calculability",[62] which is perceived as a desirable priority, especially in a historical context that seems dominated and overwhelmed by an unstable and perennially jurisprudential change. On the other hand, the use of machines raises the risk of a system that is not only incapable of creating new legal arguments but is also reduced to the tired and stagnant repetition of itself, losing that capacity to adapt to reality, with the danger of leading to the devaluation of the peculiarities of individual cases and to standardised formulas for handling legal issues.

Another aspect relates to possible discrimination and bias, similar to what happened years ago with the COMPAS algorithm developed by Northpointe, which was used to assess the risk of repeat offenders.[63] Although the algorithm was not

N. Maleki, B. Padmanabhan and K. Dutta, *AI Hallucinations: A Misnomer Worth Clarifying*, in *arXiv preprint arXiv:2401.06796*, vol. 2024.

[61] S. Wachter, B. Mittelstadt and C. Russell, *Do large language models have a legal duty to tell the truth?*.

[62] C.V. Giabardo, *The Judge and the Algorithm (in Defense of the Human Nature of Judging),* available at SSRN: https://ssrn.com/abstract=3859922; N. Irti, *Per un dialogo sulla calcolabilità giuridica*, in *Calcolabilità giuridica*, A. Carleo (ed.), Bologna, il Mulino, 2017; G. Alpa, *Note sulla calcolabilità nel diritto nordamericano*, in ibid., ; M. De Felice, *Calcolabilità e probabilità. Per discutere di 'incontrollabile soggettivismo della decisione'*, in *ibid.;*

[63] The Loomis case was decided in 2016. In assessing the conviction of the defendant Eric Loomis for the offences ascribed to him, the Wisconsin court took into account both the criminal record of the defendant and another rather peculiar parameter, namely the outcome rendered by the software COMPAS, an acronym for "Correctional offender management profiling for alternative sanctions", when commuting the sentence. Specifically, the assessment of the merits as to the defendant's propensity to commit a crime in the future was not based on a methodological criterion of ascertainment of the fact, nor even on a precise prescription of the law, but was entrusted to a risk assessment algorithm patented and produced by the company Northpointe, which, moreover, boasted industrial secrecy on the source code, database and data processing techniques. It was only after an in-depth journalistic investigation that the software was publicly accused of being discriminatory, particularly against African-American individuals. Nevertheless, the Wisconsin Supreme Court considered it not legitimate to re-examine the case in the light of the critical remarks raised, nor did it consider that these critical elements were at least sufficient to reform the decision made in the first instance, effectively ignoring the allegations of discriminatory effects--all on the assumption that no due process violation can be recognised if the system has been correctly applied and if other factors

declared "unconstitutional",[64] the use of algorithms to process information and handle complex calculations, such as recidivism rates, was considered racially discriminatory.

GenAI is used to reveal a critical issue concerning legal authority and legitimacy, a matter long recognised by legal scholars. In general, the complexity of law defies computational reduction;[65] legal norms resist encapsulation in mathematical terms, and their interpretation is inherently keen to ambiguity. Machines reason by inference and not by causal deduction, and since they are not human, they lack empathy and emotional understanding, thus empathy. Perhaps the most significant distinction is that machines currently lack the ability to comprehend their actions and show empathy, a trait exclusive to humans.[66]

The usual juxtaposition of algorithms and legal reasoning is reflected in the notion that both concepts employ an approach based on logical sequences. Consequently, the belief that judicial logic can be aligned with the methodology of digital technologies should serve as the ideal framework for structuring forensic discourse. However, algorithms apply the precision of their automatisms to an imperfect human world, where calculation encounters the uncertainty and unpredictability inherent in human behaviour.

On the other hand, the object of legal reasoning is not exhausted in the serial and typical repetition of facts, which, although susceptible to quantification, are still characterised by the peculiarity of being, in any case, adherent to a concrete case. Hence, in making decisions, the need to be guided by the qualitative criterion expressed by the merits of the interest emerges rather than by quantitative criteria that prove insufficient.

New questions open up concerning the accessibility of algorithms to adversary proceedings, the preliminary investigation, the formation of evidence, the independence of the judge and his or her ability to take precedent into account—without ever losing sight of what is new in reality. It is no longer a matter of passionate theoretical discussions but of scholarly perspectives aimed at probing the practicality of a justice system governed with the help of progressively refined AIs.

Entering judicial information to prompt the machine is not a neutral endeavour. Once information is transmitted to the system provider, it may be used as training data for future users and to generate future outputs. This might produce a breach of personal data used in the prompt or, worse, unintended disclosure of classified or otherwise sensitive information (cybersecurity). Servers or clouds of GenAI providers are unknown and generally placed outside the EU territory.

The need to strike a balance between the need to protect and incentivise technological development, and the need to respect the rights of individuals assumes fundamental importance today. It is, in fact, precisely here that one can see the importance of developing conditions in favour of respecting the rights of the individual through accurate information, transparency, and independent and effective supervision of technologies without creating obstacles to the identification of responsibilities and remedies in the event of a violation of rights. Equally necessary is a process that respects the principles of non-discrimination, equality, and personal data protection.

corroborate the automated assessment. For a recap of the case see J. Larson, S. Mattu, L. Kirchner and J. Angwin, *How we analyzed the COMPAS recidivism algorithm*, in *ProPublica*, 23/5/2023. Retrieved from: https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm; D.L. Kehl and S.A. Kessler, *Algorithms in the criminal justice system: Assessing the use of risk assessments in sentencing*, Berkman Klein Center for Internet & Society, 2017. Retrieved from: https://dash.harvard.edu/handle/1/33746041; A.Z. Huq, *Racial equity in algorithmic criminal justice*, in *Duke LJ*, vol. 68, 2018; T. Brennan and W. Dieterich, *Correctional offender management profiles for alternative sanctions (COMPAS)*, in J.P. Singh, D.G. Kroner, J.S. Wormith, S.L. Desmarais and Z. Hamilton (eds.), *Handbook of recidivism risk/needs assessment tools*, New York, Wiley, 2018; G. Sartor and F. Lagioia, *Il sistema COMPAS: algoritmi, previsioni, iniquità*, in *XXVI Lezioni di Diritto dell'Intelligenza Artificiale*, U. Ruffolo (ed.), Torino, Giappichelli, 2021.

[64] A. Simoncini, *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*, in *Biolaw Journal*, vol. 2019.

[65] A. Garapon and J. Lassègue, *Justice digitale: Révolution Graphique et Rupture Anthropologique*, 1re édition., Paris, Presses universitaires de France, 2018.

[66] There is a broad consensus among computer scientists and even jurists that machines have computing and data accumulation powers incomparable to humans. Yet, they do not reason like humans: they amass and process data whose meaning they do not understand. This is also true of software such as ChatGPT and the rest of the so-called generative AI: they mimic human texts without grasping the sense of what they do. E.J. Larson, *The Myth of Artificial Intelligence: Why computers can't think the way we do*, Cambridge, MA, Harvard University Press, 2021; G. Gigerenzer, *How to Stay Smart in a Smart World: Why Human Intelligence Still Beats Algorithms*, Cambridge, MA, The MIT Press, 2022.

These objectives also highlight the urgency of legal rules and ethical principles that are called upon to promote the development of a technology that is now indispensable for economic growth and social welfare, on the one hand, and to ensure the protection of fundamental rights and the principles of freedom and democracy on which the constitutional state is founded, on the other.

Therefore, the intimate structure of the law, coupled with specific limitations associated with LLMs, strongly discourages unauthorised and amateurish deployment of these tools. Consequently, various judicial bodies have promulgated guidelines advising judges to only use GenAI in their work with proper authorisation.

Hence, the following section of this paper will examine key documents and guidelines already issued by some European judicial authorities on this matter.

4. *From the Ethical Charter of the CEPEJ to Documents and Guidelines on the Use of GenAI in Judicial Activities*

As we mentioned before, the story of the use of AI in law is nothing but new. The possible spread of algorithmic judicial decisions, especially in criminal justice, has long since prompted European institutions to adopt several ethical and normative documents to defend against AI's possible negative effects in trials. Among many, the most important is the European Ethical Charter for the Use of Artificial Intelligence in Criminal Justice Systems and Related Environments, adopted in 2018 by the Council of Europe Commission for the Efficiency of Justice (CEPEJ). The Ethical Charter is the first important document at the international level to address the issues of using AI tools in the judiciary.[67] For the first time, the importance of artificial intelligence in justice, in terms of efficiency and quality of management, is recognised for its benefits and negative effects.[68]

The Charter was drafted before the advent of GenAI. Therefore, the situation has slightly changed, and the possibilities of using AI in the legal profession have multiplied and are within (almost) everybody's ability. AI chatbots powered by LLMs generate text by stitching together sequences of linguistic forms detected in its training data "according to probabilistic information about how they combine but without any reference to meaning".[69]

As we have seen, AI tools are far from reasoning, judging, and even tailoring legal "advice". In addition, the output generated by these systems may exhibit an appearance of authority and coherence, it may simultaneously present inaccurate facts and responses inconsistent with established legal principles and facts. The judiciary's adoption of intrinsically defective instruments and negligent use of AI systems may undermine fundamental rights, such as the right to a fair trial and due process, access to justice and effective remedy, privacy, data protection, data security, equality before the law, and non-discrimination. Additionally, it may compromise constitutional values such as impartiality, independence, and accountability, eroding society's trust in the judicial system.

However, from what we have experienced in these two years, the idea of an autonomous machine deciding a case without a human is a reality that we did not experience in the past. Consequently, many have called for principles and specific guidance from the judiciary to prevent the adoption of AI that does not comply with fundamental rights, prevent the misuse of these systems, and enhance the benefits they can deliver.

Numerous institutions are trying to analyse the effects of GenAI on justice, and particularly on the judicial process, issuing documents that are more "ethical" than legal.[70] With the "Opinion on Regulating and Strengthening the Judicial Application of Artificial Intelligence" issued on 8 December

[67] The Charter provides a framework of principles that can guide policymakers, legislators, and justice professionals as they confront the rapid development of AI in national judicial processes. The Charter has identified five core principles: respect of fundamental rights; non-discrimination; quality and security with respect to data; transparency, impartiality and fairness; "under user control". CEPEJ, *European Ethical Charter on the use of artificial intelligence in judicial systems*, Strasbourg, 2019. Retrieved from: https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c.

[68] S. Quattrocolo, *Artificial Intelligence, Computational Modelling and Criminal Proceedings*.

[69] E.M. Bender, T. Gebru, A. McMillan-Major and S. Shmitchell, *On the Dangers of Stochastic Parrots*

[70] R. Schworer, *Ethical Implications of Utilizing Generative Artificial Intelligence in the Legal Profession: A Cautionary Note on Potential Model Rules of Professional Conduct Violations*.

2022, China's Supreme People's Court has clearly established principles for AI-assisted trials and emphasised that, regardless of the level of technological development, AI should not replace judges in adjudication, thus ensuring that adjudicative powers are always exercised by the judicial organisation and not by AI decision-making.[71]

In July of 2024, the American Bar Association (Standing Committee on Ethics and Professional Responsibility) issued a "Formal Opinion" on Generative Artificial Intelligence Tools aimed at ensuring that clients are protected when lawyers use GenAI.[72] They have to consider "their applicable ethical obligations, including their duties to provide competent legal representation, to protect client information, to communicate with clients, to supervise their employees and agents, to advance only meritorious claims and contentions, to ensure candor toward the tribunal, and to charge reasonable fees."

This initiative is only the last of a long list. In January 2024 the New York State Bar Association has established a Task Force on Artificial Intelligence, which has produced a long report with a profound study and recommendations for lawyers on how to use these tools.[73] It describes the legal, social and ethical impact of artificial intelligence (AI) and generative AI on the legal profession. The Report also reviews AI-based software, generative AI technology and other machine-learning tools that may enhance the profession but also generate unforeseen risks for individual attorneys' inability to understand this unfamiliar technology and issues for courts concerned about the integrity of the judicial process. Further, the Report makes recommendations, including proposed guidelines for responsible AI uses.

Early documents have emerged in Europe, elucidating the potential applications of such instruments by the judiciary while highlighting their associated risks and limitations. Although these documents were specifically crafted to guide the judiciary in adopting and utilising AI systems, their relevance extends to the broader legal profession. Consequently, many of the principles and recommendations articulated in the Guidelines could be suitably adapted for prosecutors, arbitrators, lawyers, bar associations, law societies, civil servants in legal administration, and researchers. Civil-society organisations, including professional associations and universities, have recently issued analogous guidance.

The top English and Welsh judicial bodies drew up a guide in 2023.[74] The guide warns judges and strongly emphasises the responsible use of technology,[75] especially considering issues of accuracy (of documents produced by lawyers and judges),[76] confidentiality, cybersecurity (secrecy of information and cybersecurity),[77] and data

[71] W. Zheng, *Research on Generative Artificial Intelligence Legal Profession Substitution*, in *Mod. L. Rsch.*, vol. 4, 2023.

[72] A.B. Association, *Formal Opinion 512 "Generative AI Tools"*, Chicago, 2024. Retrieved from: https://www.americanbar.org/groups/professional_responsibility/committees_commissions/ethicsandprofessionalresponsibility/.

[73] NYSBA, *Report and Recommendations of the New York State Bar Association Task Force on Artificial Intelligence*, New York, 2024. Retrieved from: https://nysba.org/app/uploads/2022/03/2024-April-Report-and-Recommendations-of-the-Task-Force-on-Artificial-Intelligence.pdf.

[74] Courts and Tribunals Judiciary, *Artificial Intelligence (AI). Guidance for Judicial Office Holders*, London, 2023.

[75] The first part of the document is devoted to the 'responsible use of AI (especially those that interact colloquially) in the courts and tribunals.' The initial statement already demonstrates the requirement that chatbots built through AI should only be used by those aware of these technologies' capabilities and potential limits. Although these limitations are identified throughout the document, at the outset, a few are selected as the 'main' ones: the lack of authority, the incomplete verifiability of the information, the poor quality of the responses, the overabundance of data from the American system.

[76] A specific section is devoted to the need to use 'accurate and responsible information.' With regard to the first aspect, information acquired through an AI tool must always be verified before its use. AI tools may contain inaccurate, incomplete, misleading or out-of-date data. Even if they claim to represent full legal knowledge, this is not necessarily the case. As seen, LLMs may 'hallucinate' and invent fictitious cases, quotes or phrases, refer to non-existent laws, articles or legal texts, provide incorrect or misleading information on legal sources or their application, or make factual errors. One of the document's most important points is the correct use of answers. As mentioned earlier, one needs to be aware that, at given the 'unsupervised' nature of these tools, AI training datasets can reflect errors and biases in the training data. This possibility must be fully recognised, and efforts must be made to correct any possible errors.

[77] In the document, there is also a section on the fundamentals of "computer security". As it is well known, one of the main elements to guarantee against possible cyber abuse is the use of certified equipment that is designed to guarantee a high level of cyber security. In contrast, using hardware and software not provided by the justice system can greatly increase the dangers or be not fully controllable by those in charge of cybersecuri-

protection.[78]

The document by the top bodies of the English and Welsh judiciary warns judicial officeholders of material produced in their name. The starting assumption concerns that "judges are generally not obliged to describe the research or preparatory work done to produce a judgment".

An important element of the Guide concerns the relationship with external parties, the 'users' of the courts, i.e. legal professionals, consultants and the parties themselves.[79]

The guide represents just the initial phase of a series of forthcoming efforts to assist the judiciary in dealing with AI. It has been noted that these initiatives will be reassessed as the technology progresses. It is emphasized that "the use of AI by or for the judiciary must align with the judiciary's overall duty to uphold the integrity of the administration of justice."

The document also devotes a special section to respecting the privacy and confidentiality of information. The last warning concerns the possible use of false material (texts, images and videos) and the additional and more pervasive truthfulness check to which documents produced in court must be subjected.

The document of the English and Welsh supreme judiciaries helps us to focus on an assumption from which it is necessary to move to carry out a broader reasoning about the use of GAI within trials. Given the impact of disruptive technological discoveries and their diffusion, no state can pretend that GAI is not a possible aid and, simultaneously, a problem for exercising its powers. This general consideration applies as much to the guarantee of public services as to security or collective health care, and it is also valid for the area of justice and the guarantee of fundamental rights.

The modernisation of administration and services is not only a factor to consider regarding one's own citizens but also an element that positions each state within world geopolitics and is a very important component of sovereignty. States' administrations will, therefore, have to reorganise themselves because those who control the production and use of AI will be able to shape the global (political) future.

At the individual level, GenAIs are different from the AIs we have used so far, both in their uses and in the power they express vis-à-vis the agent subject, who may not be able to understand the consequences they generate. The English judiciary is aware of this problem, indeed implicitly asks in the Guide not how to adapt GenAI to trials, but how to instruct judges to its conscious use, now that such applications are freely available to all.

Also the Court of Justice of the European Union has issued a "Strategy on Artificial Intelligence". The Strategy indicates that these technologies can be harnessed to achieve multiple objectives, focusing on three key areas: improving the efficiency and effectiveness of judicial proceedings, improving the quality and consistency of judicial decisions, and increasing access and transparency for EU citizens. The Strategy does not stop at a simple 'performative' analysis, as it describes a set of 'principles' on AI for the CJEU, risks and how to manage them, as well as some indications on the governance of these technologies for the future of justice at the EU level.

Already from the executive summary, the

---

ty in the document.

[78] In an interview with the Financial Times, Master of Rolls, Geoffrey Vos anticipated the arrival of the Guide by pointing out that AIs would not be officially allowed to aid decision-making until the judiciary was absolutely sure of the goodness of these tools and citizens were confident in their use. A. Gray, *Judges in England and Wales told to restrict their use of AI in cases*.

[79] The document observes that certain types of AI have been used for a long time and without difficulty in the judicial sphere and many areas of life in general: for example, in "search engines for auto-compilation of applications, in social media to select content for distribution, in image recognition and predictive text". With respect to these uses, all professionals working in the legal field have a specific responsibility for the material they produce in court and a specific professional obligation to ensure its accuracy and appropriateness, as we have seen in the New York case and the Texas court decision. Even in these cases, if an IA is used responsibly, it is considered that there is no reason for a professional to refer to its use. However, it is suggested that until the legal profession becomes familiar with these new technologies, "lawyers may need to be reminded of their obligations and confirm that they have independently verified the accuracy of any research or citation of cases generated with the assistance of an AI chatbot". The need for the conscious use of chatbots created with LLMs must also apply to those who are in court without professional representation. In these cases, when a patient turns to traditional search engines to acquire medical information, it is suggested that the use of these tools is not the only source of advice or assistance for the parties. Judges, therefore, need to exercise greater scrutiny to verify the accuracy of the information provided by the parties.

Strategy emphasises that AI holds significant potential for the CJEU, both by enabling the automation of simple tasks in judicial and administrative settings and by offering new possibilities for legal research, translation, and interpretation, improving accessibility, and providing a modernised approach to accessing information.

These technologies can be exploited to achieve multiple objectives, focusing on three key areas: improving the efficiency and effectiveness of judicial proceedings, improving the quality and consistency of judicial decisions, and increasing access and transparency for EU citizens.

Guidelines on how judges must use GenAI have appeared in other jurisdictions, such as Australia,[80] Canada,[81] and New Zealand. [8283]

Also the Council of Europe was quick to make its voice heard. After enacting the first "code" encompassing the ethics of digital justice, the CEPEJ Working Group on Cyberjustice and Artificial Intelligence has also issued an "information note" on the "Use of Generative Artificial Intelligence (AI) by judicial professionals in a work-related context". The document provides guidelines and considerations for appropriately using tools powered by GenAI: it outlines best practices for ensuring they are used responsibly and effectively while highlighting situations where their use may be inappropriate or prohibited.

One of the most active international institutions studying and advocating for the responsible use of AI systems by judicial operators is UNESCO. After having issued a "Global Toolkit on the AI and Rule of Law for the Judiciary"[84] and delivered the results of a survey on the AI literacy of judges,[85] where it was found that nearly ten per cent of judicial organisations have issued appropriate guidelines or AI-related training, the UNESCO "AI and the Rule of Law programme", under the "Global Judges Initiative", has developed a draft "Guidelines for the Use of AI Systems in Courts and Tribunals".[86] The document (now under consultation),[87] aims to "offer comprehensive guidance to courts and tribunals to ensure that the deployment of AI technologies aligns with the fundamental principles of justice, human rights, and the rule of law".

The draft Guidelines highlight the benefits and limitations of using AI systems in decision-making processes. On one side, AI could improve the efficiency of judicial proceedings,[88] making access to justice and the right to a fair trial within a reasonable time more effective. On the other side, AI may have some harmful impact because of the effects we have seen above regarding hallucinations and possible discriminatory effects.[89]

While these AI tools cannot substitute

---

[80] Q. Courts, *The use of Generative Artificial Intelligence (AI): Guidelines for responsible use by non-lawyers*, Brisbane, 2024. Retrieved from: https://www.courts.qld.gov.au/going-to-court/using-generative-ai; S.C.o. Victoria, *Guidelines for litigants. Responsible use of artificial intelligence in litigation*, Melbourne, 2024. Retrieved from: https://www.supremecourt.vic.gov.au/forms-fees-and-services/forms-templates-and-guidelines/guideline-responsible-use-of-ai-in-litigation.

[81] F.C.o. Canada, *Notice to the Parties and the Profession. The Use of Artificial Intelligence in Court Proceedings*, Toronto, 2024. Retrieved from: https://www.fct-cf.gc.ca/Content/assets/pdf/base/FC-Updated-AI-Notice-EN.pdf; F.C.o. Canada, *Interim Principles and Guidelines on the Court's Use of Artificial Intelligence*, Toronto, 2023. Retrieved from: https://www.fct-cf.gc.ca/en/pages/law-and-practice/artificial-intelligence.

[82] C.o.N. Zeland, *Guidelines for use of generative artificial intelligence in Courts and Tribunals (Judges, Judicial Officers, Tribunal Members and Judicial Support Staff)*, Wellington, 2023. Retrieved from: https://www.courtsofnz.govt.nz/assets/6-Going-to-Court/practice-directions/practice-guidelines/all-benches/20231207-GenAI-Guidelines-Judicial.pdf.

[83] For more information, see UNESCO, *Draft UNESCO Guidelines for the Use of AI Systems in Courts and Tribunals*, Paris, 2024.

[84] UNESCO, *Global Toolkit on AI and the Rule of Law for the Judiciary*, Paris, 2023. Retrieved from: https://unesdoc.unesco.org/ark:/48223/pf0000387331.

[85] UNESCO, *Global Judges' Initiative: Survey on the use of AI systems by judicial operators*. On the issue of judges literacy see also B. Custers, *A fair trial in complex technology cases: Why courts and judges need a basic understanding of complex technologies*, in *Computer Law & Security Review*, vol. 52, 2024.

[86] UNESCO, *Draft UNESCO Guidelines for the Use of AI Systems in Courts and Tribunals*.

[87] The guidelines were open for public consultation in English until 5 September 2024.

[88] The Draft Guidelines underline that "These tools can help improve the quality of the judiciary's work by facilitating court management, document drafting, exploring specific topics, automating tasks, and supporting decision-making processes. Moreover, one of the key features of LLMs is that they facilitate the interaction between computers and users through natural language."

[89] The Draft Guidelines underline that "the adoption of defective instruments and the negligent use of AI systems by the judiciary may also undermine human rights, such as fair trial and due process, access to justice and effective remedy, privacy and data protection, equality before the law, and non-discrimination, as well as judicial values such as impartiality, independence, and accountability. Moreover, the misuse of AI systems may undermine society's trust in the judicial system."

qualified legal reasoning because they have no rationale or contextual understanding of a legal problem, they represent a plurality of technologies that can potentially transform and disrupt the whole practice of law and judicial adjudication.[90] Their use should be managed cautiously by non-lawyers.

The Guidelines follow a tripartite framework. The first section provides thirteen principles, mainly related to the need to protect fundamental rights, that organisations and individual judiciary members should follow when adopting and using any AI system.[91] The second section comprises recommendations and guidance for organisations that are part of the judiciary (bodies that govern the judiciary, courts, and tribunals). This guidance covers: the adoption of systems for impact assessment of AI tools; internal procedures and standards (risk management, data governance, cybersecurity); training sessions for judges; and specific elements for GenAI. The third part contains guidance for individual judiciary members. In this part, there is a specific sub-section on "the use of generative AI systems", namely "protect personal and confidential data", "main uses for LLMs", "Unreliability as search engines and for legal analysis", "Awareness of LLMs' limitations and risks", "Verify outputs before using them", "Transparent use", "Integrity", and "Responsibility".

5. *The AI Act*

The need for guidance on the use of AI in the field of justice is accentuated by the new AI regulation issued by the European Union in July 2024. At the moment, the EU AI Act is the most advanced law on AI, which also addresses issues related to the uses of these systems in the judicial context.[92]

Formally speaking, the Regulation (EU) No. 2024/1689, also called the "AI Act," mainly introduces a "product safety framework" formed around four risk categories: unacceptable, high, limited, and minimal.[93] We can find prohibited systems, labelled as unacceptable for the risk they poe, in Article 5.[94]

The AI Act categorises high-risk AI systems into two distinct groups. The first category includes systems that serve as safety components within products already governed by existing ex-ante conformity assessments, as mandated by the comprehensive EU harmonisation legislation on health and safety outlined in Annex I.[95] The second category

[90] M. Dahl, V. Magesh, M. Suzgun and D.E. Ho, *Large legal fictions: Profiling legal hallucinations in large language models*, arXiv preprint arXiv:2401.01301, 2024. Retrieved from: https://arxiv.org/abs/2401.01301

[91] These principles are mainly related to "fairness", "non-discrimination", "personal data protection", "safety", "information security", "awareness", "transparency", "explainability", "human oversight".

[92] C. Gallese, *Predictive justice in light of the new AI Act proposal*, Available at SSRN 4286023. Available at SSRN: https://ssrn.com/abstract=4286023 or http://dx.doi.org/10.2139/ssrn.4286023; F. Donati, *The use of artificial intelligence in judicial systems: Ethics and efficiency*.

[93] The AI Act considers only some aspects of the possible uses of AI. See B. Prainsack and N. Forgó, *New AI regulation in the EU seeks to reduce risk without assessing public benefit*, in *Nature Medicine*, vol. 30, 2024; C. Novelli, F. Casolari, A. Rotolo, M. Taddeo and L. Floridi, *AI Risk Assessment: A Scenario-Based, Proportional Methodology for the AI Act*, in *Digital Society*, vol. 3, 2024; G. De Minico, *Giustizia e intelligenza artificiale: un equilibrio mutevole*, in *Rivista AIC*, vol. 15, 2024; G. Mazzini and S. Scalzo, *The Proposal for the Artificial Intelligence Act: Considerations around Some Key Concepts*, in *La via europea per l'Intelligenza Artificiale*, C. Camardi (ed), Milano, Wolters Kluwer, 2022.

[94] These encompass the deployment of artificial intelligence to manipulate human behaviour in ways that undermine individual autonomy and exploit the vulnerabilities of natural persons, particularly considering people that, for their age, disability, or socioeconomic status, can be vulnerable. Additionally, it includes using AI systems for criminal risk assessments, predicting individuals' behaviour without human oversight, and evaluating or classifying individuals based on their social behaviour or personal characteristics (social scoring). However, such practices are specifically addressed when they result in adverse or unfavourable treatment in social contexts unrelated to the original context in which the data was collected or when such treatment is unjustified or disproportionate. The deployment of emotion recognition technology within the workplace and educational institutions is expressly prohibited, thereby allowing its application in other contexts despite the inherent ethical concerns it raises. Similarly, the indiscriminate scraping of facial images from the internet or CCTV footage to construct facial recognition databases is also prohibited. Moreover, the legal framework does not permit categorising biometric data based on sensitive attributes, including political, religious, or philosophical beliefs, sexual orientation, or race. The AI Act permits law enforcement agencies to employ live facial recognition technology in public spaces, but only under strict conditions, such as acquiring prior judicial authorisation, and in some specific circumstances: conducting targeted searches for victims, preventing particular and imminent terrorist threats, or locating or identifying individuals who have been convicted or are suspected of committing specified serious crimes.

[95] Within this category, it is possible to find a wide range of products, including medical devices, toys, aviation equipment, automobiles, and lifts.

consists of stand-alone high-risk AI systems, which are primarily of concern due to their potential to adversely affect fundamental rights.[96] Regarding the focus on fundamental rights, the AI Act, in Article 27, provides for a Fundamental Rights Impact Assessment (FRIA) for certain high-risk AI systems.

Recital No. 4 considers "justice" a sector where AI brings evident improvement by providing "key competitive advantages to undertakings and supporting socially and environmentally beneficial outcomes".[97] Recital No. 61 says that some "AI systems intended for the administration of justice and democratic processes should be classified as high-risk considering their potentially significant impact on democracy, the rule of law, individual freedoms as well as the right to an effective remedy and to a fair trial". In particular, Recital No. 61 continues to explain that "to address the risks of potential biases, errors and opacity", the "AI systems intended to be used by a judicial authority or on its behalf to assist judicial authorities in researching and interpreting facts and the law and in applying the law to a concrete set of facts" must be qualified as "high-risk". At the same time, those AI systems "used by alternative dispute-resolution bodies" should also be considered high-risk "when the outcomes of the alternative dispute-resolution proceedings produce legal effects for the parties".

Recital No. 61 also points out that "AI tools can support the decision-making power of judges or judicial independence, but should not replace it". In any case, "the final decision-making must remain a human-driven activity". Not all AI systems are considered high-risk. Those intended for "purely ancillary administrative activities that do not affect the actual administration of justice in individual cases" are excluded. The Recital provides examples of exclusion, listing the AI systems intended for "anonymisation or pseudonymisation of judicial decisions, documents or data, communication between personnel, administrative tasks" as those not included among high risk.

Annex III further specifies "AI systems intended to be used by a judicial authority or on their behalf to assist a judicial authority in researching and interpreting facts and the law and in applying the law to a concrete set of facts, or to be used in a similar way in alternative dispute resolution";

It will be important to understand properly the activities covered by the expression "research and interpret facts and the law and in apply the law to a concrete set of facts".[98]

In addition, Article 86 grants affected persons who are subjected to (most) high-risk AI systems listed in Annex III an (additional) "right to an explanation", covering the "right to obtain from the deployer clear and meaningful explanations of the role of the AI system in the decision-making procedure and the main elements of the decision taken". This right, however, only applies if the decision "produces legal effects or similarly significantly affects that person in a way that they consider to have an adverse impact on

[96] These systems are comprehensively enumerated in Annex III and span eight specific application domains. This latter category represents a particularly significant focus of the AI Act, as the systems in Annex I are already subject to dedicated regulatory frameworks. The second category includes systems used to access employment or systems employed in education, justice, and public services, for which conformity assessment requirements are established. High-risk systems must adhere to the "essential requirements" delineated in Articles 8 to 15 of the AI Act, which concern the risk-management system, data quality and data governance measures, technical documentation and (automated) logging capabilities, transparency provisions, human oversight, cybersecurity, and quality management system. The requirements of the AI Act are differentiated according to the type of operators (i.e. suppliers, deployers, distributors, importers or authorised representatives).

[97] See the text of Recital No. 4 "AI is a fast evolving family of technologies that contributes to a wide array of economic, environmental and societal benefits across the entire spectrum of industries and social activities. By improving prediction, optimising operations and resource allocation, and personalising digital solutions available for individuals and organisations, the use of AI can provide key competitive advantages to undertakings and support socially and environmentally beneficial outcomes, for example in healthcare, agriculture, food safety, education and training, media, sports, culture, infrastructure management, energy, transport and logistics, public services, security, justice, resource and energy efficiency, environmental monitoring, the conservation and restoration of biodiversity and ecosystems and climate change mitigation and adaptation".

[98] As for the enforcement of these provisions, Article 74, paragraph 8, provides that Member State should use data protection supervisory authorities (under Regulation (EU) 2016/679 or Directive (EU) 2016/680), like the Italian "Garante per la protezione dei dati personali" as "market surveillance authorities for the purposes of this Regulation" when are considered "high-risk AI systems listed in point 1 of Annex III to this Regulation, in so far as the systems are used for law enforcement purposes, border management and justice and democracy," and "high-risk AI systems listed in points 6, 7 and 8 of Annex III to this Regulation".

their health, safety or fundamental rights", and national or Union law can provide exceptions to this right.

The AI Act establishes transparency obligations for systems of limited risk, such as chatbots (Article 50). Finally, systems of minimal risk can be covered by voluntary codes of conduct. The AI Act mandates the European Commission, Member States, and the AI Office (a supervisory entity that we discuss in the next section) to encourage and facilitate the development of codes of conduct to promote the voluntary application of the high-risk requirements to no-risk AI systems.

Generative foundation models are theoretically outside the risk-based classification but are regulated within the category of "general-purpose AI" (GPAI), also listed in the AI Act.[99] Generative AI was not part of the original proposal; it was added by the Council and Parliament during the trilogue and approval in order to impose additional transparency requirements on providers of generative foundation models.[100] Indeed, Article 53 imposes obligations on GPAI model providers to develop and maintain technical documentation, ensure compliance with EU copyright law, and disseminate "sufficiently detailed" summaries concerning the content utilised for training the models before their placement on the market.[101]

Provided that an AI system based on large language models (LLMs) can be utilised within the "administration of justice", as outlined in Annex III, item 8, particularly in assisting the judiciary with research, interpretation of facts, and application of law to specific factual scenarios, if the intended purpose of that AI system is indeed to have a decisive role in the decision-making process (Article 6, para 3), it shall be considered high-risk.

Conversely, if the system should not be developed or deployed in a manner that could materially influence decision-making, the intended purpose of the AI system falls within the exemption outlined in Article 6, para 3 of the AI Act. However, given the usual over-reliance on automation, it is not easy for the AI system provider to appeal for the exemption. They should foresee the potential roles of automation in the system's design. In this case, any predictable uses should be considered as producing risks and then be mitigated through comprehensive and effective instructions for use, which should also be designed to prevent any material influence that could heighten the risk of adverse impacts on fundamental rights.[102] In these cases, the right to an explanation should also be considered, as provided in Article 86 (see above).

In conclusion, the AI Act includes many rules for using AI in the legal field, and all present and future applications – following the grace period for enforcement provided by Article 113 – using these models must adhere to its rules.

6. *Constitutional profiles relating to the use of LLM in the judicial process*

The unstoppable development of GenAI warns us that a radical change is underway

[99] G. Olivato, *Paving the path towards general purpose AI systems regulation in the AI Act: an analysis of the Parliament's and Council's proposals*, in *Medialaws*, vol. 2023.

[100] P. Hacker, A. Engel and M. Mauer, *Regulating ChatGPT and other Large Generative AI Models*, presented at 2023 ACM Conference on Fairness, Accountability, and Transparency.

[101] Obligations for GPAI include "to draw up and keep up-to-date the technical documentation of the model, including the training and testing process and the results of its evaluation", "to draw up, keep up-to-date and make available information and documentation to providers of AI systems who intend to integrate the general-purpose AI model into their AI systems", "to draw up and make available to the public a sufficiently detailed summary of the content used for training the general-purpose AI model". While these minimum standards apply to all models, GPAI models identified as posing a "systemic risk" due to their "high impact capabilities" are subject to additional obligations. These obligations include conducting model evaluations, performing adversarial testing, assessing and mitigating systemic risks, reporting severe incidents, and ensuring an adequate level of cybersecurity. However, it is essential to note that providers of GPAI models classified as posing systemic risk can conduct their audits and evaluations rather than relying on external, independent third-party audits. Furthermore, there is no requirement for ex-ante public licensing. See N.A. Smuha and K. Yeung, *The European Union's AI Act: beyond motherhood and apple pie?*, Forthcoming in "The Cambridge Handbook on the Law, Ethics and Policy of Artificial Intelligence". Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4874852; M. Ebers, *Truly Risk-Based Regulation of Artificial Intelligence-How to Implement the EU's AI Act*, SSRN 4870387Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4870387.

[102] For a similar insight, see I. Carnat, *Addressing the Risks of Generative Ai for the Judiciary: The Accountability Framework (S) Under the EU Ai Act*, SSRN 4887438. Available at SSRN: https://ssrn.com/abstract=4887438 or http://dx.doi.org/10.2139/ssrn.4887438.

with far-reaching effects on both the practice and theory of law. Quantitatively, AI works compared to semi-automatic reactions for which the human mind uses heuristics and bias. Yet, from a qualitative point of view, AI has not made, and perhaps never will, the leap from data processing to its human use, which means being capable of conscious, reflective, responsible reasoning. This imitation of human intelligence bewitches those who use the latest versions of ChatGPT. However, they should reckon that such software results are derived from mere word associations obtained by an algorithm unaware of their meaning or relevance.

This problem goes directly to the crucial points we face when using AI in judicial activities, such as the process: What concept of justice did computer scientists have in mind when they programmed the algorithm running beneath ChatGPT? Who is in charge of selecting the data to be used in the training of the AI model? Can we assess the operations of an algorithm using a specific concept of justice? Can we use these AI tools in judicial and legal activities without knowing the answer to the questions above?

Because of these problems, the employment of AI tools in justice has garnered both support and opposition for a long time.[103] On the one hand, the numerous endeavours to establish justice through artificial intelligence illustrate only one of the use cases today available for automating law. On the other hand, as we have seen, several arguments caution against these uses, citing the limitations of machines in comprehending context, their propensity for unlawful generalisations, and the tendency of many AI models to "entrench or even exacerbate existing discriminatory practices and inequalities".[104] These concerns raise significant doubts about the feasibility of administering justice through algorithms.[105]

The advent of GenAI has produced a leap in the study and application of AI to legal activities. AI has become more and more capable of recognising and extracting information from texts, such as statutes and court decisions, contracts, legal essays, etc. Unquestionably, the speed of text processing, data extraction accuracy, and subsequent forming of coherent and meaningful responses are overwhelming. As legal scholars have highlighted, the information extracted by AI is also suitable for further efficient processing since it is often structured data.[106] Such structured data are produced as an extraction output, even in natural-language processing, designed to get closer to understanding information as humans do. Although machines do not possess true comprehension of the information they process, they excel in extracting, interpreting, and integrating data. The responses they generate are often indistinguishable from those produced by humans, and in some cases, they may even surpass human outputs in quality.[107] The advent of GenAI has enabled the transition from auxiliary to autonomous AI, rendering the latter a more attainable and accessible prospect for all, especially in the legal field.

With these elements in mind, we now understand the impact on constitutional principles regarding justice in a situation where AI is largely used by legal professionals inside and outside the process as a tool that completely or partially makes autonomous decisions.

The gradual, invisible, yet profound infiltration of algorithms into both the public and private spheres of life has reached a critical juncture.[108] They are now beginning to acquire "legitimacy". People stop questioning if they are justified based on the assumption that efficiency means secure legitimation. Many of the documents issued by Institutions worldwide that we have analysed above consider the intrusion of AI into justice as irrevocable. Individuals perceive AI as a tool to use normally, as it was in the past for typewriters, computers, and the Internet – just a part of their lives and society. It now appears

---

[103] P. Rubim Borges Fortes, *Paths to Digital Justice: Judicial Robots, Algorithmic Decision-Making, and Due Process*, in *Asian Journal of Law and Society*, vol. 1, 2020, 1.

[104] I. Strümke, M. Slavkovik and V.I. Madai, *The social dilemma in artificial intelligence development and why we have to solve it*, in *AI and Ethics*, vol. 2, 2022.

[105] Y. Razmetaeva, *Artificial intelligence and the end of justice*, in *Biolaw Journal*, vol. 2024; S. Fraley, *Digital Justice: Shining Future or Illusory Chimera?*, in *Pólemos*, vol. 14, 2020, 47.

[106] Y. Razmetaeva, *Artificial intelligence and the end of justice*, 355.

[107] J. Collenette, K. Atkinson and T. Bench-Capon, *An explainable approach to deducing outcomes in European Court of Human Rights cases using ADFs*, in *Frontiers in Artificial Intelligence and Applications*, vol. 326, 2020.

[108] G. De Gregorio, *The Normative Power of Artificial Intelligence*, in *Ind. J. Global Legal Stud.*, vol. 30, 2023, 55.

next to impossible to separate technology from humans and abandon the benefits people are accustomed to relying on.[109] In the justice system today, efficiency and effectiveness principles appear to begin to subsume "democratic legitimation". Therefore, adding AI to justice cannot be considered neutral and apolitical. AI is intrinsically incontrollable in many ways, and can perform better than humans in fulfilling constitutional principles regarding justice.

These loopholes certainly do not mean that humanity must abandon artificial intelligence; rather, it must pay careful attention to its implementation and not only in the moment of application to cases. Nonetheless, the debate over AI is more than theoretical; it is practical.[110]

The emergence of GenAI represents a profound shift from the concept of "digital justice" we have previously examined.[111] The deployment of such tools introduces a dual risk. Firstly, when conducted through machine circuitry, the adjudicative process lacks the intrinsic human qualities essential to judicial decision-making. This raises concerns about the loss of humanity in judgment. Secondly, judicial processes are inherently anchored to specific physical locales, such as courtrooms, where the presence of a judge imparts irreplaceable legitimacy to the proceedings.[112] The authority of an artificial entity, irrespective of its claimed intelligence, is insufficient to fulfil this indispensable human function.

In our prior analysis, we explored the constitutional principles impacted by the advent of digital justice. At that time, we noted that, despite entering a transformative era, the degree to which digitisation could infiltrate the judicial process remained uncertain. The potential legislative safeguards designed to mitigate this shift, the recontextualisation of constitutional norms in response to the increasing digitisation of judicial processes, and the implications for maintaining a "fair" process amidst the growing dependence on advanced technological tools within the judiciary—all these issues remained unresolved.[113]

Firstly, a fundamental distinction exists between the judicial process and other state functions. The judicial process focuses on applying legal norms to particular cases, distinguished by a strict adherence to established procedural rules. Unlike the procedural norms governing administrative and parliamentary functions, the judicial process primarily deals with resolving disputes between opposing parties. As a result, activities in this field often involve legal disputes rooted in unique sets of facts.

Secondly, the judicial process is sustained through dialogue and relationships, serving as a rational mechanism for exercising public authority. This process is actualised through a series of formal acts designed to facilitate judicial decision-making in executing the judicial function.

Thirdly, the judicial function necessitates a clear delineation of jurisdictions, particularly concerning the distinct natures of protectable subjective situations, whether these pertain to subjective rights or legitimate interests. Moreover, applying a rigid, universal trial framework does not guarantee the expeditious resolution of all types of disputes and may compromise the objectivity of judicial determination. Therefore, it is imperative to differentiate case-handling procedures according to each case's unique characteristics.

Fourthly, the judicial process holds significance in both an ideal and material sense. Ideally, it is vital to discern when a recognised right is also protected. Materially, it is crucial to identify the concrete manifestation of an objective right which is enforced. The effectiveness of legal protection is contingent upon applying the law within the judicial process and realising the underlying "political" objectives. Given the substantial public interests, the trial cannot be considered private.

[109] A. Rachovitsa and N. Johann, *The human rights implications of the use of AI in the digital welfare state: Lessons learned from the Dutch SyRI case*, in *Human Rights Law Review*, vol. 22, 2022.

[110] S.P. De Souza, *The Spread of Legal Tech Solutionism and the Need for Legal Design*, in *European Journal of Risk Regulation*, vol. 13, 2022, 373; L.A. DiMatteo, A. Janssen, P. Ortolani, F. de Elizalde, M. Cannarsa and M. Durovic (eds.), *The Cambridge Handbook of Lawyering in the Digital Age*; D. Szostek and M. Załucki (eds.), *Legal Tech*.

[111] E. Longo, *Giustizia digitale e Costituzione: riflessioni sulla trasformazione tecnica della funzione giurisdizionale*, 45.

[112] A. Garapon and J. Lassègue, *Justice digitale: Révolution Graphique et Rupture Anthropologique*, 200.

[113] A. Pajno, *L'uso dell'intelligenza artificiale nel processo tra problemi nuovi e questioni antiche*, in *Biolaw Journal*, vol. 9, 2022, 205.

The formula that encapsulates the essence of these guarantees within the broader constitutional framework is "due process." The evolution of this concept is of contemporary significance, particularly in the context of the rise of digitisation and the integration of new forms of artificial intelligence.[114]

As mentioned earlier, legal norms and ethical and moral standards safeguard the approach to the judicial process. The judicial process is grounded in numerous ethical and legal principles constituting its foundational values. Among these are the judges' independence and impartiality, the exercise of discretionary power, the right to an effective remedy, and the guarantees of cross-examination.[115] Moreover, the criminal process is governed by additional principles, such as the presumption of innocence and the right to immediate appeal against measures that infringe upon personal liberty.

The constitutional principles applicable to humans must equally apply to GenAI. Consequently, the failure of such instruments to uphold the guarantees mentioned earlier would render them unconstitutional.

The foundation of this line of reasoning lies in the principle of judicial protection, which is recognised as one of the supreme principles of our constitutional order and is intrinsically linked to the principle of democracy. It ensures a judge and a judgment for every individual in any dispute. Moreover, these rights are robustly protected and guaranteed at the European level by Article 47 of the Charter of Fundamental Rights of the European Union (CDFUE),[116] and Article 6 of the European Convention on Human Rights (ECHR).

Digital justice's distinctive characteristic is that it provides a novel venue for seeking judicial redress in cases of infringements of rights or interests. This is especially pertinent when digital tools, particularly those employing algorithmic decision-making, are used. The efficacy of judicial protection in such contexts depends on the fairness in designing and applying these technologies.

In this novel technological landscape, the practical and contemporary application of the constitutional principle of the defence of rights fundamentally depends on ensuring fairness in the design and use of these technologies. This requirement becomes particularly critical when an entity—whether public or private, natural or legal—utilises artificial intelligence to make decisions that could potentially interfere with an individual's rights or interests. In such cases, any individual subjected to these decisions must be afforded the right to be informed about whether a technological tool was employed in the decision-making process. This right to be informed also extends to situations where an individual interacts with a machine.

Andrea Simoncini introduces the concept of "transparency", which is defined in this context as the fairness of the algorithms used to elucidate the situation.[117] Within a constitutional framework grounded in the principles of legality and the protection of rights, it is imperative that any entity or individual utilising automated decision-making tools explicitly declares their intention to do so. He uses the expression "loyalty" as the bond that arises from this duty.[118] Such transparency is essential to ensure that all potentially affected parties are fully aware of the specific tool employed in the decision-making process. With the advent of generative AI, "transparency" assumes even greater importance, especially in scenarios where such technology is "delegated" — even informally — to make autonomous decisions.

Given that the judicial process is the

[114] See *infra multis* M. Ebers, *Automating due process- the promise and challenges of AI-based techniques in consumer online dispute resolution*, in *Frontiers in Civil Justice*, X. Kramer, J. Hoevenaars, B. Kas and E. Themeli (eds.), Cheltenham, Edward Elgar Publishing, 2022, 142; F. Pasquale, *Inalienable Due Process in an Age of AI: Limiting the Contractual Creep toward Automated Adjudication*, in *Constitutional Challenges in the Algorithmic Society*, H.-W. Micklitz, A. Reichman, A. Simoncini, G. Sartor, O. Pollicino and G. De Gregorio (eds.), Cambridge, Cambridge University Press, 2022, 42; J. Villasenor and V. Foggo, *Artificial Intelligence, Due Process, and Criminal Sentencing*, in *Michigan State Law Review*, vol. 2020, 2020; D.K. Citron, *Technological due process*, in *Wash. UL Rev.*, vol. 85, 2008, 1249.

[115] K. Terzidou, *The use of artificial intelligence in the judiciary and its compliance with the right to a fair trial*, in *Journal of Judicial Administration*, vol. 31, 2022, 154.

[116] M. Almada and M. Dymitruk, *Data protection and judicial automation*, in *Research Handbook on EU Data Protection Law*, E. Kosta, R. Leenes and I. Kamara (eds), London, Edward Elgar Publishing, 2022, 282.

[117] A. Simoncini, *La dimensione costituzionale della giustizia predittiva. Riflessioni su intelligenza artificiale e processo*, 406.

[118] This interpretation reminds the argument of W. Hartzog and N. Richards, *The Surprising Virtues of Data Loyalty*, in *Emory LJ*, vol. 71, 2021, 985.

"factor upon which the degree of the concrete realisation of rights depends, and the measure of their practical enforcement,"[119] any intervention that introduces technological elements impacting the protection of rights within the process must be evaluated following the "Due Process" clause.[120] The principle of technological transparency as a form of algorithmic fairness is intrinsically linked to the broader respect for the law. When addressing the legal community, the judge must thoroughly consider the arguments presented and strive to ensure that all community members accept and uphold the law.

The concept of "AI loyalty" has also been incorporated into the AI Act, particularly in cases where AI systems are said to "interact directly" with humans. Specifically, Article 50 mandates that such systems must be designed and developed to ensure natural persons are informed when interacting with an AI system. While these two principles are not identical, they can be understood as arising from a common underlying concern. As more decisions are being delegated to machines rather than made by humans, ensuring that this shift is perceived as fair is imperative.

Establishing whether a decision has been made through automated processes can help us determine how to engage with the machine freely and informally. Furthermore, the legal system's application of legal principles and rules for such acts is contingent upon the awareness that a decision has been made automatically.[121] For instance, Article 22 of the General Data Protection Regulation (GDPR),[122] which addresses automated decision-making, has recently been interpreted broadly by the Court of Justice of the European Union.[123]

As previously noted, it is easy to foresee a scenario where lawyers, consultants, and judges increasingly use these systems in the legal process.

However, more than digital allegiance is required. It is a fundamental right of every individual to be informed of the reasons underpinning a decision. This principle is derived from Article 24 of the Italian Constitution and the necessity of ensuring internal rationality in decision-making through a high degree of "readability" of the judicial function. This requirement is further supported by Article 111 of the Italian Constitution, which mandates that "all judicial measures must state the reasons on which they are based." The reasoning provided by a judge within the judicial system is essential for legitimising their authority and ensuring accountability to the parties, lawyers, and colleagues in accordance with the law.

While this principle generally applies to all public or private acts that exercise power—serving as a logical prerequisite for their reviewability and, consequently, conditioning the exercise of the constitutional right to defence—it applies with particular force to judicial acts. In such acts, the reasoning must always be accessible, especially for those decisions with decisive content.

Moreover, the statement of reasons constitutes the mechanism through which a judge establishes legitimacy by demonstrating adherence to the law within the broader judicial system rather than within an individual trial. This accountability is directed not only toward the parties and lawyers involved but also toward their judicial colleagues.

The obligation to provide a rationale should be understood as both a safeguard and a conduit for the judge's internal autonomy and as an essential element through which the exercise of discretion in identifying, interpreting, and applying the law must manifest. To judge well, therefore, is to articulate the rationale for exercising power within the boundaries established by the law, clearly indicating all factors that influenced the judge's final decision.

Those who employ decisional technologies with the potential to affect rights and freedoms are responsible for disclosing their use and ensuring that the language employed is accessible to humans. This is a prerequisite for assuring those rights and liberties, whether judicial or otherwise.

The convergence of these principles leads to a clear conclusion: only "explicable" or,

119 A. Simoncini, *La dimensione costituzionale della giustizia predittiva. Riflessioni su intelligenza artificiale e processo*, 408.
120 Also the "fair trial" of the ECHR and the EU Charter of Fundamental Rights.
121 M. Almada and M. Dymitruk, *Data protection and judicial automation*.
122 A. Simoncini, *Art. 22*, in *Codice della privacy e data protection*, R. D'Orazio, G. Finocchiaro, O. Pollicino and G. Resta (eds.), Milano, Giuffrè, 2021.
123 F. Palmiotto, *When Is a Decision Automated? A Taxonomy for a Fundamental Rights Analysis*, in *German Law Journal*, vol. 25, 2024.

better yet, comprehensible technologies should be admitted in decisions that culminate in judicial measures. Consequently, every judge must comprehend and agree with the machine's output.[124]

Undoubtedly, a well-articulated rationale allows for the review of an act's legitimacy, as well as its reasonableness, proportionality, and adequacy. However, within the constitutional framework, all judicial measures, particularly those with a decisional character, must convey an intelligible reason to be considered legally valid acts, irrespective of their potential appealability.

As we increasingly rely on AI systems for decision-making, there may be instances where the boundaries between permissible and impermissible become blurred due to the previously outlined structural features. There is the potential for creating zones where technological apparatuses, inherently incomprehensible to human reason, could replace or "capture" decision-makers. This would render exceedingly difficult, if not impossible, to verify the appropriateness of decisions, their legality, and, most critically, their legitimacy.

This principle of comprehensibility should not be confined to the exercise of "public" powers, where the need for transparency is particularly pressing (in addition to Article 97 of the It. Constitution, Article 41, paragraph 2, letter c), of the EU Charter of Fundamental Rights is also pertinent). Instead, this principle must be extended to private entities, especially those that utilise these algorithms to gain a significant advantage or, at the very least, create a profound asymmetry over individual private subjects or even other private legal entities. This principle is recognised in the GDPR (Articles 13 et seq. and Article 22) and, more recently, in the Digital Services Act (DSA), which explicitly mandates transparency and comprehensibility when decision-making algorithms, such as "recommendation algorithms," are employed, as articulated in Articles 15, 17, and 27.[125]

The interplay between transparency and justice gives rise not only to the principles of fairness and comprehensibility but also to several other significant considerations. In this context, another critical requirement emerges, which is enshrined in various forms within the constitutions of many States and in the Bills of Rights: the necessity for decisions to be made by human agents. The *Magna Carta Libertatum* may be cited as an early expression of this principle, but also the U.S. Constitution Art. III §1 vests the judicial power of the United States in its federal courts and their duly appointed judges.[126] In the interpretation of our judiciary, this guarantee has been articulated as non-exclusivity, which is understood as the prohibition against entrusting judicial decisions solely to automated decision-making systems. Such a practice would unduly constrain the interpretative freedom of judges, subject "only" to the law—a constitutional prerogative of the judicial function. Moreover, the knowability and intelligibility of algorithms should serve as safeguards for the autonomy and independence of the judiciary. These guarantees protect the judicial function, extending beyond the individual judge to encompass the judiciary as a collective entity.

The principle of internal functional independence precludes a judge in our system from being legally bound by precedent, with the only exception being the obligation to conform to the principles of law established by the Supreme Court for judges to whom the Supreme Court refers cases. The Italian Constitution subjects judges to the law, obligating them to adhere to a specific rule that contemplates the case. Additionally, judges must evaluate the law's implications regarding relationships, acts, and facts and comply with the legal consequences that the law establishes.

Consequently, constitutional provisions safeguard not only the individuals administering justice but also the judicial function itself, including the procedural forms and structural elements of the process. The traditional concepts of autonomy and independence thereby find their complete realisation in the characteristics that must define the judicial process and the structure of

[124] R. Crootof, *"Cyborg Justice" and the Risk of Technological-Legal Lock-in*, in *Colum. L. Rev. Online*, vol. 119, 2019, 233.

[125] It is important to distinguish the issue of the comprehensibility of decision-making algorithms from the closely related and widely debated issue of civil liability for damages caused by AI algorithms in private law. The comprehensibility of technology must be established as a prerequisite before determining liability, whether in civil or criminal contexts.

[126] M.R. Grossman, P.W. Grimm, D. Brown and M. Xu, *The GPTJudge: Justice in a Generative AI World*, in *Duke Law & Technology Review*, vol. 23, 2023, 33.

the process. This framework is essential for upholding the rule of law, which mandates, among other things, the absence of external influences not derived from the law, the maintenance of neutrality, the involvement of an impartial third party, and a procedural structure that ensures decisions are based on an objective assessment of the facts. Most importantly, it requires respect for fundamental rights through the conduct of a fair trial, characterised by an adversarial process between the parties.[127]

In the Italian legal system, the concept of case law interpretation has been established, largely due to the decisions of the Constitutional Court, as a methodology that transcends the mere logical-deductive application of legal postulates. Consequently, interpretation is recognised as a more complex intellectual process involving discretionary decisions and autonomous assessments by judges. It is not simply about the "discovery" of a text's meaning or identifying the most appropriate hermeneutic approach or a new and superior model of rationality; it also entails an "allocation of powers." Judicial interpretation, therefore, cannot be confined solely to the hermeneutic context but must be understood within the broader framework of the relationship between "words and institutions."

The aforementioned Italian constitutional provisions endow judicial bodies and proceedings with essential characteristics, without which there can be neither a judge nor the exercise of judicial function in its various forms. However, judges' roles and duties are fully articulated within the context of law. Additionally, it is important to recognise that the interpretation of legal texts is profoundly influenced by the jurisdiction's structure and the legal process's dynamics. Similarly, a judge's decision is shaped by many factors, including political, ideological, and legal cultural considerations, as well as procedural strategy.

Implementing decision-making technologies is likely to pose radical challenges for these principles. As previously noted, such technologies generate future predictions based on existing data processed through highly sophisticated "neural networks."

In the context of predictive justice primarily driven by statistical algorithms, the potential for a shift in judicial norms is constrained. Based on the data used to train the decision-making algorithms, the decisions rendered in the past can significantly influence the outcomes of new cases.

However, in this scenario, the judge would no longer be free to be subject only to the law, as stipulated by the Constitution, but would instead become bound by a series of previous decisions, thereby contradicting the principle enshrined in the second paragraph of Article 101 of the Italian Constitution. The precedent established by earlier judicial decisions would exert a substantial and immutable influence on future rulings.

In light of these considerations, our constitutional framework precludes the possibility of a judge being entirely supplanted in the decision-making process by an algorithmic system that relies on prior decisions. It is clear, however, that the situation differs when the output generated by such software is utilised in an auxiliary capacity, serving merely as one of the tools or sources of data available to the judge for discretionary decision-making.

The prohibition against any form of predetermination of judicial decisions by imposing constraints on the judge beyond adherence to the rules established by the legal system necessitates considering the additional power wielded by technology. The advent of new technological tools capable of autonomous decision-making has brought about profound transformations in the administration of justice and the act of judging itself. The digitisation process enhances the potential for predicting and correlating data, introduces novel forms of automation, and establishes new technical norms that may exert more significant influence than legal statutes.

The constitutional principles outlined above unequivocally preclude any form of "prescriptive" justice—that is, a system that dictates the content of a judicial decision to the judge.[128] The drafters of the CEPEJ Ethics

[127] J. Ulenaers, *The Impact of Artificial Intelligence on the Right to a Fair Trial: Towards a Robot Judge?*, in *Asian Journal of Law and Economics*, vol. 11, 2020, 1.

[128] A.F. da Silva and D. Mirante, *The Constitutional Limits of Digital Justice*, in *The IT Revolution and its Impact on State, Constitutionalism and Public Law*, M. Belov (ed.), London, Bloomsbury Publishing, 2021, 200.

Charter also addressed this concern.[129]

In this context, the constitutional parameter invoked is that of the reasonableness of the law, understood as the absence of arbitrariness, as delineated in Article 3 of the Italian Constitution. This principle allows us to refine the argument under consideration further. The principle that the judge is subject only to the law presupposes that the law itself is reasonable, meaning that it must not arbitrarily constrain the judicial function in its application to specific cases.

Artificial intelligence can support the justice field, not by creating a "robot judge" but by replicating certain cognitive skills required of judges and lawyers. This capability can help mitigate typical errors these professionals may encounter.

The final point derives from the collective provisions of Articles 3 and 111 of the Italian Constitution, which stipulates that "every trial shall be conducted before a third and impartial judge." The impartiality of the judge and their position as a neutral third party with respect to the parties involved are fundamental prerequisites for a fair trial. These requirements are essential to ensure that the process upholds the principles of equality and non-discrimination on the one hand and the independence of the judicial function on the other.

It is a notable development in the field of science and technology that algorithms have exhibited discriminatory tendencies in many instances, with greater frequency than might be expected.[130] This phenomenon is often attributed to a factor in technical terminology known as bias.[131]

Constitutional law prohibits using algorithms trained on discriminatory or incomplete data sets, as such algorithms are inherently incapable of producing unbiased judgments.

Notably, the AI Act mandates that developers or those intending to market high-risk AI systems, including those utilised in the justice sector, must ensure that such systems are trained, validated, and tested on datasets that are "relevant, representative, error-free, and complete." This requirement underscores the critical necessity for these systems to be free from discrimination.

These principles are particularly salient given the unprecedented potential of generative AI applications, for which numerous calls for caution and vigilance have been issued.

From this additional perspective, the role of technology is not merely to produce assessments, estimates, or indices to be utilised and weighed in the decision-making process but rather to generate texts, arguments, and writings — in essence, to compose the judgment itself. In doing so, technical systems would entirely supplant the traditional way judges communicate their decisions, thereby explicitly contravening the previously discussed constitutional stipulations.

Regardless of the future trajectory, it is imperative to elevate the level of scrutiny when considering integrating these tools into the judicial function. It is crucial to recognise that, in practice, the widespread use of such technologies in legal and quasi-legal professions, particularly for drafting case documents or legal counsel, is inevitable.[132]

Moreover, beyond the requisite adherence to the Italian Constitution and the laws, it is of paramount importance to consider each magistrate's ethical conscience and professional ethics. By affixing their signature to a judgment, the magistrate formally and definitively assumes responsibility for the contents expressed therein, even if, in reality, those contents were generated by an artificial intelligence system. Such ethical consciousness and professional ethics presuppose a high level of competence and education in using these tools.

Unfortunately, the "digital divide" persists, particularly in the context of cultural and educational disparities affecting our country. Given the potential risks associated with these technologies, all users must thoroughly understand their capabilities to prevent accidental or unintended harm. This necessity underscores the urgent need for more significant investment in digital literacy, a principle explicitly addressed in the AI Act

[129] CEPEJ, *European Ethical Charter on the use of artificial intelligence in judicial systems*.

[130] F.J. Zuiderveen Borgesius, *Strengthening legal protection against discrimination by algorithms and artificial intelligence*, in *The International Journal of Human Rights*, vol. 24, 2020, 1572.

[131] M.A. Malek, *Criminal courts' artificial intelligence: the way it reinforces bias and discrimination*, in *AI and Ethics*, vol. 2, 2022, 233.

[132] T. Sourdin, *Robo Justice: Constitutional Issues with Judge AI*, in *Ind. J. Global Legal Stud.*, vol. 30, 2023, 293.

through a dedicated article (Article 4).

The integration of humans and machines in determining the outcomes of practical actions represents a phenomenon far more extensive and profound than the concept of predictive justice alone. Whether this combination will lead to mutual aid remains uncertain. What is clear, however, is that humanity is rapidly approaching a critical juncture between a future characterised by cooperation and competition.[133] Therefore, it is essential to maintain a broad perspective and avoid an exclusive focus on specific, narrower aspects, which often typify the approach of jurists. This wider view is necessary to ensure that the constitutional transformations at stake are noticed.

Most of the principles and tenets indicated are now also part of the documents that we have indicated above, which complement the Charter of the CEPEJ by adding other concerns and legal guidance for using GenAI in judicial activity. These elements can be considered a specific and iterative risk assessment of AI tools before and during their employment in legal tasks. Of course, this evaluation cannot be a replacement for the constitutionality of these (laws allowing or permitting the use of these) tools.

Artificial intelligence can do wonderful things. But rights cannot exist in a world of hidden calculations. Just as with a lawyer or a doctor, we must have an artificial intelligence that acts in our self-interest. AI that is developed in these areas, as well as in other areas (I am thinking not only of justice but also of health or work), needs a "constitution" or, more precisely, a law that defines rules on how to use artificial intelligence that acts exclusively on behalf of humankind and therefore is trustworthy.[134]

7. *Conclusions*

The rapid advances in AI, exemplified by LLMs such as GPT, signal significant changes in law and the legal profession.[135] As these systems become more sophisticated, they are likely to reshape the way legal work is done and require a considered approach to ensure they are used responsibly and ethically.[136]

Although these AI tools have acquired unprecedented capabilities for understanding, analysing and creating draft legal documents compared to just a few years ago, legal practitioners must use them carefully and responsibly. These systems can sometimes inadvertently perpetuate bias from training data. Moreover, the inner workings of these systems are sometimes opaque, making it difficult to understand why they produced the results or the text they did.

The need for more transparency and interpretability can raise concerns about accountability and trust in legal decision-making.

Therefore, to ensure the appropriate use of LLMs in law, users must develop a basic understanding of their inner workings, capabilities, and limitations. However, while taking these concerns into account, if used appropriately, the remarkable advances in AI in recent years also hold out the hope of improving law practice and increasing access to justice.[137]

Are we at the dawn of a new form of justice?

If we look at the legal professions, we already have to recognise that the GenIA and the legal profession will integrate and develop, creating complementary-career forms within the legal professions;[138] at the same time, new legal professional groups will arise from non-legal professions such as legal technology companies.[139] These two realities

[133] J. Morison and T. McInerney, *When should a computer decide? Judicial decision-making in the age of automation, algorithms and generative artificial intelligence*, in S. Turenne and M. Moussa (eds.), *Research handbook on judging and the judiciary*, London, Edward Elgar Publishing, 2024.

[134] It must be remembered that the history of the AI Act started several years ago with the work of the High-Level Expert Group on Artificial Intelligence, *Ethics Guidelines for Trustworthy AI*, Brussels, 2019. See A. Renda, *Europe: toward a policy framework for trustworthy AI*, in *The Oxford Handbook of Ethics of AI*, M.D. Dubber, F. Pasquale and S. Das (eds.), Oxford, Oxford University Press, 2020, 6W. Hartzog and N. Richards, *The Surprising Virtues of Data Loyalty*.

[135] M.R. Grossman, P.W. Grimm, D. Brown and M. Xu, *The GPTJudge: Justice in a Generative AI World*, 20.

[136] J. Regalia, *From Briefs to Bytes: How Generative AI is Transforming Legal Writing and Practice*, in *Tulsa L. Rev.*, vol. 59, 2024, 193.

[137] E. Wright, *Professionals beware: The opportunities and risks of generative ai in legal practice*, in *Intellectual Property Forum: Jjournal of the Intellectual and Industrial Property Society of Australia and New Zealand*, vol. 135, 2024, 33.

[138] G. Gentile, *Human Law, Human Lawyers and the Emerging AI Faith*, LSE Public Policy Review. Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4918632.

[139] K. Frazier, *The Rise of the Interdisciplinary Lawyer: Defending the Rule of Law in the Age of AI*, in *USFL*

should be regulated through the law, the ethical rules of science and technology (the ethics of artificial intelligence) and the ethical rules of the legal professions. Within a few years, new ways of formulating regulatory provisions and applying rules, introducing new legal concepts, and testing the resilience of a long series of established principles will be necessary. We have, therefore, reached the point where the fundamental connection between the two fields is taken for granted without exploiting every aspect of it.

The possibility of using "machines" in law has definitively abandoned the utopian or dystopian dimension of science fiction to become the subject of ethical, legal, and political debate. In the near future, following the resolution of issues related to inaccuracies and biases, judges may transform their decision-making methods by utilizing advanced GenAI systems to simplify the composition of their opinions. When this happens, the expression "there is still a judge in Berlin" will acquire a fresh and restyled meaning.

Rev. F., vol. 58, 2023, 19.